Automated Assistance for Breast Cancer Identification on Mammograms Using Computer Vision Algorithms

K. Nagaiah[a], Dr.K. Manjunathachari[b], Dr.T.V. Rajinikanth[c]

[a]Assistant Professor ECE, ICFAI University Raipur, Raipur CG, India. E-mail: nagaiah.k@iuraipur.edu.in
[b]Prof & Head ECE, GITAM University, Hydarabad, India. E-mail: manjunath4005@rediffmail.com
[c]Prof & Dean CSE, R&D, Sreenidhi Institute of Science and Technology, Hyderabad, TG, India.
E-mail: rajinitv@gmail.com

Abstract: One of the greatest health problems in the world is breast cancer. If these breast cancer abnormalities are identified early, there is a maximum chance of recovery. We can go for this early prediction. It is one of the most effective detection and screening strategies and is widely used. The basic goal of CAD systems is to support physicians in the process of diagnosis. CAD systems, however, are very expensive. Our emphasis is on developing a CAD system that is low-cost and effective. To categorize breast cancer as either benign or malignant, a computer-aided detection approach is suggested. The standard mammogram image corpus, Digital Database used for Screening Mammography, images are used for enhancement, segmented and GLCM, intensity and histogram methods are used to extract features. The work is carried out by effective multilayer perceptron classifier (MLP) and support vector machine (SVM). Compare the performance of the classifiers. The proposed approach achieved 96 % accuracy and 8% improvement in accuracy compared to previous approaches with same dataset [4].

1. Introduction

Breast cancer detection is a more focused research area now. Early detection of breast cancer is very important to save women life. Breast cancer is the world second disease of women death rate. Breast cancer can affect men also But 100 time common in women. According to American cancer society survey report mentioned in 2017 men also diagnosed 2470 new cases of invasive breast cancer. According to national cancer institute one out of 8 women has a chance of being diagnosed by breast cancer. Women who have blood relatives like sisters brothers they have high risk factor [1-5].

This paper therefore proposes an integrated computer-aided diagnostic device that helps radiologists diagnose and identify digital mammograms for breast cancer. Section II explains about the literature review and existing methods. Section III gives proposed method, section IV explains results and analysis and section V explains conclusion.

2. Related Work

Assessed affectability of the radiologist in screening for breast cancer. It is only around 70%, so execution will be improved. IN the off chance that the conceivable region was incited to them of the norm's variants. CAD System of breast malignancy they can provide such assistance and are necessary and valuable for that controlling breast development. Mammography provides a philosophy to assist radiologists to identify the majority of mammogram photographs. And to arrange them as rare or ordinary [6]. Likewise, it should differentiate the growth of cells that are small effectively. Multi-day classifiers currently presume a noteworthy therapeutic conclusion role. The Area of Interest (ROI) was viewed as the improved sharp Edge or mammogram picture boundaries. Segmentation is now segmentation, using common statistical-driven morphology based on ROI Approximations. Three key steps were performed in this work. First Method the unnecessary marks and labels in the picture should be removed [9 – 13]. Then one, then Segmentation based on strength is performed to remove pectoral muscles. The transformation of Hough after pectoral muscle segmentation is for extracting functions, done on ROI. This is an important way for a trend to be remembered. It is a method of image transformation in which it is possible to obtain a particular image with a specific shape within an image. It also turns an original image into a 2D function. The extraction of features plays an important part in classification. The function extracted from the transform of Hough is used to explain them as being natural or unusual. Multiple strategies like Bricial Neural networks, LDA-Linear analysis of discriminant, methods of closest neighbor were used. SVM classifiers are used in this work to identify the data obtained from the extraction of features [7-15].

Mammogram picture ROI is classified into the highest possible. Amount of small squared shape area non-overlapping of fixed region size for further research to obtain a broad dataset. A traditional one in general, the mammogram classification scheme consists of four Sequential steps: (1) Enhancement of image (2) Segmentation, (3) extraction of characteristics, and (4) classification of mammograms.
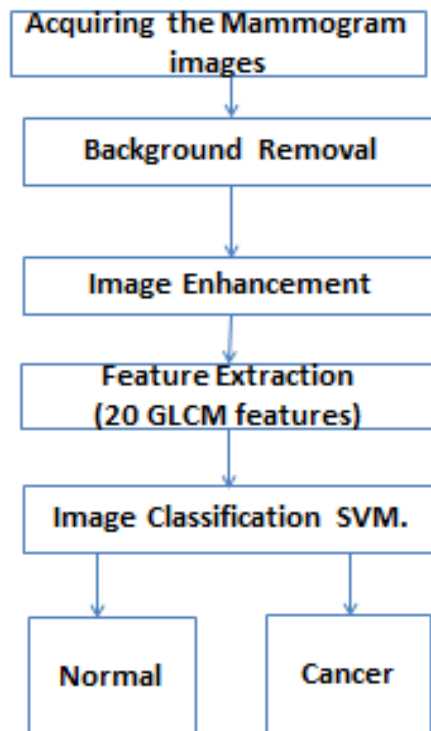
Existing Approach-I
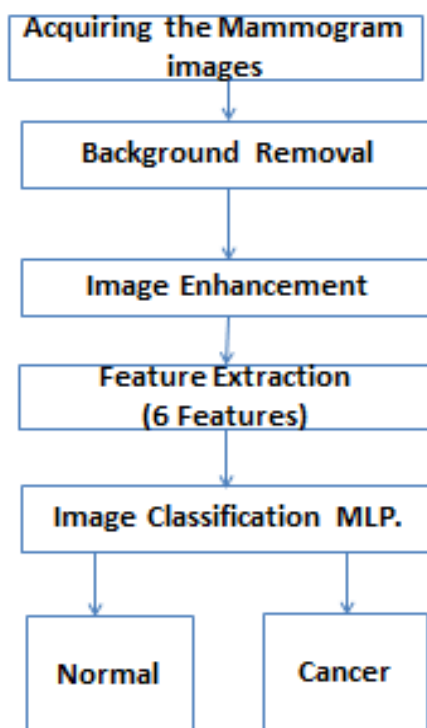

Fig. 1. SVM Classifier

Existing Approach-II

Fig. 2. MLP Classifier

3. Proposed Methodology

In the medical area, improving image quality in computing is the use of computers to explain images [1]. Noise reduction, point rise, and contrast enhancement comprise the forms of image quality improvements. Quality enhancement may be used to restore a bad image or to boost those image characteristics [2]. Mammography is the testing of the breast gland that is used to detect early breast cancer by using X-rays. The goal of digital mammography is to apply digital device techniques to digital mammograms. There is a capacity for automated systems to detect breast cancer. In order to generate high-resolution digital mammograms without losing information from original mammograms, 12 bit resolution detection is typically needed. Due to the low illumination and high noise in the image that can exceed 10-15 percent of the full pixel intensity, mammography analysis is a challenge. The most difficult images to analyze and interpret are mammograms [2].

This proposed method:
1. Getting the images from DDSM. 2 Improved image quality, at this stage the image quality is improved so that it is easy to process mammography images. 3 Image segmentation, the segmentation stage aims to divide the image into segments so that the form is suitable for processing. 4 Extraction of features, this is done to get the features that exist in the image. 5 Classification, after getting information from the image, existing images are classified into two, whether normal or cancerous.
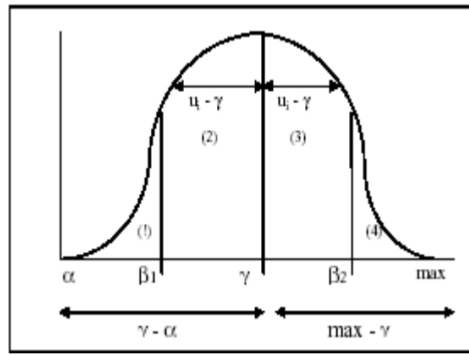

Fig. 4. Fuzzy based Image Enhancement

We are using Gaussian function to find out the membership function. We are using five points minimum equal to ∞, maximum equal to max, mean equal to (min + max)/2, b1 equal to (mean + ∞)/2, b2 equal to (max + ∞)/2. We will get all values. Based on the fuzzy rules we will get enhanced image. The enhanced image is further processed for segmentation. After segmentation we are going to find the 6 features. Those features contrast, standard deviation, mean, kurtosis, variance and smoothness. These features are given to the two classifiers that is SVM and MLP [10-15].
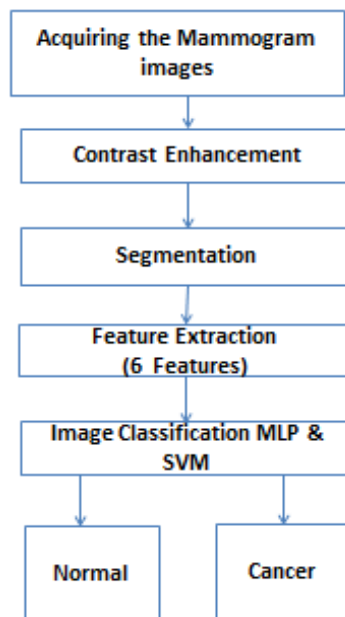

Fig. 3. MLP & SVM Classifier

Feature Extraction

Algorithm of GLCM (Gray Level Co-Occurrence Matrix) is a two-dimensional matrix representing gray level variations in the image. The GLCMM measurement is a measure of a sort of correlation, energy, contrast, entropy, homogeinity.

1. Mean:

$$\mu = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} p(i,j)$$

2. Variant Equations:

$$\text{Varian} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (p(i,j) - \mu)^2$$

3. Standard Deviation Equation:

$$\sigma = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (p(i,j) - \mu)^2}$$

4. Contrast Equations:

$$\text{Contrast} = \sum_{i,j=0}^{G-1} (i-j)^2 P(i,j)$$
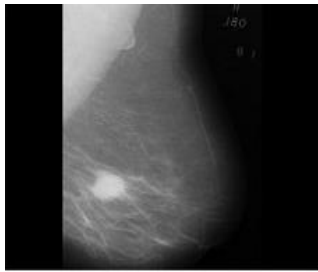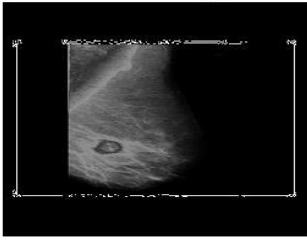
5. Kurtosis equation:

$$K = \left\{ \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left[\frac{p(i,j) - \mu}{\sigma} \right]^4 \right\} - 3$$

6. Smoothness Equation:

$$R = 1 - 1/1 + (\text{variance})^2$$

4. Experimental Results and Discussions

Image Enhancement Output

IMAGES	FBS	
	Original Images	Enhanced Images
Mdb 028		

Mdb 105		Enhanced Image
Mdb 155		Enhanced Image
Mdb 158		Enhanced Image
Mdb 170		Enhanced Image
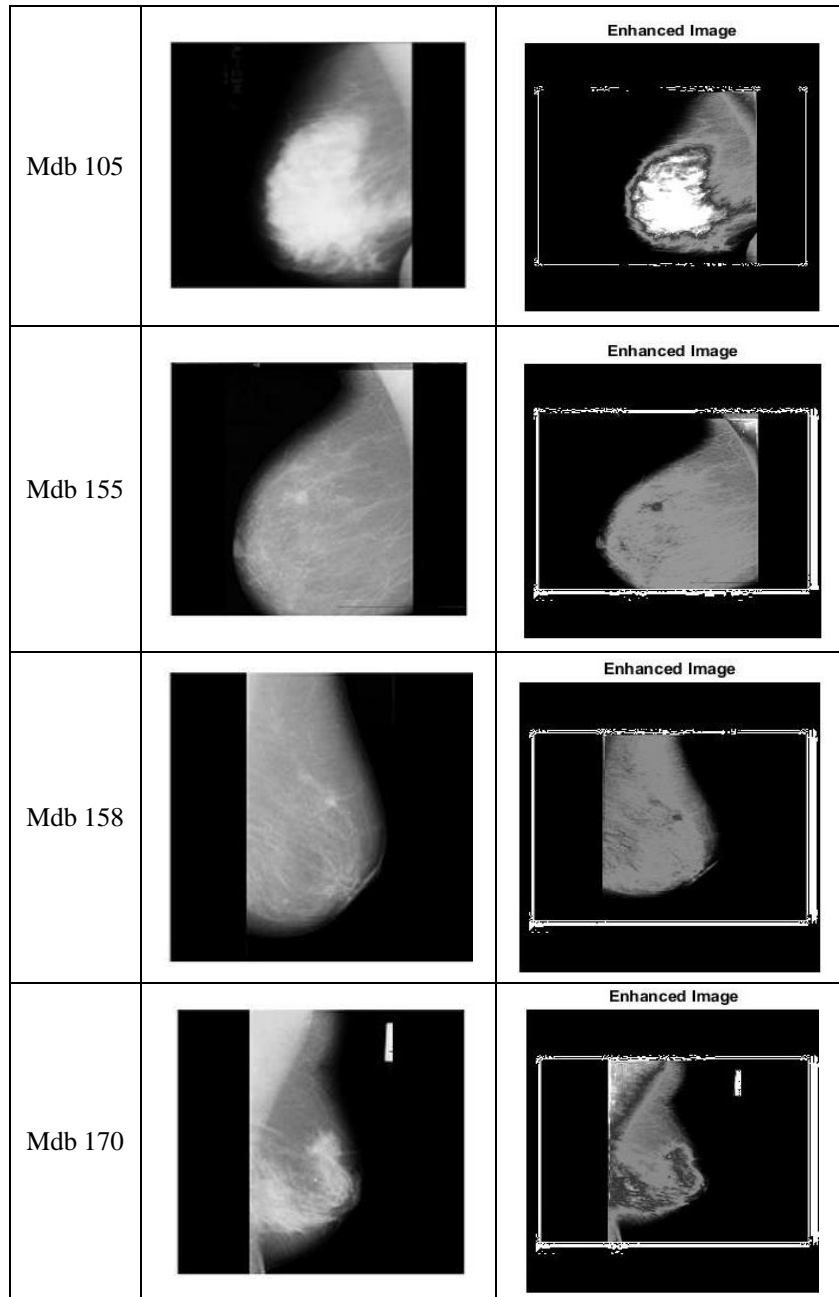
Table 1. Sample Features of Cancer Images

Features	Cancer 1	Cancer 2	Cancer 3	Cancer 4	Cancer 5
Contrast	0.9621	1.3375	0.7068	0.6331	0.4965
Variance	0.3954	0.4793	0.3686	0.9089	0.4213
Std dev	0.6288	0.6923	0.6071	0.9534	0.649
Kurtosis	9.4513	9.4692	7.0472	3.0506	5.0567
Mean	0.5397	0.5926	0.5338	0.9692	0.4942
Smoothness	0.2834	0.324	0.2693	0.4761	0.2964

Table 2. Sample Features of Normal Images

Features	Normal 1	Normal 2	Normal 3	Normal 4	Normal 5
Contrast	0.2982	0.6969	5.293	0.6274	0.5679
Variance	0.0517	0.109	0.1575	0.072	0.0737
Std dev	0.2274	0.3301	0.3968	0.2683	0.2715
Kurtosis	10.4964	26.0741	22.625	10.4646	6.4397
Mean	0.218	0.3033	0.3693	0.2622	0.2559
Smoothness	0.0491	0.0983	0.136	0.0671	0.0687
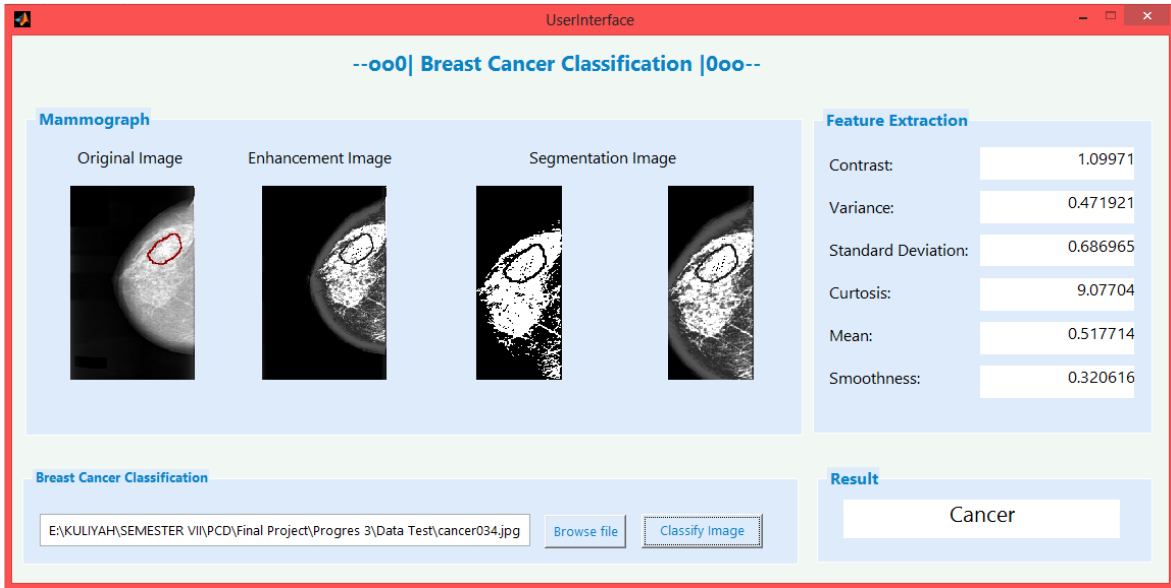
Fig. 5. Classifier Output Cancer
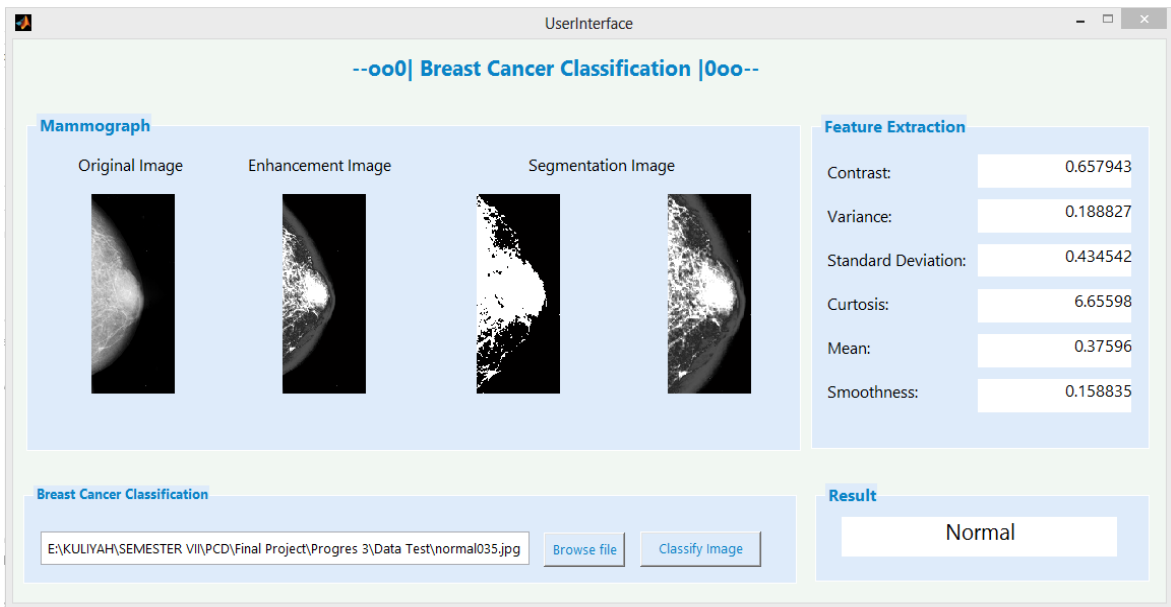
Fig. 6. Classifier Output Normal

Performance Analysis

The metric are calculated based on the above figure below formulas
Negative Rate Matric (NRM): (FP Rate + FN Rate)/2.

All Confusion Matrix
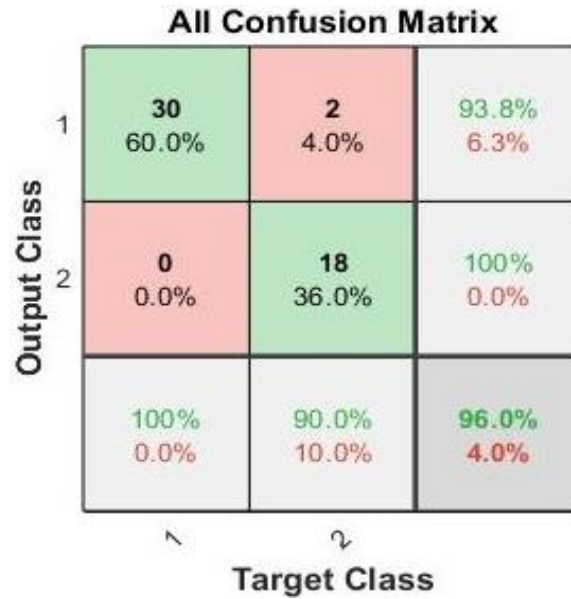
In this work we have tested two different databases.

1. MIAS DATABASE it contains 322 image.
2. DDSM data base for this experimental analysis we have used 120 images.
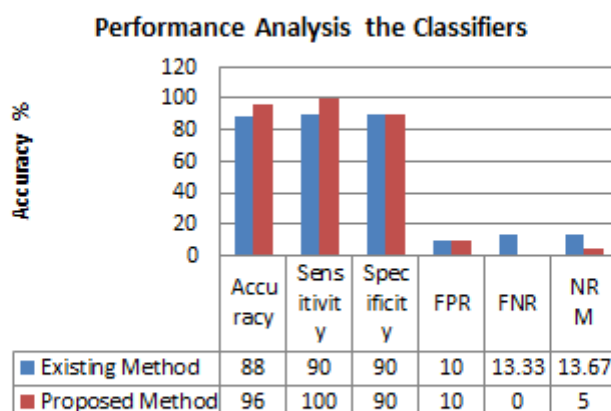3. Real images with the help of Dr. Suresh Kumar Radiologist.

Method	No of Images	Normal Images	Cancer Images
MLP	50	20	30
SVM	50	20	30

Existing Methods			
TP	FP	TN	FN
26	2	18	4
22	4	16	8

Proposed Method			
TP	FP	TN	FN
30	2	18	0
29	1	18	2

Existing Methods						
Method	SN	SP	ACCURACY	FP Rate	FN Rate	NRM
MLP	86.6	90	88	10	13.33	11.67
SVM	73.3	80	76	20	26.67	23.35

Proposed Method						
Method	SN	SP	ACCURACY	FP Rate	FN Rate	NRM
MLP	100	90	96	10	0	5
SVM	93.54	94.73	94	5.2	6.4	5.8

Performance Analysis the Classifiers

	Accuracy	Sensitivity	Specificity	FPR	FNR	NRM
Existing Method	88	90	90	10	13.33	13.67
Proposed Method	96	100	90	10	0	5

5. Conclusion and Future Scope

Breast cancer classification is a critical stage for the success of a machine has helped diagnose breast cancer. This decreases the by lowering the excessive biopsy and the false positive rate the cost of hospitals as well.

Various classifiers have been used in Application of biomedical imaging, such as breast cancer diagnosis using mammograms. ANN demonstrates really fine in medical diagnostic systems, performance. In this paper. The standard of the proposed algorithm in terms of performance and adaptability, structure is assessed. On 100 images containing normal and cancer, it was evaluated. The accuracy is increased 88% to 96% that is 8% and negative rate of the metrics is reduced from 11.67 to 5 that is 6.67. The proposed methodology works well.

References

1. Eason, G., Noble, B., & Sneddon, I.N. (2013). *On certain integrals of [1] Survey by Indian cancer society,* Indian Cancer Society.
2. Padmanabhan, S., & Sundararajan, R. (2012). Enhanced accuracy of breast cancer detection in digital mammograms using wavelet analysis. In 2012 *International Conference on Machine Vision and Image Processing (MVIP),* 153-156.
3. Spandana, P., & Rao, K.M. (2013). Novel image processing techniques for early detection of breast cancer, mat lab and lab view implementation. In 2013 *IEEE Point-of-Care Healthcare Technologies (PHT),* 16-18.
4. Kumari, V., & Rani, K.U. (2016). Mammogram classification using multilayer perceptron and support vector machine. *International Journal of Computational Science, Mathematics and Engineering IJCSME, 3*(1), 237-244.
5. Karssemeijer, N. (1998). Automated classification of parenchymal patterns in mammograms. *Physics in medicine & biology, 43*(2), 365-389.
6. Petroudi, S., Kadir, T., & Brady, M. (2003). Automatic classification of mammographic parenchymal patterns: A statistical approach. In *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (IEEE Cat. No. 03CH37439), *1,* 798-801.
7. Oliver, A., Freixenet, J., & Zwiggelaar, R. (2005). Automatic Classification of Breast Density, in: In *Proceedings of the IEEE International Conference on Image Processing, ICIP,* 1258–1261.
8. Bovis, K., & Singh, S. (2002). Classification of mammographic breast density using a combined classifier paradigm. In 4th *international workshop on digital mammography,* 177-180.
9. Muštra, M., Grgić, M., & Delač, K. (2012). Breast density classification using multiple feature selection. *Automatika, 53*(4), 362-372.

10. Padmavathy, T.V., Vimalkumar, M.N., Nagarajan, S., Babu, G.C., & Parthasarathy, P. (2018). Performance analysis of pre-cancerous mammographic image enhancement feature using non-subsampled shearlet transform. *Multimedia Tools and Applications,* 1-16.

11. Chen, Z., Denton, E., & Zwiggelaar, R. (2011). Local feature based mammographic tissue pattern modelling and breast density classification. In 2011 4th *International Conference on Biomedical Engineering and Informatics (BMEI), 1,* 351-355.

12. Qu, Y., Shang, C., Wu, W., & Shen, Q. (2011). Evolutionary Fuzzy Extreme Learning Machine for Mammographic Risk Analysis. *International Journal of Fuzzy Systems, 13*(4), 282–291.

13. Parthasarathy, P., & Vivekanandan, S. (2018). Urate crystal deposition, prevention and various diagnosis techniques of GOUT arthritis disease: a comprehensive review. *Health information science and systems, 6*(1), 1-13.