Elm Based Detection of Micro-Calcification in Mammogram using Glcm Features

Jayesh George Melekoodappattu, Perumal Sankar Subbian

*Abstract***:** *The breast is made up of many different types of tissue and cells. When the cells in the breast grow or change abnormally and it is called breast cancer. Most breast cancers occur in women who are over the age of fifty. Calcification is the main causes of breast cancer. The deposit of calcium in breast tissue is known as calcification. And it is two types, Micro-calcification and macro-calcification. Large calcium deposits represent the macro-calcification which may relate to non-cancerous. The tiny white dots on mammogram represent the micro-calcification which is the earliest stage of breast cancer and the calcification can be found in different shapes. Mammography is the one of the method to determine the breast cancer. In this paper we are determining the micro- calcification in mammogram using different steps which include preprocessing, enhancement, feature extraction, feature selection and the classification.*

*Index Terms***:** *Extreme learning machine, Global Swarm Optimization ,Gray level co- occurrence matrix ,Mammography.*

I. INTRODUCTION

Each year in the world lots of women are learned that they have breast malignancy. The incidence of breast malignancy decreasing so as the death rate also decreases. The increased cure rate can be attributed to effective surgical and medical treatment also early find out of breast cancer plays an important role[1]. The size of the breast tumor is detected by annual mammogram is very small compared to the size of a cancer detected by first mammogram and the size of the breast cancer discovered by a physician examining a women's breast. The size of the breast cancer detected by the self examination is even larger. To understand breast cancer every women should have to know a little about normal breast tissue and it is shown in figure1.[2]

Female breast contains a number of pipes or ducts that lead from the area where milk is made. Under normal circumstances the inside of those ducts contains a single layer of cells that responsible for making milk. Cancer occurs when a cell divides over and over again in a process that runs out control. The fatty tissues is used to made the breast and the milk is produced in the lobules which is a small chamber can be seen in the fatty tissue[3,4]. The milk passes in tiny channels known as ducts to arrive at the nipples. Breast

cancer is developed inside this milk channel. non-invasive and invasive are the two types of breast malignancy. In non-invasive cancers the ducts fill with the cancer cells but there is no evidence of the cancer outside of the ducts. The medical term for this is Ductal carcinoma insitu [5].

A surgical strategy to remove the cancer is highly likely to result in a cure with invasive breast cancer cells have broken down the wall of the duct and are found in the fatty part of the breast.
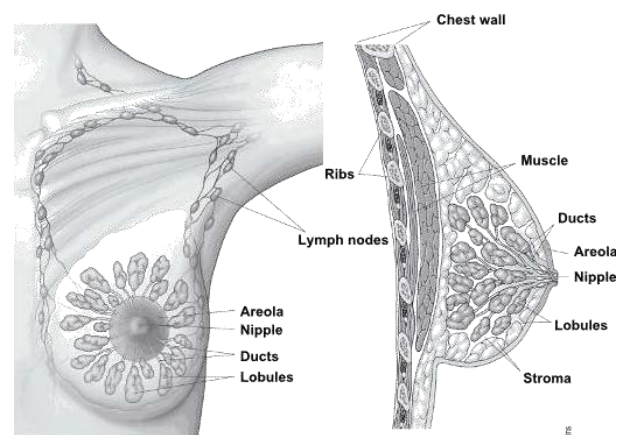

Fig. 1: Breast Tissue

cancer has access to the circulation of the lymph system within the breast. Cancer may travel to the lymph glands beneath the arm or enter the bloodstream and move into other parts of body. The strategy for invasive breast cancer involves surgery as well as a search for cancer cells in other parts of the body[6,7]. The cancer health care providers assigns stages. Breast cancer stages use a TNM classification system. T is the size of the primary malignancy, N notes how many lymph nodes are involved. M addresses whether any cancer is beyond the breast and lymph nodes . Stage 0 breast cancer is confined to the duct in the breast, it has not spread to the surrounding breast tissue and is unlikely to spread beyond the breast. Stages 1, 2 and 3 are based on the malignancy size and lymph node participation. The cancer which is spread to the other areas of the body is comes under the stage 4 category. The main signs of breast malignancy are a lump in the breast, a change in breast size or shape, nipple change and nipple discharge and skin alteration includes redness, inflammation and skin dimpling [8]. Most cancers occur by accident one cell divides into two cells and creates the only abnormal cell in the whole body this is called sporadic or somatic mutation. Every other cell in the breast is fine so the risk of reoccurrence once this cancer is treated is relatively low.

Retrieval Number: A3198058119/19©BEIESP
Journal Website: www.ijrte.org

1146

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Elm Based Detection of Micro-Calcification in Mammogram using Glcm Features

In some cases breast cancer runs in the family for example a patient's mother, aunt and grandmother were diagnosed with cancer of the breast. In those circumstances testing and genetic counseling can help to determine if a patient inherited a genetic mutation from one of her parents. Some clues to a genetic predisposition are cancer at a young unusual types of cancers or a clustering of cancers in the family. For example one women has breast cancer her mother has ovarian cancer and her uncle has prostate cancer these women receive genetic counseling while they are treated for breast cancer. Unfortunately a women who is diagnosed with breast cancer is at higher risk of developing a new breast cancer for the rest of her lifetime regardless of the family history[9,10]. The chance of developing a new breast cancer in the other breast is approximately a half a percent per year which means about one in 200 women a year will develop a new breast cancer. The main risk factors for breast cancer may include older age, dense breast and having breast cancer before, parent or sibling with breast cancer and having specific mutations on gene called as BRCA1 and BRCA2[11].

II. PREPROCESSING

The steps involved for detecting micro-calcification in mammogram is illustrated in figure 2. In the first step the input images are collected and preprocessed. The input images are obtained from MIAS(Mammographic Image Analysis Society) and DDSM(Digital Database for Screening Mammography). MIAS is developed by a UK research group to understand the mammography. It consist of total of 322 mammogram images from 161 patients and these mammogram are representing three types such as normal, benign and malignant. In that 322 images 208 mammogram images are represent normal, 63 images represents benign and 51 represent the malignant. The DDSM consist of around 2500 studies which contain images of each breast and have some information about the patients. DDSM has two parts that is cases and volume.

(PSNR) and low mean square error (MSE). And from the denoised mammogram the pectoral muscles are removed[12]. To remove the pectoral muscles from the mammogram images, First the image enhanced using the contrast enhancement adaptive histogram equalization method. Then change the gray picture into binary picture by using threshold method that is replace the input image with luminance greater than level with value 1(white) and replace all the other pixel with value 0. Then determined the connected components found in the binary image and compute the area of the connected components. After that remove all the small objects from the binary image and extract the ROI. Trace the exterior boundaries of the object to extract the pectoral muscles[13,14].

III. ENHANCEMENT

For the ROI enhancement Contrast Limited Adaptive Histogram Equalization(CLAHE) method is used. CLAHE is a method which is developed to overcome the limitations of adaptive histogram equalization that is the over amplification of the noise. A transformation function is derived by applying the contrast limiting procedure for each neighborhood[15]. Whereas histogram is used to determine how many pixels will be present in an image at a specific gray level. The algorithm is explained in the following steps

- Divide the every input picture into a number of non overlapping contextual areas of same size of the 8x8 blocks, each of them have neighborhood of 64 pixels.

- Calculate the intensity histogram of each contextual regions.

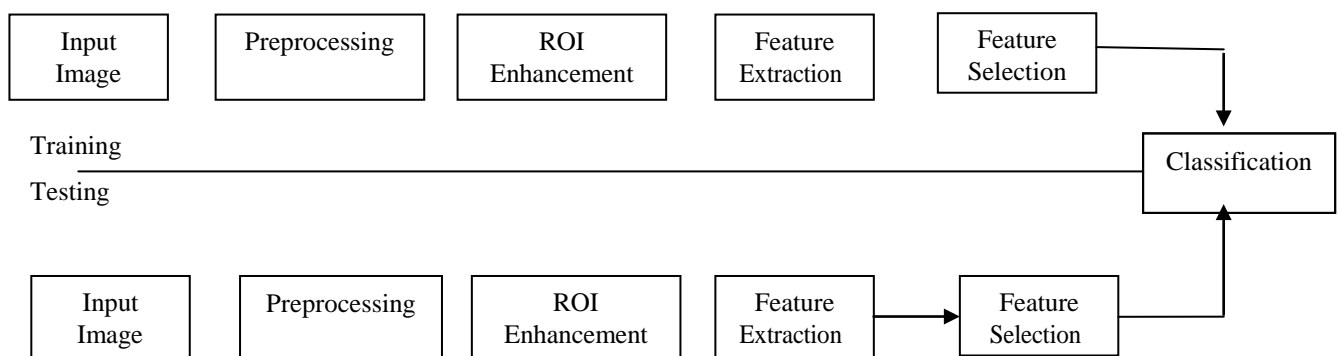- For clipping the histograms, set the clip limits (for example c=0.002).

Fig. 2: The steps involved for detecting micro calcification.

A case which represents the collection of a mammogram image and patient's information from a single mammogram exam of a patient .where as volume is the collection of cases for the distribution purpose.

After collecting the input images from the database next step is the preprocessing. In the preprocessing step, firstly the noises are removed from the image. In this paper we are using wiener filter to remove the noise. Wiener filter is the best one to denoise the images because it has high peak to signal ratio

- The threshold parameter for effectively altering the contrast of the image is decided by clip limit. Large clip limits improves the contrast of local picture areas. Thus it have to be set to minimum optimal rate.
- By selecting a transformation functions, modify each histogram

- Every histogram is changed in such a way that its height doesn't surpass the picked out clip limit

IV. FEATURE EXTRACTON

After the image enhancement next step is the feature extraction. In feature extraction we have a n dimensional features x1,x2…xn and map it to a lower dimensional space which is m dimensional and get the features z1,z2…zm where m is less than n[16,17]. And each of these features is sum function of the original feature set x1,x2...xn. In feature extraction taking the original features and mapping it to a lower dimensional space and each feature is obtained as a function of the feature set. That is it is a projection of a high dimensional feature space to less dimensional feature space[18].

In this paper we are using Gray Level Co-occurrence Matrix (GLCM) Features. It is used to extract the second order texture information. In the GLCM matrix the row and column number is equal to the distinct gray levels. he GLCMs are kept in a i x j x n matrix, in which n represent the count of GLCMs evaluated usually due to the orientation difference and displacements considered in the algorithm. The values i and j are using to reprent 'NumLevels' measure of the GLCM calculating function graycomatrix and which will generate a matrix. And the size of this GLCM matrix is the count of gray grade in the picture. The gray level scaling can be controlled by using the parameters Numlevels and GrayLimits Table 1 shows the different features of different images that are extracted by using GLCM

V. FEATURE SELECTION

From the GLCM features appropriate features are selected by using Global Swarm Optimization(GSO) method. It is a swarm intelligent optimization algorithm developed based on the behavior of glow worm. That is it has the ability to change the Lucifer in emission intensity.
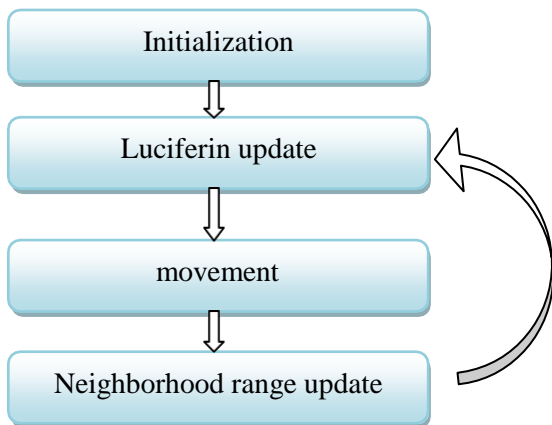

Fig. 3 : GSO flowchart

The glow interact with each other by using the luciferin. More glow worm will attracted to a glow worm which emit more light There is one circular range for each glow worm to determine the neighborhood. And the distance between two glow worm is determined by using the Euclidian distance. Figure 3 shows the GSO flowchart.

VI. CLASSIFICATION

For the classification Extreme Learning Machine(ELM) is used and it is a one layered feed forward neural network consist of input layer ,hidden layer and output layers. Figure 4 shows the ELM network structure it consist of p input layer neuron , q hidden layer neuron and r output layer neuron. Then firstly the ELM defines the number of hidden layer nodes and then randomly it will give the input neuron weight and hidden layer neuron bias. Then the hidden layer output matrix is calculated. Finally by using the moore-penrose generalization the output is calculated from the hidden neural layer output matrix.
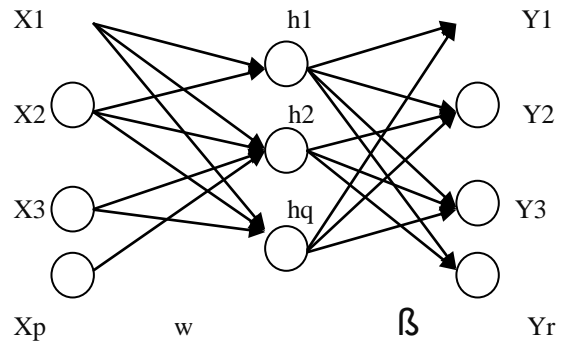

Fig. 4 : Network layout of ELM

If an ELM with a single hidden layer whose output function of the pth hidden node is represented by equation 1.

$$h_p(x) = G(a_p, b_p, x) \tag{1}$$

Where a_p and b_p is the p-th hidden node parameters. Then equation 2 is the output of the ELM used for SLFNs with L hidden node

$$f_L(x) = \sum_{p=1}^{L} \beta_p \ h_p(x) \tag{2}$$

Where β_p represents the output weight of p-th hidden node

$$h(x) = [G(h_p(x), \ldots \ldots, h_L(x))] \tag{3}$$

Equation 3 is the hidden layer output mapping . The hidden layer output matrix H For a given M training samples of ELM is given as in equation 4

$$H = \begin{bmatrix} h(x1) \\ \vdots \\ h(xM) \end{bmatrix} = \begin{bmatrix} G(a_1, b_1, x_1) & \ldots & G(a_L, b_L, x_L) \\ \vdots & \vdots & \vdots \\ G(a_1, b_1, x_M) & \ldots & G(a_L, b_L, x_M) \end{bmatrix} \tag{4}$$

$$T = \begin{bmatrix} t_1 \\ \vdots \\ t_M \end{bmatrix} \tag{5}$$

Here T is the training data target matrix.

ELM is also known as a regularization neural network with non tuned hidden layer which will minimize the term represented in equation 6

$$\|\beta\|_p^{\sigma_1} + C\|H\beta - T\|_q^{\sigma_2} \qquad (6)$$

$$\text{Where } \sigma_1 > 0 , \sigma_2 > 0 , p, q = 0, \tfrac{1}{2}, 1, 2, \dots .. +\infty$$

The different learning algorithm for the clustering, classification, feature learning and regression uses different combinations of σ_1, σ_2, p and q .

VII. RESULT AND DISCUSSION

In preprocessing the image is denoised by using a filter and to identify the good filter for removing the noise a comparison is performed. It is shown in table 1 .The comparison is based on the peak to signal ratio(PSNR), Signal to Noise Ratio(SNR), Mean Square Error(SNR) and the Similarity Index(SI).

MSE is the one of the method to examine the difference between true value and the value decided by a filter.

Table 1: comparison of different filters for different noises

$$MSE = \frac{1}{A \times B} \sum_{i=1}^{n} \sum_{j=1}^{n} (n_{i,j} - m_{i,j})^2 \qquad (7)$$

where A x B is the image size and $n_{i,j}, m_{i,j}$ are the initial image and the reconstructed image respectively.

PSNR is the possible value(maximum) and corrupting noise

Filter	Salt & pepper				Speckle noise				Gaussian noise				poisson noise			
	PSNR	SNR	MSE	SI	PSNR	SNR	MSE	SI	PSNR	SNR	MSE	SI	PSNR	SNR	MSE	SI
Wiener filter	27.85	19.13	106.6	0.91	32.03	22.77	40.69	0.94	23.71	15.29	277	0.47	38.99	29.61	12.97	0.96
median filter	16.67	7.95	1399	0.28	22.72	13.45	347.5	0.75	21.49	13.07	461	0.25	31	21.62	51.57	0.8
Gaussian Filter	24.38	15.66	237	0.89	30.24	20.98	61.44	0.83	29.06	20.64	80.6	0.88	36.5	27.13	19.16	0.93

value ratio of a signal that has dependency on the quality of its representation .

$$PSNR = 20 \log_{10}\left(\frac{MAX_{IM}}{\sqrt{MSE}}\right) \qquad (8)$$

Signal-to-noise ratio is a measure that describes how much noise is in the detected image, in relation to the original image.

$$SNR = 20 \log\left(\frac{Intensity\ of\ imge}{Intensity\ of\ noise}\right) \qquad (9)$$

Structural Similarity Index (SSIM) is a metric that measures picture quality degradation. Based on the comparison wiener filter is used to remove the noise from the image because it has high PSNR and low MSE for different noises like salt noise and pepper noise , speckle noise, Gaussian noise and Poisson noise . The de- noised image using the wiener filter is shown in figure 5 b

In the enhancement stage different method like Unsharp masking ,Histogram equalization and Contrast Limited

Adaptive Histogram Equalization is compared to determine the best method. The comparison is based on the correlation coefficient. Correlation measures the linear influence of grey stages of neighboring pixels. And it is shown in table 2. The unsharp masking method is mainly used for sharpening the image by utilizing the blurred version of the original image.and by subtracting this image from the original image to make the insharp mask. Where as histogram equalization is a straight forward method used to improve the global contrast of the image and it will improve the contrast that is the area having local contrast will become high contrasted area. Based on the comparison the Contrast limited adaptive histogram equalization is the best method for the enhancement of image because it has high correlation value than the other methods. The enhanced image using CLAHE is shown in figure 5c.

From the enhanced image the features are extracted by using the GLCM Method. The values of the features for different images is shown in table 3. There are sixteen features are extracted by using the GLCM method.

The images obtained from each steps is shown in figure 5. The figure 5 d) shown the detected micro-calcification in the mammogram image.

Table 2: comparison of different enhancement methods

Method	Correlation coefficient
Wiener filter (row from table above)	
Unsharp masking	0.991
Histogram	0.986
CLAHE	0.993

Table 3 : feature values for different images

Features	Image 1	image 2	image 3	image 4	image 5	image 6	image 7	image 8	image 9	image 10
Autocorrelation	7.1373	9.8032	11.3622	904891	10.319	6.4977	8.9275	7.9821	9.6661	16.9308
Contrast	0.0246	0.0418	0.0454	0.0468	0.0389	0.0328	0.0313	0.0251	0.0433	0.0581
Correlation	0.9962	0.9952	0.9958	0.9941	0.9955	0.9947	0.9964	0.9967	0.9946	0.9955
Cluster prominence	703.73	963.81	963.25	585.99	688.85	723.23	812.55	869.66	649.205	1.07E+03
Cluster Shade	76.648	94.144	122.24	62.125	70.674	78.517	104.3407	93.3706	65.7877	67.632
Dissimilarity	0.0199	0.0319	0.0289	0.0303	0.0312	0.0201	0.0192	0.0196	0.0255	0.0355
Energy	0.5367	0.4417	0.0794	0.4438	0.4055	0.6074	0.5471	0.5817	0.4275	0.2997
Entropy	1.1536	1.4159	1.3162	1.3259	1.4463	1.0001	1.116	1.0219	1.3679	1.7086
Homogeneity	0.9905	0.985	0.9874	0.9863	0.9851	0.9913	0.9916	0.9908	0.9889	0.9843
Maximum probability	0.7243	0.6515	0.6807	0.6441	0.6169	0.7742	0.7307	0.7578	0.6336	0.5087
Sum of squares: Variance	7.0882	9.7515	11.308	9.4391	10.262	6.4563	8.8762	7.9313	9.6136	16.8591
Sum average	3.9479	4.6592	4.8876	4.7139	4.9045	3.7104	4.2999	4.0674	4.7635	6.4618
Sum variance	20.891	28.260	34.5712	27.455	29.392	19.692	27.4845	24.779	27.6936	48.9015
Sum entropy	1.1366	1.3859	1.2868	1.2973	1.4214	0.9835	1.1011	1.0065	1.3464	1.679
Difference variance	0.0246	0.0418	0.0454	0.0468	0.0389	0.0328	0.0313	0.0251	0.0433	0.0581
Difference entropy	0.096	0.1393	0.125	0.1297	0.1372	0.0909	0.0886	0.0942	0.1111	0.146

Fig. 5: (a) Input image ,(b) De-noised image ,(c)
Enhanced Image, (d) calcification detected Image

Retrieval Number: A3198058119/19©BEIESP
Journal Website: www.ijrte.org

1150

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

VIII. CONCLUSION

Breast malignancy is the one of the most commonly seen cancers in women. A women born in today has about one in eight chance of being diagnosed with breast cancer at sometime during her life. Early diagnosis is the one of the solution to increase the survival rate and mammography is the best available screening method. In this paper we used the image processing tools for mammography image to enhance contrast to improve visualization ,identify micro-calcification clusters to highlight suspicious areas and improve diagnosis. In the preprocessing we compared different filters for removing the noises and determined that wiener filter is the best one. Also a comparison performed to determine the best enhancement method and find out that CLAHE is the best method for enhancing the image. Sixteen features are drawn out from the image by using GLCM method and finally the micro-calcification is determined by using ELM.

REFERENCES

1. Y.C. Cheung, Y.C. Lin, Y.L. Wan, et al., Diagnostic performance of dual-energy contrast-enhanced subtracted mammography in dense breasts compared to mammography alone: interobserver blind-reading analysis, Eur. Radiol. 24 (10) (2014) 2394–2403.
2. Buciu, I. , & Gacsadi, A. (2011). Directional features for automatic tumor classification of mammogram images. Biomedical Signal Processing and Control, 6 (4), 370–378 .
3. Huang, G.-B. , Zhu, Q.-Y. , & Siew, C.-K. (2004). Extreme learning machine: A new learning scheme of feedforward neural networks. In Proceeding of the IEEE inter- national joint conference on neural networks: 2 (pp. 985–990) .
4. R. Maria, D. Matteo, and C. Beniamino, A Review: Health Care CAD Systems for Breast Microcalcification Cluster Detection.Journal of Medical and Biological Engineering, pp. 147156, 2012.
5. O. Arnau, T. Albert, L. Xavier, and T. Meritxell, T. Lidia, S. Melcior, F. Jordi and Z. Reyer , Automatic microcalcification and cluster detection for digital and digitised mammograms. Elsevier, vol.28, pp. 6875, April 2012.
6. Melekoodappattu, J.G. & Subbian, P. J Med Syst (2019) 43: 183. https://doi.org/10.1007/s10916-019-1316-3.
7. B.B Simon, V. Thomas, J. Kumar," Algorithm for the Detection of Microcalcification in Mammogram on an Embedded Platform," International Journal of Scientific and Engineering Research,vol. 4, no.4, pp. 1167–1175, 2013.
8. Rabidas, R., Midya, A., & Chakraborty, J. (2017). Neighborhood structural similar- ity mapping for the classification of masses in mammograms. IEEE Journal of Biomedical and Health Informatics . doi: 10.1109/JBHI.2017.2715021
9. Ramos, R. P. , do Nascimento, M. Z. , & Pereira, D. C. (2012). Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms. Expert Systems with Applications, 39 (12), 11036–11047
10. Rocha, S. V. , Junior, G. B. , Silva, A. C. , de Paiva, A. C. , & Gattass, M. (2016). Tex- ture analysis of masses malignant in mammograms images using a combined approach of diversity index and local binary patterns distribution. Expert Sys- tems with Applications, 66 , 7–19 .
11. Rouhi, R. , Jafari, M. , Kasaei, S. , & Keshavarzian, P. (2015). Benign and malignant breast tumors classification based on region growing and cnn segmentation. Ex- pert Systems with Applications, 42 (3), 990–1002 .
12. Sahiner, B. S. , Chan, H. P. , Petrick, N. , Helvie, M. A. , & Goodsitt, M. M. (1998). Computerized characterization of masses on mammograms: The rubber band straightening transform and texture analysis. Medical Physics, 25 (4), 516–526 .
13. E. Malar, A. Kandaswamy ,D .Chakravarthy, A .Giri Dharan "A novel approach for detection and classification of mammographic micro-calcification using wavelet analysis and extreme learning machine" Computers in Biology and Medicine 42 (2012) 898–905
14. Bhupendra Singh, Amit Verma , R. C. Tripathi "Contrast Enhancement and Micro- Calcification Detection using Statistical and Wavelet Features in Digital Mammograms" 2017 Fourth International Conference on Image Information Processing (ICIIP)
15. J. Haukka, G. Byrnes, M. Boniol, and P. Autier, "Trends in breast cancer mortality in sweden before and after implementation of mammography screening," PLoS One, vol. 6, no. 9, p. e22422, 2011.
16. M. Jiang, S. Zhang, H. Li, and D. N. Metaxas, "Computer-aided diagnosis of mammographic masses using scalable image retrieval," IEEE Transactions on Biomedical Engineering, vol. 62, no. 2, pp. 783– 792, 2015.
17. C.Abiram, R.Harikumarand, S.R.Sannasi Chakravarthy "Performance Analysis and Detection of Micro calcification in Digital Mammograms using Wavelet Features" IEEE WiSPNET 2016 conference.
18. F. L. S. Nunes, H. Schiabel, et al., "Contrast enhancement in dense breast images to aid clustered micro-calcifications detection," J. Digit. Imaging, vol. 1, no. 20, pp. 53–66, 2007.

AUTHORS PROFILE

Jayesh George Melekoodappattu is Assistant Professor in Department of Electronics and Communication, Vimal Jyothi Engineering College, Kannur, Kerala, India. He has 6 years of experience processing. He is a graduate and post graduate of Anna University, Chennai. He is pursuing PhD from APJ Abdul Kalam Technological University, Kerala. He has published more than 15 publications in various national, international journals and conferences in image processing biomedical image and signal and other related fields.in teaching and his area of specialization is

Dr. Perumal Sankar Subbian, received PhD degree in Communication Engineering from Anna University, Chennai. He has been a professor with the department of Electronics and Communication Engineering in various engineering colleges in India. Now he is working in Department of Electronics and Communication, TocH Institute of Science and Technology, Ernakulam, Kerala, India. His current research interests include Image Processing, Classification Signal, Image and Video Processing, Feature Extraction, Optics and Photonics, Optics, Image Data Analysis.