Received: 1 October 2021 | Revised: 27 January 2022 | Accepted: 21 March 2022
DOI: 10.1002/cpe.7008

RESEARCH ARTICLE

WILEY

A Wrapper based feature extraction framework based on AlexNet deep convolutional neural network parameters optimized using gradient-based optimizer for mammogram images

N. K. Sakthivel[1] | S. Subasree[2] | Shaveta Malik[3] | Amit Kumar Tyagi[4]

[1]Nehru Institute of Engineering and Technology, Coimbatore, India

[2]Department of Computer Science and Engineering, Nehru Institute of Engineering and Technology, Coimbatore, Tamil Nadu, India

[3]Department of Computer Science and Engineering, Terna Engineering College, University of Mumbai, Mumbai, India

[4]School of Computer Science and Engineering, Vellore Institute of Technology, 600127, Chennai, India

Correspondence
N. K. Sakthivel, Dean, Nehru Institute of Engineering and Technology, Coimbatore, Tamil Nadu, India.
Email: drnksakthivelphd@gmail.com

Abstract

In this manuscript, a Wrapper based feature extraction framework based on AlexNet deep convolutional neural network (ADCNN) with gradient-based optimizer (GBO) is proposed for early detection of breast cancer. In this input images are occupied as mini-mammography image analysis society (MIAS) database. Then the images are pre-processed to eliminate that noises using Markov random field (MRF) method. Image features are extracted by the process of Wrapper based feature extraction framework with ADCNN. Then, the weight parameters of ADCNN are optimizing through the aid of GBO. Then the mammogram images are characterized as normal or abnormal (malignant and benign) with SVM classifier. The simulation process is implemented on MATLAB platform. The proposed ADCNN-SVM-GBO attains higher accuracy 34.64%, 28.86%, 19.86%, 24.64%, 32.86%, higher Precision 28.07%, 18.96%, 16.86%, 25.86%, 26.86%, higher recall 27.86%, 32.54%, 27.86%, 23.95%, 19.97%, and the efficiency of the proposed method FE-ADCNN-GBO-SVM is likened with the existing processes. Classification of mammograms depends features removal processes with support vector machine (FE-LBP-GLCM-SVM), breast cancer classification with global discriminate features on mammographic images (FE-GLCM-ANN), enhancing breast cancer classification with (SMOTE) technique and pectoral muscle removal on mammographic images (FE-SMOTE-RF), application of artificial intelligence depends deep learning on breast cancer screening and imaging diagnosis (FE-CNN-CDCNN) respectively.

KEYWORDS

AlexNet deep convolutional neural network, breast cancer, gradient-based optimizer Wrapper-based feature extraction unit, SVM classifier mammogram images

1 | INTRODUCTION

Nowadays, breast cancer is becoming more serious disease that affects women and carries a higher mortality rate.[1,2] Based on the WHO (World Health Organization), 450,000 patients die every year universal.[3,4] The mortality rate from breast cancer is detected using an effective screening process at initial stage of cancer.[5] The main measure for detection is to take an x-ray of the breast region known mammogram.[4] Mammography is very

operative for initial diagnosis as it is able to detect lesser change on tissues. This minor change can designate that presence of cancer.[5,6] The most recycled processes to diagnose breast cancer are mammography, biopsy, thermography and ultrasound.[7,8] A biopsy is standard clinic methodology utilized to diagnose early-stage cancer on microscope.[9,10] Based on subjective nature of human clarification, radiologists can have dissimilar opinions about related mammograms.[7] A false negative diagnosis phase can have serious penalties of patient.[11,12] The infected cells spread to other parts of the body.[13–18] Conversely, a false-positive clarification can present an needless biopsy and thus lead to terminated painful process. The growth of an efficient CAD (computer aided diagnosis) system could aid pathologist, as it will help improve confidence in manual diagnosis.[19,20] This scheme is potential candidate for an automated support system in conjunction through manual diagnosis.[21–25]

1.1 | Problem statement

There are numerous breast cancer classification processes that are mentioned on literature review,[13,26–36] which methods have some limitations, like low detection rate of benign samples likened with high fraction, the accuracy of Benign is obviously categorize and diminish that accuracy of malignant, some processes do not obviously categorize malignant and benign display the accuracy of normal region, certain approaches display the accuracy of image. In this process it overcomes all these problems and provides more precision.

In this manuscript, a Wrapper based feature extraction framework based on AlexNet deep convolutional neural network (ADCNN) with gradient-based optimizer (GBO) is proposed. In this the input images are occupied as mini-mammography image analysis society (MIAS) database. Then the images are pre-processed to eliminate that noises using Markov random field (MRF) method. Image features are mined by the process of Wrapper based feature extraction framework with ADCNN. Then, the weight parameters of ADCNN are optimizing with the aid of GBO.[37,38] Then the mammogram images are characterized as normal or abnormal (malignant and benign) with SVM classifier.

The key contributions of this manuscript are abridged beneath:

- In this manuscript, a Wrapper based feature extraction framework depend ADCNN with GBO is proposed for early detection of breast cancer.
- The input images are occupied as mini-MIAS database.
- Then the images are pre-processed to eliminate that noises using MRF method.
- Image features are extracted by the process of Wrapper based feature extraction framework with ADCNN.
- Then, the weight parameters of ADCNN are optimizing with the aid of GBO.[37,38]
- Then the mammogram images are characterized as normal or abnormal (malignant and benign) with SVM classifier.
- The aim of this work is to detect abnormal tissue present in breast cancer mammogram images through process of screening programs and increase the accuracy.
- The simulation process is implemented on MATLAB platform.
- The competence of proposed FE-ADCNN-GBO-SVM approach is evaluated using number of programs that evaluate metrics such as accuracy, sensitivity, specificity and f-score.
- Then the efficiency of proposed method is likened with existing methods such as Digital mammogram images utilizing 2D-BDWT and GLCM features through FOA-based feature selection method (FE-2-D-BDWT-GLCM-FOA-SVM),[39] classification of mammograms depend features removal processes with support vector machine (FE-LBP-GLCM-SVM),[33] breast cancer classification with global discriminate on mammographic images (FE-GLCM-ANN),[34] pectoral muscle removal in mammographic images (FE-SMOTE-RF),[35] application of artificial intelligence depend deep learning on breast cancer and imaging diagnosis (FE-CNN-CDCNN)[36] respectively.

Remaining manuscript is stated beneath. Section 2 outlines literature survey. Section 3 explains Wrapper Based Feature Extraction Framework Based on ADCNN-GBO, and then the images are classified using SVM classifier for the Mammogram Images. Section 4 illustrates outcome and discussion. At last, Section 5 finishes that manuscript.

2 | LITERATURE SURVEY

In this segment, earlier work field of feature extraction and classification in breast cancer data is revised. Such works focused on feature extraction, absorbed on classification system, and several of them have contributions under both stages. Previous researches reveal that CAD systems can streamline the process of translating mammograms and also offering precise result. The result of a CAD is utilized to assist the radiologist in the

discovery of breast cancer. Over the previous two decades, scientists presented various sorts of classifiers in order to create an efficient and also optimum CAD for mammograms.

Mohanty et al.,[39] have presented Mammogram images utilizing contour let features along forest optimization-based feature selection method. Here, the Mammography images were utilized to diagnose that Breast cancer. Then, the images were pre-processed and the features were extracted using kernel extreme learning machine. Therefore, a CAD system was assumed to improve the results of the radiologists. The ROI was created by cropping function, texture feature extraction with the help of Contourlet transformation including a wrapper-based feature selection method. Then features were extracted by using the Contourlet method and the weight parameters were optimized by using the FOA. The disadvantages were (i) the computational time was high (ii) the parameters deemed to detect the near-term risk was not labeled to be adequate.

Mohanty et al.,[40] have presented Digital mammogram images utilizing 2D-BDWT and GLCM features through FOA-based feature selection method. Here, the Mammography images were used to diagnose that Breast cancer. Then, the images were the features depending on the consolidation of 2-D block DWT (2D-BDWT) with GLCM. Furthermore, the FOA was considered to handpicked that optimal features as set of minimum features. The disadvantages of the presented method were (i) multitudes imbrications have been not sensed capably, (ii) the edge segment does not comprise adequate statistics.

Khan et al.,[41] have presented an enhanced Gabor feature extraction for mass classification with cuckoo search. Here, the mammography images were used to detect the breast cancer. The features were extracted by using the CAD schemes for healthcare could be effectual tool for automatically processed such big data. The weight parameters optimized with cuckoo search optimization algorithm (CSO). Here, the experimental result portrays that sensitivity, specificity, accuracy, and area below the curve. The disadvantages of the presented method were (i) the cost of technique was high, (ii) the computation time appears to be more, and (iii) validation of the consequence's requirements modification.

Shree and Kumar[42] have presented an identification, segmentation and detection of infecting area of brain tumor MRI images through feature removal with DWT and probabilistic neural network. Here, the mammography images were utilized for detecting the Breast cancer. The images were pre-processed and filtered with the morphological filtering. Then the features are extracted by using the GLCM features, DWT-based brain tumor region segmentation for diminishing that complexity. The accuracy was estimated as 8.9%. The disadvantages were, 1. The massive data examination receipts extra time and 2. Evidentiary cognitive develops additional complex.

Liu et al.,[43] have presented an organized optimum graph based sparse feature removal. Here, the mammography images were utilized to diagnose the Breast cancer. The images were pre-processed and filtered using the morphological filtering which removes the noise. Then the features were extracted by using the structured optimum graph based sparse feature extraction (SOGSFE) method. The experimental results showed that accuracy 90%, efficiency 7.2%. The disadvantage was early alterations of tissues could not be trapped by an insistently controlled classifier.

Horn et al.,[44] have presented to perform the convolutional neural networks (CNNs) for feature removal on froth flotation sensing. Here, mammography images were used to notice that Breast cancer. Then the features are removed by the CNNs, which reduce the curse of dimensionality inherent on completely associated networks. The experimental results show the accuracy 8.0%, efficiency 9.17%. The disadvantage was Small change in training data basis big change in model.

Zhang et al.,[26] have presented a new behavior fault diagnosis method depending on modified CNN. Here, the mammography images were used to notice that breast cancer. Here, the images were pre-processed and filtered removes the noise. The features were extracted by using the CNNs, moderates the profanity of dimensional characteristic in fully connected networks and trained with feature extractors. The experimental results show the accuracy 8.99%, efficiency 9.2%. The presented method was used to diagnosis. The disadvantage was the accuracy does not compared with the accuracy of Image Net outcomes as the image database for detecting breast cancer has higher feature dependence.

Lahoura et al.,[29] have presented the cloud computing-based framework for breast cancer diagnosis with extreme learning machine. The input images are occupied as Wisconsin Diagnostic Breast Cancer (WBCD) dataset. The framework of this document amalgamates three investigation domains: initially, ELM is used to diagnosis of breast cancer. Second, to remove irrelevant features, the gain ratio feature selection process. Finally, a system based on cloud computing is used to remote diagnosis of breast cancer with ELM. The efficiency of cloud-based ELM was likened to certain state-of-art technologies. Important findings from the experimental outcomes denote that achieved precision is 0.9869, recall is 0.9131, precision is 0.9055, and the F1 score is 0.8128. The limitation of this paper was low true negative rate, low accuracy.

Irfan et al.,[30] have presented dilated semantic segmentation for the detection of ultrasonic breast lesions by merging parallel features. In this to categorize breast cancer ultrasonic images are taken. The presented algorithm was used to segment images of ultrasonic breast lesions with dilated semantic segmentation network (Di-CNN). Here, the 24-layer CNN usages transfer learning-based feature extraction. The accuracy of the CNN-powered feature vectors and the DenseNet201-powered feature vectors joined with support vector machine (SVM) classifier was 90.11% and 98.45%, correspondingly. The limitation of this article was the low detection rate of benign samples likened with high fraction of malignant samples.

Kadry et al.,[31] have reported tumor removal on breast MRI using threshold and joint segmentation. In this, images of the breast cancer are taken as MRI slices. The clinical diagnosis of BC will be made with; (i) image-assisted detection and (ii) central needle biopsy (CNB)-assisted confirmation. This work performs a joint thresholding and segmentation procedure to improve and remove BTS. A three-level threshold depend slime-mold-algorithm and Shannon's entropy (SMA + SE) is realized to improve that BTS and watershed-segmentation (WS) to extract the BTS. The limitation of this article was the low detection rate for benign samples.

Rajinikanth et al.,[32] have suggested the breast cancer detection by screening system using breast thermal imaging (BTI) with selected features of the marine predators algorithm. The implemented methodology is below; (i) image recording for several breast orientations, (ii) extraction of healthy/DCIS image patches, (iii) treatment of patches through image processing, (iv) feature extraction, (v) feature optimization through marine predators-algorithm (MPA), and (vi) classification and authentication under two classes. The consequence of this study confirms that decision-tree (DT) classifier aids to accomplish higher accuracy (>92%) likened to other approaches.

Saeed et al.,[33] have presented the classification of mammograms depend feature removal performances with support vector machine. In this, breast cancer mammography images are taken as MIAS database. In this system, a mean filter and binary image through global threshold have been realistic. Second, the segmentation phase, an HBBRG (hybrid bounding box and region growing) algorithm is used. In the feature removal phase, method three was recycled to prepare the texture features: first order (statistic features), local binary patterns (LBP), and feature level matching matrix. Gray (GLCM), lastly, SVM has been realistic at two levels to organize mammography images at initial level into normal or abnormal.

Tariq et al.,[34] have suggested a classification of breast cancer through global discrimination features on mammographic images. An efficient computer aided diagnostic (CAD) system for breast cancer detection with mammography images is presented in this paper. The CAD system removes highly discriminating features at a worldwide level: the 20 removed features and seven best ranked features between them. Robust results are accomplished and offered after rotating the data up to five times, showing greater than 99% accuracy of target groups, thus outperforming existing solutions.

Abdulla et al.,[35] have presented an development of classification of breast cancer using the technique and the removal of the pectoralis muscle in mammographic images. In this, breast cancer mammography images are considered as Mini-MIAS dataset. The breast area was then removed after disregarding the empty regions around the breast on mammogram images. The mammogram image is overturned and the inverted image is deducted as initial image. For pectoralis muscle removal, a region growth process with K-means clustering process is recycled. Subsequently, the suspected ROI is segmented using K-means technique. To achieve much better classifier efficiency, the SMOTE algorithm is recycled to introduce novel samples of minority classes to obtain balanced dataset. The experimental consequences obtained a precision of 97.1%, the sensitivity of 95.1%,

Wang et al.,[36] have presented an application of artificial intelligence based on deep learning on breast cancer and diagnostic imaging. In this, breast cancer mammography images are considered as mini-MIAS dataset. In this breast cancer screening model depends convolution and deconvolution neural network (CDNN) through convolution neural network (CNN). The fuzzy C-means (FCM) clustering algorithm is recycled to recover and optimize breast cancer imaging, and the experimental consequences are deliberated. The optimized kernel fuzzy C-means clustering algorithm was verified on communal dataset.

3 | PROPOSED WRAPPER BASED FEATURE EXTRACTION FRAMEWORK BASED ON ALEXNET DEEP CONVOLUTIONAL NEURAL NETWORK PARAMETERS OPTIMIZED USING GRADIENT-BASED OPTIMIZER FOR MAMMOGRAM IMAGES

In Figure 1 portray the flow diagram of a Wrapper based feature extraction framework depending on ADCNN method and optimized using GBO for mammogram images is proposed to early detection the breast cancer with great accuracy. Here, the input mammogram images[13,27,28] are taken, and then the images are pre-processed the noises is removed. ROI is extracted to extract features of breast cancer from input images. Then, the features are removed with Wrapper based feature extraction framework depending on the ADCNN. The image features are removed by the process of extracting ROI with the help of Wrapper based feature extraction framework according to ADCNN. Then, the weight parameters of ADCNN are optimized through the aid of GBO. Mammogram images are classified to diagnose the breast cancer as normal or abnormal by SVM classifier. The detailed block explanations are given below,

3.1 | Image acquisition

In this, the input images are considered as mini-MIAS database. In the proposed system, the mammogram mini-MIAS database is recycled. This data set has 322 mammograms: 270 sample images as normal (non-cancerous) and remaining 52 samples as malignant (cancerous). Every sample refers

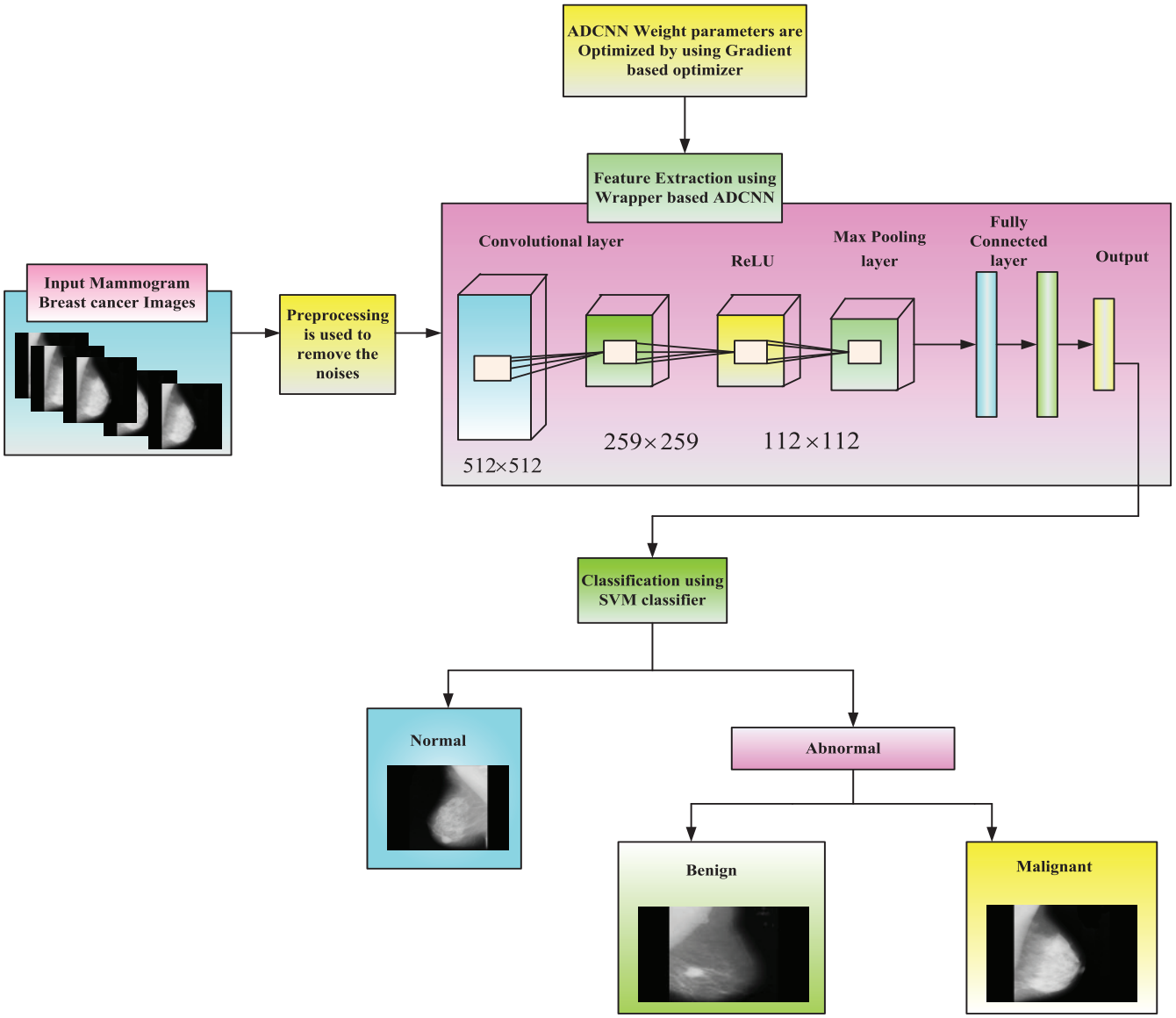

FIGURE 1 Block diagram for Wrapper based feature extraction framework based on AlexNet deep convolutional neural network (ADCNN)–gradient-based optimizer (GBO)-support vector machine (SVM) using mammogram images

24-bit RGB image using standard resolution of 1024 × 1024 pixels. Then, the images are sent to the pre-processing stage to remove the noise. Then image sizes are reduced by the feature extraction process and to avoid matching problems before to the classification process.

3.2 | Pre-processing using Markov random field method for removing noises

Mammograms are noisy and inconsistent to interpret on its original form. To reduce noise from mammography images because numerous images have existing artifacts such as written labels that must to be eliminated and this may be done by cropping the images. Image pruning removes the entire background noise. In this, the mammography images are affected by Gaussian noise and these noises are removed using the MRF method. The MRF is common graphical model for state prediction. Naturally, an image may be viewed as an MRF, with the intensity values being the state and the coordinates being the positions. Thus, MRF may be exploited to remove noise as image. The images affected by noise with the Gaussian noise equation are given in Equation (1)

$$G(n) = \sum_{a=1}^{a} \beta_{a,b}, N\left(X_b^t a(x); 0, h_{b,a} \times \sigma_{b,a}^2\right) \tag{1}$$

where $\beta_{a,b}$ refers weight of Gaussian component, X refers filter bank, $\sigma^2_{b,a}$ refers noise level, h refers Gaussian scale, b, a denotes input image, N denotes the MRF, x denotes the largest clique, t is represented as the time, g is represented as the additive noise in the filter. Then these noises are removed by using e prior probability function prob(m) to rob(n, m) in MRF and its equation is given in Equation (2):

$$p(m)\,\alpha \oplus_m p(n,m)\alpha \prod_a \prod_{x\in X_j} \bigoplus_{m_{ix}} \beta_b \left(g_i^t c(x), h_{ax}\right) \tag{2}$$

where β_b refers half-quadratic expression of potential function, m refers independent auxiliary variable. Here Equation (2) is known as the noise removed mammogram breast image and the overall computational complexity is removed and this image is given to the AlexNet model for extracting the image features for detecting the Breast cancer.

3.3 | ROI extraction

The mammogram breast cancer image consists of full portion of the breast images. From this, the noises are removed by using the pre-processing stage and then extract the particular cancer region with the help of ROI extraction. Here, the background of the image contains the pectoral muscles with the undesired regions. Using this ROI extraction procedure, the system can crop the cancer affected area. Then, the features are extracted to that portion and to find the optimal accuracy.

3.4 | Wrapper-based feature extraction using AlexNet deep convolutional neural network

The feature extraction stage several features are removed depending spectral, texture and contextual features. Here, image features are extracted and the image size is reduced with the help of ADCNN feature extraction method. Here, wrapper based feature removal with ADCNN is used to remove that features to avoid over fitting during classification process.

The Wrapper based feature extraction unit (WFEU) is used to remove the features as input data set. At first the dataset is pre-processed and the features are extracted. In WFEU system consists of three units. From this, the 1st step the datasets images are taken and important features are removed that affected regions and it is noted as $YG_{features}$ array. By using this array then the second array is created and it is given as YG_{thre} threshold function. In 2nd step, the input feature is divided into two unit for selecting the wrapper feature extraction unit. The input vectors are attributed to get the normalized values. The final step, wrapper feature extraction unit trained image is given to ADCNN. By using the WFEU-ADCNN features are removed.

By using ADCNN method the features are removed. From the ROI extracted images, the features are extracted to detect the cancer. AlexNet classifier is well-organized classifiers which are used to identify the issues in the image classification when extracting the features of the diseases. By using ADCNN various wrapper based features are extracted they are given as F1–F14. Then the 14 features with notations are given in Table 1.

Here, the input mammogram breast cancer image 512×512 is send to the convolutional layer. The images are resized to 259×259 pixels with respect to the breadth and height. The convolutional layer contains the L count of layers used to represent feature maps. Then, the feature map is given as $G_k^l\ (h = 1, 2, 3, \ldots, V^l)$. The images are intended in first layer and then classified with the help of 2-D filters X_{lkl}^l is united with the cth feature map I_1^{m-1}. The $(l-1)$th layer and joints the neurons with the pth feature map G_k^l present in lth layer and the bias f_1. Here, the convoluted process is provided among fully the input features map position through filter to detect the features and evaluate the desired features. I_j^m is attained from the $G_k^{l-1}\ (L = 1, 2, 3, \ldots, V^{l-1})$. Here, the input features are convoluted with respect to the filter G_{km}^l. The bias f_K^l is added with the outcomes and applied in the non-linear activity function $\phi(\cdot)$ is provided in the component manner. Then, the feature maps of the convolutional layer is given as below:

$$G_k^l = \beta\left(\sum_{L=1}^{V^{l-1}} G_1^{l-1} \times H_{hl}^l + F_k^l\right),\quad h = 1, 2, \ldots, V^l \tag{3}$$

where $\times$ represents the convolution operation.

Here, ADCNN comprises so many layers, the each layer contains several neurons and some neurons are hidden. Then, the selected features in the hidden neurons of the latter layer may be assumed from vector of features. The framework of the CN deals with the deep-learning algorithm and it consists of 32 filters with the size $3 \times 3 \times 3$ pixels with many neurons. The neuron values are convoluted to the windows present on convolutional layer. The window size in the convolutional layer is represented as 5×5. Here, the count of weight parameters changes accordingly. The features of the images are removed based on weight of the input images. The priority given to the highest weight neurons that is first extracted and the low weight neurons are secondly extracted. Thus, the feature extracted images are stored in the BN layer of the AlexNet. Here the data's are

TABLE 1 Various features extracted from WFEU-ADCNN

Notation	Name	Description
F1	Contrast	The variance on luminance or color creates an object unique
F2	Correlation	A particular number denotes degree of relationship among two variables
F3	Dissimilarity	Variation of pairs of gray levels on image
F4	Energy	Energy returns the sum of the square elements on ADCNN. The energy range is [0 1]
F5	Entropy	Entropy is statistic measure of unpredictability that may be recycled to describe that texture of input image
F6	Maximum probability	It computes gray-level contain maximal probability
F7	Sum of square: variance	Variance assigns moderately high weights vary as average value
F8	Sum average	The relation among pure and dense areas on image
F9	Sum variance	Reveals the spatial heterogeneity
F10	Sum entropy	Sum of micro (local) variances on image
F11	Difference variance	Local changeability measure
F12	Difference entropy	Variability of the micro variances
F13	Inverse difference normalized	Another measure of the local homogeneity of image
F14	Inverse difference moment normalized	It is predictable to be great if gray levels of every pair of pixels are related

Abbreviation: ADCNN, AlexNet deep convolutional neural network.

separated into batches and the image data generator can provide the batch-wise images of iterations in an epoch. ReLU layer is used to employ the element—wise activity functions. This neural architecture used the activation function as $d(L) - \max i(0, l)$ than other common sigmoid functions that is utilized for enhancing the performance and training speed of the proposed model. This function is also known as non-saturated function, which can control the issues in the architecture like gradient departure as well as explosion. Moreover, this function is necessary for large dataset for easy training a deeper network. This layer is used to recharge the negative activity by applying 0 and then applying the non-linearity to the system and the equation is given as.

$$d(L) - \max i(0, l) \;\; \text{FC} = 1000 \tag{4}$$

Then to avoid over fitting problem the convolutional network is computed by the process of the down sampling process with the feature maps using the pooling layer that means samples are reduced with the size of 2×2 spatial parameters. Max-pooling barely accounts for covered representation by removing full non-maximal qualities in the non-overlapping subspace and recovers the performance of images. Here, a maximum pooling layer is used for each convolution layer for the entire purposes, each layer is labeled as 3×3. Moreover, the increased parameters in the convolution layers are improved the feature extraction process to reduce that local features. In the other hand, the parameter of the convolution layer extracts the features. Additionally, the connection layer input is diminished because of the feature graph value simplification process that can reduce the total quantity of nodes in the layer, thus to diminish the parameter ratio in the network.

Additionally, it normalizes the dimensional features of the images with the use of computing variance as well as mean value of the input image samples. In this, two parameters β, φ are introduced for recovering the feature distribution function $(\widetilde{t})$ using parameter training. The calculation formula is mentioned as Equation (5) and (6):

$$p_i = \alpha \widetilde{t_i} + \varphi \tag{5}$$

$$\widetilde{t_i} = \frac{t_i - \mu}{\sqrt{\sigma^2 + \Upsilon}} \tag{6}$$

where μ, σ denotes the sample mean value and variance value. Υ denotes constant value and i represents total number of samples. By this process overfitting problem is avoided in the feature extraction process. Overfitting problems are avoided to improve the classification accuracy. Then to recover the classification accuracy the weight parameters of ADCNN are optimized using the GBO optimization algorithm. Here, (α, φ) represent the parameter used for feature removal. The cancer region is removed as Mammogram breast cancer images features.

3.5 | Optimized ADCNN for breast cancer feature extraction using GBO

Here, GBO is employed to optimize that weight parameters of ADCNN in Feature extraction process. The GPO algorithm is used to obtain optimal accuracy and efficiency, and then decreasing the calculation time. In this, GBO is a metaheuristic optimization algorithm and it is deals with the Gradient based Newton's method and it consists of two important operators such as gradient search rule (GSR) and local escaping operators (LEO). Form a group of vectors to estimate the search space. This GSR algorithm is used to improve the exploration occupancy and get better positions on convergence rate on search space. The LEO is used to get local optimum accuracy and it reduces the computational time.

3.6 | Step by step procedure of proposed ADCNN-GBO algorithm

The Gradient based optimizer is recycled to tune the parameters of the ADCNN. It can tune the parameters using the GSR and LEO and then, form the group of vectors to estimate the search space with the help of Newton's method in the GBO algorithm. The flow chart for GBO algorithm shows a Figure 2. The step-by-step process of GBO algorithm trained ADCNN and SVM are assumed beneath.

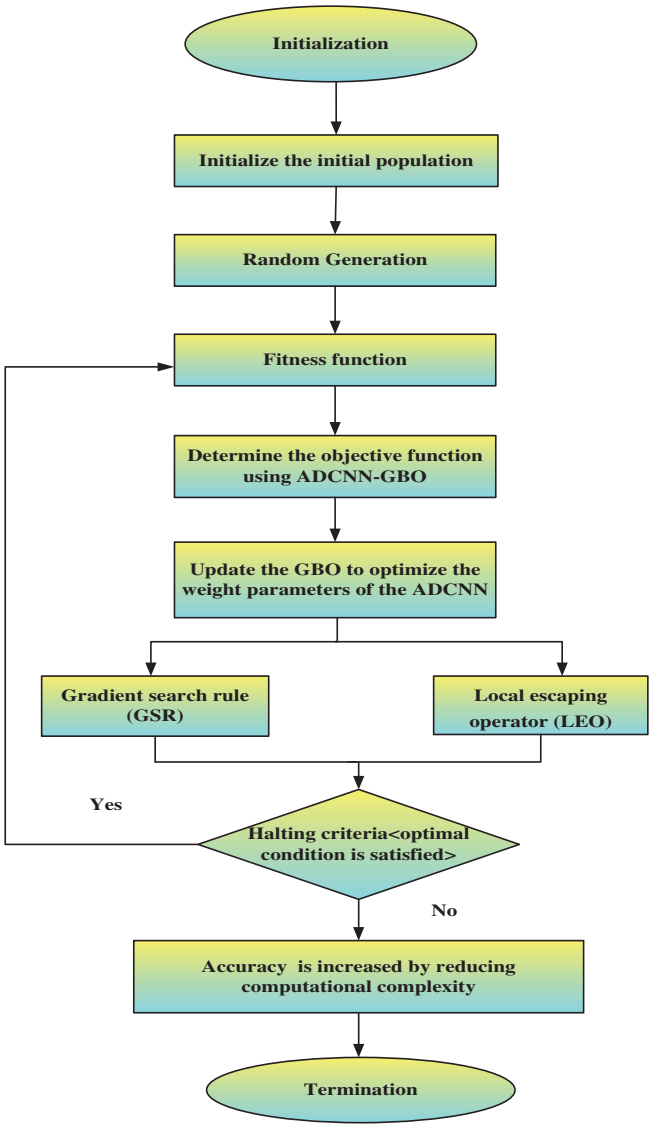

FIGURE 2 Flow chart for gradient based optimization algorithm

***Step 1*:** Initialization.

In this step, initialize the initial parameters based on the decision variables, constraints. The parameters are controlled by the transformations of the explorations to exploitation (α). Then, the iterations and population sizes are estimated and the complexity problem is reduced. In the population, the person represents the vector form. Then, the vectors are mentioned as M within dimension L_D with in the search space. The vector equation is formulated as below:

$$A_{m,f} = \left[A_{m,1}, A_{m,2}, \ldots A_{m,L_D}\right], \quad m = 1, 2, \ldots, M, \quad f = 1, 2, \ldots, F \tag{7}$$

***Step 2*:** Random generation.

From the initialization vectors present in the GBO, the constraints are randomly generated in the dimension L_D with the process of domains in the search space and the equation is given as

$$A_m = A_{\min i} + random(0, 1) \times (A_{\max i} - A_{\min i}) \tag{8}$$

where $A_{\min i}$ and $A_{\max i}$ gives the bounds and gives the decision variables A and the random$(0, 1)$ is represented as the random number [0, 1].

***Step 3*:** Fitness function (FF)

In this, the solution of the random generation is given as the FF values. The FF used to maximize the accuracy and minimizing the computational time and attain the objective function by using GSR, and LEO. Here, (α, φ) represent the parameter used for feature extraction. From this, the fitness value is estimated as $(X(a_n))$ and it is used to the time- consuming in the computations. Then, the GSR is derived as:

$$\mathrm{GSR}_{\mathrm{GBO}} = \mathrm{random} \times \frac{2\Delta x \times a_m}{(\alpha_{\mathrm{worse}} - \alpha_{\mathrm{bst}} + \varphi)} \tag{9}$$

***Step 4*:** Gradient search rule.

In this, GSR to get the better search positions is attained by the motion of the vectors. Here, the GSR is used to improve the exploration phase and by accelerating the conjunction GBO. Where, random is denoted as normally distributed in the random number and the γ is represented as the little amounts within the range of [0,0.1]. a_{bst} and a_{wors} is represented as the best and the worst-case solutions defined in the optimization process and also support the recent solutions to update the positions. Then improve the search capability of the Algorithm and obtaining stable the exploration (global) as well as exploitation (local) and GSR is altered and with the random parameter α_1 and the derived in the below equations:

$$\alpha_1 = 2 \times \mathrm{random} \times \varphi - \varphi \tag{10}$$

Here, to get the stable solutions and global optimum solutions. Then, the GSR is replaced by using the adaptive parameters and the equation is given as

$$\varphi = \left|\alpha \times \sin\left(\frac{3\pi}{2} + \sin\left(\alpha \times \frac{3\pi}{2}\right)\right)\right| \tag{11}$$

$$\alpha = \alpha_{\min i} + \alpha_{\max i} - \alpha_{\min i}) \times \left(1 - \left(\frac{n}{N}\right)^3\right)^2 \tag{12}$$

where $\alpha_{\min i}$ and $\alpha_{\max i}$ is 0.2 and 1.2, individually, n is represented as number of iterations, then N is represented as whole amounts in the iterations. Then for obtaining stable exploration as well as exploitation phase, by considering the parameter α_1 is replaced as the sine function φ. Here, to get the local optimum and by maximizing the diversity of populations to search around to get the best solutions.

From this, GSR helps to explore the GPO and escape the local optima with the help of local generations. Here, Δa is estimated by taking difference among the best solutions and the randomly selected positions a^n_{t1}. To enhance the Δa by changing every iterations with the parameter coefficient μ and the equation is derived in Equation (17) Then, to enhance the exploration a random number is included and the equation is derived as:

$$\Delta a = \mathrm{random}(1 : M) \times |\mathrm{steps}| \tag{13}$$

$$\mathrm{steps} = \frac{\left(a_{\mathrm{bst}} - a^n_{t1}\right) + \mu}{2} \tag{14}$$

$$\mu = 2 \times \mathrm{random} \times \left(\left|\frac{a^n_{s1} + a^n_{s2} + a^n_{s3} + a^n_{s4}}{4} - a^n_{sm}\right|\right) \tag{15}$$

where random$(1:M)$ is represented as the random number with M dimensions, $s1, s2, s3, s4 (s1 \neq s2 \neq s3 \neq s4)$ are the randomly chooses various integer from $(1, M)$, steps is a step size, and it is estimated as $a_{\text{bst}}, a_{t1}^{n}$.

Then, the direction of the movement (DM) is included to get better solutions of the a_m. Then, term chose the best vector and move towards current vector a_m to direction of the $(a_{\text{worse}} - a_m)$. To increase the speed with the search tendency the direction of the movement is used and the equation is given as

$$\text{DM} = \text{random} \times \lambda_2 \times (a_{\text{bst}} - a_m) \tag{16}$$

$$\lambda_2 = 2 \times \text{random} \times \beta - \beta \tag{17}$$

The above equation is known as the GSR and the DM equations. Then to update the current positions $\left(a_m^n\right)$.

$$A1_m^n = a_m^n - \text{GSR} + \text{DM} \tag{18}$$

$$A1_m^n = a_m^n - \text{random} \times \lambda_1 \times \frac{2\Delta a \times a_m^n}{a_{\text{worse}} - a_{\text{bst}} + \mu} + \text{random} \times \lambda_2 \times (a_{\text{bst}} - a_m) \tag{19}$$

where $A1_m^n$ is represented as the new vector generated to update the a_m^n. $A1_m^n$ is generated as the random point and forms the new search space from the GSR and DM.

From the Newton's method to enhance the GSR the equation is given as:

$$a_{m+1} = a_m - \frac{2\Delta a \times X(a_m)}{X(b_m + \Delta a) - X(b_m - \Delta a)} \tag{20}$$

Then, by re-changing the best positions of the vector (a_{bst}) with the vector a_m^n and by substituting the new vector and the equation is given as $A2_m^n$

$$A2_m^n = a_{\text{bst}} - \text{random} \times \lambda_1 \times \frac{2\Delta a \times a_m^n}{\left(bl_m^n - bk_m^n + \gamma\right)} + \text{random} \times \lambda_2 \times \left(a_{t1}^n - a_{t2}^n\right) \tag{21}$$

$$A3_m^n = a_m^n - \alpha_1 \times \left(A2_m^n - A1_m^n\right) \tag{22}$$

where, $A1_m^n$, $A2_m^n$, $A3_m^n$ is known as the 2-D search space. The GBO algorithm is used to improve the exploration and exploitation criteria and to enhance the global search in the exploration criteria and the local search in the exploitation criteria. Then, the solutions are generated in the Equations (21), (23), (24). The accuracy is increased and it is reduces the complexity and time.

***Step 4*:** Local escaping operators.

LEO is recycled to raise that effectiveness and diminish the complexity of system. Then, the positions of the LEO give perfect solutions to estimate the efficiency. To generate the best solutions (a_{bst}) with the solutions $A1_m^n$ and $A2_m^n$ with the two random generation solutions a_{t1}^n and a_{t2}^n with the newly generated solutions a_l^n. Then, the LEO is given as A_{LEO}^n with the uniformly distributed variables and the random variables are given as

$$h_1 = \begin{cases} 2 \times \text{random} & \text{if } \lambda_1 < 0.5 \\ 1, & \text{otherwise} \end{cases} \tag{23}$$

$$h_2, h_3 = \begin{cases} \text{random} & \text{if } \lambda_1 < 0.5 \\ 1, & \text{otherwise} \end{cases} \tag{24}$$

where random is a randomized numbers are represented in the range of [0,1] and λ_1 is represented as the numbers in the range of [0,1]. Then, the above Equation (24) is simplified as

$$\begin{aligned} h_1 &= K_1 \times 2 \times \text{random} + (1 - K_1) \\ h_2, h_3 &= K_1 \times \text{random} + (1 - K_1) \end{aligned} \tag{25}$$

where, K_1 is represented as binary parameter, which takes the value of 0,1.

$$a_k^n = \begin{cases} \text{random} & \text{if } \lambda_1 < 0.5 \\ a_l^n, & \text{otherwise} \end{cases} \tag{26}$$

$$a_l^n = K_2 \times a_u^n + (1 - K_2)\, a_{\text{random}} \tag{27}$$

where K_2 is represented as the binary parameter with [0,1]. Then, λ_1 is lower than the 0.5, $K_2 = 1$ or it is 0. Then, the random behavior with the parameters is h_1, h_2, h_3 setted to get the optimum values.

$$a_{\text{random}} = A_{\min} + \text{random}(0, 1) \times (A_{\max} - A_{\min}) \tag{28}$$

***Step 5*:** Termination.

The optimal weight parameters such as, β, λ, n, A of the ADCNN features are optimized with the aid of GBO. At last, get the objective function efficiency is increased by estimating accuracy, sensitivity, specificity, *F*-score, recall, precision and reduces the computational complexity. The various features are classified on the basis of the area, perimeter, compactness, uniformity, standard deviation, smoothness, entropy, skewness analysis, mean and variance.

3.7 | SVM classifier

Here, the features removed values are assumed with SVM classifier to get more accurate values to detect the breast cancer. To categorize the mammogram images as three classes benign, malignant and normal. This method is used to evade the overfitting problem. It is used to increase the classification accuracy, speed, and parameters are tuned by using the GBO Algorithm. The support vectors of data points are nearer to separate that hyper plane. Then, to select the best hyper plane a kernel function are used and parameters are tuned by using GBO Algorithm. Then, the SVM equation is given as:

$$S(A, B) = \left(\beta \times A^R B + t\right)^f \times \beta > 0 \tag{29}$$

where A, B is represented as the vectors in the input space, t, f, β are Kernel SVM parameters. Then, the classification accuracy is estimated by using this classifier.

4 | RESULT AND DISCUSSION

The simulation performance of Breast cancer is detected in the early stage with mammogram images. The ROI features are removed. The extracted ROI features given to the Wrapper based feature extraction using ADCNN. Then, the removed features are given with SVM classifier, to categorize that mammogram images as three classes: benign, malignant and normal. The classifiers and the neural Network weight parameters are optimized with the help of GBO. The MATLAB simulations are performed on PC with Intel Core i5, 2.50 GHz CPU, 8GB RAM and Windows 7. The proposed system is replicated by MATLAB. Here, assessment metrics as balanced error rate, precision, recall, *F*-score, specificity, accuracy are analyzed.

4.1 | Dataset description

In this, the input images are considered as mini-MIAS database. In the proposed system, the Mammogram mini-MIAS database is used. This data set has 322 mammograms: 270 sample images are normal and remaining 52 samples are malignant. In this the abnormal images may be benign or malignant. Eighty percentage of the data set is used for training and 20% for evaluation.

4.2 | Performance metrics

To measure precision, recovery, *F*-measure, accuracy, specificity, error rate, the confusion matrix is utilized. To measure that confusion matrix, true negative, true positive, false negative, and false positive values are necessary.

- True positive (ψ): Breast cancer features are categorized as abnormal and abnormal.
- True negative (ζ): Breast cancer features are categorized as normal and normal.
- False positive (φ): Breast cancer features are categorized as abnormal and normal.
- False negative (η): Breast cancer features are categorized as normal and abnormal.

4.2.1 | Precision

The precision are named positive predictive values, it may be resolute through the aid of Equation (30).

$$\text{Precesion value} = \frac{\psi}{\psi + \phi} \tag{30}$$

4.2.2 | *F*-score

F score may be resolute through the aid of subsequent Equation (31)

$$F - \text{score value} = \frac{\psi}{\psi + \frac{1}{2(\varphi+\zeta)}} \tag{31}$$

4.2.3 | Accuracy

The accuracy values resolute with the support of subsequent Equation (32)

$$\text{Accuracy} = \frac{\psi + \zeta}{\psi + \zeta + \varphi + \theta} \tag{32}$$

4.2.4 | Balanced error rate

The balanced error rate may be resolute through the aid of subsequent Equation (33)

$$\text{Balanced error rate value} = 1 - 0.5 \times \frac{\text{Recall value} + \text{Specificity value}}{100} \tag{33}$$

4.3 | Comparison of performance analysis through numerous methods

In Figure 3 shows, the of mammogram images with and without breast cancer. In Here, the input images are pre-processed. Then, the noises are removed using pre-processing stage. The features are extracted using Wrapper based ADCNN method. The ROI Features are extracted.

It is cropping particular affected regions and removes other regions. Then, the features are extracted to get the exact affected area. Here, the features are extracted to detect the breast cancer on primary stage from mammogram images and it improves the development of the medical field. The wrapper based ADCNN is used to extract the features. The feature extracted values are mentioned in the Figure 3 for the maximum and minimum range value. Here, the trained data is given to the classifier to get the accurate output benign, malignant or normal image.

4.3.1 | Comparison of performance analysis through numerous classifiers

The performance as accuracy, precision, recall and *f*-measure, specificity, sensitivity analysis is likened with several feature extraction methods. The proposed FE-ADCNN-GBO-SVM is likened with numerous feature extraction existing methods such as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively. Tables 1–7 demonstrates the performance analysis of several feature extraction methods recycled for breast cancer, Figure 4 demonstrates the receiver operating characteristics (ROC) curve for feature extraction methods.

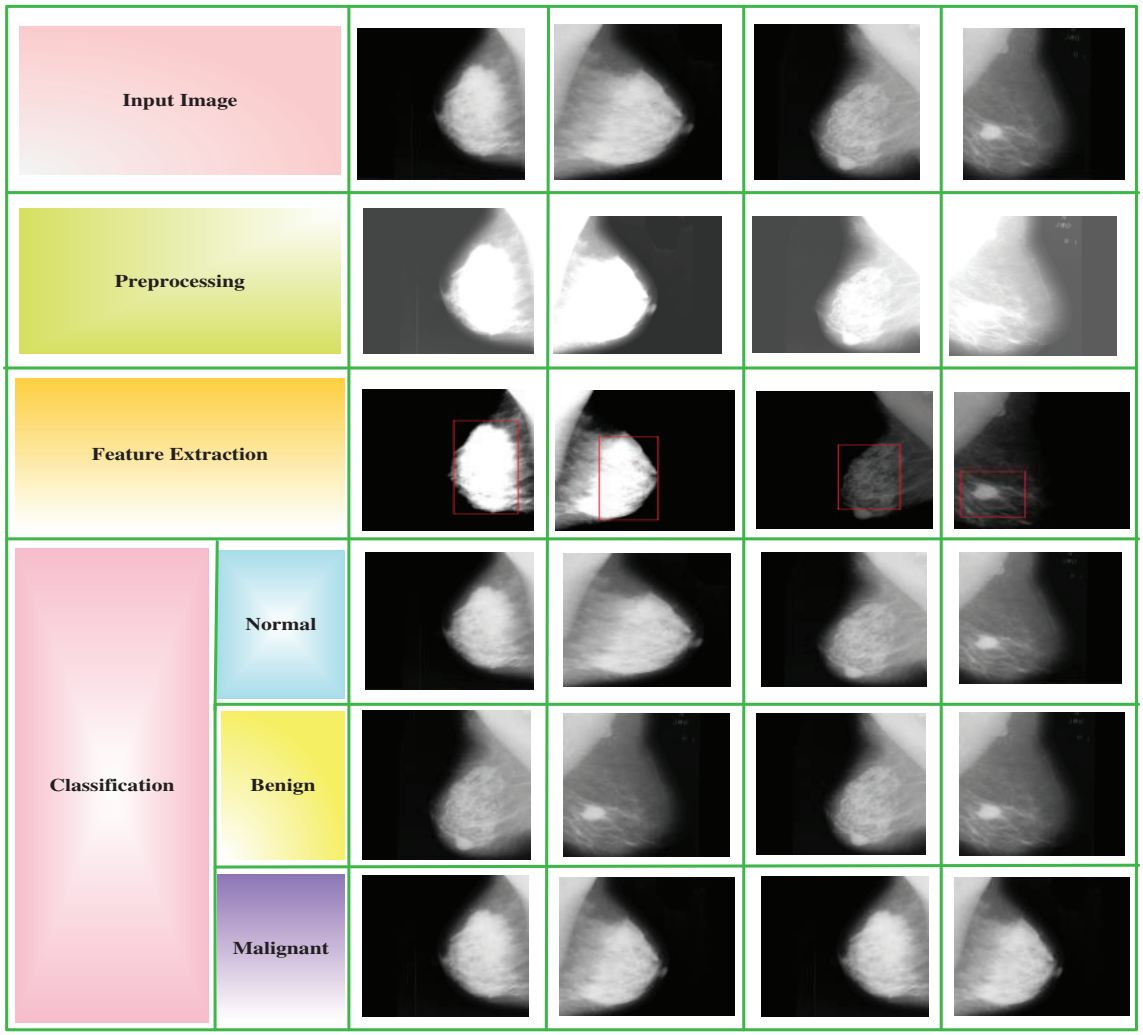

FIGURE 3 Feature extracted and classification output images using mammogram images

TABLE 2 Performance analysis of using mammogram breast cancer images

Diseases	FE-2-D-BDWT-GLCM-FOA-SVM	FE-LBP-GLCM-SVM	FE-GLCM-ANN	FE-SMOTE-RF	FE-CNN-CDCNN	FE-ADCNN-GBO-SVM (proposed)
Normal	73	48	62	57	65	97
Malignant	61	57	59	76	77	98
Benign	55	64	76	62	85	99

From Table 2 shows that performance metrics of Accuracy for detecting breast cancer using mammogram images as normal, malignant, benign and proposed FE-ADCNN-GBO-SVM method is comparing with the numerous existing process as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For normal analysis, accuracy of proposed system portrays 25.97%, 33.42%, 23.86%, 33.09%, and 25.64% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For malignant analysis, accuracy of proposed method shows 36.84%, 45.42%, 15.94%, 47.94%, and 29.04% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For benign analysis, accuracy of proposed system portrays 37.02%, 29.46%, 20.94%, 39.22%, and 32.03% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively.

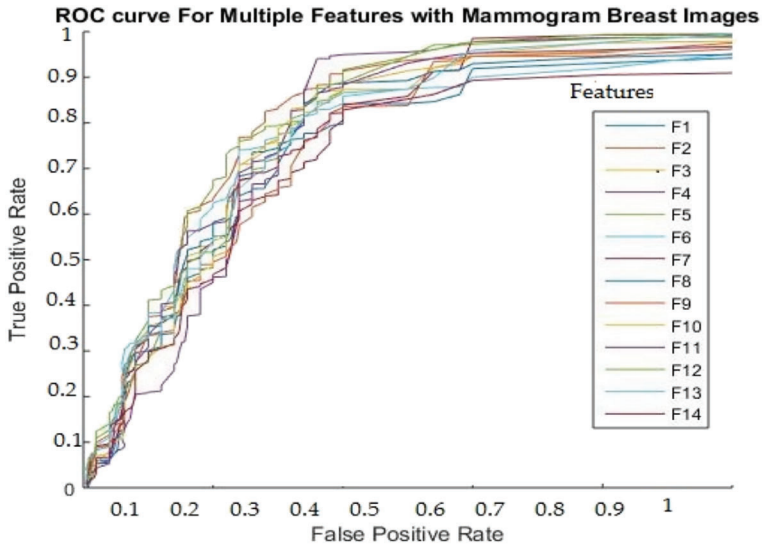

FIGURE 4 Receiver operating characteristics (ROC) curve multiple features using mammogram breast cancer image

TABLE 3 Performance analysis *F*-score using mammogram breast cancer images

Diseases	FE-2-D-BDWT-GLCM-FOA-SVM	FE-LBP-GLCM-SVM	FE-GLCM-ANN	FE-SMOTE-RF	FE-CNN-CDCNN	FE-ADCNN-GBO-SVM (proposed)
Normal	77	56	63	50	56	92
Malignant	59	63	75	76	71	91
Benign	55	56	56	43	54	95

From Table 3 shows that performance metrics of the *F*-score for detecting breast cancer using mammogram images as normal, malignant, benign and proposed FE-ADCNN-GBO-SVM method is comparing with the several existing systemlikeFE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For normal analysis, *F*-score of proposed system portrays 23.55%, 39.83%, 21.09%, 22.43%, and 31.04% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For malignant analysis, *F*-score of the proposed system portrays 29.03%, 31.43%, 24.32%, 28.42%, and 38.94% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For benign analysis, *F*-score of proposed system portrays 28.94%, 38.92%, 32.93%, 19.03%, and 32.93% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively.

From Table 4 shows that performance metrics of the precision for detecting breast cancer using mammogram images as normal, malignant, benign and proposed FE-ADCNN-GBO-SVM method is comparing with the several existing process as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For normal analysis, the precision of proposed system portrays 34.98%, 35.08%, 42.97%, 27.98%, and 26.97% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For malignant analysis, precision of proposed system portrays 25.09%, 32.94%, 25.04%, 21.87%, and 33.97% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For benign analysis, precision of proposed system portrays 32.09%, 24.09%, 25.03%, 26.98%, and 31.22% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively.

From Table 5 shows that performance metrics of sensitivity for detecting breast cancer using mammogram images as normal, malignant, benign and the proposed FE-ADCNN-GBO-SVM method is comparing with the several existing process as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively. For normal analysis, sensitivity of proposed system displays 27.84%, 33.84%, 27.84%, 17.38%, and 27.94% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For malignant analysis, sensitivity of proposed system displays 27.94%, 24.63%, 31.32%, 34.67%, and 17.34% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN,

TABLE 4 Performance analysis precision using mammogram breast cancer images

Diseases	FE-2-D-BDWT-GLCM-FOA-SVM	FE-LBP-GLCM-SVM	FE-GLCM-ANN	FE-SMOTE-RF	FE-CNN-CDCNN	FE-ADCNN-GBO-SVM (proposed)
Normal	78	64	71	62	59	93
Malignant	65	67	64	79	74	90
Benign	47	75	59	58	56	96

TABLE 5 Performance analysis sensitivity using mammogram breast cancer images

Diseases	FE-2-D-BDWT-GLCM-FOA-SVM	FE-LBP-GLCM-SVM	FE-GLCM-ANN	FE-SMOTE-RF	FE-CNN-CDCNN	FE-ADCNN-GBO-SVM (proposed)
Normal	48	53	72	63	52	95
Malignant	69	51	53	72	64	94
Benign	77	48	59	63	72	98

TABLE 6 Performance analysis specificity using mammogram breast cancer images

Diseases	FE-2-D-BDWT-GLCM-FOA-SVM	FE-LBP-GLCM-SVM	FE-GLCM-ANN	FE-SMOTE-RF	FE-CNN-CDCNN	FE-ADCNN-GBO-SVM (proposed)
Normal	55	64	72	49	63	90
Malignant	64	72	63	63	64	91
Benign	65	64	52	51	59	97

FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For benign analysis, sensitivity of proposed system displays 34.96%, 2.04%, 42.44%, 22.94%, and 19.03% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively.

From Table 6 shows that the performance metrics of the specificity for detecting breast cancer using mammogram images as normal, malignant, benign and proposed FE-ADCNN-GBO-SVM method is comparing with several existing process as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For normal analysis, specificity of proposed system displays 32.76%, 31.87%, 19.08%, 27.86%, and 29.76% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For malignant analysis, specificity of proposed system displays 28.06%, 20.97%, 32.75%, 28.65%, and 26.44% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For benign analysis, specificity of proposed system displays 37.09%, 27.76%, 42.56%, 31.98%, and 33.76% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively.

From Table 7 shows that the performance metrics of the recall for detecting breast cancer using mammogram images as normal, malignant, benign and the proposed FE-ADCNN-GBO-SVM method is comparing with the several existing process as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For normal analysis, the recall of proposed system displays 23.55%, 39.83%, 21.09%, 22.43%, and 31.04% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For malignant analysis, recall of proposed system displays 29.03%, 31.43%, 24.32%, 28.42%, and 38.94% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For benign analysis, recall of proposed system displays 32.03%, 37.32%, 24.93%, 31.04%, and 32.09% greater than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively.

From Table 8 shows that the performance metrics of the error rate for detecting breast cancer using mammogram images as normal, malignant, benign and the proposed FE-ADCNN-GBO-SVM method is comparing with the several existing process as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively. For normal analysis, error rate of proposed system displays 25.96%, 43.92%, 27.09%, 28.69%, and 34.87% lower than the existing such as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively. For malignant analysis, error rate of proposed system displays 34.92%, 32.09%, 28.78%, 34.82%, and 26.98% lower than the existing such as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly. For benign analysis, error rate of proposed system displays 24.84%, 17.04%, 27.94%,

TABLE 7 Performance analysis recall using mammogram breast cancer images

Diseases	FE-2-D-BDWT-GLCM-FOA-SVM	FE-LBP-GLCM-SVM	FE-GLCM-ANN	FE-SMOTE-RF	FE-CNN-CDCNN	FE-ADCNN-GBO-SVM (Proposed)
Normal	45	63	55	67	59	93
Malignant	65	47	70	74	68	96
Benign	61	41	62	54	64	98

TABLE 8 Performance analysis error rate using mammogram breast cancer images

Diseases	FE-2-D-BDWT-GLCM-FOA-SVM	FE-LBP-GLCM-SVM	FE-GLCM-ANN	FE-SMOTE-RF	FE-CNN-CDCNN	FE-ADCNN-GBO-SVM (Proposed)
Normal	0.049	0.035	0.042	0.035	0.042	0.02
Malignant	0.043	0.053	0.035	0.043	0.031	0.021
Benign	0.032	0.042	0.038	0.053	0.039	0.019

31.02%, and 33.12% lower than the existing such as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively.

Figure 4 shows the ROC curve multiple features through mammogram breast cancer image. The 14 features are F1 represents the contrast, F2 represents the correlation, F3 represents the dissimilarity, F4 represents the energy, F5 represents the entropy, F6 represents the maximum probability, F7 represents the sum of square: variance, F8 represents the sum average, F9 represents the sum variance, F10 represents the sum entropy, F11 represents the difference variance, F12 represents the difference entropy, F13 represents the inverse difference normalized, F14 represents the inverse difference moment normalized. It is calculated based on total count of true positive rate splits that total count of false positive rate. The values 0.0 and 1.0. TPR is demarcated as proportion of positive data points are properly forecast as positive. It is also referred as recall. FPR is demarcated as proportion of negative data points falsely forecast as positive.

4.4 | Discussion

Table 2 shows that performance of accuracy for detecting the mammogram breast cancer images on basis of normal, malignant, and benign and proposed FE-ADCNN-GBO-SVM method is comparing with the numerous existing process as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively.. Here the accuracy of normal shows 23.55%, 35.97%, 28.97%, 25.73%, 23.13%, the accuracy of malignant, shows 25.96%, 32.97%, 36.91%, 26.97%, 15.97% the accuracy of benign shows 26.97%, 43.97%, 24.54%, 25.74%, 31.83% greater than existing approaches. Table 3 shows that performance of *F*-score for detecting the mammogram breast cancer images on normal, malignant, and benign. Here the *F*-score of normal shows 24.86%, 17.97%, 32.97%, 27.97%, 42.84%, the *F*-score of malignant, shows 24.86%, 29.08%, 34.97%, 39.75%, 19.08% the *F*-score of benign shows 23.86%, 27.93%, 43.97%, 27.97%, 29.96% greater than existing approaches. Table 4 shows that performance of precision for detecting mammogram breast cancer images based on normal, malignant, and benign. Here the precision of normal shows 25.97%, 34.86%, 32.97%, 31.97%, 26.97%, the precision of malignant, shows 18.97%, 54.53%, 33.64%, 26.86%, 31.97% the precision of benign shows 25.97%, 29.08%, 28.97%, 26.97%, 42.97% greater than existing approaches. Table 5 shows the performance of sensitivity for detecting the mammogram breast cancer images on normal, malignant, and benign. Here the sensitivity of normal shows 32.86%, 26.98%, 32.98%, 29.08%, 42.98%, the sensitivity of malignant, shows 25.97%, 26.97%, 36.86%, 25.07%, 17.86% the sensitivity of benign shows 32.97%, 30.94%, 24.94%, 31.93%, and 31.94% greater than existing approaches. Table 6 shows that performance of specificity for detecting the mammogram breast cancer images on normal, malignant, and benign. Here the specificity of normal shows 26.97%, 33.92%, 33.34%, 25.97%, 31.65%, the specificity of malignant, shows 21.97%, 32.56%, 28.97%, 32.54%, 19.08% the specificity of benign shows 32.97%, 31.54%, 29.08%, 33.86%, 16.97% greater than existing approaches. Table 6 shows that performance of recall for detecting the mammogram breast cancer images on normal, malignant, and benign. Here the recall of normal shows 32.86%, 31.98%, 42.87%, 37.97%, 28.97%, the recall of malignant, shows 25.97%, 18.96%, 33.97%, 26.97%, 27.97% the recall of benign shows 32.86%, 21.86%, 26.97%, 26.71%, 33.86% greater than the existing approaches. Table 7 displays the performance of error rate for detecting the mammogram breast cancer images on normal, malignant, and benign. Here the error rate of normal shows 34.92%, 32.09%, 28.78%, 34.82%, and 26.98% the error rate of malignant, shows 32.03%, 37.32%, 24.93%, 31.04%, and 32.09% the error rate of benign shows 23.55%, 39.83%, 21.09%, 22.43%, and 31.04% lower than existing as FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN respectively.

5 | CONCLUSION

In this manuscript, a Wrapper based feature extraction framework based on ADCNN parameters optimized using GBO for mammogram images is proposed for early detection of breast cancer is successfully implemented. The simulation process is accomplished on MATLAB platform. The proposed ADCNN and SVM with GBO attains higher recall 32.86%, 31.98%, 42.87%, 37.97%, 28.97%, higher *F*-measure 21.97%, 32.56%, 28.97%, 32.54%, 19.08% and the proposed system is compared with existing processes as, FE-2-D-BDWT-GLCM-FOA-SVM, FE-LBP-GLCM-SVM, FE-GLCM-ANN, FE-SMOTE-RF and FE-CNN-CDCNN correspondingly.

The restriction of this work is classification efficiency of proposed model is proportionate with number of training images; a small image dataset will disturb their enactment. Though, the proposed system accomplishes high classification on great image dataset has a high computational cost. Breast cancer may be weakly restricted, so future work is suggested to recover accuracy through lower computational cost.

In future work, improve the classification accuracy of present work and observe that selective scheme of classifier by combining more than one classifier and feature selection procedures.

DATA AVAILABILITY STATEMENT

Data sharing is not appropriate to this article as no novel data are generated or analysed under this study.

ORCID

N. K. Sakthivel https://orcid.org/0000-0003-1772-9204

REFERENCES

1. Akram M, Iqbal M, Daniyal M, Khan A. Awareness and current knowledge of breast cancer. *Biol Res*. 2017;50(1):33.
2. Sakthivel NK, Gopalan NP, Subasree S. G-HR: gene signature based HRF cluster for predicting human diseases. *Int J Pure Appl Math*. 2017;117(9):157-161.
3. Sharma R. Breast cancer incidence, mortality and mortality-to-incidence ratio (MIR) are associated with human development, 1990–2016: evidence from global burden of disease study 2016. *Breast Cancer*. 2019;26(4):428-445.
4. Sopik V, Narod S. The relationship between tumour size, nodal status and distant metastases: on the origins of breast cancer. *Breast Cancer Res Treat*. 2018;170(3):647-656.
5. Sakthivel N, Gopalan N, Subasree S. IHDGAP: deep learning based intelligent human diseases-gene association prediction technique for high dimensional human diseases data sets. *J Eng Appl Sci*. 2019;14(21):8072-8079.
6. Carioli G, Malvezzi M, Rodriguez T, Bertuccio P, Negri E, La Vecchia C. Trends and predictions to 2020 in breast cancer mortality in Europe. *The Breast*. 2017;36:89-95.
7. Nir G, Hor S, Karimi D, et al. Automatic grading of prostate cancer in digitized histopathology images: learning from multiple experts. *Med Image Anal*. 2018;50:167-180.
8. Sheeba P, Murugan S. Fuzzy dragon deep belief neural network for activity recognition using hierarchical skeleton features. *Evol Intell*. 2019;1-8.
9. Zakirov T, Galeev A, Statsenko E, Khaidarova L. Calculation of filtration characteristics of porous media by their digitized images with the use of lattice Boltzmann equations. *J Eng Phys Thermophys*. 2018;91(4):1069-1078.
10. Sakthivel NK, Gopalan NP, Subasree S. Deep learning based human diseases pattern prediction technique for high dimensional human diseases data sets. *J Eng Appl Sci*. 2019;14:8072-8079.
11. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48(4):441-446.
12. Toğaçar M, Ergen B, Cömert Z. Detection of lung cancer on chest CT images using minimum redundancy maximum relevance feature selection method with convolutional neural networks. *Biocybern Biomed Eng*. 2020;40(1):23-39.
13. Subasree S, Gopalan NP, Sakthivel NK. Smart multi-objective particle swarm optimizer for cancer patterns classification and prediction. *Int J Eng Res Technol*. 2018;11(4):661-673.
14. Varuna Shree N, Kumar T. Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network. *Brain Inform*. 2018;5(1):23-30.
15. Mythili S, Thiyagarajah K, Rajesh P, Shajin FH. Ideal position and size selection of unified power flow controllers (UPFCs) to upgrade the dynamic stability of systems: an antlion optimiser and invasive weed optimisation algorithm. *HKIE Trans*. 2020;27(1):25-37.
16. Rajesh P, Shajin F. A multi-objective hybrid algorithm for planning electrical distribution system. *Eur J Electr Eng*. 2020;22(4–5):224-509.
17. Shajin FH, Rajesh P. Trusted secure geographic routing protocol: outsider attack detection in mobile ad hoc networks by adopting trusted secure geographic routing protocol. *Int J Pervasive Comput Commun*. 2020.
18. Thota MK, Shajin FH, Rajesh P. Survey on software defect prediction techniques. *Int J Appl Sci Eng*. 2020;17(4):331-344.
19. Khan S, Islam N, Jan Z, Ud Din I, Rodrigues J. A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. *Pattern RecognitLett*. 2019;125:1-6.
20. Amin J, Sharif M, Yasmin M, Fernandes S. Big data analysis for brain tumor detection: deep convolutional neural networks. *Future Gener Comput Syst*. 2018;87:290-297.
21. Wang C, Elazab A, Wu J, Hu Q. Lung nodule classification using deep feature fusion in chest radiography. *Comput Med Imaging Graph*. 2017;57:10-18.
22. Wahab N, Khan A. Multifaceted fused-CNN based scoring of breast cancer whole-slide histopathology images. *Appl Soft Comput*. 2020;97:106808.
23. Subasree S, Gopalan NP, Sakthivel NK. A comparative study and analysis of data mining classifiers for microarray based cancer pattern diagnostics. *Proceedings of the International Conference on Informatics and Analytics*. Association for Computing Machinery; 2016:1-5.
24. Kasongo S, Sun Y. A deep learning method with wrapper based feature extraction for wireless intrusion detection system. *Comput Secur*. 2020;92:101752.

25. Shanthi T, Sabeenian R. Modified Alexnet architecture for classification of diabetic retinopathy images. *Comput Electr Eng*. 2019;76:56-64.
26. Zhang J, Sun Y, Guo L, Gao H, Hong X, Song H. A new bearing fault diagnosis method based on modified convolutional neural networks. *Chin J Aeronaut*. 2020;33(2):439-447.
27. Arora R, Rai P, Raman B. Deep feature–based automatic classification of mammograms. *Med Biol Eng Comput*. 2020;58(6):1199-1211.
28. Rajathi G. Optimized radial basis neural network for classification of breast cancer images. *Current Medical Imaging*. 2021;17(1):97-108.
29. Lahoura V, Singh H, Aggarwal A, et al. Cloud computing-based framework for breast cancer diagnosis using extreme learning machine. *Diagnostics*. 2021;11(2):241.
30. Irfan R, Almazroi AA, Rauf HT, Damaševičius R, Nasr EA, Abdelgawad AE. Dilated semantic segmentation for breast ultrasonic lesion detection using parallel feature fusion. *Diagnostics*. 2021;11(7):1212.
31. Kadry S, Damaševičius R, Taniar D, Rajinikanth V, Lawal IA. Extraction of tumour in breast MRI using joint thresholding and segmentation–a study. *2021 Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII)*. IEEE; 2021:1-5.
32. Rajinikanth V, Kadry S, Taniar D, Damaševičius R, Rauf HT. Breast-cancer detection using thermal images with marine-predators-algorithm selected features. *2021 Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII)*. IEEE; 2021:1-6.
33. Saeed EM, Saleh HA, Khalel EA. Classification of mammograms based on features extraction techniques using support vector machine. *Comput Sci Inf Technol*. 2021;2(3):121-131.
34. Tariq N, Abid B, Qadeer KA, Hashim I, Ali Z, Khosa I. Breast cancer classification using global discriminate features in mammographic images. *Breast Cancer*. 2019;10(2):1-7.
35. Abdulla SH, Sagheer AM, Veisi H. Improving breast cancer classification using (SMOTE) technique and pectoral muscle removal in mammographic images. *MENDEL*. 2021;27(2):36-43.
36. Wang Y, Yang F, Zhang J, Wang H, Yue X, Liu S. Application of artificial intelligence based on deep learning in breast cancer screening and imaging diagnosis. *Neural Comput Appl*. 2021;1-1:9637-9647.
37. Ahmadianfar I, Bozorg-Haddad O, Chu X. Gradient-based optimizer: a new metaheuristic optimization algorithm. *Inf Sci (NY)*. 2020;540:131-159.
38. Ismaeel A, Houssein E, Oliva D, Said M. Gradient-based optimizer for parameter extraction in photovoltaic models. *IEEE Access*. 2021;9:13403-13416.
39. Mohanty F, Rup S, Dash B, Majhi B, Swamy M. Mammogram classification using contourlet features with forest optimization-based feature selection approach. *Multimed Tools Appl*. 2018;78(10):12805-12834.
40. Mohanty F, Rup S, Dash B, Majhi B, Swamy M. Digital mammogram classification using 2D-BDWT and GLCM features with FOA-based feature selection approach. *Neural Comput Appl*. 2019;32(11):7029-7043.
41. Khan S, Khan A, Maqsood M, Aadil F, Ghazanfar M. Optimized Gabor feature extraction for mass classification using cuckoo search for big data E-healthcare. *J Grid Comput*. 2018;17(2):239-254.
42. Shree NV, Kumar TN. Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network. *Brain Inf*. 2018;5(1):23-30.
43. Liu Z, Lai Z, Ou W, Zhang K, Zheng R. Structured optimal graph based sparse feature extraction for semi-supervised learning. *Signal Process*. 2020;170:107456.
44. Horn Z, Auret L, McCoy J, Aldrich C, Herbst B. Performance of convolutional neural networks for feature extraction in froth flotation sensing. *IFAC-Papers OnLine*. 2017;50(2):13-18.

How to cite this article: Sakthivel NK, Subasree S, Malik S, Tyagi AK. A Wrapper based feature extraction framework based on AlexNet deep convolutional neural network parameters optimized using gradient-based optimizer for mammogram images. *Concurrency Computat Pract Exper*. 2022;e7008. doi: 10.1002/cpe.7008