

Gaussian Optimized Deep Learning-based Belief Classification Model for Breast Cancer Detection

Areej A. Malibari¹, Marwa Obayya², Mohamed K. Nour³, Amal S. Mehanna⁴,
Manar Ahmed Hamza^{5,*}, Abu Sarwar Zamani⁵, Ishfaq Yaseen⁵ and Abdelwahed Motwakel⁵

¹Department of Industrial and Systems Engineering, College of Engineering, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

²Department of Biomedical Engineering, College of Engineering, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh, 11671, Saudi Arabia

³Department of Computer Sciences, College of Computing and Information System, Umm Al-Qura University, Saudi Arabia

⁴Department of Digital Media, Faculty of Computers and Information Technology, Future University in Egypt, New Cairo, 11845, Egypt

⁵Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam Bin Abdulaziz University, AlKharj, Saudi Arabia

*Corresponding Author: Manar Ahmed Hamza. Email: ma.hamza@psau.edu.sa

Received: 27 March 2022; Accepted: 09 May 2022

Abstract: With the rapid increase of new cases with an increased mortality rate, cancer is considered the second and most deadly disease globally. Breast cancer is the most widely affected cancer worldwide, with an increased death rate percentage. Due to radiologists' processing of mammogram images, many computer-aided diagnoses have been developed to detect breast cancer. Early detection of breast cancer will reduce the death rate worldwide. The early diagnosis of breast cancer using the developed computer-aided diagnosis (CAD) systems still needed to be enhanced by incorporating innovative deep learning technologies to improve the accuracy and sensitivity of the detection system with a reduced false positive rate. This paper proposed an efficient and optimized deep learning-based feature selection approach with this consideration. This model selects the relevant features from the mammogram images that can improve the accuracy of malignant detection and reduce the false alarm rate. Transfer learning is used in the extraction of features initially. Next, a convolution neural network, is used to extract the features. The two feature vectors are fused and optimized with enhanced Butterfly Optimization with Gaussian function (TL-CNN-EBOG) to select the final most relevant features. The optimized features are applied to the classifier called Deep belief network (DBN) to classify the benign and malignant images. The feature extraction and classification process used two datasets, breast, and MIAS. Compared to the existing methods, the optimized deep learning-based model secured 98.6% of improved accuracy on the breast dataset and 98.85% of improved accuracy on the MIAS dataset.



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Keywords: Breast cancer detection; computer-aided diagnosis (CAD); deep learning; CNN; entropy; butterfly optimization

1 Introduction

Among the various types of cancers in the works, breast cancer is the one with an increasing death rate. Based on the Global cancer observatory report, approximately 0.68 million persons died due to breast cancer, and Asia is the most affected place with 50.5% [1]. Breast cancer is not diagnosed in the advanced stage, and it needs to be prevented from reducing the death risk. It affects women's health and leads to death. Various imaging modalities such as mammography, magnetic resonance imaging (MRI), digital breast tomosynthesis, and ultrasound have been used to diagnose breast cancer. The recommended image modality is mammography, an affordable, less radiation test suggested for breast cancer diagnosis [1,2].

Deep and machine learning approaches have been used for various applications such as medical imaging [3], renewable energy [4], agriculture [5], optimization [6], and so on. In recent years, medical imaging with deep learning has provided various advancements. The diagnosed results have different stages and types classified by the radiologist with the double reading. For the human observer fault and to reduce the false-negative effects, it is recommended that the hospitals double read. Due to time constraints, double reading is difficult [7]. Among the two tasks, any one of the readings may be inaccurate.

Medical industries support radiologists' CAD systems to overcome this, which reduces perceptual errors. Various deep learning models such as AlexNet, ResNet, GoogleNet, EfficientNet, and MobileNet[8] have been used for multiple tasks in object detection and classification. This proposed work uses the deep learning-based model for feature extraction from mammography images and categories. It will help the radiologist to detect normal (benign) and abnormal (Malignant) tissues [9]. The contribution of this paper is as follows: The Proposed fused feature selection optimization. The two deep learning models, Transfer learning and Convolution Neural network (CNN), are used to extract the features from input preprocessed datasets.

- The extracted features are combined to produce the final feature vector, which is optimized using enhanced butterfly optimization with Gaussian algorithm (EBOG) to make the optimized fused feature set. This optimization-based feature selection improves the classification accuracy.
- Proposed model is evaluated using the breast and MIAS dataset with the evaluation metrics. The comparative analysis proves that the proposed model secured improved accuracy with a reduced false alarm rate.

The remaining section of this paper is as follows: the related work is discussed in Section 2. Proposed methods are introduced in Section 3. Section 4 discussed the experimented results and evaluation. Section 5 concludes the proposed work with its future extension.

2 Related Works

The mammographic image feature extraction and classification of malignant and benign are performed by CAD-based Deep convolutional neural network and AlexNet by Nawaz et al. [10]. Ragab et al., [11] developed a support vector machine (SVM) based model with a fully connected layer to obtain good detection accuracy. They secured 0.94% of AUC and 87.2% of accuracy. Falconi et al., [12] applied VGG, Xception, and Resnet to the CBIS-DDSM dataset, and they used transfer learning to avoid overfitting issues. The obtained AUC value was 0.84. Khan et al., [13] developed multi-view

feature fusion on the MIAS dataset and CBIS-DDSM dataset with the obtained AUC value of 0.932%. Ansar et al., [14] used MobilenetV2 with transfer learning for classification. The obtained accuracy was 74.5% with data augmentation. Lbachir et al., [15] developed a CAD system using DDSM and MIAS datasets. They used a histogram for segmentation with a k-means algorithm. The texture and shape of the image features are extracted and classified using an SVM classifier. The obtained accuracy of the MIAS dataset is 94.2%, with an AUC value of 0.95%, and the CBIS-DDSM dataset secured 90.44% of accuracy and 0.90% of AUC value.

Antari et al., [16] used ResNet-50, CNN, and Inception-ResNetV2 to classify breast and DDSM datasets. The mammogram images were classified with 88.74% and 95.32% accuracy. Agnes et al. [17] used an augmentation method called flipping and rotation on MIAS datasets. 450000 images were augmented and resized into $192 * 192$ size. They developed Multiscale CNN to classify normal, benign, and malignant results. The obtained AUC value is 0.99 with a sensitivity of 96%. Jagtap et al. [18] proposed Kronecker neural network with an adaptive activation function. The Kronecker product reduces the network parameters. Compared to the feed-forward Neural network, Kronecker NN establishes global convergence. The saturation regions are removed with a rowdy activation function using sinusoidal fluctuations.

3 Proposed Methods and Materials

The proposed breast cancer detection system has five processing stages to detect benign and malignant. The input data is preprocessed to remove the noise, and the training samples are increased with the data augmentation method as the first stage. In the second stage, feature extraction is performed with two deep learning models, transfer learning (TL) and convolution neural network (CNN). These two extracted features are combined to form a final selected feature vector optimized using enhanced Butterfly optimization with Gaussian algorithm (EBOG). The improved, optimized feature set is used for cancer cell detection using the deep learning classifier called Deep Belief Network (DBN). It classifies the benign and malignant. This efficient fused feature selection with EBOG enhances the classifier accuracy and reduces the false alarm rate. The stages are explained briefly in the following sections.

3.1 Dataset Description

The proposed breast cancer detection system used two datasets as INbreast and MIAS. The dataset is divided into 50% for training and 50% for testing samples for evaluation.

Breast dataset: this dataset has been generated by the Portugal breast research organization. The breast dataset consists of 410 images of 115 patients. There are 108 mass mammogram images that include BIRADS data and 107 images with mass annotations data are available in Dicom format. The size of the images is $2238 * 4084 * \text{pixels}$. For our experiment, the mammogram 108 images are considered in the Fig. 1.

MIAS Dataset: (Mammographic image analysis society) It consists of 322 images of the size $1024 * 1024$. These images are condensed to 200-micron pixels and it is available in PGM (portable Gray Map). The benign, normal, and malignant sample images of this dataset are shown in Fig. 2. Among the 322 images, the images without calcification of 300 images are considered for this experiment.

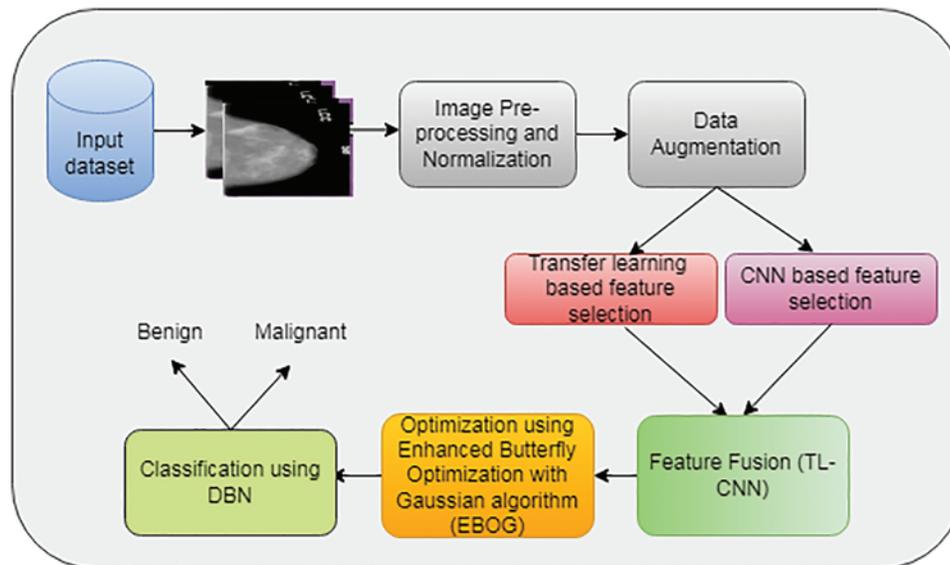


Figure 1: Proposed deep learning-based feature selection and detection architecture

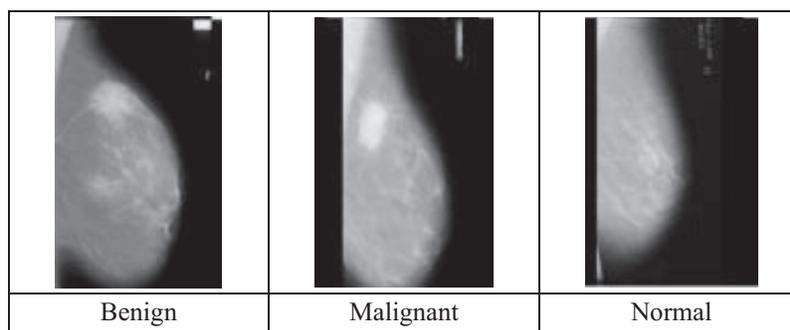


Figure 2: MIAS dataset sample image

3.2 Preprocessing

The proposed feature selection approach is based on CNN's deep learning model. The input images are converted into the same size. CNN works well on the same size images in the dataset. The sample images from the dataset have various sizes with their squared ROI region. The inter-cubic interpolation method is converted into an equal size of $299 * 299$ as the standard size.

3.3 Normalization

Normalization is the process that changes the pixel intensity values into the range of the input image. The n-dimensional grayscale image (I) is converted into a clearer RGB image within the scale 0 to 255 using the Eq. (1)

$$I = 255 * \left(\frac{I - \min}{\max - \min} \right) \quad (1)$$

3.4 Data Augmentation

The sample images are increased using the data augmentation [19] method before applying the deep learning approach. For large amounts of data, DL models perform better. In this work, the ematical actions such as horizontal flip, vertical flip and 90 degrees random rotation, random brightness, saturation, and contrast adjustment are performed. The data augmentsation process is stated in Fig. 3.

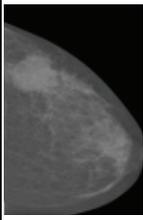
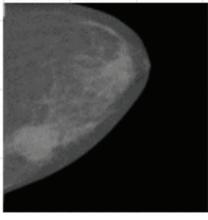
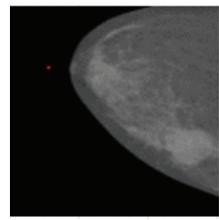
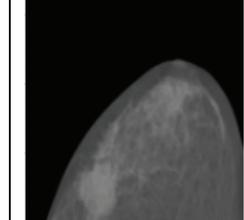
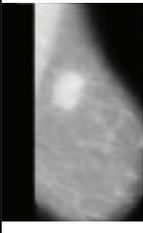
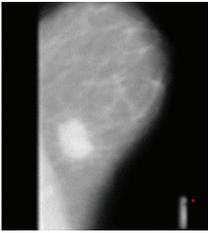
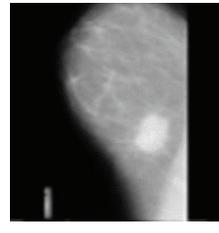
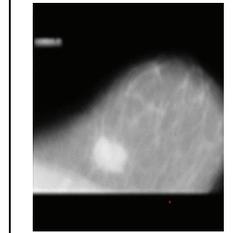
Datasets	Sample image	Left to right	Up to down	Rotate at 90°
INbreast dataset sample image				
MIAS dataset sample image				

Figure 3: Data augmentation process

After augmentation, the breast dataset consists of the original 108 images, and the augmented images are 7200 images. The IAS dataset consists of 14400 augmented images from 300 original images.

3.5 Feature Selection Using Fused TL and CNN Approaches

The feature extraction process used the preprocessed and augmented input data images. This proposed feature selection approach consists of two methods to extract the features, and then the extracted features are combined to form a final feature vector. Transfer learning is used to remove the features. The transfer learning-based model training is helpful in a small amount of data, and it saves time with improved results [20–25].

This TL approach can transfer the input mammogram image (I_s) into the target mass mammogram image (I_t). The target classifier (T_c) called DBN is trained with the input mammogram image to target mammogram image to produce the prediction (P_{Ti}) of breast cancer as benign, regular, and malignant as in Eq. (2).

$$P_{Ti} = T_c(MI_t) \quad (2)$$

The feature is extracted through the transfer layer. The classifier top layer is retrained with new classes, and the remaining classes are kept by themselves.

This transfer learning has been used on the DenseNet network. The fully connected layer of this network is used to extract the features from both datasets. The feature vector is generated row-wise from the input benign and malignant images.

CNN is the structural Neural Network (NN) with many layers, including convolution, pooling, and classification. It consists of two parts such as feature extraction and classification. The feature extraction process involved the convolution layer, pooling layer, and activation function. The classification process involved fully connected layers. For practical use, LeNet-5 [26] is the first proposed CNN. This paper uses CNN to extract the features from the datasets. Initially, the input preprocessed image is transferred to a couple of convolution layers and pooling layers for feature extraction. The initial simple features are extracted efficiently. Then all the elements are combined to form a resulting feature vector. The layers of CNN are explained as follows:

Preprocessing Layers: The preprocessing operations such as sizing, augmentation, and normalization are performed in this layer. Compared to other networks, CNN required less preprocessing operations.

Convolution Layers: This layer extracts the features and produces the feature map. The first convolutional layer extracts the features at the edge which are filtered by the network neurons. These neurons are trained and learned about the image and pass that information to the upcoming convolution layers for high order feature extraction. The kernel in the convolution has extracted the features throughout the input plane and each neuron is assigned to handle the various parts of the input image to provide the feature maps with equal size. The weights are shared based on Time-delay neural network [27], which denotes the feature map stack. The number of neurons is related to the stack depth. Every neuron shares the same bias and weights to form a stack. Each convolution defined various parameters such as input size, feature map stack depth, kernel size, stride, and zero padding. The output is computed as in Eqs. (3) and (4)

$$F_x = \frac{I_x - K_x}{S_x} \quad (3)$$

$$F_y = \frac{I_y - K_y}{S_y} \quad (4)$$

where (F_x, F_y) -feature map size, (I_x, I_y) -input size, (K_x, K_y) -kernel size, (S_x, S_y) -row and column of the stride. The non-breast activation function is used with weights and bias which is used to make the linear combination into a universal approximate of the neural network. The sigmoid activation function is used with the output of the pooling layer [28]. Yet, Rectified linear unit (ReLU) [29] into CNN will improve the network performance stated by Jarret et al., which is stated in Eq. (5).

$$f(X) = \max(0, X) \quad (5)$$

3.6 Fused Feature Vector and Optimization Using EBOG

The extracted features from both DL approaches, such as transfer learning and CNN, are fused to form a final feature vector using parallel concatenation. This combined feature vector (F) is optimized with an enhanced butterfly optimization algorithm to optimize the feature selection process and improve the classification performance. A butterfly algorithm is a meta-heuristic approach developed by Wang et al. [30] based on butterfly mating and foraging nature. In Butterfly optimization (BO), three hypotheses were followed such as (i) fragrance has been relieved by all butterflies to attract each other (ii) towards the butterfly with relevant scent, each butterfly randomly moves. (iii) The butterfly stimulus intensity has been determined based on the fitness function. While the butterfly moves, its fragrance is changed, forming a network based on fragrance. If the butterfly does not feel this network, it can fly randomly, which is called global search space. The butterfly approach defines the local search space as the one with the most incredible fragrance concentration. Based on this local and international

search space, BO can solve the optimization problems. The butterfly fragrance is computed stimulus physical intensity function as denoted in Eq. (6)

$$F_i = mS^a i = 1, 2, \dots N \tag{6}$$

where, F –butterfly fragrance, m –modality of sensory, S –intensity of stimulus, a –random value in the range $[0,1]$ and N –number of butterflies. The global and local search space of BO is defined be in Eqs. (7) and (8)

$$D_i^{t+1} = D_i^t + (rd^2 \times D_{best}^t - D_i^t) \times F_i \tag{7}$$

$$D_i^{t+1} = D_i^t + (rd^2 \times D_j^t - D_k^t) \times F_i \tag{8}$$

where, D_i^t –denotes the position of t^{th} iteration i^{th} butter fly, D_{best}^t –global optimal solution, rd –random number in the range $[0,1]$, j and k are the j^{th} and k^{th} butterfly randomly selected. The main drawback of BO is unbalanced exploitation and exploration capability. To overcome this issue, Gaussian function is added to BO. Gaussian estimation denotes inter individual link between via probabilistic model. The probability can be computed through current population and produce new offspring through sampling and obtains optimal solution. The estimation of this proposed study used weighted maximum likelihood method as shown in Eq. (9)

$$D_i^{t+1} = \sigma + x \tag{9}$$

where $x \sim (N, c)$ c -weighted covariance matrix

$$\sigma = \frac{D_{best} + D_{mean}^t + D_i^t}{3} \tag{10}$$

$$c(i) = \frac{1}{N/2} \sum_{i=1}^{N/2} (D_i^{t+1} - D_{mean}^t) \times (D_i^t - D_{mean}^t)^T \tag{11}$$

$$D_{mean}^t = \sum_{i=1}^{N/2} w_i \times D_i^t \tag{12}$$

$$w_i = \frac{\ln(0.4N + 0.4) - \ln(i)}{\sum_{i=1}^{N/3} (\ln(0.4N + 0.4) - \ln(i))} \tag{13}$$

where, D_{mean}^t –dominant pollution weighted position, w –weight coefficient based on fitness value in descending order. The algorithm for proposed GBOA is shown in following steps. The workflow of the proposed classification system is shown in Fig. 8.

3.7 Classification Using DBN

The optimized feature set detects the benign and malignant input dataset using a deep belief network (DBN). DBN has been used for faster implication with a more extensive network structure [31] which consists of various hidden units and one visible layer to provide generalization. The visual layer is responsible for transferring the input features into an invisible layer for processing [32] based on RBM (Restricted Boltzmann Machine) [33]. This RBN can communicate through its restricted hidden and visible sub-layers to previous and subsequent layers. The layering process is activated through the sigmoid activation function as Eq. (5) based on the RBM learning rule. The DBN architecture is shown in Fig. 8, consisting of stacked RBM where RBM 1 has visible and hidden layers [34,35], RBM2 has hidden layers 1 and 2, RBM3 has hidden layers 1,2 and 3, and RBM4 have hidden layer three and

output layer. The DBN is trained with learning rules and parameters such as layer synaptic weight, states and bias of each neuron. The state of a neuron is based on bias and previous layer neuron weight.

$$P(state_i = 1) = \frac{1}{1 + \exp(-b_i - \sum_j state_j w_{ij})} \quad (14)$$

The training data have positive and negative steps. First positive step will convert the visible layer data to hidden layer and negative step convert hidden layer data to visible layer data. Further using the individual activation function, the positive and negative steps are stated in Eqs. (15) and (16). The weight parameters are optimized until the maximum training epochs as in Eq. (17).

$$P(V_i = 1|H) = \text{sigm}(-b_i - \sum_j H_j w_{ij}) \quad (15)$$

$$P(H_i = 1|V) = \text{sigm}(-c_i - \sum_j V_j w_{ij}) \quad (16)$$

$$W' = \text{update} \left(w_{ij} + \frac{\eta}{2} \times (\text{positive}(Ed_{ij}) - \text{negative}(Ed_{ij})) \right) \quad (17)$$

where,

positive (Ed_{ij})-Positive statistics of edge $Ed_{ij} = (H_j = 1|V)$

negative (Ed_{ij})-Positive statistics of edge $Ed_{ij} = P(vd_j = 1|H)$

η -learning rate that ranges [0,1]

The same training is executed for all the RBM to detect the benign and malignant.

4 Results and Discussions

The experimental evaluation using the proposed feature selection-based classification using two datasets such as breast and MIAS, is discussed in this section. The details of the dataset are discussed in Section 3.1. Each dataset's results are measured using deep learning models, and various classifiers are applied for validation with 10-fold cross-validation. The number of subsets is created as training and testing, and the process is repeated until the models get trained. To avoid over fitting and under-sampling, ten-fold cross-validation has been used. The deep learning models on the input images are experimented with using MATLAB2020A.

4.1 Evaluation

The evaluation is performed using the evaluation metrics such as Accuracy, Sensitivity, Specificity, Precision, F1-Score, AUC, False Prediction Rate (FPR) and computation time.

$$Acc = \frac{TP \cup TN}{TP \cup FP \cup TN \cup FN} \quad (18)$$

$$SN = \frac{TP}{TP \cup FN} \quad (19)$$

$$SP = \frac{TN}{TN \cup FP} \quad (20)$$

$$Precision = \frac{TP}{TP \cup FP} \tag{21}$$

$$TPR = \frac{TP}{TP \cup FN} \tag{22}$$

$$FPR = \frac{FP}{FP \cup TN} \tag{23}$$

AUC is plotted based on true positive and false positive rates at various thresholds.

4.2 Evaluation and Comparison of Proposed Model Performance Using INbreast Dataset

The AUC of the proposed fused feature selection with EBOG based DBN is shown in Fig. 4 for INbreast dataset. With the deep learning based fused feature selection and EBOG optimization-based classification system confusion matrix shows that the TPR of proposed model is 0.986 on detecting the benign and 0.984 on detecting the malignant. This model reduced the false prediction rate and obtained minimum false prediction.

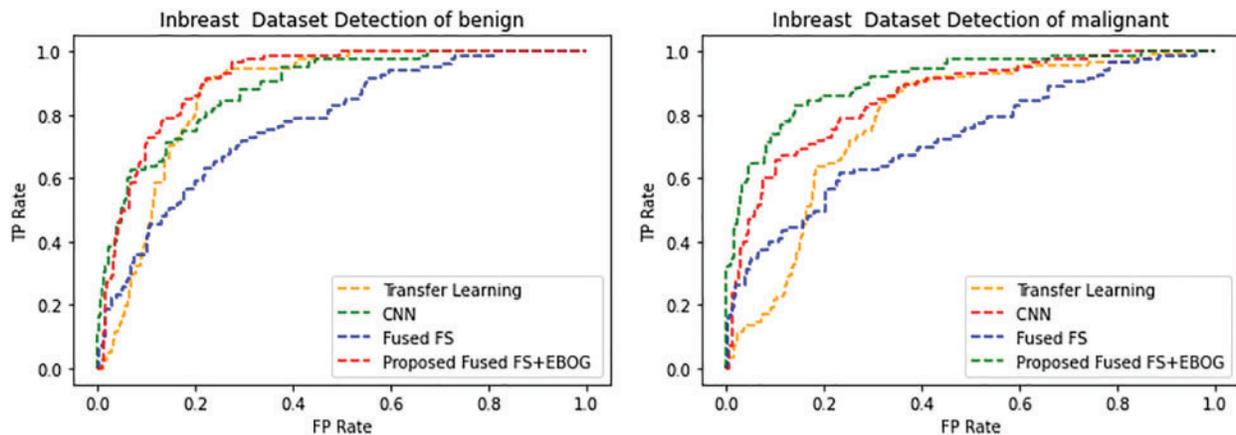


Figure 4: AUC of proposed fused feature selection based DBN classification model–INbreast dataset

Tab. 1 shows the evaluated results of proposed DBN classification model with the optimized fused feature selection model, transfer learning based feature selection and CNN based feature selection. The evaluation of proposed EBOG optimization based fused feature selection with DBN secured improved accuracy, sensitivity, specificity, precision, F1-Score and AUC of 98.6%, 98.4%, 97.5%, 98.3%, 98.42% and 0.98%. Compared to transfer learning and CNN based feature selection, proposed model secured notable increase in the evaluation metrics. Compared to the fused selection without optimization, the proposed EBOG improves the percentage of accuracy as 2%. The False alarm rate is reduced in proposed model compared to TL, CNN and without optimization. Computation time of fused model is 278.1 and with the implementation of optimization it is increased in proposed model. Hence, the fusion of TL and CNN based feature selection optimized with EBOG improves the classification result on INbreast dataset.

Table 1: Proposed optimized fused feature selection based classification–INbreast dataset

Models	Accuracy	SN	SP	Precision	F1-score	AUC	FPR	Computation time (s)
Transfer learning	89.3	88.5	87.3	88.12	87.6	0.83	0.421	291.6
CNN	91.4	91.32	90.5	91.02	90.61	0.88	0.283	286.3
Fused FS	94.6	94.3	93.6	94.2	94.04	0.94	0.142	278.1
Proposed fused FS + EBOG	98.6	98.4	97.5	98.3	98.42	0.98	0.101	282.14

The proposed fused feature selection based classification is compared with MobileNet, Densenet and CNN activated feature based classification. This experimental evaluation is shown in [Tab. 2](#). The deep learning based classifier DBN is applied on MobileNet and DenseNet activated feature selection models. These obtained results are compared with proposed optimized fused feature selection model for evaluation. From the observation of [Tab. 2](#) it is clearly noted that the proposed optimization based fused feature selection secured improved performance than traditional deep features model such as MobileNet and DenseNet.

Table 2: Performance comparison of various feature selection approaches with DBN–INbreast dataset

Models	Accuracy	SN	SP	Precision	F1-score	AUC	FPR	Computation time (s)
MobileNet	91.2	89.4	88.5	89.03	88.8	0.91	0.321	312.04
DenseNet	93.3	92.41	91.4	90.3	91.72	0.93	0.426	286.3
Proposed fused FS + EBOG	98.6	98.4	97.5	98.3	98.42	0.98	0.101	282.14

The DBN classifier performance with the proposed model is evaluated by comparing various classifiers such as SVM, cubic SVM, Gaussian NB and KNN to classify the breast cancer mammography images. In terms of accuracy, AUC and FPR, this evaluation is shown in [Figs. 5–7](#). The classifiers with the proposed fused FS + EBOG are evaluated and compared. [Fig. 5](#) illustrates the accuracy comparison, and DBN with the proposed model secured improved accuracy of 98.6%. Various the traditional classifiers such as SVM, Cubic SVM, KNN, G-NB secured 86.3%, 91.3%, 89.5% and 90.2%, respectively. From [Fig. 6](#), the AUC comparison of proposed FS with DBN and other classifiers. This illustrates that the DBN secured an improved AUC value of 0.98 compared to SVM, Cubic SVM, KNN, G-NB assured the accuracy of 0.81%, 0.89%, 0.87% and 0.86%, respectively. The illustration in [Fig. 7](#) shows the FPR comparison. The proposed model secured a reduced false prediction rate of 6% compared to other approaches. Hence, for the INbreast dataset, the proposed optimized fused feature selection with DBN classifiers performs better than different approaches to detecting benign and malignant.

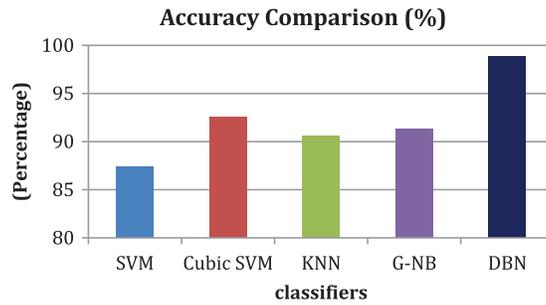


Figure 5: Performance comparison proposed detection model in terms of accuracy-INbreast dataset

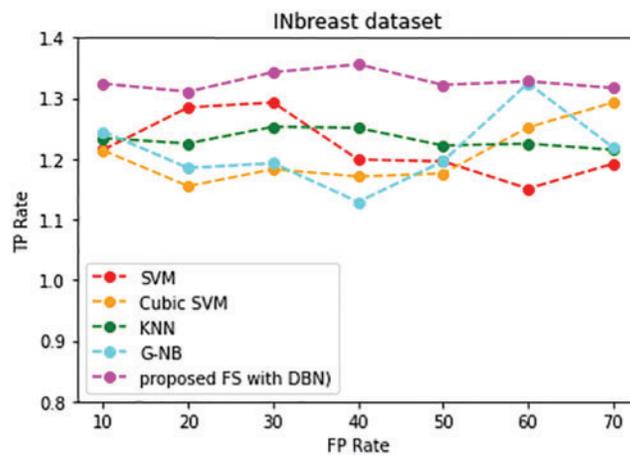


Figure 6: Performance comparison proposed detection model in terms of AUC-INbreast dataset

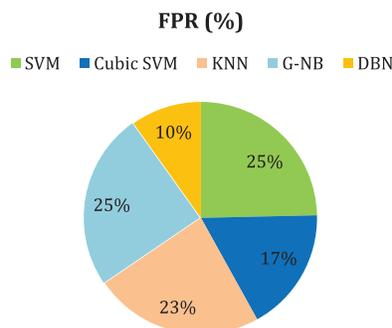


Figure 7: Performance comparison proposed detection model in terms of FPR-INbreast dataset

4.3 Evaluation and Comparison of Proposed Model Performance Using MIAS Dataset

The confusion matrix of the proposed fused feature selection with EBOG based DBN is shown in Fig. 8 for MIAS dataset. With the deep learning based fused feature selection and EBOG optimization based classification system confusion matrix shows that the TPR of proposed model is 98.9% on detecting the benign, 98.8% on detecting the malignant and 99.8% on detecting the normal

mammography images. This model reduced the false prediction rate and obtained minimum false prediction rate.



Figure 8: Confusion matrix of proposed fused feature selection based DBN classification model–MIAS dataset

Based on this confusion matrix, the evaluation metrics are calculated, and the results are shown in [Tab. 3](#). And it shows the evaluated results of the proposed DBN classification model with the optimized fused feature selection model, transfer learning-based feature selection and CNN-based feature selection using the MIAS dataset. The proposed EBOG optimization-based fused feature selection evaluation with DBN secured improved accuracy, sensitivity, specificity, precision, F1-Score and AUC of 98.85%, 98.5%, 97.8%, 98.4%, 98.51% and 0.985%. Compared to transfer learning and CNN-based feature selection, the proposed model increased the evaluation metrics. Compared to the fused section without optimization, the proposed EBOG improves the percentage of accuracy by 2.5%. The False alarm rate is reduced in the proposed model compared to TL, CNN and without optimization. The computation time of the fused model is 102.1 minimum compared to other approaches. Hence, the fusion of TL and CNN-based feature selection optimized with EBOG improves the classification result on the MIAS dataset.

Table 3: Proposed optimized fused feature selection based classification–MIAS dataset

Models	Accuracy	SN	SP	Precision	F1-score	AUC	FPR	Computation time (s)
Transfer learning	90.4	89.43	88.91	89.02	88.5	0.84	0.204	194.3
CNN	92.4	92.2	91.5	90.05	90.06	0.89	0.184	187.5
Fused FS	96.4	96.3	94.8	94.92	95.2	0.95	0.108	143.2
Proposed fused FS + EBOG	98.85	98.5	97.8	98.4	98.51	0.985	0.098	102.1

The deep learning-based classifier DBN is applied on MobileNet, and DenseNet activated feature selection models. The proposed fused feature selection-based classification is compared with

MobileNet, Densenet, and CNN started feature-based classification. These results are compared with the proposed optimized converged feature selection model for evaluation. This experimental evaluation is shown in [Tab. 4](#). From [Tab. 4](#), it is noted that the proposed optimization-based fused feature selection secured improved performance than traditional deep features models such as MobileNet and DenseNet.

Table 4: Performance comparison of various feature selection approaches with DBN–MIAS dataset

Models	Accuracy	SN	SP	Precision	F1-score	AUC	FPR	Computation time (s)
MobileNet	93.4	90.45	91.2	90.23	90.61	0.92	0.148	187.4
DenseNet	95.3	93.5	92.5	91.6	91.52	0.94	0.144	153.2
Proposed Fused FS + EBOG	98.85	98.5	97.8	98.4	98.51	0.985	0.098	102.1

The DBN classifier performance with the proposed model is evaluated by comparing various classifiers such as SVM, cubic SVM, Gaussian NB and KNN to classify the breast cancer mammography images. In terms of accuracy, AUC and FPR, this evaluation is shown in [Figs. 9–11](#). The classifiers with the proposed fused FS + EBOG are evaluated and compared. [Fig. 9](#) illustrates the accuracy comparison, and DBN with the proposed model secured improved accuracy of 98.85%. Various the traditional classifiers such as SVM, Cubic SVM, KNN, and G-NB secured 87.4%, 92.5%, 90.56% and 91.3%, respectively. From [Fig. 10](#), the AUC comparison of proposed FS with DBN and other classifiers is evaluated. This illustrates that the DBN secured an improved AUC value of 0.985 compared to SVM, Cubic SVM, KNN, and G-NB secured the accuracy of 0.83%, 0.91%, 0.9% and 0.88%, respectively. The illustration in [Fig. 11](#) shows the FPR comparison. The proposed model secured a reduced false prediction rate of 10% compared to other approaches. Hence, for the INbreast dataset, the proposed optimized fused feature selection with DBN classifiers performs better than different approaches to detecting benign and malignant.

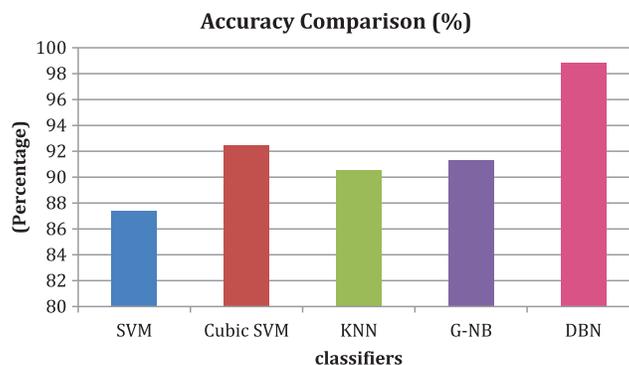


Figure 9: Performance comparison proposed detection model in terms of accuracy-MIAS dataset

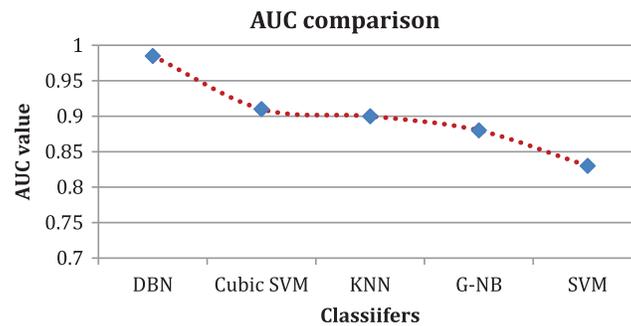


Figure 10: Performance comparison proposed detection model in terms of AUC-MIAS dataset

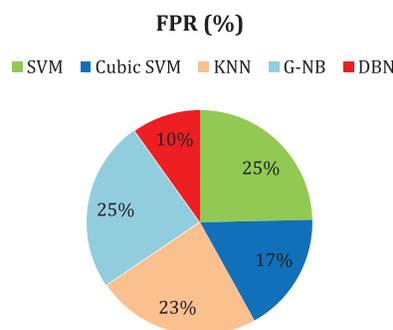


Figure 11: Performance comparison proposed detection model in terms of FPR-MIAS dataset

5 Conclusion

In this paper, a fused feature selection approach has been proposed for breast cancer detection. The two deep learning models were implemented to extract the features from the mammogram images. Before feature selection, the input data is preprocessed to resize the images, and the augmentation method is used to increase the number of training samples. Deep learning algorithms perform better on larger datasets, so preprocessing step helps to get a large dataset. The two deep learning models such as transfer learning and convolution neural networks, are used to extract the features from the preprocessed augmented images. These extracted features are combined/fused to form a final feature vector for processing. To enhance the convergence of the model, fused features are optimized using enhanced Butterfly optimization with a Gaussian algorithm (EBOG). This optimized fused feature is then fed as input to the deep learning classifier called deep belief network (DBN) for breast cancer diagnosis of benign or malignant. The proposed model has been implemented using MATLAB with INbreast and MIAS datasets. These evaluated results are analyzed and compared with various traditional approaches. The proposed optimized fused feature selection-based DBN performs better than other approaches by obtaining an accuracy of 98.6% for the INbreast dataset and 98.85% for the MIAS dataset. The computation time and FPR are reduced for the proposed approach by obtaining 282.14s for the INbreast dataset and 102.1s for the MIAS dataset. The AUC value of the proposed approach on the INbreast dataset is 0.98, and for the MIAS dataset, it is 0.985. It has been noted that, based on the evaluated and compared results, the proposed feature selection with deep learning-based classification performs better with improved accuracy and reduced false alarm rate. The proposed model will implement with larger datasets and real-time applications in the future.

Funding Statement: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R151), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: (22UQU4310373DSR12).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram *et al.*, “Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A cancer*,” *Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.
- [2] J. Hemalatha, S. A. Roseline, S. Geetha, S. Kadry and R. Damaševičius, “An efficient densenet-based deep learning model for malware detection,” *Entropy*, vol. 23, no. 3, 2021.
- [3] S. Albahli, H. T. Rauf, M. Arif, M. T. Nafis and A. Algosaibi, “Identification of thoracic diseases by exploiting deep neural networks,” *Neural Networks*, vol. 5, no. 6, 2021.
- [4] J. Gao, H. Wang and H. Shen, “Smartly handling renewable energy instability in supporting a cloud datacenter,” in *Proc. IEEE Int. Parallel and Distributed Processing Symp. (IPDPS)*, New Orleans, LA, USA, pp. 769–778, 2020.
- [5] D. Oyewola, O. Dada, S. Misra and R. Damaševičius, “Detecting cassava mosaic disease using a deep residual convolutional neural network with distinct block processing,” *PeerJ Computer Science*, vol. 7, pp. 1–15, 2021.
- [6] H. T. Rauf, S. Malik, U. Shoaib, M. N. Irfan and M. Lali, “Adaptive inertia weight bat algorithm with sugeno-function fuzzy search,” *Applied Soft Computing*, vol. 90, pp. 106159, 2021.
- [7] N. Y. Jung, B. J. Kang, H. S. Kim, E. S. Cha, E. Lee *et al.*, “Who could benefit the most from using a computer-aided detection system in full-field digital mammography?,” *World Journal of Surgical Oncology*, vol. 12, no. 1, pp. 1–9, 2014.
- [8] J. Chai, H. Zeng, A. Li and E. W. Ngai, “Deep learning in computer vision: A critical review of emerging techniques and application scenarios,” *Machine Learning with Applications*, vol. 6, 2021.
- [9] S. Zahoor, I. U. Lali, M. A. Khan, K. Javed and W. Mehmood, “Breast cancer detection and classification using traditional computer vision techniques: A comprehensive review,” *Current Medical Imaging*, vol. 16, no. 10, pp. 1187–1200, 2020.
- [10] M. Nawaz, T. Nazir, M. Masood and R. Mahum, “Analysis of brain MRI images using improved cornernet approach,” *Diagnostics*, vol. 11, no. 10, 2021.
- [11] D. A. Ragab, M. Sharkas, S. Marshall and J. Ren, “Breast cancer detection using deep convolutional neural networks and support vector machines,” *PeerJ*, vol. 7, pp. e6201, 2019.
- [12] L. G. Falconi, M. Perez, W. G. Aguilar and A. Conci, “Transfer learning and fine tuning in breast mammogram abnormalities classification on CBIS-DDSM database,” *Advances in Science, Technology and Engineering Systems*, vol. 5, no. 2, pp. 154–165, 2020.
- [13] H. N. Khan, A. R. Shahid, B. Raza, A. H. Dar and H. Alquhayz, “Multi view feature fusion based four views model for mammogram classification using convolutional neural network,” *IEEE Access*, vol. 7, pp. 165724–165733, 2019.
- [14] W. Ansar, A. R. Shahid, B. Raza and A. H. Dar, “Breast cancer detection and localization using mobilenet based transfer learning for mammograms,” in *Proc. Int. Symp. on Intelligent Computing Systems*, Dubai, pp. 11–21, 2020.
- [15] I. A. Lbachir, I. Daoudi and S. Tallal, “Automatic computer aided diagnosis system for mass detection and classification in mammography,” *Multimedia Tools and Applications*, vol. 80, no. 6, pp. 9493–9525, 2021.

- [16] A. Antari, M. A. Han and T. S. Kim, "Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms," *Computer Methods and Programs in Biomedicine*, vol. 196, pp. 105584, 2020.
- [17] S. A. Agnes, J. Anitha, S. Pandian and J. D. Peter, "Classification of mammogram images using multiscale all convolutional neural network (MA-CNN)," *Journal of Medical Systems*, vol. 44, no. 1, pp. 1–9, 2020.
- [18] A. D. Jagtap, Y. Shin, K. Kawaguchi and G. E. Karniadakis, "Deep kronecker neural networks: A general framework for neural networks with adaptive activation functions," *Neuro Computing*, vol. 468, pp. 165–180, 2022.
- [19] R. Yan, F. Ren, Z. Wang, L. Wang, T. Zhang *et al.*, "Breast cancer histopathological image classification using a hybrid deep neural network," *Methods*, vol. 173, pp. 52–60, 2020.
- [20] P. Wang, J. Wang, Y. Li, P. Li and L. Li, "Automatic classification of breast cancer histopathological images based on deep feature fusion and enhanced routing," *Biomedical Signal Processing and Control*, vol. 65, pp. 102341, 2021.
- [21] A. Kumar, S. K. Singh, S. Saxena, K. Lakshmanan, A. K. Sangaiah *et al.*, "Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer," *Information Sciences*, vol. 508, no. 1, pp. 405–421, 2020.
- [22] K. Yu, L. Tan, L. Lin, X. Cheng, X. Yi *et al.*, "Deep learning empowered breast cancer auxiliary diagnosis for 5GB remote E-health," *IEEE Wireless Communications*, vol. 28, no. 3, pp. 54–61, 2021.
- [23] Q. Hu, H. M. Whitney and M. L. Giger, "A deep learning methodology for improved breast cancer diagnosis using multiparametric MRI," *Scientific Reports*, vol. 10, no. 1, pp. 1–11, 2020.
- [24] M. Toğaçar, K. B. Özkurt, B. Ergen and Z. Cömert, "BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer," *Physica A: Statistical Mechanics and its Applications*, vol. 545, pp. 123592, 2020.
- [25] 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.
- [26] C. Hu, X. Sun, Z. Yuan and Y. Wu, "Classification of breast cancer histopathological image with deep residual learning," *International Journal of Imaging Systems and Technology*, vol. 31, no. 3, pp. 1583–1594, 2021.
- [27] S. Singh, T. P. Matthews, M. Shah, B. Mombourquette, T. Tsue *et al.*, "Adaptation of a deep learning malignancy model from full-field digital mammography to digital breast tomosynthesis. In medical imaging 2020," *Computer-Aided Diagnosis*, vol. 11314, pp. 1131406, 2020.
- [28] C. Li, J. Xu, Q. Liu, Y. Zhou, L. Mou *et al.*, "Multi view mammographic density classification by dilated and attention-guided residual learning," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 3, pp. 1003–1013, 2020.
- [29] V. Lahoura, H. Singh, A. Aggarwal, B. Sharma, M. A. Mohammed *et al.*, "Cloud computing based framework for breast cancer diagnosis using extreme learning machine," *Diagnostics*, vol. 11, no. 1, 2021.
- [30] P. Wang, Q. Song, Y. Li, J. Wang and H. Zhang, "Cross task extreme learning machine for breast cancer image classification with deep convolutional features," *Biomedical Signal Processing and Control*, vol. 57, 2020.
- [31] V. K. Singh, H. A. Rashwan, S. Romani, F. Akram, N. Pandey *et al.*, "Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network," *Expert Systems with Applications*, vol. 139, pp. 112855, 2020.
- [32] N. Dhungel, G. Carneiro and A. P. Bradley, "Deep structured learning for mass segmentation from mammograms," in *Proc. Int. Conf. on Image Processing (ICIP)*, France, pp. 2950–2954, 2015.
- [33] X. R. Zhang, X. Sun, W. Sun, T. Xu and P. P. Wang, "Deformation expression of soft tissue based on BP neural network," *Intelligent Automation & Soft Computing*, vol. 32, no. 2, pp. 1041–1053, 2022.
- [34] P. Yang, G. Liu, X. Li, L. Qin and X. Liu, "An intelligent tumors coding method based on drools," *Journal of New Media*, vol. 2, no. 3, pp. 111–119, 2020.
- [35] R. F. Mansour, "A robust deep neural network based breast cancer detection and classification," *International Journal of Computational Intelligence and Applications*, vol. 19, no. 1, pp. 2050007, 2020.