TuijinJishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol.44 No. 6 (2023)

Enhanced Multi-Resolution Feature Extraction for the Early Detection of Breast Cancer

B.Sathiyaprasad[1,*], D.Menaka[2], R.Asha[3], M.Gowri[4]

[1-4]*Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Rajivee Gandhi Salai, Jeppiaar Nagar, Chennai- 600119, Tamil Nadu. India.*

Abstract:In recent years,breast cancer has been one among leading cause for cancer deaths among the female population in the world. Digital mammograms are the most common reliable tool for the elimination and detection of breast cancer. Mammography is an important tool for visualization and detects breast cancer using low-level x-rays. X-ray images of the breast must be accurately evaluated to fetching the earliest stages of cancer growth. In this paper, we proposed a Multi Resolution based Random Forest (MRRF) Technique by prior detection of Mammogram Image Based on K-means Cluster as well as Random Forest Classification. The proposed approach is depending on a multi resolution basedfeature extraction which can be used to extract features at multi-level of decomposed images. The four stages of the proposed strategy include preprocessing to redesign an image, K-means clustering for segmentation, for the feature extraction, discrete wavelet transform as well as the random-forest classification techniques are used for prediction of regular or malignant image.The proposed methods have been implemented in python and the results show that our proposed work is possible and effective for an accurate classification as well as identification of breast cancer.

Keyword:*Mammograms, detection, classification, k-means, ROI, malignant, GLCM, DWT, RF.*

1. Introduction

Breast malignant growth is the most averagefemaledisease worldwide representing almost a quarter (25%) of all tumors with an expected 1.67 million new disease cases perceived in 2012. It is additionally the most lethal kind of canceramong women. Breast cancer has positioned number one disease among Indian. The occurrence of breast cancer increments with age and this is legitimate in India like rest of the world. Except for 5-10% breast malignant growths where the essential hazard factor is hereditary inclination, in the last 90% of irregular breast tumors, the perceived hazard factors are both conceptive, lifestyle or ecological variables, explicitly through their affect the hormonal milieu.

Breast cancerprojection for India all through year 2020 proposes the number to go as unnecessary as 1797900. Reports from the Indian Council for Medical Research In India, there are 1.5 lakh new instances of bosom disease annually, with 70,000 deaths.

Breast cancer, has become the commonest form of cancer in developing countries. It is characterized as a heterogeneous disease, including in its different forms and different stages of cancer. According to the World Health Organization (WHO) [2], it is found to occupy the second position in causing cancer deaths in women worldwide. Though it is a disease of females, a 1% chance of occurrence in men has been reported. It is predominantly found to affect women 45-55 years of age in the post-menopausal phase. The causative factors for breast cancer are familial history, genetic risk factors and lifestyle risk factors such as obesity, physical inactivity, diet and environmental factors [3]. The percentile division constitutes 10% due to genetic factors. Approximately 75% belongs to the Estrogen receptor (ER) expression as an increased concentration of estrogens in circulation promotes tumorigenesis.

Human cells will be isolated in a methodical manner. Every new cell will be regenerated after the existing cells worn out or die. The limitless growth of new or existing cells leads to Cancer. This process of continuous growth affects normal cell structures and results in increased malignancy. This causes health issues in the human body, very specifically to a concern part or organs, where cancer originated [4].

Cancer cells are vulnerable to spread across other portions of the human physiological system. For instance, malignant cells in the lungs can easily travel to the skeletal system and grow vigorously. Spreading of cancer cells is called Metastasis and some spread fast. The treatment also varies from stage to stage and the response to the treatment is the major concern towards iterative curing procedure of cancer diseases. Few types of malignance are treated with surgery and drugs called chemotherapy [5].

The philosophy that drives breast cancer screening is "Early detection is the greatest 23 protection". The need for the enhanced Computer-Aided Diagnosis (CAD) has been used over the years for early identification and evaluation of breast cancer. Cancer Images Processing Method has motivated the researcher to develop ICT aided mechanism in theprior prediction of malignant cancer.

The aim is to decrease the death rate resulting from breast cancer by enhancing the precision of early breast cancer diagnosis and attaining an efficient Remote Diagnosis and Decision Support System using enhanced Machine Learning categorization [6] Mechanism. This publication aims to identify novel feature extraction and classification algorithms for early breast cancer diagnosis. The study primarily focuses on the analysis of machine learning techniques to assist hybrid approaches for quickly identifying malignant cancer cells.

2. Literature survey

The SVM classifier, which is able to achieve great efficiency in the categorization of mammograms and diagnostic outcomes, was proposed by S.M.A. Beheshti et al. [7]. Additionally, they have investigated how picture enrichment is implemented overall in each data set's categorization, demonstrating the various effects of enrichment on various lesion types.

In the meantime, in an effort to achieve better accuracy, JeklinHarefa [8] has suggested doing a comparison study between the "k-Nearest Neighbour (k-NN) and Support Vector Machine (SVM)" classifier. The final result shows that, with 93.88% accuracy, SVM performs better than KNN in the categorization of breast cancer anomalies. In order to differentiate between the two tumour classes using the obtained elements, Ahmed M. Sayed et al. [9] proposed Bi-classification methods like KNN and Linear Discriminate Analysis (LDA). With the potential to reduce false positive errors associated with standard MRI diagnosis, the suggested automatic classification approach can be used as noninvasive diagnostic tools for breast cancer.

A four-class breast density classification framework using a combination of neural network classifiers was proposed by Indrajeet Kumar et al. [10]. They note that the framework can be automatically used in medical settings for differential diagnosis between various breast denseness templates. Pedro Henrique BandeiraDiniz [11] and colleagues presented an improved computational scenario aimed at automatically identifying structures that may be associated with tumours during mammography assessment. This article describes a convolution neural system-based computational approach to automatically predict mass regions in mammography.

Sathiyaprasad[12] et al proposed an automatic image recognition technique to solve the problem of video retrieval which reduces the system memory and maintenance time to solve the problem of identification of duplicate images using R-ratio analysis technique with <1000ms recognition time.

S.Punitha [13] et al developed pixel based methodologies are contrasted with other regions growing techniques utilizing the analysis of ROC. The accuracy of the suggested technique was up to 98%, and the specificity was 97.9%, with three hundred mammography pictures being used for preparation and testing.

Mai S. Mabrouk [14] proposed et al proposed coordination of various components, for example, shape, texture and invariant moment aspects on mammography. Instead of using one type of perspective to classify breast cancer, this coordination produced findings with acceptable sensitivity and specificity. The accuracy of our integration tool reached 96% using ANN's programmed technique, whereas the best precision achieved by using aspects was shown by invariant moments, which reached 97% using ANN's programmed way.

The proposed framework utilize the wavelet device to decay the cancer image as sub groups to apply the GLCM on each sub groups for separating the texture features dependent on the texture we form the feature vector. So as to classify the feature vector we embraced random forest calculation to characterize the mammogram images as benign and cancerous.

3. Proposed system

The main goal of the proposed method is to identify and categorise mammography images through a number of steps, such as pre-processing, extraction of features, selection of characters as well as classification with segmentation. Initially, the breast image's excess noise is eliminated using the median filter. After omission of noise, segregate the mass region utilizing K-means clustering algorithm. Following this; the significant texture features are extracted with the help of GLCM. In order to identify an image as benign or cancerous, those features are finally fed to the RF classifier. The ensuing sections detail the suggested breast cancer detection system's step-by-step procedure.

3.1 Pre processing

Pre-processing aimed to reduce undesired indications in the image, resulting in higher-quality and more consistent outcomes. A useful medium for exposing visual data is a picture [15]. When a picture is transferred from one device to another using a network cable and a satellite remote, any disruption that weakens the image signal is referred to as noise in the image.

In computerized Image Processing, the expulsion of noise is an exceptionally requested area of research. There are a few types of noise can be occurring during the image acquisition. For example, Gaussian noise, Salt and Pepper noise, (Poisson noise), Speckle noise andquantum noise [16]. All medical images comprise some visual noise. In this proposed work we use mammogram (x-ray) images.

In this image is mainly affected through quantum noise which is caused due to the alternate in the number of photons in the mammogram unit.A number of researchers have used different filters to improve the image. The median filter is used in this suggested technique to eliminate noise from the mammography image. The median filter is a simple yet efficient nonlinear filter that replaces a pixel value with its median. Utilising a 3 x 3 sampling window for 2D median filtering: maintaining the boundary values shown in fig. 2.

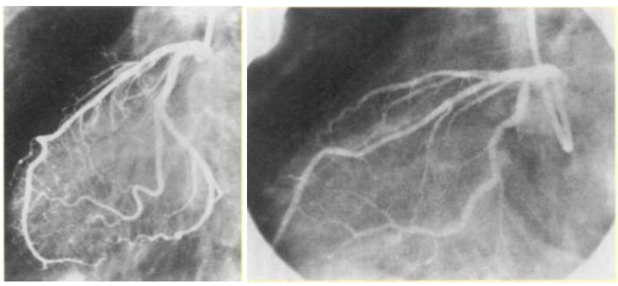


Fig 1: The image with low and high noise

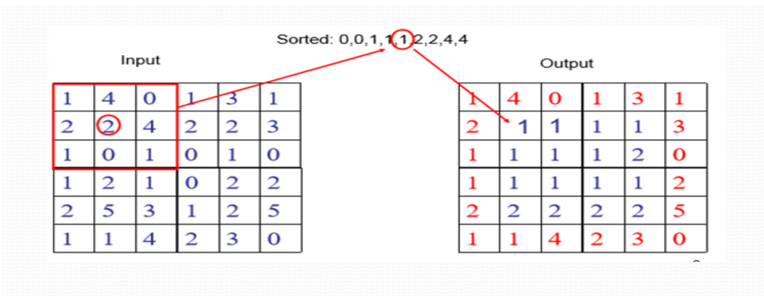


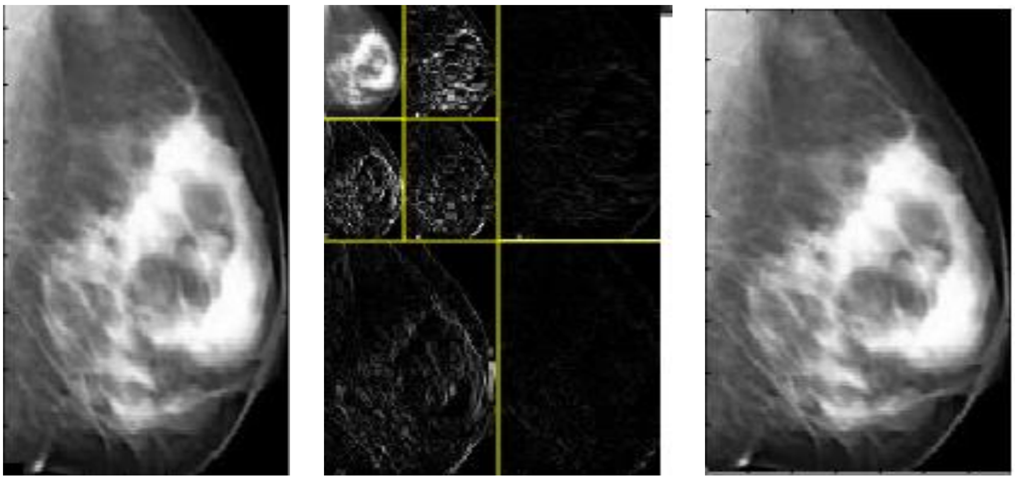
Fig 2: A 3 x 3 sampling window based 2D Median filtering :Maintaining constant border values

3.1.1 Enhancement of mammogram image using Wavelet transform

Wavelet is a power mathematical tool which is used to subdivide the image into different sub bands which provides balanced frequency and localizing information as well. Here, a discrete wavelet transform (DWT) method is used to enhance the image with each illumination and sharpening. In addition, DWT has multi resolution characteristics to shape image in various degree of resolution [18], each degree reduces the size and orientation. The given fig.3 shows that discrete wavelet transform structure.

<table>
<tr><td>$W^c_{LL,3}\,(m,n)$</td><td>$W^c_{HL,3}\,(m,n)$</td><td rowspan="2">$W^c_{HL,2}\,(m,n)$</td><td rowspan="3">$W^c_{HL,1}\,(m,n)$</td></tr>
<tr><td>$W^c_{LH,3}\,(m,n)$</td><td>$W^c_{HH,3}\,(m,n)$</td></tr>
<tr><td colspan="2">$W^c_{LH,2}\,(m,n)$</td><td>$W^c_{HH,2}\,(m,n)$</td></tr>
<tr><td colspan="3">$W^c_{LH,1}\,(m,n)$</td><td>$W^c_{HH,1}\,(m,n)$</td></tr>
</table>

Fig. 3.The 2-level decomposition image structure of DWT.



Input mammogram image | 2 –level decomposition of input image using DWT | Output image (image sharpening)

Fig. 4Mammogram image sharpening based discrete wavelet transform

Fig.4 shows that mammogram image decompose into two level of decomposition based on discrete wavelet transform. The given input mammogram image has the poor illumination due to image capturing, conversion , storage etc., as a result the facts are poorly presented it which will affect prediction of output. In order to enhance illumination and sharpening of input image we used DWT to decompose input into various levels such as LL, HL, LH and HH. The LL subband which consists of the low frequency information i.e smoothening or average of an image and HH subband contains high frequency data i.eedge of the image. To enhance the illumination of image to include LL subbands with inputimage and then mix with high frequency information(HH) into the LL subbands in order to preserve the edgeinformation which is shown in fig. 5.

3.2 K- means clustering

The algorithm works repetitively to set asides every data set to one of K groups based totally on the elements that are provided. Data sets are clustered based on feature similarity [19]. n objects are divided into k clusters using the K-Means clustering algorithm [20], where each object is assigned to the cluster that has the closest mean.

K-means use distance-based measurements to classify the similarity between data sets. Clusters the data into k groups where k is predefined.

1. Choose k points at random to serve as the cluster midpoints.

2. Using the Euclidean distance function, assign the objects to the closest cluster centre.

Total distance= $\sum_{j=1}^{k}\sum_{i=1}^{n}\left\|x_i^{(j)}-c_j\right\|^2$ (1)

3. An Enumerate the centroid or mean of all objects in each cluster.

$\frac{\sum_{i=1}^{n}x_i}{n}$ (2)

4. Repeat the steps 2, 3 and 4 until the same data sets are assigned in to each cluster in successive rounds.

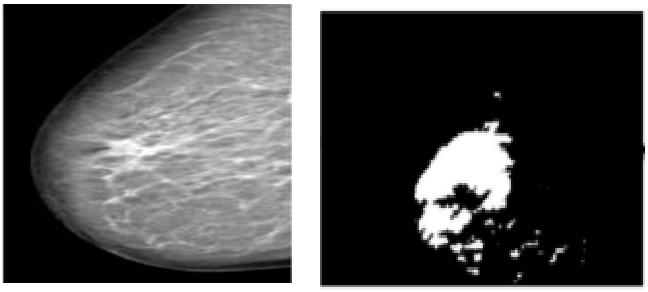


Fig.5 (a) Pre processing image (b) Segmented image

3.3 Feature extraction

In Mammogram image classification, the features of texture is the supreme role for contrasting the usual and abnormal bosom malignant growth cell.After complete the pre-processing next need to extract the texture features by using “Gray level co-occurrence matrix (GLCM)” which was proposed by Haralick (1973). These features are broadly utilized for various gray scale images. Still it’s better than other various descriptors. This method determines the frequency of pixel pairs in an image with exact values and in a particular spatial relationship. The frequency of a pixel with gray- value i occurring vertically, diagonally or horizontally, next to neighboring pixels with value j can be calculated using the GLCM. In this proposed work feature extraction is used to isolate the normal and unusual cell classification on the mammogram image. This strategy is utilized to analysis the textures through considering the spatial relationship of the pixels.

Here, the feature extraction is utilized to segregate the normal and anomalous cell classification on the mammogram. Several prior studies hasdiscovered that GLCM is a favorable methodology in theanalysis of image texture. In various researches, the GLCM is utilized to analyze the image texture so it still becomes the important thing for the research [21] [22]. Therefore, in our proposed work we have use GLCM method. This methodology can be used in the feature extraction of computerized mammogram images.

Certain matrices are built at d = 1, and the directions of θ are indicated as 0°, 45°, 90°, and 135°. Afterwards, the texture information for each normal and aberrant cell location in the mammography pictures is extracted using these four directions.

125	128	128	126
125	125	126	127
126	128	127	127
128	125	128	125

The GLCM's contrast, , homogeneity, correlation and energy of grey level values serve as descriptors, measuring how near the GLCM diagonal the component dispersion is. There are fourteen textural elements, according to Haralick. Correlation, Sum Entropy, Sum Entropy and Sum Variance contrast, are the four primary leading aspects of GLCM that depend on the t-test. This experiment will therefore solely make use of those four aspects.

For mammogram images classification, texture features assumes a significant job for seperating the usual and abnormal breast cancer cell.

- Correlation

The size of the local pixel's linear relationship in grayscale is indicated by the correlation.

Correlation(f1) $=\sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j}$ (3)

- Sum Entropy

The size of an image's irregular shape is displayed by its entropy.

Sum Entropy (f2)$=\sum_{i=2}^{2ng} p_{x+y}(i) \log\{p_{x+y}(i)\}$ (4)

- ASM(Angular Second Moment)

The ASM displays the image size homogeneity attributes or the proximity size of each element in the occurrence matrix.

ASM(f3) $=\sum_{i,j} \frac{p(i,j)}{1+|i-j|}$ (5)

- Sum Variance

The co-occurrence matrix's most notable and minimal cells exhibit an equal concentration of high recurrence of occurrences, as indicated by the Sum variance. $Sum_{var}(f4)=\sum_{i=2}^{2ng}(i-SumEntro)^2 p_{x+y}$ (6)

- Contrast

A pixel's intensity relative to its neighbour across the image is called contrast (Con). When an object and other objects in the same field of view differ in terms of colour and brightness, contrast is created.

Con (f5)$=\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i-j)^2 M(i,j)$ (7)

Feature vector V= [f1, f2, f3, f4, f5] (8)

3.4 Random forest classification

The DT forest algorithm trains multiple decision trees that are powered by slightly different subsets of data. Finding the root node and splitting feature nodes at random are two of the approaches employed in random forest, which can be utilized to extract the most important features from the training data set [23].

Random forest algorithm:

1. From the total number of "m" features, select "k" features at random. In any case k << m
2. Choose the best split technique in order to calculate the node "d".
3. Split the node into leaf nodes using the optimal split.
4. Once the "l" nodes are reached, carry out steps one through three once more.
5. To create a forest, go through steps 1 through 4 "n" times, creating "n" trees each time.

Here, each randomly generated decision tree's rules are used to move forward with the features taken from the mammography images, forecast the result, and store the conclusion as a classed outcome (targets).

3.5 The performance of Classification

Three measures, namely Specificity and Accuracy, are used to assess the work's classification performance. These metrics are quantified using the equations (9), (10), and (11).

- **Sensitivity:**Sensitivity is a metric used to determine the likelihood of events such as "that the woman has the tumour" occurring.

$$\text{Sensitivity}=\frac{Tr_{po}}{Tr_{po}+Fa_{ne}} \quad (9)$$

- **Specificity:**The Specificity is a metric used to determine the likelihood of true negative outcomes, such as "that the woman does not have the tumour."

$$\text{Sensitivity}=\frac{Tr_{ne}}{Tr_{ne}+Tr_{po}} \quad (10)$$

- **Accuracy:**The Accuracy is a metric to assess the likelihood of how many samples are correctly identified.

$$\text{Accuracy}=\frac{Tr_{po}+\boldsymbol{Tr_{ne}}}{Tr_{po}+Tr_{ne}+Fa_{po}+Fa_{ne}} \quad (11)$$

A number of parameters are used to express these evaluations, such as True Negative (Tr_{ne}), Negative (Fa_{ne}), False Positive (Fa_{po}) as well asTrue Positive (Tr_{po}). We can infer from the aforementioned Equations (9) and (10) that specificity indicates the total number of patients that receive a negative test result, but sensitivity indicates the number of diseases that are accurately predicted by a positive test. As a result, it serves as a gauge of test performance to distinguish between patients with and without breast cancer.

4. Results

In this proposed work a several tests has performed to assess the breast cancer detection. In this section, has to present the result and discussion generated by the various approaches are used. Initially the mammogram images are selected randomly from the DDMS. Which contain 2,620 scanned film mammography.Images in the database were digitized and the resolution with 1024*1024 at a 256 grey level. The raw mammography images has lesion, architectural distortion, mass lesion, normal tissues. At the stage of decomposition and feature extraction has been assessed by verifying whether the region in the mammogram associated with mass and any other suspected areas are detected as depicted in fig. 6.

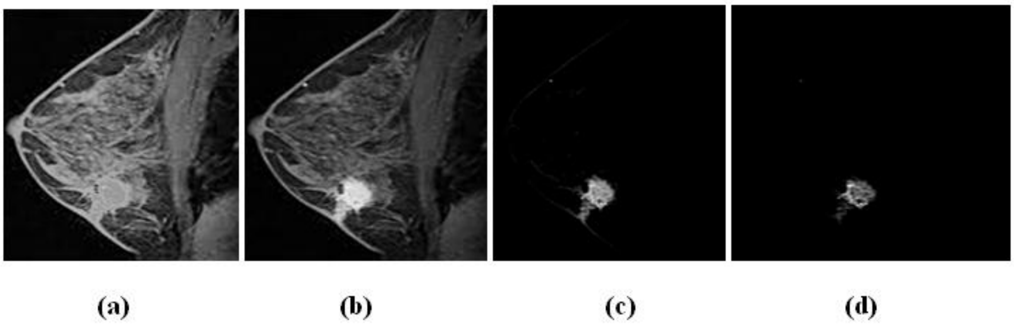


Fig.6 (a) Original image (b)Pre processed image (c) Segmented image (d) Resultant segmented image

Table 1:Comparison between the proposed MRRF algorithm and the prevailing algorithm

Model	Compared model	Source	Metrics	Reference
RF	Multi Neural Network (MNN) and Gradient Boosting (GB)and Genetic Algorithm (GT)	Mohamed cancer institute	Accuracy, sensitivity and specificity	[22]
SVM, Logistic Regression, DT and KNN	RF	"Wisconsin Breast Cancer Diagnostic dataset (WBCD)"	Accuracy, Area under roc curve (AUC), Precision, sensitivity, F-Measure	[23]
"Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA)"	ML approaches	Swiss clinic-based dataset	Accuracy	[24]
Convolutional Neural Network (CNN)	"Multimodal deep neural network by integrating multi-dimensional data (MDNNMD)"	METABRIC dataset	AUC, Accuracy, sensitivity	[25]
XGBOOST	LR, KNN, DT, RF, SVM, GB	Indo-American Cancer Hospital and Research Institute (BIACH & RI)	Accuracy, Recall, precision , F1 score and AUC	[26]
RF, KNN,DT and LR		Wisconsin Breast Cancer Diagnostic (WBCD)	Accuracy	[27]
MRRF (Proposed Algorithm)	SVM,RF, KNN,DT, LR, CNN, XGBOOST	Digital database screening mammography (DDSM) database and Mammogram image analysis society (MIAS)	Accuracy, Sensitivity and specificity	

Table 1 depicts the comparative study of the existing algorithm with the proposed MRRF algorithm and table 2 depicts the Performance comparison results of proposed MRRF algorithm compared to the other existing algorithms. From this table it in inferred that, the proposed MRRF algorithm performs better identification of breast cancer in the segmented images for the same dataset. The resultant performance graph is represented in fig 7.

Table 2. Performance comparison results of proposed MRRF algorithm compared to the other existing algorithms.

Algorithm	Metrics		
	Accuracy	**Sensitivity**	**Specificity**
SVM	0.91	0.9	0.91
KNN	0.92	0.91	0.91
DT	0.93	0.92	0.93
LR	0.95	0.94	0.93
RF	0.95	0.94	0.95
CNN	0.95	0.95	0.96
XGBOOST	0.96	0.96	0.97
MRRF (Proposed Algorithm)	0.97	0.98	0.98

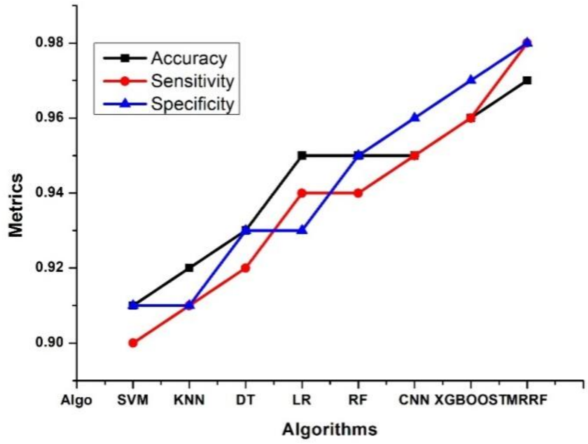


Fig.7 comparative results of performance metrics accuracy, sensitivity and specificity of proposed MRRF algorithm

5. Discussion

This paper focused on the accurate identification of mammography pictures using the proposed MRRF technique. Using median filter, the unwanted noise is removed and the images are sharpened using DWT. The GLCM as well as the proposed MRRF method extracts local statistical features: sum variance, correlation, contrast, and entropy at various sub bands utilizing discrete wavelet transform. Further, it classifies an image into benign or malignant.The proposed technique attains high performance results with 97.1% specificity and sensitivity is 98.1% when compared to the of art-of- art techniques.

Reference

[1] Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in globocan 2012. Int J Cancer 2015;136:E359–86

[2] Seely, J.M., and Alhassan, T., (2018). “Screening for breast cancer in 2018— what should we be doing today?” Current Oncology, 25(s1), pp.115-124.

[3] Ataollahi, M.R., Sharifi, J., Paknahad, M.R., and Paknahad, A.,(2015). “Breast cancer and associated factors: a review”, Journal of Medicine and Life, 8(4), p.6.

[4] Travis,R.C., and Key, T.J., (2003). “Oestrogen exposure and breast cancer risk”, Breast Cancer Research, 5, pp.239–47.

[5] Lakhtakia,R.A.,(2014).“A brief history of breast cancer: Part I: Surgical domination reinvented”,SultanQaboos University Medical Journal,14(2),pp.166- 169

[6] Sathiyaprasad, B. "Ontology-based video retrieval using modified classification technique by learning in smart surveillance applications." *International Journal of Cognitive Computing in Engineering* 4 ,2023: 55-64

[7] M.A. Beheshti a, *, H. AhmadiNoubari b , E. Fatemizadeh c , M. Khalili d “Classification of abnormalities in mammograms by new asymmetric fractal features” Biocybern Biomed Eng (2015),

[8] Jeklin Harefa1, Alexander2, Mellisa Pratiwi 3 “Comparison Classifier: Support Vector Machine (SVM) and K-Nearest Neighbor (K-NN) In Digital Mammogram Images” JUISI, Vol. 02, No. 02, Agustus 2016.

[9] Ahmed M. Sayeda, EmanZaghloulb, Tamer M. Nassef c, d” Automatic Classification of Breast Tumors Using Features Extracted from Magnetic Resonance Images” 1877-0509 © 2016.

[10] Indrajeet Kumar a, *, H.S. Bhadauria a , Jitendra Virmani b , Shruti Thakur “ A classification framework for prediction of breast density using an ensemble of neural network classifiers” Biocybern Biomed Eng (2017), http://dx.doi.org/10.1016/j.bbe.2017.01.001

[11] Jo˜aoOt´avioBandeiraDiniza, Pedro Henrique BandeiraDinizb, Thales Levi AzevedoValentea, Arist´ofanesCorrˆeaSilvaa, Anselmo Cardoso de Paivaa, Marcelo Gattassb” Detection of mass regions in mammograms by bilateral analysis adapted to breast density using similarity indexes and convolutional neural networks” 10 January 2018.

[12] Sathiyaprasad.B, and SeetharamanKrishnamoorthy. "Unsupervised learning- based recognition and extraction for intelligent automatic video retrieval." *The Photogrammetric Record* ,2022

[13] S. Punitha, A. Amuthan, K.Suresh Joseph Benign and Malignant Breast Cancer Segmentation Using Optimized Region Growing Technique “Future Computing and Informatics Journal 24 October 2018.

[14] Mai S. Mabrouk a,⇑ , Heba M. Afify b , Samir Y. Marzouk c” Fully automated computer-aided diagnosis system for micro calcifications cancer based on improved mammographic image techniques” Ain Shams Engineering Journal, https://doi.org/10.1016/j.asej.2019.01.009

[15] Athira P1, Fasna K.K1, Anjaly Krishnan “An Overview Of Mammogram Noise And Denoising Techniques” International Journal of Engineering Research and General Science Volume 4, Issue 2, March-April, 2016 ISSN 2091-2730

[16] Mr. Mandar D. Sontakke1, Dr. Mrs. Meghana S. Kulkarni 2 DIFFERENT TYPES OF NOISES IN IMAGES AND NOISE REMOVING TECHNIQUE International Journal of Advanced Technology in Engineering and Science Volume No.03, Issue No. 01, January 2015

[17] VibhaBafna Bora1, A.G. Kothari2, and A.G. Keskar3” Mammogram Segmentation Using Rough k-Means and Mass Lesion Classification with Artificial Neural Network” A. Ell Hassanien et al. (Eds.): AMLTA 2012, CCIS 322, pp. 60–69, 2012. © Springer-Verlag Berlin Heidelberg 2012.

[18] Satheesh Kumar, B, and Sathiyaprasad B. "Bone Cancer Detection Using Feature Extraction with Classification Using K-Nearest Neighbor and Decision Tree Algorithm." 2021.

[19] 19.Mellisa Pratiwia, Alexandera, JeklinHarefaa, SakkaNandaa “Mammograms Classification using Gray-level Co-occurrence Matrix and Radial Basis Function Neural Network “ International Conference on Computer Science and Computational Intelligence (ICCSCI 2015)

[20] Nalini Singh, Ambarish G Mohapatra, GurukalyanKanungo” Breast Cancer Mass Detection in Mammograms using K-means and Fuzzy C-means Clustering “International Journal of Computer Applications (0975 – 8887) Volume 22– No.2, May 2011.

[21] Yassine AL AMRANIa, *, Mohamed LAZAARb , Kamal Eddine EL KADIRIa “Random Forest and Support Vector Machine based Hybrid Approach to Sentiment Analysis” Procedia Computer Science 127 , 511–520, 2018.

[22] Rabiei, Reza, et al. "Prediction of breast cancer using machine learning approaches." *Journal of Biomedical Physics & Engineering* 12.3 (2022): 297.

[23] Naji, Mohammed Amine, et al. "Machine learning algorithms for breast cancer prediction and diagnosis." *Procedia Computer Science* 191 (2021): 487-492.

[24] Ming, Chang, et al. "Machine learning techniques for personalized breast cancer risk prediction: comparison with the BCRAT and BOADICEA models." *Breast Cancer Research* 21.1 (2019): 1-11.

[25] Kayikci, Safak, and Taghi M. Khoshgoftaar. "Breast cancer prediction using gated attentive multimodal deep learning." *Journal of Big Data* 10.1 (2023): 1-11.

[26] Botlagunta, Mahendran, et al. "Classification and diagnostic prediction of breast cancer metastasis on clinical data using machine learning algorithms." *Scientific Reports* 13.1 (2023): 485.

[27] Monirujjaman Khan, Mohammad, et al. "Machine learning based comparative analysis for breast cancer prediction." *Journal of Healthcare Engineering* 2022 (2022).