Received: December 12, 2022. Revised: February 27, 2023.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Effective Feature Selection Using Multi-Objective Improved Ant Colony Optimization for Breast Cancer Classification

Nelli Sreevidya[1]* **Penamakuru Siva Jyothi[1]** **Gaddam Sumalatha[2]** **Shanthi Pannala[1]**

[1]*Department of IT, Sreenidhi Institute of Science and Technology, Hyderabad, India*
[2]*Department of Computer Science,*
Telangana Social Welfare Residential Degree College for Women, Telangana, India
* Corresponding author's Email: sreevidyan@sreenidhi.edu.in

Abstract: Breast cancer is considered one of the main reasons for the mortality of females around the world, therefore the recognition and classification of the initial stage of breast cancer are essential to help the patients to perform the required action. However, the precise classification of breast cancer is a difficult task in medical imaging because of the complexity of breast tissues. In this paper, Otsu's thresholding based segmentation (OTS) is used for segmenting the cancer regions from the images of mammographic image analysis society (MIAS) dataset. Subsequently, feature extraction using Alexnet is done for extracting the features. The feature selection using multi-objective improved ant colony optimization (MIACO) is proposed in this paper to select the optimal features from the overall features acquired using Alexnet. The MIACO is mainly selects the features based on correlation coefficient to choose the features with higher correlation. Further, the stacked autoencoder (SAE) is used to perform an effective breast cancer classification using the features selected by MIACO. The performance of MIACO-SAE method is analyzed by means of accuracy, precision, recall, characteristic stability index (CSI) and F-measure. The existing research namely threshold-based and trainable segmentation (TTS) with machine learning (ML), infinite feature selection with genetic algorithm (IFS-GA) and deep convolutional neural networks (DCNN) are used to evaluate the MIACO-SAE method. The classification accuracy of MIACO-SAE is 99.36%, which is high when compared to the TTS-ML, IFS-GA and DCNN.

Keywords: Alexnet, Breast cancer classification, Feature selection, Multi-objective improved ant colony optimization, Otsu's thresholding based segmentation, Stacked autoencoder.

1. Introduction

Cancer is one of the deadliest diseases affecting human lives in recent times. The international agency for research on cancer states that around 14 million new cases in 2014 in that 8 million people led to death and also almost more than 30 million people will get affected in the upcoming time. This motivates the researchers to perform the investigation on imaging, diagnosing, and curing affected patients. Generally, the tumor affects various parts of the human, however, breast cancer is the most common type in women [1]. Breast cancer occurrence is the second highest cause in women among other cancer types, except lung cancer. Additionally, the death due to breast cancer is extremely large than the other cancer types [2]. Additional growth of cell mass in the breast region of women is referred to as breast cancer. This breast tissue creates a tumor that is categorized as benign or malignant where the benign is the non-cancerous region and the malignant is a highly affected cancerous region [3-5]. The report given by the American cancer society stated that one in eight women is affected by breast cancer in their lifetime [6, 7].

There is no prevention approach is developed for breast cancer, however, the prior identification significantly improves human lives. Moreover, this prior identification also minimizes the treatment cost [8]. The visual examination of diagnostic pathology is a boring, error-prone task that is subjected to taking wrong decisions. Therefore, computer-aided diagnosis (CAD) is developed to minimize the

pathologist's workload and enhance the efficiency of diagnosis [9-11]. The deep learning and machine learning approaches play an essential role in CAD based breast cancer classification [12]. The feature-based approaches are extensively employed in CAD for diagnosing cancer by segmenting the nucleus and cytoplasm, acquiring the features, then utilizing the features for training the classifier [13]. In classification, the feature selection is used, once the feature extraction is completed. This feature selection process chooses the relevant features from overall features that are additionally used to enhance the classification accuracy [14, 15].

The contributions of this work are concise as follows:

- The OTS is used for segmenting the cancer portions from the mammogram images, then an effective feature extraction is performed by using the alexnet with a huge amount of hidden layers.
- After feature extraction, optimal features are selected by using the proposed MIACO. The developed MIACO is optimized by considering the correlation coefficient, group constraint and accuracy which helps to avoid invalid features while selecting the features. The correlation coefficient considered in MIACO is used to for choosing the feature based on its predictive ability which helps to improve the classification.

The remaining paper is ordered as follows: Section 2 delivers existing works performed in the classification of breast cancer masses. A detailed explanation of the MIACO-SAE method is provided in section 3. Section 4 presents the outcomes of the MIACO-SAE whereas the conclusion is given in section 5.

2. Related work

Zebari [16] implemented an enhanced multi-fractal dimension (M-FD) and feature fusion for effective detection of breast cancer. The region of interest (ROI) from the mammogram images was acquired by using hybrid thresholding and machine learning (ML). Next, the ROI was divided into five blocks followed by a wavelet transform used for reducing the noise. The M-FD approach was used to obtain the features and the genetic algorithm was used to select the features. Further, the artificial neural network (ANN) was used to perform the classification, but the ANN requires more amount of data for training than the ML.

Zebari [17] developed an enhanced threshold-based and trainable segmentation (TTS) for deriving the ROI from breast cancer images. The thresholding and ML were combined to develop the hybrid segmentation. The bands of wavelet transform were avoided for highlighting the breast region in the estimation of breast boundary. The overrated border of the breast region was corrected by using masking and morphological operations. This work was mainly concentrated on segmenting the ROI, however, for a better diagnosis of breast cancer, the feature extraction and selection have to be concentrated further during the classification.

Ittannavar, and Havaldar [18] implemented the infinite feature selection with genetic algorithm (IFS-GA) for enhancing the recognition of breast cancer. The features from the breast cancer were extracted from the segmented images by using the histogram of oriented gradients, Haralick texture features, and local directional ternary pattern. Subsequently, the IFS-GA was used to choose the appropriate features according to the entropy value followed by these features which were applied to the deep neural network for accomplishing the classification. The developed IFS-GA used only entropy for optimizing the feature selection.

Ragab [19] developed feature extraction and deep earning based classification for implementing the CAD to categorize the lesions of breast cancer. This work was accomplished in four experiments to identify the optimum method. In the first experiment, the DCNN was utilized whereas the features from DCNN were extracted and given as input to the support vector machine (SVM). The deep features fusion was done in 3rd experiment to improve the accuracy of SVM. Further, PCA was used in the fourth experiment to reduce the huge feature vector. However, the segmentation of ROI wasn't done in this breast cancer classification.

Rajendran [20] implemented the hybrid optimization approach which was the combination of the grasshopper optimization algorithm (GOA) and crow search algorithm (CSA) for eliminating the optimal features. Here, the multilayer perceptron (MLP) neural network was used to perform the classification. The feature extraction of the gray level co-occurrence matrix used in this classification was extracted only from the statistical textures.

Thirunavukkarasu [21] presented DCNN for classifying the breast cancer. The DCNN was parallelized at the mappers of map reduce (MR) in that network weights were altered through identification of fractional gradients. Hence, the training phase was parallelized by Mapper based on the distribution of preprocessed normal, benign, and

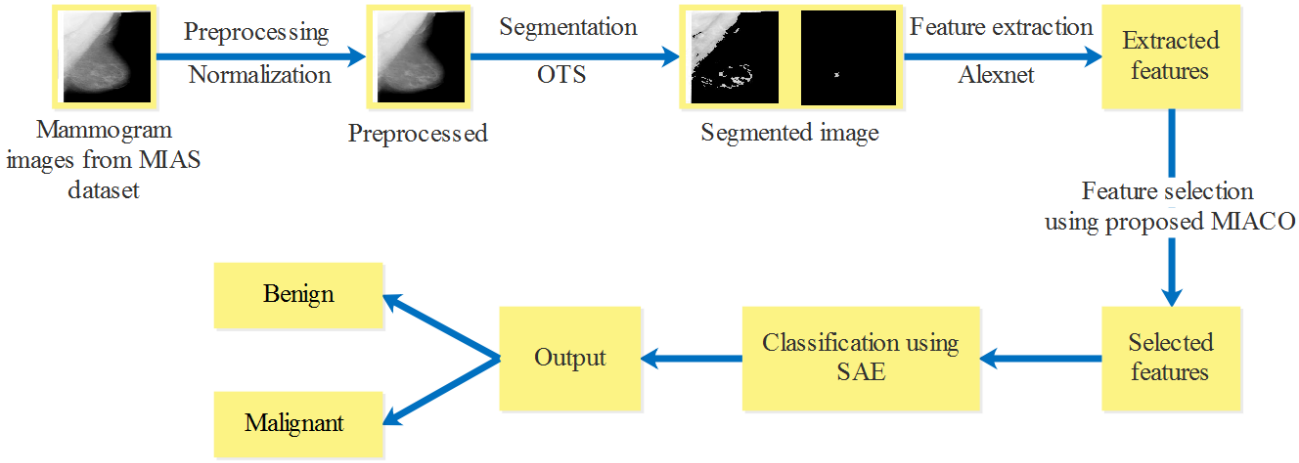


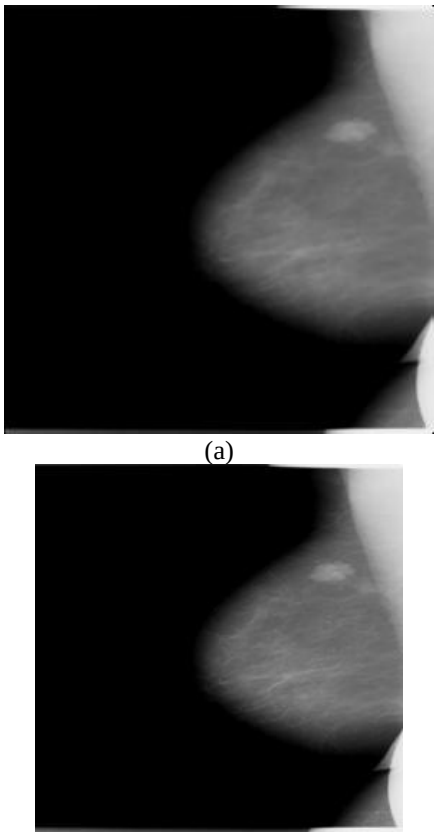
Figure. 1 Block diagram of MIACO-SAE method

malignant images using K-means (KM) method. Further, the distributed DCNN was used to classify the breast cancer images using the updated weight value. An equal distribution using KM was used to minimize the false measure of classification. The classification was affected because it does not consider an effective feature extraction and selection before performing the breast cancer analysis.

Purba Daru Kusuma [22] developed the guided pelican algorithm (GPA) which has the developments in the deficiency approach i.e., pelican optimization algorithm (POA). The GPA was imitated the pelican bird's behavior while hunting the prey where POA was enhanced in three ways in GPA. Initially, the random target was interchanged with global solution in first phase. Next, current position of pelican was interchanged with search space size to determine the size of local search space in second phase. Further, a huge amount of candidates in both phases was designed by GPA than the single candidate as it was utilized in POA. However, the multiple objectives were required to be considered for further optimizing the solutions from the GPA.

3. Proposed method

The MIACO-SAE based breast cancer classification includes the following stages: 1) image collection, 2) preprocessing, 3) segmentation using OTS, 4) alexnet based feature extraction, 5) MIACO based feature selection and 6) classification using SAE. The proposed MIACO is used to select effective features according to the correlation coefficient, classification accuracy and amount of features. Therefore, the elimination of irrelevant features using MIACO is used to enhance the classification using SAE. The block diagram of MIACO-SAE method is shown in Fig. 1.



(a)

(b)

Figure. 2 Mammogram images: (a) Sample image and (b) Pre-processed image

3.1 Image acquisition and preprocessing

At first, the mammogram images are collected from the MIAS dataset [23] which includes 322 mammogram images where each image has the size of 1024×1024. The preprocessing using normalization is applied over the images acquired from the MIAS dataset. The pixel intensity of input is enhanced by changing the pixel limits using the normalization as shown in Eq. (1).

$$I' = (I - min)\frac{newmax - newmin}{max - min} + newmin \quad (1)$$

Where, I is the input image; the minimum and maximum intensity values of the input image are min and max respectively; I' is a normalized image with $newmin$ and $newmax$ intensity values. The sample and pre-processed mammogram images are shown in Figs. 2 (a) and (b) respectively.

3.2 Otsu's thresholding-based segmentation

The OTS [24] is used to improve the class difference between the cancer images by segmenting the cancer portions from the preprocessed image. The otsu is used to classify the image pixel by making the maximum variance among the background and the target region. Consider, the preprocessed image I' has grey levels from 0 to $L-1$, where the total amount of pixels in the image is N and the amount of grey level for each pixel i is denoted as n_i. Eq. (2) represented the possible distribution of the histogram.

$$p_i = n_i/N \quad (2)$$

Consider that there are m thresholds and the image is divided into $m+1$ classes. The threshold levels of the image is denoted as $Th_1, Th_2, \dots, Th_m$. The mean value of overall image I' is expressed in Eq. (3).

$$u_T = w_0 u_0 + w_1 u_1 + \cdots + w_m u_m \quad (3)$$

Where, $u_0 - u_m$ denotes the mean value; $w_0(Th) = \sum_{i=0}^{Th_1-1} p_i$, $w_1(Th) = \sum_{i=t_1}^{Th_2-1} p_i$, … , $w_m(Th) = \sum_{i=Th_m}^{L-1} p_i$. The image segmented from the preprocessed image is denoted as SI. The segmented image using OTS is shown in Fig. 3.

3.3 Alexnet based feature extraction

After segmenting the target region, the AlexNet is used to perform effective feature extraction, because more amount of hidden layers exists in the network. The Alexnet [25] generally has 8 successive layers where the first 5 layers are convolution layers and the remaining 3 are fully connected layers. The base layer of CNN is the convolutional layer where the segmented image SI is given as input to the convolutional filter that provides the characteristic map. The kernel denotes the 5×5 shaped matrix that needs to be transformed as an input pattern matrix. The output of the convolutional layer is shown in Eq. (4).

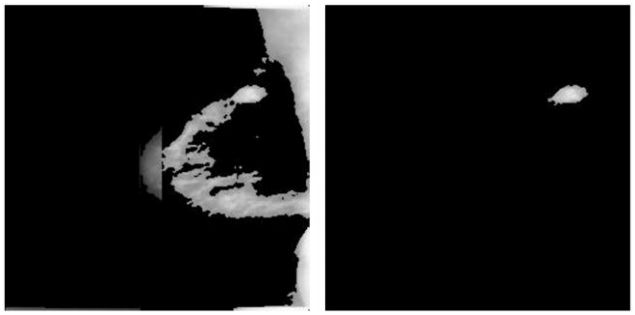


Figure. 3 Segmented image using OTS

$$x_j^l = f\left(\sum_{a=1}^{M} W_j^{l-1} * Y_a^{l-1} + b_j^l\right) \quad (4)$$

Where the feature map j from the layer l is denoted as x_j^l; $f()$ denotes the activation function; M represents the total features at $l-1$ layer; W_j^{l-1} denotes the kernel; Y_a^{l-1} indicates the feature map of $l-1$ layer; $(*)$ represents the convolution mode and b_j^l is bias value. The 2nd layer followed by the convolution layer is the pooling layer that is utilized for performing the down-sample to input feature data. In addition, the operation of pooling diminishes the space size of data and parameter quantity for each layer of the model. The highest value with a definite array size is chosen by using max pooling in each feature map that returns in reduced output neurons. This process is used to eliminate network divergence and overfitting. Eq. (5) expresses the computation of the pooling layer.

$$xs_j^l = \beta_j^l down\left(x_j^{l-1}\right) + b_j^l \quad (5)$$

where, β_j^l denotes as jth multiplication of layer l and $down(.)$ denotes the pooling function.

The fully connected layer behaves as multi layered perception which is an essential layer of CNN. The rectified linear unit (ReLU) is used by the connected layer which is shown in Eq. (6). This ReLU is defined as an activation function that is used to eliminate the gradient vanishing issue.

$$ReLU\ (xs) = \begin{cases} 0, & xs < 0 \\ xs, & xs \geq 0 \end{cases} \quad (6)$$

The feature extracted using Alexnet is further processed under feature selection using MIACO to choose the appropriate features.

3.4 MIACO based feature selection

In this phase, an effective feature from the overall features is selected by using the MIACO. The proposed MIACO considered the correlation

coefficient, group constraint for avoiding invalid feature sets during the feature selection. However, the conventional ACO [26] and IACO doesn't consider the multiple distinct fitness values, so it leads to choosing the invalid feature sets as the final selected features which leads to a decrease in the efficiency of the classification. The primary fitness value of MIACO is correlation coefficient which leads to choose the feature with higher predictive capacity. The separation of feature space is according to the group constraint that corresponds to the groups of multi-character feature set (MCFS) and offers the feature's group information to the probability transition rule (PTR) and generation of feature subset. MCFS has various feature groups along with totally diverse features. Features in various groups provide diverse merits to recognition without any group redundancy. Hence, the best set has at least one feature from each group.

The process of feature selection using MIACO is stated as follows:

Initialize the ant population c, the pheromone trial level is τ linked with a feature i by fixing identical value and highest number of iterations t. Next, each feature's heuristic information η is fixed by information gain (IG) whereas IG is used to offer the statistical measure of the feature's connection to the heuristic value of the feature. Eq. (7) shows $IG(D, xs_i)$ for a feature xs_i $(i = 1,2, \dots, e)$ for amount of examples D of a given input data.

$$IG(D, xs_i) = Entropy(D) - \sum_{v \in vals(xs_i)} \frac{D_v}{D} Entropy(D_v) \quad (7)$$

Where, group of possible values for feature xs_i is denoted as $vals(xs_i)$; D_v is the subset of D for that xs_i has value v. In this MIACO, c amount of ants are created and subsets S_j are constructed based on group constraints. In subset construction, the subspace is selected that is not participated as the feature to be a subset in the iteration. The PTR shown in Eq. (8) is used to choose the feature from the possible features and is included in the subset. The inclusion chosen feature is decided based on the mean square error of the classifier.

$$PTR_i^j(t) = \begin{cases} \frac{[\tau_i(t)]^\alpha [\eta_i(t)]^\beta}{\sum_{c=J_k^j} [\tau_c]^\alpha [\eta_c]^\beta} & if\ i \in J_k^j \\ 0 & Otherwise \end{cases} \quad (8)$$

Where the group of possible features exists in the current group k which is included in the subset is denoted as J_k^j; The heuristic and pheromone values are denoted as η_i and τ_i respectively; the related value of heuristic and pheromone value are defined by using β and α. The optimal subset S_{best} is discovered by computing the fitness of subsets S_j that used to compute the degree of goodness.

The correlation coefficient is used as one of the important measures in the fitness function. The principle followed in correlation coefficient is that optimal feature subset is greatly correlated with the predictive class. This correlation coefficient (CC) expressed in Eq. (9) is used to estimate the feature subset based on the predictive capacity of individual feature as well as its degree of redundancy.

$$CC = \frac{\sum_{i=1}^{M} \rho_{ic}}{\sum_{i=1}^{M} \sum_{j=i+1}^{M} \rho_{ij}} \quad (9)$$

The fitness function of MIACO is expressed in Eq. (10) where it is analyzed by classification accuracy and total number of features.

$$Fitness\left(fea_j\right) = \gamma(CC) + \delta J\left(fea_j\right) + \psi\left(\frac{1}{|fea_j|}\right) \quad (10)$$

Where, the subset chosen by ant j is denoted as fea_j; classification accuracy is represented as $J\left(fea_j\right)$; $|fea_j|$ denotes the number of features in fea_j (i.e., group constraint); δ, ψ and γ are the random numbers among [0,1].

This feature selection using MIACO is terminated once it reached maximum iterations. Further, the optimal feature set S_{best} is taken as an output from the feature selection process, and it is given as input to the SAE for classification.

3.5 SAE based classification

An effective classification among the benign and malignant mammograms is done by using SAE [27] classifier using the optimal features S_{best} selected using MIACO. The SAE is formed by using numerous layers of autoencoder and logistic regression layer. The autoencoder is the base unit of the SAE which includes the encoder and decoder/reconstruction processes. The process of encoder and reconstruction is expressed in Eqs. (11) and (12) respectively, where SAE's weight matrices are denoted as Q and Q' i.e., transpose of Q; bias vectors are denoted as g and g'; non-linearity function (i.e., sigmoid function) is represented as NL; y denotes the hidden variable indication of the input

layer S_{best} and the prediction/ classification is denoted as z.

$$y = NL(QS_{best} + g) \tag{11}$$

$$z = NL(Q'y + g') \tag{12}$$

From the unsupervised pre-training stage, numerous layers of autoencoder are assembled in SAE. The hidden value y is calculated from the autoencoder which is utilized as input to successive autoencoder layers. The reduction in reconstructing error is used to train each layer as an autoencoder, here the cross-entropy is used to compute the reconstruction error.

4. Simulation results and discussion

The outcomes of the MIACO-SAE based classification of breast cancer masses are provided in this section. The design and evaluation of the MIACO-SAE method are accomplished by using MATLAB R2020a software. Here, the system is operated with 8GB RAM and an i5 processor. The dataset used to analyze the MIACO-SAE method is MIAS where 80% of data is used to train and 20% of data is used to test processes. The data is randomly taken from the dataset to perform the training and testing based on the iterations. The MIACO-SAE method is evaluated in terms of accuracy, precision, recall, CSI and F-measure that are shown in Eqs. (13)-(17).

$$Accuracy = \frac{TP+TN}{TN+TP+FN+FP} \times 100 \tag{13}$$

$$Precision = \frac{TP}{TP+FP} \times 100 \tag{14}$$

$$Recall = \frac{TP}{TP+FN} \times 100 \tag{15}$$

$$CSI = \frac{TP}{TP+FN+FP} \times 100 \tag{16}$$

$$F - measure = \frac{2TP}{2TP+FN+FP} \times 100 \tag{17}$$

4.1 Performance analysis of MIACO-SAE method

This section delivers the performance of breast cancer classification using the MIAS dataset. The performances are evaluated for MIACO-SAE method with different classifiers and different optimization-based feature selection approaches. The different classifiers used to analyze the MIACO-SAE are K-nearest neighbour (KNN), random forest classifier (RFC) and multiclass SVM (MSVM). Table 1 shows the performance evaluation of the MIACO-SAE method for different classifiers with and without MIACO. The graphical comparison of the SAE with different classifiers are shown in Figs. 4 and 5. Here, the classifiers without MIACO and with MIACO are shown in Figs. 4 and 5 respectively. From the evaluation, it is known that the SAE with MIACO provides improved classification accuracy than the SAE without MIACO. Besides, the SAE provides better classification among benign and malignant mammograms for both the with and without MIACO. For example, the SAE with MIACO achieves the accuracy of 99.36% whereas the KNN, RFC and MSVM obtain the accuracy of 92.79%, 95.31% and 97.10% respectively. The SAE provides better classification because of the following merits such as 1) the stacking of multiple autoencoders and 2) the overfitting of SAE is avoided because of the weight penalty value.

Table 2 shows the performance of MIACO-SAE with different optimization-based feature selection methods. The different optimizations used for the comparison are artificial bee colony (ABC), grey wolf optimization (GWO), conventional ACO, IACO and GPA [22]. Additionally, Fig. 6 shows the graphical comparison of MIACO with different feature selection methods. From Table 2 and Fig. 6, it is known that the MIACO provides higher classification accuracy of 99.36% than the ABC, GWO, ACO, IACO and GPA. The MIACO selects the optimal features by considering the correlation coefficient, accuracy and number of features extracted from Alexnet. The features from MIACO have higher predictive ability which used to improve the classification accuracy.

4.2 Comparative analysis of the MIACO-SAE method

The comparative analysis of MIACO-SAE with existing research are shown in this section. Existing research such as TTS-ML [17], IFS-GA [18] and DCNN [21] are used to evaluate the efficiency of the MIACO-SAE method. Table 3 shows the comparison of MIACO-SAE with TTS-ML [17], IFS-GA [18] and DCNN [21]. From Table 3, it is known that the classification using MIACO-SAE is improved than the TTS-ML [17], IFS-GA [18] and DCNN [21]. The graphical comparison of classification accuracy is shown in Fig. 7. The classification accuracy of the MIACO-SAE is 99.36% whereas the TTS-ML [17] has 98.13%, IFS-GA [18] has 97.43% and DCNN [21] has 88.35%. The MIACO method selects the optimal features with better correlation that helps to

Table 1. Performance evaluation of MIACO-SAE with different classifiers

Feature selection	Classifiers	Accuracy (%)	Precision (%)	Recall (%)	CSI (%)	F-measure (%)
Without MIACO	KNN	91.39	90.54	89.04	90.18	89.11
	RFC	93.17	94.11	93.37	94.33	95.44
	MSVM	94.28	94.91	95.12	95.07	95.83
	SAE	96.23	95.12	96.11	95.37	96.08
With MIACO	KNN	92.79	93.81	94.37	94.22	94.12
	RFC	95.31	96.18	95.28	95.79	95.11
	MSVM	97.10	97.16	96.75	97.20	96.42
	SAE	99.36	99.22	99.20	99.50	99.66

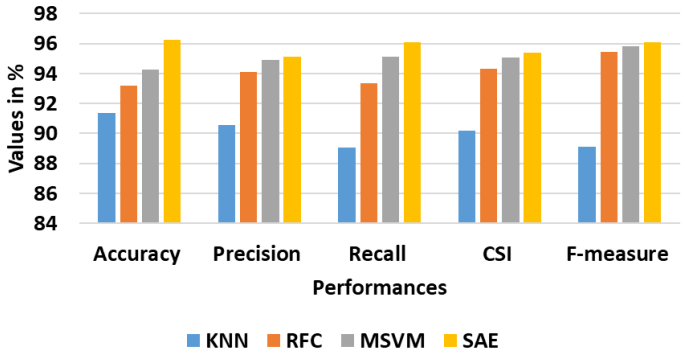


Figure. 4 Graphical comparison of the classifiers without MIACO

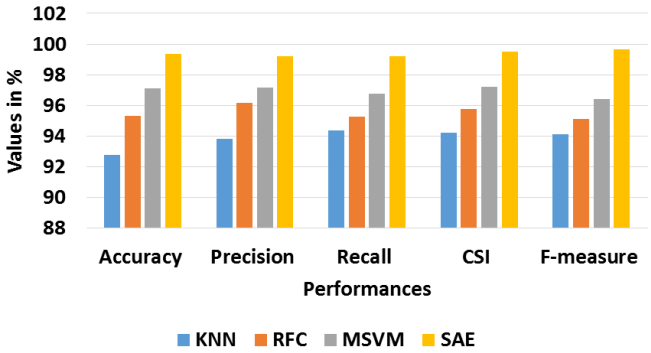


Figure. 5 Graphical comparison of the classifiers with MIACO

Table 2. Performance evaluation of MIACO-SAE with different feature selection methods

Feature selection methods	Accuracy (%)	Precision (%)	Recall (%)	CSI (%)	F-measure (%)
ABC	93.48	92.11	93.77	92.04	92.86
GWO	94.05	93.29	94.17	93.08	94.55
ACO	96.27	95.73	96.04	95.11	95.91
IACO	99.24	99.07	98.93	99.44	99.64
GPA	97.47	98.03	97.15	97.94	97.05
MIACO	99.36	99.22	99.20	99.50	99.66

Table 3. Comparative analysis of MIACO-SAE

Performances	TTS-ML [17]	IFS-GA [18]	DCNN [21]	MIACO-SAE
Accuracy (%)	98.13	97.43	88.35	99.36
Precision (%)	NA	98.93	85.27	99.22
Recall (%)	98.9	97.85	86.55	99.20

improve the classification accuracy using SAE. Moreover, the SAE based classification is improved based on the stacking of multiple autoencoders.

5. Conclusion

The selection of suitable segmentation, feature extraction and selection, and classification is significant in the accurate cancer diagnosis on

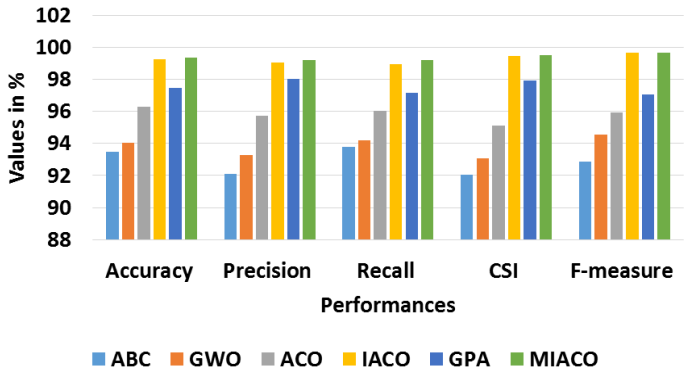


Figure. 6 Graphical comparison of MIACO with different feature selection methods

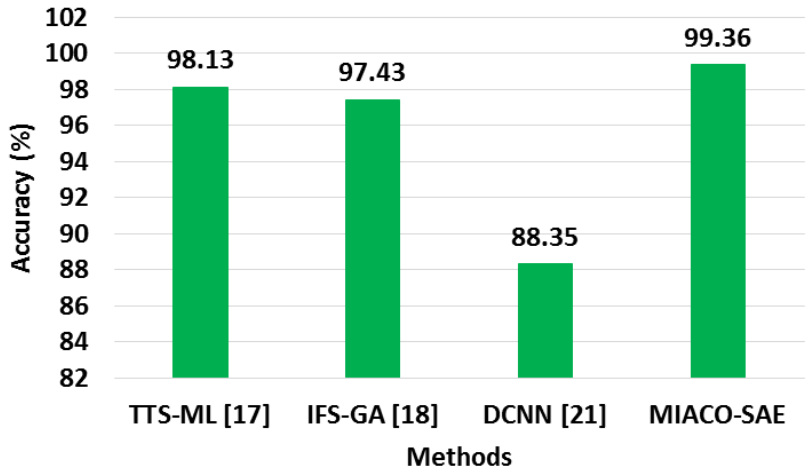


Figure. 7 Graphical comparison of accuracy for MIACO-SAE

mammogram images from the MIAS dataset. The images from the MIAS are preprocessed using normalization followed by the OTS used to segment the cancer regions. Subsequently, the Alexnet with a huge amount of hidden layers effectively extracts the features. Further, the proposed feature selection using MIACO is used to select optimal features according to the correlation coefficient, group constraint and accuracy. Therefore, the optimal selection of features creates an effective classification between benign and malignant breast cancer masses using the SAE classifier. From the results, it is determined that the MIACO-SAE delivers improved classification than the TTS-ML and IFS-GA. The classification accuracy of MIACO-SAE is 99.36%, which is high when compared to the TTS-ML, IFS-GA and DCNN. In the future, the convolutional features from deep learning are combined with the statistical and structural texture features for improving the performance of breast cancer classification.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st, 2nd and 4th author. The supervision and project administration, have been done by 3rd author.

References

[1] S. Alkassar, B. A. Jebur, M. A. Abdullah, J. H. A. Khalidy, and J. A. Chambers, "Going deeper: a magnification-invariant approach for breast cancer classification using histopathological images", *IET Computer Vision*, Vol. 15, No. 2, pp. 151-164, 2021.

[2] R. Yan, F. Ren, Z. Wang, L. Wang, T. Zhang, Y. Liu, X. Rao, C. Zheng, and F. Zhang, "Breast cancer histopathological image classification using a hybrid deep neural network", *Methods*, Vol. 173, pp. 52-60, 2020.

[3] M. Thilagaraj, N. Arunkumar, and P. Govindan,

"Classification of breast cancer images by implementing improved dcnn with artificial fish school model", *Computational Intelligence and Neuroscience*, Vol. 2022, p. 6785707, 2022.

[4] L. Alzubaidi, O. A. Shamma, M. A. Fadhel, L. Farhan, J. Zhang, and Y. Duan, "Optimizing the performance of breast cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model", *Electronics*, Vol. 9, No. 3, p. 445, 2020.

[5] K. K. Dewangan, D. K. Dewangan, S. P. Sahu, and R. Janghel, "Breast cancer diagnosis in an early stage using novel deep learning with hybrid optimization technique", *Multimedia Tools and Applications*, Vol. 81, No. 10, pp. 13935-13960, 2022.

[6] I. Hirra, M. Ahmad, A. Hussain, M. U. Ashraf, I. A. Saeed, S. F. Qadri, A. M. Alghamdi, and A. S. Alfakeeh, "Breast cancer classification from histopathological images using patch-based deep learning modelling", *IEEE Access*, Vol. 9, pp. 24273-24287, 2021.

[7] M. M. Ghiasi and S. Zendehboudi, "Application of decision tree-based ensemble learning in the classification of breast cancer", *Computers in Biology and Medicine*, Vol. 128, p. 104089, 2021.

[8] A. R. Vaka, B. Soni, and S. Reddy, "Breast cancer detection by leveraging Machine Learning", *ICT Express*, Vol. 6, No. 4, pp. 320-324, 2020.

[9] S. Boumaraf, X. Liu, Z. Zheng, X. Ma, and C. Ferkous, "A new transfer learning based approach to magnification dependent and independent classification of breast cancer in histopathological images", *Biomedical Signal Processing and Control*, Vol. 63, p. 102192, 2021.

[10] M. M. Gour, S. Jain, and T. Sunil Kumar, "Residual learning based CNN for breast cancer histopathological image classification", *International Journal of Imaging Systems and Technology*, Vol. 30, No. 3, pp. 621-635, 2020.

[11] S. A. A. Hassan, M. S. Sayed, M. I. Abdalla, and M. A. Rashwan, "Breast cancer masses classification using deep convolutional neural networks and transfer learning", *Multimedia Tools and Applications*, Vol. 79, No. 41, pp. 30735-30768, 2020.

[12] S. Muduli, R. Dash, and B. Majhi, "Automated diagnosis of breast cancer using multi-modal datasets: A deep convolution neural network based approach", *Biomedical Signal Processing and Control*, Vol. 71, Part B, p. 102825, 2022.

[13] P. Wang, J. Wang, Y. Li, P. Li, L. Li, and M. Jiang, "Automatic classification of breast cancer histopathological images based on deep feature fusion and enhanced routing", *Biomedical Signal Processing and Control*, Vol. 65, p. 102341, 2021.

[14] J. G. Melekoodappattu and P. S. Subbian, "Automated breast cancer detection using hybrid extreme learning machine classifier", *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-10, 2020.

[15] M. A. Rahman and R. C. Muniyandi, "An enhancement in cancer classification accuracy using a two-step feature selection method based on artificial neural networks with 15 neurons", *Symmetry*, Vol. 12, No. 2, p. 271, 2020.

[16] D. A. Zebari, D. A. Ibrahim, D. Q. Zeebaree, M. A. Mohammed, H. Haron, N. A. Zebari, R. Damaševičius, and R. Maskeliūnas, "Breast cancer detection using mammogram images with improved multi-fractal dimension approach and feature fusion", *Applied Sciences*, Vol. 11, No. 24, p. 12122, 2021.

[17] D. A. Zebari, Zeebaree, D. Q. Abdulazeez, A. M. Haron, and H. N. A. Hamed, "Improved threshold based and trainable fully automated segmentation for breast cancer boundary and pectoral muscle in mammogram images", *IEEE Access*, Vol. 8, pp. 203097-203116, 2020.

[18] S. S. Ittannavar and R. H. Havaldar, "Detection of breast cancer using the infinite feature selection with genetic algorithm and deep neural network", *Distributed and Parallel Databases*, Vol. 40, No. 4, pp. 1-23, 2022.

[19] D. A. Ragab, O. Attallah, M. Sharkas, J. Ren, and S. Marshall, "A framework for breast cancer classification using multi-DCNNs", *Computers in Biology and Medicine*, Vol. 131, p. 104245, 2021.

[20] R. Rajendran, S. Balasubramaniam, V. Ravi, and S. Sennan, "Hybrid optimization algorithm based feature selection for mammogram images and detecting the breast mass using multilayer perceptron classifier", *Computational Intelligence*, Vol. 38, No. 4, pp. 1559-1593, 2022.

[21] N. Thirunavukkarasu, G. Muthukumarasamy, and R. Murugesan, "Breast Cancer Detection in Mammogram Images by MapReduce Based Deep Convolutional Neural Networks", *International Journal of Intelligent Engineering and Systems*, Vol. 15, No. 1, pp. 477-489, 2022, doi: 10.22266/ijies2022.0228.43.

[22] P. D. Kusuma, and A. L. Prasasti1, "Guided Pelican Algorithm", *International Journal of Intelligent Engineering and Systems*, Vol. 15,

No. 6, pp. 179-190, 2022, doi: 10.22266/ijies2022.1231.18.

[23] J. Suckling, "The mammographic images analysis society digital mammogram database", *Exerpta Medica. International Congress Series*, Vol. 1069, pp. 375-378, 1994.

[24] X. Yue and H. Zhang, "A multi-level image thresholding approach using Otsu based on the improved invasive weed optimization algorithm", *Signal, Image and Video Processing*, Vol. 14, No. 3, pp. 575-582, 2020.

[25] S. B. Akbar, K. Thanupillai, and S. Sundararaj, "Combining the advantages of AlexNet convolutional deep neural network optimized with anopheles search algorithm based feature extraction and random forest classifier for COVID-19 classification", *Concurrency and Computation: Practice and Experience*, Vol. 34, No. 10, p. e6958, 2022.

[26] H. Huang, H. B. Xie, J. Y. Guo, and H. J. Chen, "Ant colony optimization-based feature selection method for surface electromyography signals classification", *Computers in Biology and Medicine*, Vol. 42, No. 1, pp. 30-38, 2012.

[27] W. Li, H. Fu, L. Yu, P. Gong, D. Feng, C. Li, and N. Clinton, "Stacked Autoencoder-based deep learning for remote-sensing image classification: a case study of African land-cover mapping", *International Journal of Remote Sensing*, Vol. 37, No. 23, pp. 5632-5646, 2016.