



A New Modified Recurrent Extreme Learning with PSO Machine Based on Feature Fusion with CNN Deep Features for Breast Cancer Detection

¹Surendra Prasad M, ²Manimurugan S

¹ Annapoorana Medical College and Hospitals. Salem - Kochi Highway, Kombadipatti, Tamil Nadu, India

²Faculty of Computers and Information Technology, University of Tabuk, KSA.

Corresponding Author: semanimurugan@gmail.com

Received: 20.08.2020,
Revised: 14.11.2020,
Accepted: 16.12.2020,
Published: 22.12.2020

DOI:
10.53409/mnaa.jcsit20201303

Abstract: Breast cancer is a prevalent cause of death, and is the only form of cancer that is common among women worldwide and mammograms-based computer-aided diagnosis (CAD) program that allows early detection, diagnosis and treatment of breast cancer. But the performance of the current CAD systems is still unsatisfactory. Early recognition of lumps will reduce overall breast cancer mortality. This study investigates a method of breast CAD, focused on feature fusion with deep features of the Convolutional Neural Network (CNN). First, present a scheme of mass detection based on CNN deep features and modified clustering of the Extreme Learning Machine (MRELM). It forecasts load through Recurrent Extreme Learning Machine (RELM) and utilizes Artificial Bee Colony (ABC) to optimize weights and biases. Second, a collection of features is constructed that relays deep features, morphological features, texture features, and density features. Third, MRELM classifier is developed to distinguish benign and malignant breast masses using the fused feature set. Extensive studies show the precision and efficacy of the proposed method of mass diagnosis and classification of breast cancer.

Keywords: Breast cancer, Computer-Aided Diagnosis, Deep Learning, Fusion Feature, Recurrent Extreme Learning Machine, Artificial Bee Colony.

I. INTRODUCTION

The malignant tumor activating in the breast cells is breast cancer (BC). A tumor is likely to spread to other body areas [1]. BC is a universal disease that typically hammers women's lives in the 25-50 age category. The potential increase in the number of BC cases in India is alarming. The BC survival rate in the last five years is approximately 90% in the United States, compared to approximately 60% in India [2]. For India in 2020, BC's projections suggest that the statistic is as high as 2 million [3]. Doctors have identified hormones, lifestyle and environmental factors that can increase the chance of developing BC of a person. More than 5%–6% of patients in BC have been associated with gene mutations through the ages of the family. The other factors that cause BC are obesity, an increasing age, postmenosal hormonal imbalances.

As such, there is no mechanism for preventing BC, however, early detection can improve the results considerably. In addition, the cost of

treatment may also be significantly reduced. However, cancer symptoms can often be unusual, so it is difficult to detect them early. In order to detect any early irregularities before it develops, mammograms and self-breasts are essential [4].

The demand for machine learning nowadays increases until it is a service. Machine learning unfortunately continues to be a field of high barriers and often requires skills. A number of skills and expertise are required for an efficient machine learning model including the phases of preprocessing, selection of features and classification processes.

In addition, advances in CNNs can help radiologists as well as eventually diagnosis systems in the near future to read mammograms independently. Propose a method using CNN and MRELM for the extraction and clustering of features, respectively. First, a mammogram will be divided into various sub-regions. Then, CNN is used to extract functions according to individual sub-regions, followed by the use of MRELM with ABC in the cluster of sub-regions where the breast tumor region eventually is located. During the diagnostic period, each

mammogram used as MRELM input for classification is based on its fusion-deep features. The result directly determines if the patient has a benign or malevolent tumor of the breast. Finally, the experimental results show that the best performance is achieved by our proposed methods, the sub-regional ELM clustering and the MRELM clustering with the deep fusion features.

The structuring of the rest of paper is as follows: In Section 2, discuss breast cancer detection methods. Section 3 describes the proposed methodology. Section 4 describes the results and discussion; in Section 5, conclude the paper and define the future work.

II. RELATED WORK

In this study, the application of the K-Particle Swarm Optimization (KPSO) variation is proposed for a new hybrid for breast cancer detection. [5] KPSO is used to initialize and then update the centers and variances of the radial functional neural network with back propagation.

Texture features were removed from the co-occurrence matrix and matrix for runtime lengths and features for automatic classification of normal and malignant breast conditions were added to the Support Vector Machine (SVM) classifier in [6].

Compared to [7] for the automatic classifying of images of breast cancer histology in two machine-learning approaches, benign and malignant subclasses. The first approach is to extract a set of man-made characteristics encoded by two coding models (wordsack and locality-restricted linear codification) and trained in the use of vector support machines.

A CAD system is developed to characterize the breast nodules as either benign, or malignant on

an ultrasound image. The system aims at a high performance classifier. In [8] is developed a Fuzzy Cerebellar Model CAD Network (FCMNN).

The traditional way to diagnose the disease is based on the experience of human beings to identify certain patterns from the database. It is likely to be mistaken, to take time and to work hard. Therefore an automatic breast cancer diagnostic technique was proposed for the parameter optimization of the Artificial Neural Net Network (ANN) using a genetic algorithm (GA).

In [10], a new classification of NB (weighted NB) was proposed, with its application on detection of breast cancer submitted. Several tests on the weighted NB on the breast cancer database were conducted to evaluate its performance.

III. PROPOSED METHODOLOGY

In this research, consider five steps in the detection of breast-cancer: pre-processing of the breast image, mass detection, extraction of features, generation of data and classification training. Delusions and increased contrast processes on the original mammograms have been used in breast image pre-processing to increase the difference between the masses and the surrounding tissue. The ROI is then located for mass detection. Then the ROI extracts features such as deep characteristics, morphological characteristics, textures and density characteristics. Each image from the breast image dataset with the extracted features and corresponding labels was trained by classifiers during the training process. Thus, the diagnosed mammograms can be identified by the well-trained classifications. Fig. 1 presents the flow diagram of the entire diagnosis process.

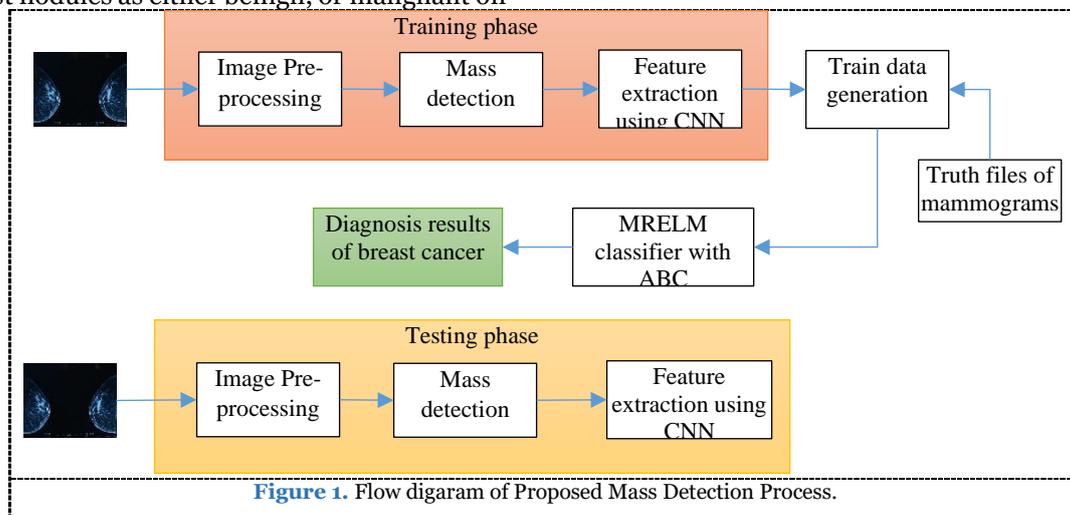


Figure 1. Flow digaram of Proposed Mass Detection Process.

3.1. Image Pre-processing

In order to avoid the effect of noise on the subsequent auxiliary diagnosis, the adaptive mean filter algorithm [11] is chosen to eliminate noise from the original mammogram. The main idea of this is to use a window with a fixed size which slides in the line direction to ascertain whether the noise is available in the window, calculating the mean, variance and spatial correlation values for each sliding window. If detect noise, start replacing the mean value of the pixel of the chosen window. In this work, an algorithm for improving the contrast between the suspected masses and the surrounding tissues has been employed. The key idea is to uniformly distribute the histogram of the original image. The gray scale of the picture is expanded following that process, while also improving the contrast and clarifying the details of the image.

3.2. Mass Detection and extract ROI

Mass detection is intended to remove the mass area from normal tissues. The more precise the mass division is, the more precise the features extracted. A mass detection process based on deep features of sub-domain CNN and ELM is proposed in this paper. After pre-processing, the first step is to take ROI from the images. The ROI is then divided by sliding window into several unsurfacing sub-regions. Determine if all sub-regions have been crossed successfully. If it is yes, the deep characteristics of the sub-regions can be extracted, otherwise the deep characteristics of the sub-regions will be clustered and the mass detection process will be completed.

There are many 0 gray-value areas in mammography, which do not have an effect on breast-CAD. The mammographic area must be separated from the whole mammogram by ROI, in order to improve mammographic processing efficiency or ensure the precise monitoring diagnosis. A mass-recovery adaptive algorithm was used in this work to extract the region of breast mass. Explicitly, in a mammogram, all rows are scanned consecutively to discover the first nonzero pixel by means of abscissa signified as $x_{firstabs}$ and the last nonzero pixel with abscissa signified as $x_{lastabs}$, then all columns are formerly scanned sequentially to find the first nonzero pixel through ordinate denoted as $y_{firstord}$ and the last nonzero pixel (with ordinate denoted as $y_{lastord}$).

In this part, a method to distribute the ROI into several non-overlapping sub-regions is planned. The searching area to manage the masses from the ROI is unchanging in a rectangular area $[x_{firstabs}, x_{lastabs}, y_{firstord}, y_{lastord}]$, wherever the length of the searching rectangular is $L = x_{lastabs} - x_{firstabs}$ and width is $W = y_{lastord} - y_{firstord}$. The rectangular penetrating area is segmented by means

of a sliding window with length w and width h ($L \geq l, W \geq w$). In the rectangular searching area ($L \times W$), the sliding window ($l \times w$) is enthused with a certain step size, negotiating the searching area deprived of crossing the ROI boundary. Consequently, the ROI is alienated into several equal size ($l \times w$), non-overlapping sub-regions and such sub-regions will help as the foundation for subsequent feature extraction. In this work, the size of the sliding window is fixed as 48×48 and the searching step size is equal to 48. In conclusion, the ROI has been alienated into N non-overlapping sub-region such as (s_1, s_2, \dots, s_N) .

3.3. Extract Deep Features Using CNN

This work uses CNN to extract deep characteristics from the sub-regions of ROI. The CNN input is a sub-regional image of 48×48 dimensions which is taken from previous steps. The first coalescing layer filters $48 \times 48 \times 3$ images input to 12 kernels size 9×9 and receives a $40 \times 40 \times 12$ output.

$$Conv_{(i,j)}^k = \sum_{u,v} W^{k,l}(u,v) \cdot input^j(i-u, j-v) + b^{k,l} \quad (1)$$

where $W^{k,l}$ signifies the weight of the k th kernel and $b^{k,l}$ signifies the bias of k th layer. The activation value is controlled in the range $[-1,1]$ using tanh as the activation function.

$$Output_{(i,j)}^k = \tanh(Conv_{(i,j)}^k) \quad (2)$$

The output of the first convolution layer is related with a max-pooling layer. At that point the second and third convolution/max-pooling layer are linked to one another until take the output with size $2 \times 2 \times 6$. The fully-connected layer has $2 \times 2 \times 6 = 24$ neurons which are the features for the next clustering analysis.

3.4. Clustering Deep Features Using MRELM

The MRELM algorithm in this section is used to cluster deep characteristics extracted from the previous CNN architecture. There are two cluster numbers and sub-regional features are divided into two categories: suspect mass areas and unsuspecting mass areas. The effect of supervised learning on the model cannot satisfy the demand when the volume of data is small. Semi-controlled learning is therefore used to improve the effect, but can also accomplish certain tasks of clustering. The ERELM algorithm is a semi-supervised learning algorithm, so that the internal relationships of structure between the unlabeled data set can be identified.

The input is the deeper matrix X of the algorithm and the output is the cluster results of the feature. In particular, Laplacian operator LO was first built from the X training set, followed by the random generation of a cloaked layer neuron output matrix.

In particular, Laplacian L operator is first constructed from the training set X, followed by a randomly generated hidden layer neuron output matrix. If the number of hidden neuron is less than the number of input neuron, we use eq.(3) in the calculation of output weight

$$\min_{\beta \in R^{n_h \times n_o}} \|\omega\|^2 + \lambda \text{Tr}(\beta^T H^T L O \omega) \quad (3)$$

where ω represents the weights between hidden layer and output layer. Then, use the equation $(I_0 + LH^T LH)v = \gamma H^T H v$ to compute output weights. After that, compute the embedding matrix and use k-means algorithm clustering N points into K categories.

3.5. Modified Recurrent Extreme Learning Machine with ABC

A new medical diagnostic framework MRELM-ABC, consisting of two main phases, has been suggested for this study. The first stage is to search for the best feature combo in the medical data by using ABC to filter out redundant and insignificant information. Three steps are the basis of the proposed procedure:

Step1: Optimal network parameter, such as network approximation feature, will be finalized, as will neuron and context neurons no. For the first time, ABC learning algorithms have been used in combination with the RELM to optimize weights and biases to enhance forecast accuracy.

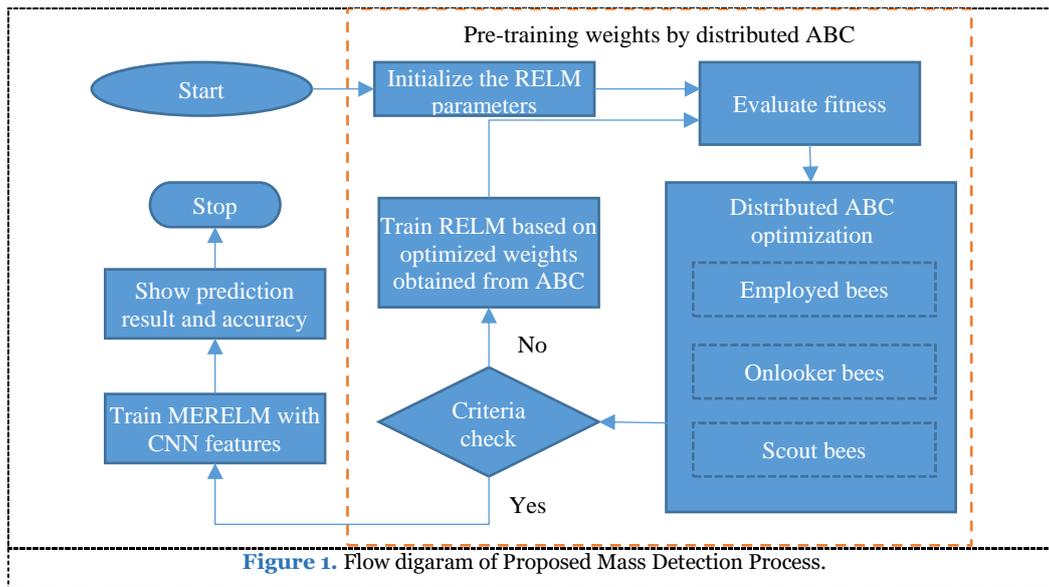
Step 2: ERELM accuracy of forecasts is calculated using the fitness feature for ABC, RMSE for MAE and MSE measurement as shown below is calculated as proposed technology.

$$MSE = \frac{\sum_{t=1}^N [a(t) - p(t)]^2}{N} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^N [a(t) - p(t)]^2}{N}} \quad (5)$$

$$MAE = \sum_{t=1}^N [a(t) - p(t)]^2 \quad (6)$$

Where t is current iteration, N indicates number of iamgess, a is actual value and p is predicted value. Weights are optimized in the work using ABC technology when search algorithms with network structure and learning rate are given input data. ABC searches for the best weight and bias value.



The following algorithmic steps explain the mechanism for breast cancer predictions in detail.

1. Input N image sample, objective function original image dataset
2. Output Predicted Segmented part of the desired value
3. Start
4. Assign the w_i and biases b received by input weights after ABC optimisation.

5. Compute the hidden-layer output matrix H, where $H = h(i = 1, \dots, N)$ and $j = (1, \dots, K)$ and $h_{ij} = g(w_j \cdot x_i + b_j)$
6. Calculate the output weight matrix as $\beta = H^+ T$, where H^+ shows Moore-Penrose generalized inverse of H matrix
7. Updated weights are provided for input and hidden layer as context neurons.
8. End

The MRELM algorithm is a proposed single hidden layer feeder network which has good overall performance, quick learning speed and manual

parameter setup insensitive. In this piece, MRELM was used to obtain benign and malignant diagnostic results for breast cancer.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The performance of the proposed MRELM-ABC is evaluated in this section, and the performance results are compared with existing SVM [6], SVM-FCMNN [8] and weighted NB [10] image compression schemes. The Lung Image Database Consortium image database (LIDC-IDRI) in real time consists of diagnostic and lung cancer screening thoracic computed tomography (CT) scans of annotated marked-up lesions. Seven academic centers and eight medical imaging companies partnered to create this collection of data that contains 1018 cases. That subject contains images from a clinical thoracic CT scan and an accompanying XML file that documents the findings of four experienced thoracic radiologists conducting a two-phase image annotation procedure. The figures given below show that the device proposed has achieved better performance in terms of precision, f-measurement, recall and accuracy.

Precision: It reflects the proportion of positive samples correctly classified as expected in equation (6):

$$Precision = \frac{TP}{FP+TP} \quad (6)$$

Recall: The recall of a classifier reflects the positive samples properly assigned to the total number of positive samples and is calculated as in equation (7):

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

F-measure: this is also referred to as F 1-score, and as in equation (8) is the harmonic mean of precision and recall:

(8):

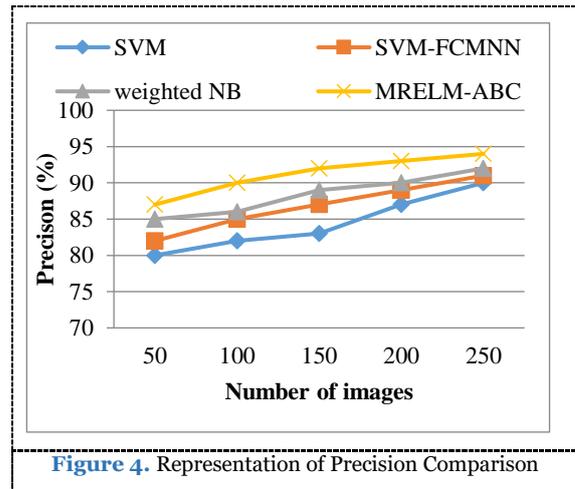
$$F - measure = \frac{2*(Recall * Precision)}{(Recall + Precision)} \quad (8)$$

Accuracy: This is one of the most frequently used performance classification measures and is defined as a ratio between the correctly classified samples and the total number of samples as in equation (9):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

Where true positive (TP) samples are properly classified as natural, false positive (FP) samples are incorrectly classified as irregular, True negative (TN) samples are properly classified as irregular, and false negatives (FN) are incorrectly classified as natural.

4.1. Precision Rate comparison



From the above Figure 4, the graph shows how accurate the number of images in the specified datasets is compared. These methods are implemented as SVM, SVM-FCMNN, weighted NB, and MRELM-ABC. When the number of records increases according to the precision value. From this graph, it is learned that the proposed MRELM-ABC offers 94% higher precision than previous methods that yield better results in the classification of Breast cancer mass due to prior segmentation of the Breast cancer mass using k-means technique. The numerical results of Precision Comparison is shown in Table 1.

Table 1. The numerical results of Precision Comparison

No.of images	SVM	SVM-FCMNN	weighted NB	MRELM-ABC
50	80	82	85	87
100	82	85	86	90
150	83	87	89	92
200	87	89	90	93
250	90	91	92	94

4.2. Recall comparison

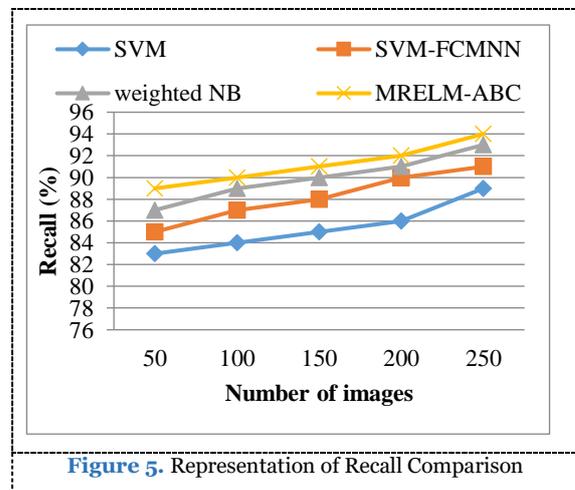


Figure 5. Representation of Recall Comparison

From the above Figure 5 the graph illustrates the recall relation for the number of images in the listed datasets. These methods are implemented as SVM, SVM-FCMNN, weighted NB, and MRELM-ABC. Increasing the number of photographs often increases the correct value for the recall. Through this graph, it is discovered that the current MRELM-ABC offers recall 94% higher than previous methods. The explanation for this is that the MRELM-ABC extracts the features directly which will enhance the Breast cancer classification tests. The numerical results of Recall Comparison is shown in Table 2.

Table 2. The numerical results of Recall Comparison

No.of images	SVM	SVM-FCMNN	weighted NB	MRELM-ABC
50	83	85	87	89
100	84	87	89	90
150	85	88	90	91
200	86	90	91	92
250	89	91	93	94

4.3. F-measure Rate comparison

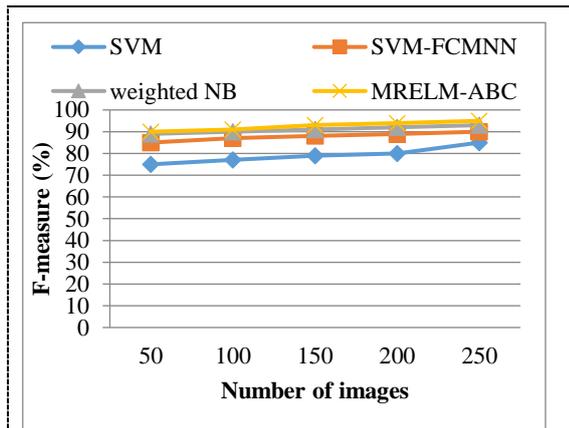


Figure 6. Representation of F-measure Comparison

From the above Figure 6, the graph explains the f-measure relation for the number of images in the given datasets. These methods are implemented as SVM, SVM-FCMNN, weighted NB, and MRELM-ABC. When the number of data is increased, and the f-measure value is increased accordingly. From this graph it is learned that the proposed MRELM-ABC offers 95% higher f-measurement than previous methods.

Therefore the proposed MRELM-ABC algorithm is stronger than the current algorithms in terms of better performance of classifying Breast cancer. The numerical results of F-measure Comparison is shown in Table 3.

Table 3. The numerical results of F-measure Comparison

No.of images	SVM	SVM-FCMNN	weighted NB	MRELM-ABC
50	75	85	89	90
100	77	87	90	91
150	79	88	91	93
200	80	89	92	94
250	85	90	93	95

4.4. Accuracy comparison

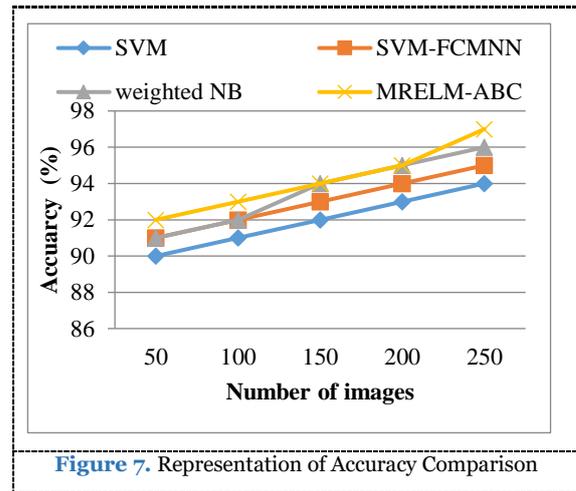


Figure 7. Representation of Accuracy Comparison

From the above Figure 7 the diagram illustrates the processing time relation for the number of images in the specified datasets. These methods are implemented as SVM, SVM-FCMNN, weighted NB, and MRELM-ABC. From this graph it is known that the proposed MRELM-ABC algorithm is higher than the existing algorithms with a high precision rate of 97% in terms of better template matching results. This is due to the automatic extraction of the features using CNN in the MRELM-ABC classification algorithm, which increases the classification precision resulting in Breast cancer. The numerical results of Accuracy Comparison is shown in Table 4.

Table 4. The numerical results of Accuracy Comparison

No.of images	SVM	SVM-FCMNN	weighted NB	MRELM-ABC
50	90	91	91	92
100	91	92	92	93
150	92	93	94	94
200	93	94	95	95
250	94	95	96	97

V. CONCLUSION AND FUTURE WORK

This work suggested a CAD breast based on deep fusion characteristics. The principal idea is the application to two phases of mass detection and mass diagnosis deep CNN extracted features. A method based on deep sub-domain CNN

characteristics and clustering is developed at the stage of mass detection. A MRELM-ABC classifier is used to classify the benign and malignant breast masses with a fused characteristic set which combines deep characteristics, morphological characteristics, texture characteristics, and density characteristics. In breast CAD, diagnostic accuracy is determined by the choice of characteristics. The classifier is used to classify the benign and malignant of the breast mass after the characteristics have been extracted. MRELM is selected as the classifier in this work, which has a better effect on multi-dimensional classification. In future work the proposed methodology will be applied to more practical issues and will plan the parallel implementation of the soft computing method using high performance tools.

REFERENCES

- [1]. Hotko, Y. S. (2013). Male breast cancer: clinical presentation, diagnosis, treatment. *Experimental oncology*, (35, No 4), 303-310.
- [2]. Statistical analysis of breast cancer in India, (2019). Available at: <https://www.biospectrumindia.com/views/21/15300/statistical-analysis-of-breast-cancer-in-india.html>.
- [3]. Malvia, S., Bagadi, S. A., Dubey, U. S., & Saxena, S. (2017). Epidemiology of breast cancer in Indian women. *Asia-Pacific Journal of Clinical Oncology*, 13(4), 289-295.
- [4]. Mehra, R. (2018). Breast cancer histology images classification: Training from scratch or transfer learning?. *ICT Express*, 4(4), 247-254.
- [5]. Senapati, M. R., Panda, G., & Dash, P. K. (2014). Hybrid approach using KPSO and RLS for RBFNN design for breast cancer detection. *Neural Computing and Applications*, 24(3-4), 745-753.
- [6]. Acharya, U. R., Ng, E. Y. K., Tan, J. H., & Sree, S. V. (2012). Thermography based breast cancer detection using texture features and support vector machine. *Journal of medical systems*, 36(3), 1503-1510.
- [7]. Bardou, D., Zhang, K., & Ahmad, S. M. (2018). Classification of breast cancer based on histology images using convolutional neural networks. *IEEE Access*, 6, 24680-24693.
- [8]. Lin, C. M., Hou, Y. L., Chen, T. Y., & Chen, K. H. (2013). Breast nodules computer-aided diagnostic system design using fuzzy cerebellar model neural networks. *IEEE Transactions on Fuzzy Systems*, 22(3), 693-699.
- [9]. Ahmad, F., Isa, N. A. M., Hussain, Z., Osman, M. K., & Sulaiman, S. N. (2015). A GA-based feature selection and parameter optimization of an ANN in diagnosing breast cancer. *Pattern Analysis and Applications*, 18(4), 861-870.
- [10]. Karabatak, M. (2015). A new classifier for breast cancer detection based on Naïve Bayesian. *Measurement*, 72, 32-36.
- [11]. Verma, K., Singh, B. K., & Thoke, A. S. (2015). An enhancement in adaptive median filter for edge preservation. *Procedia Computer Science*, 48(C), 29-36.