

Detection of breast cancer using the infinite feature selection with genetic algorithm and deep neural network

S. S. Ittannavar¹ · R. H. Havaldar²

Accepted: 21 July 2021 © The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

The breast cancer is a major health issue worldwide, so the early detection of abnormalities decreases the mortality rate. For the early detection of breast cancer, a new model is proposed in this research using mammogram images which is an effective technique used for screening and detecting the breast cancer. At first, the images are acquired from the digital database for screening mammography and mammographic image analysis society datasets. Then, the visual quality of the images is improved by using normalization, contrast limited adaptive histogram equalization and median filter. Further, multi-level multi objective electro magnetism like optimization algorithm is proposed to segment the non-cancer and cancer regions from the enhanced images. Additionally, feature vectors are extracted from the segmented regions using local directional ternary pattern, histogram of oriented gradients and Haralick texture features. Next, infinite feature selection with genetic algorithm is applied to select the active or relevant features for breast cancer classification. The genetic algorithm is applied with entropy value to measure the homogeneity to find the mutual information between the extracted features that helps in provide prominent feature values. The genetic algorithm has the advantage of probabilistic transition rule that helps to analysis the associate neighbour features. The selected feature vectors are fed to deep neural network to classify the mammogram images as malignant and benign classes. From the simulation result, the proposed infinite feature selection with genetic algorithm and deep neural network showed 0. 10-7% improvement in accuracy related to the existing models such as an extreme learning machine, conditional generative adversarial network with convolutional neural network, etc.

Keywords Breast cancer detection \cdot Deep neural network \cdot Electro magnetism like optimization \cdot Histogram of oriented gradients \cdot Infinite feature selection with genetic algorithm \cdot Local directional ternary pattern

S. S. Ittannavar ssittannavar.ece@hsit.ac.in

Extended author information available on the last page of the article

1 Introduction

In recent decades, breast cancer is the common cancer type that causes frequent death among women. Mostly, the breast cancer is found in women around forty ages and the American cancer society estimated 40,730 deaths in 2015 due to breast cancer [1, 2]. So, early recognition of breast cancer decreases the healthcare costs over time and the best way to save lives [3]. Abnormalities like the breast aches, breast mass, the difference in the color of the breast skin, change in breast shape, the dimension of the breast, etc., are the symptoms of breast cancer [4]. Currently, several imaging techniques are available in breast cancer detection; microwave imaging, photoacoustic imaging, mammogram, magnetic resonance imaging, computed tomography, positron emission tomography etc. In the available imaging techniques, mammography is the best technique to diagnose breast cancer that utilizes a low dose X-ray radiation [5, 6]. The mammogram images reveal the pronounced evidence of the abnormalities like calcification, architectural distortion, bilateral asymmetry, and masses [7, 8]. The detection of breast cancer in the early stage by classifying it as malignant and benign reduces the mortality rate of the patients. Recently, several methods are developed by the researchers in the automatic breast cancer segmentation and classification on mammogram images such as firefly updated chick-based chicken swarm optimization [9], forest optimization algorithm [10, 11], bat optimized run length networks [12], etc. In the case of young women, the developed methods tend difficult to segment the dense breast tissues, because it appears to be gray or white color in nature. The objective of the research is to develop a new model to segment and classify the mammogram images for early diagnosis of breast cancer. The contribution of the paper is discussed as follows.

- Initially, mammogram images are acquired from the Mammographic Image Analysis Society (MIAS) and Digital Database for Screening Mammography (DDSM) databases. Further, Contrast Limited Adaptive Histogram Equalization (CLAHE), normalization and median filter are employed to enhance the visibility level of the acquired mammogram images.
- Then, the cancerous and non-cancerous regions are segmented by multi-level multi objective Electro Magnetism Like (EML) optimization algorithm. Related to existing segmentation algorithms, the proposed segmentation algorithm has the benefits like adaptive to local context, enhanced robustness for holding the image details, and independent clustering parameters with low computation cost.
- Additionally, Haralick texture features, Local Directional Ternary Pattern (LDTP), and Histogram of Oriented Gradients (HOG) are undertaken to extract the features from the segmented cancerous and non-cancerous regions. Next, Infinite Feature Selection with Genetic Algorithm (IFSGA) is used to eliminate the redundant features that significantly reduces the "curse of dimensionality" concern.
- The genetic algorithm with entropy value is proposed to find the mutual information between the extracted features that helps to provide prominent features

for the classification. The genetic algorithm has the advantage of probabilistic transition rule that helps to analysis the association of neighbour features.

- The selected optimal features are fed to Deep Neural Network (DNN) classifier for classifying the malignant and benign classes. The DNN classifier comprises of two key advantages such as able to deliver high quality results, and computationally effective. Finally, the performance of proposed IFSGA-DNN model is analyzed by means of the False Positive Rate (FPR), False Discovery Rate (FDR), False Omission Rate (FOR), precision, accuracy, and recall.
- The proposed IFSGA-DNN method has the accuracy of 97.43% accuracy in the MIAS dataset and existing CNN has 90.5% accuracy in fivefold cross-validation. The proposed IFSGA-DNN model has advantage of probabilistic transition rule in the genetic algorithm with entropy value helps to find the mutual information in extracted features and provide the prominent feature value for the classification.

Some recent research papers on breast cancer segmentation and classification are surveyed in Sect. 2. The proposed IFSGA-DNN model is mathematically explained in Sect. 3. Experimental results and discussion of the proposed IFSGA-DNN model is detailed in Sect. 4. The conclusion of the research is given in Sect. 5.

2 Literature survey

Vijayarajeswari et al. [13] presented a new model for breast cancer detection on the basis of Hough transform and Support Vector Machine (SVM) classifier. The Hough transform was a 2D transform that was utilized for isolating the features of an image. In the model, Hough transform was utilized for detecting the features of mammogram images and then the features were classified by applying an SVM classifier. In this literature study, totaling 95 mammogram images were acquired and classified by using the SVM classifier. The experimental analysis showed that the developed model was effective in classifying the abnormal classes of the mammogram. However, SVM classifier was not suitable for larger databases and also it does not perform well in classification, while the target classes were over-lapped. Singh et al. [14] implemented a new model for breast cancer detection that includes two stages. Initially, conditional Generative Adversarial Network (cGAN) was applied to segment the breast cancer regions and then Convolutional Neural Network (CNN) was used for classifying four cancer shapes such as round, lobular, irregular and oval. Usually, the CNN requires more number of images to attain better performance in image classification and also it was computationally expensive. In addition, Li et al. [15] introduced DenseNet II neural network model for precise classification of malignant and benign images. Initially, normalization technique was applied to improve the visual quality of the acquired mammogram images and then DenseNet II neural network model was employed for classifying the malignant and benign images. The developed model includes two major concerns; requires more memory space and computationally expensive.

Mohanty et al. [16] developed a new model for classifying the mammograms into abnormal or normal, and further malignant or benign. In this model, principal component analysis, energy, Tsallis entropy, Renyi entropy and Shannon entropy were employed for extracting the feature from the acquired mammogram images. Next, a wrapper based kernel extreme learning machine with weighted chaotic salp swarm method was developed to classify the mammograms as abnormal or normal and malignant or benign. The developed model performance was validated on Breast Cancer Digital Repository (BCDR), DDSM, and MIAS datasets in order to evaluate its efficacy. The computational complexity of the developed framework was worst, while combining a number of methodologies. Sadad et al. [17] used a combination of region growing and fuzzy c means algorithm to segment the cancer regions from the mammogram images. Additionally, local phase quantization, local binary pattern, and Gray level co-occurrence matrix features were undertaken to extract the feature vectors from the segmented cancer regions. At last, decision tree and k-nearest neighbor classifiers were used to classify the mammogram images as malignant or benign. Generally, the decision tree classifier consumes more time for training the model, which was a major concern in the application like image classification.

Ting et al. [18] used CNN to improve the performance of breast cancer detection for assisting the clinicians in early diagnosis of breast cancer. The developed CNN model classifies the mammogram images as healthy, benign and malignant patients. Without the prior information of cancer lesion, the CNN model significantly classifies the mammogram images into malignant patient, healthy patient and benign patient. As previously mentioned, CNN was computationally expensive and it needs more number of images to attain better performance in image classification. Kaur et al. [19] developed a new system for breast cancer detection using mammogram images. Initially, the acquired mammogram images were denoised by using resizing operations and noise handling. Next, speed up robust features and k-means clustering algorithm were applied to extract the feature values from the denoised mammogram images. Finally, Multi-SVM (MSVM) classifier was used for classifying the mammogram images as normal, malignant, and benign. However, the MSVM classifier underperforms, if the number of feature points exceeds than the training samples. Additionally, Xie et al. [20] used Extreme Learning Machine (ELM) for diagnosing breast cancer. Initially, the level set model was employed to eliminate the interference in the mammogram images. Then, feature selection was accomplished using ELM for identifying the optimal features from the multi-dimensional feature vectors. Finally, SVM classifier was applied for distinguishing the benign and malignant classes. Though, SVM was a binary classifier which supports only binary classification, and it was not suitable for multi-class classification.

Mohanty et al. [11] developed a new system to detect the suspicious regions from the mammogram images, which were collected from the MIAS and DDSM databases. The developed system includes four phases; cropping operation, feature extraction utilizing contourlet transform, a forest optimization algorithm for selecting the optimal feature vectors, and classification by SVM, K-Nearest Neighbor (KNN), naive Bayes and C4.5 classifier. The experimental outcome showed that the developed system was robust and accurate in breast cancer detection. The naïve Bayes and KNN classifiers consider only independent predictor features, which was effective only in binary classification. Nirmala and Suresh Kumar [12] developed a new feature selection method; Bat Optimized Learning Machines (BORN) for selecting the optimal features. The most efficient visual saliency segmentation process was attained using BORN method. The objective of BORN was to diagnose the dissimilar stages of breast cancer with minimum error and high accuracy. The developed algorithm requires parameter tuning to achieve better search output. By highlighting the above stated issues, a new IFSGA-DNN model is proposed in this research paper for precise breast cancer segmentation and classification.

Ahmed et al. [21] developed deep learning method for the detect, segment and classification of breast cancer in mammographic images. The deep learning framework consists of DeepLab and Mask RCNN method for the classification. The transfer learning method and fine tuning method is applied in deep learning method to improve the performance of classification. Two datasets such as MIAS and DDSM were used to evaluate the performance of developed dep learning method. The validation shows that developed deep learning method has higher performance in classification. The developed model has overfitting problem due to feature handling method. Hassan et al. [22] applied transfer learning method from AlexNet and GoogleNet method for the breast cancer classification. The design and hyper-parameters are varied to find suitable parameter settings for classification. The DDSM and INbreast cancer datasets were applied to test the developed model performance. The result shows that AlexNet model has higher performance in breast cancer classification. The AlexNet model has limitation of ovefitting due to the feature handling in the classification.

Khan et al. [23] applied CNN based feature extraction and Multi-View Feature Fusion (MVFF) method for the classification of breast cancer stage. The extracted features are fused in final layer for the classification. The DDSM and mini-MIAS datasets were applied to test the performance of developed method. The developed method has the higher performance in the breast cancer classification in both datasets. The overfitting problem in the CNN classification affects the performance of the classification. Shen et al. [24] developed memory efficient deep learning model for the classification of breast cancer. The global and local information are fused in the model for the classification of breast cancer. The DDSM dataset was applied to evaluate the performance of the developed method in the breast cancer classification. The developed model has the overfitting problem due to the feature handling in the classification.

3 Methodology

The breast cancer is the 2nd leading cancer type worldwide and the most frequent cancer among women [25]. So, the early diagnosis of breast cancer plays a crucial role in reducing the mortality rate and morbidity [26]. In this research, a new IFSGA-DNN model is proposed for automatic breast cancer segmentation and classification. The proposed model includes six steps in breast cancer detection such as image collection: MIAS and DDSM, pre-processing: normalization, median filter, and CLAHE, image segmentation: multi-level, multi objective EML algorithm, feature extraction: Haralick texture features, LDTP, and HOG, feature selection: IFSGA, and classification: DNN. The flow diagram of the proposed model is graphically stated in Fig. 1.



Fig. 1 Flow diagram of proposed IFSGA-DNN model

3.1 Image collection and pre-processing

In the initial stage, the mammogram images are acquired from two datasets such as MIAS and DDSM. The MIAS dataset includes 322 mammogram images, where the size of every image is 1024×1024 [27]. Additionally, the DDSM dataset includes 2620 mammogram images with ground truth annotation of lesions [28]. The sample image of MIAS and DDSM dataset is graphically stated in Fig. 2.

After the acquisition of mammogram images, pre-processing is accomplished using image normalization, median filter, and CLAHE. Image normalization is the procedure of changing the range of mammogram pixel values [29]. The normalization is also named as a histogram or contrast stretching that effectively improves the contrast of the mammogram images. The image normalization is defined in Eq. (1).

$$Normalization = (I - Min) \times \frac{newMax - newMin}{Max - Min} + newMin$$
(1)

where, collected image is represented as *I*, maximum and minimum image pixel value is indicated as Max - Min that ranges between 0 and 255. Next, median filter is applied for denoising the noise in the normalized mammogram images. Median filter uses a square widow for filtering, where the square window size is 7×7 [30]. The median filter performs exceptionally well in the images, which are corrected by low density impulse image. In addition, CLAHE effectively improves the visibility level of the mammogram images [31]. The CLAHE divides the denoised mammogram images into many non-overlapping regions with equal size. Though, the good statistical estimation is achieved for 1024×1024 images, where the number of regions is selected to be equal to 128 by dividing the images by 8 in every direction. The output mammogram image of CLAHE is defined in Eq. (2).

$$C = \{f(X_{mod}(i,j)) | \forall X_{mod}(i,j) \in X_{mod}\}$$

$$\tag{2}$$

where, (i, j) is represented as spatial coordinates of the image pixel, and $f(X_{mod})$ is indicated as transformation function on the basis cumulative density function. The sample enhanced image of MIAS and DDSM databases are graphically stated in the Figs. 3 and 4.



Fig. 2 Sample image, a MIAS dataset, and b DDSM dataset



Fig. 3 Enhanced image of MIAS database; a normalization, b median filter and c CLAHE



Fig. 4 Enhanced image of DDSM database; a normalization, b median filter and c CLAHE

3.2 Segmentation

After improving the visibility level of the mammogram images, multi-level multi objective EML optimization algorithm is developed to solve the optimization concerns. In this scenario, the multi-level multi objective EML algorithm segments the cancerous regions from enhanced images. The EML optimization algorithm solves the issue of identifying a global solution of a non-linear optimization concern with box constraints [32], as shown in Eq. (3).

maximize
$$f(V), V = (v_1, \dots, v_n) \in \mathbb{R}^n$$
, subject to $v \in V$ (3)

where, V is represented as feasible area and R^n is denoted as nonlinear function. The feasible area V is mathematically expressed in Eq. (4).

$$V = \{ v \in \mathbb{R}^n / l_i \le v_i \le u_i, i = 1, 2, \dots n$$
(4)

where, u_i is stated as upper interval and l_i is denoted as lower interval. In EML algorithm, N and $v_{i,g}$ are used for exploring the feasible region V, where g is denoted as iteration number, g = 100. Once the population is initialized $S_g = \{v_{1,g}, v_{2,g}, v_{3,g} \dots v_{N,g}\}$, the EML optimization algorithm runs until the iteration

number is satisfied. Based on the electromagnetism theory, the charged particles are determined as $v_{i,g} \in s_g$ in the feasible areaV. In the EML optimization algorithm, the charge of every point is connected with a cost function, a point with high cost value is attracted to other points utilizing the electromagnetism theory of attraction and repulsion. In the location $y_{i,g}$, a complete force vector F_i^g is applied to each point with the direction of whole force. By utilizing a local search, the vicinity of each $y_{i,g}$ is analyzed by exploiting $y_{i,g}$ to $z_{i,g}$. In g + 1th generation, the members $v_{i,g+1}$ of S_{g+1} is mathematically expressed in Eq. (5).

$$v_{i,g+1} = \begin{cases} y_{i,g} & iff(y_{i,g}) \le f(z_{i,g}) \\ z_{i,g} & otherwise \end{cases}$$
(5)

Usually in image segmentation, thresholding process is used to classify the gray scale pixels into definite sets on the basis of pixel intensity values P. In this research, bi level thresholding rule is applied for finding the optimal threshold value for better image segmentation. The bi-level thresholding is mathematically expressed in Eq. (6).

$$G_1 \leftarrow p_i \text{ if } 0 \le p_i < th, G_2 \leftarrow p_i \text{ if } th \le p_i < P - 1 \tag{6}$$

where, G_1 and G_2 are represented as two definite sets, and p_i is stated as grayscale pixel intensity value. The bi-level thresholding is rewritten as indicated in Eq. (7).

$$G_1 \leftarrow p_i \text{ if } 0 \le p_i < th_1, G_2 \leftarrow p_i \text{ if } th_1 \le p_i < th_2, G_i \leftarrow p_i \text{ if } th_i \le p_i < th_{i+1}, G_n \leftarrow p_i \text{ if } th_n \le p_i < P - 1$$

$$(7)$$

where, $\{th_1, th_2, ..., th_k\}$ is denoted as definite thresholds. The threshold values th are selected for segmenting the cancerous regions precisely in the mammogram images. Kapurs [33] and OTSU thesholding [34] are the two extensively utilized methodologies to determine the threshold values. The kapurs and OTSU thresholding presents two distinct objective functions [35]. Next, template matching is performed to match the ground-truth image and segmented image for validating the performance of the EML optimization algorithm. The sample segmented image is graphically stated in Fig. 5.



Fig. 5 a Enhanced image, b segmented image, c template matched image, d ground truth image

3.3 Feature extraction and selection

After segmenting the cancer regions from the mammogram images, features are extracted by using LDTP, HOG, and Haralick texture features. Additionally, Haralick features (entropy, correlation, angular second moment, contrast, difference entropy, and inverse difference moment) are applied for computing Gray Level Co-occurrence Matrix (GLCM), which has desired number of gray levels [36]. The entropy features in Haralick is the measure of randomness or degree of disorder present in the image and correlation is the measure of linear dependency of gray level values in the co-occurrence matrix. The angular second moment measure the local uniformity of gray levels and contrast is the measure of intensity variation between the reference pixel and neighbor. Further, LDTP is applied to extract the features from the segmented regions, where it encodes both directional pattern and contrast information of the mammogram images on the basis of local derivative variation [37]. The LDTP conveys information about the nature of the textures by capturing the local textures using local directional pattern and local ternary pattern concepts. After extracting the feature vectors f_i , feature selection is carried out by using IFSGA.

Initially, the infinite feature selection algorithm selects the appropriate energy score $e_l(i)$ and length l for extracted feature vectors, as mentioned in Eq. (8).

$$e_{l}(i) = \sum_{j \in v} \sum_{p \in p_{i,j}^{l}} \prod_{k=0}^{l-1} a_{vk,vk+1} = \sum_{j \in v} A^{l}(i,j)$$
(8)

where, $p_{i,j}^{l}$ is indicated as the length *l* between the features *i* and *j*, A^{l} is stated as power iteration of matrix, *v* and *A* are denoted as the vertices of the feature values. Though, the probability of the feature values is normalized by expanding the path length to infinity [38]. Hence, the new energy score e(i) considers all path lengths for feature values f_{i} , as stated in Eq. (9).

$$e(i) = \left[\left(\sum_{l=0}^{\infty} A^l \right) - I \right] \overline{1}$$
(9)

where, $\overline{1}$ Is indicated as a column vector of ones, and *I* is stated as the identity matrix In the matrix algebra, $\sum_{k=0}^{\infty} X^{l}$ is represented as geometric series of matrix *X*. If p(X) < 1, the series converges to $(I - X)^{-1}$, where p(X) is indicated as maximum

magnitude of the feature vectors. By utilizing this property, regularize the energy score of feature vectors, which is defined in Eq. (10).

$$e'(i)[\left(\left(\sum_{l=0}^{\infty}r^{l}A^{l}\right)-I\right)\overline{1}]_{i}=[((I-rA)^{-1}-I)\overline{1}]_{i}$$
(10)

The computational power iteration of the matrix in the Eq. (8) is decreased by computing $((I - rA)^{-1} - I)$, where the regularized feature vectors are fed to genetic algorithm.

Genetic algorithm is utilized for feature selection to determine the most relevant features for classification. Many variations of genetic operations are developed for better feature selection. One of the extensively utilized feature selection methodology is the use of entropy value in genetic algorithms. The entropy value is utilized for measuring the homogeneity of the database in order to find the mutual information between the extracted feature values that helps in identifying the prominent feature values. The proposed IFSGA utilizes a simple entropy measure for determining the relevant feature values. Hence, the conditional entropy is estimated for both regularized and output features, where the proposed IFSGA aims in finding the maximum relevance between the regularized and output features and to diminish the redundancy with-in the regularized feature values.

In IFSGA, the initial population is set as the sub-set of the regularized features. These feature values are represented as the dominant features of the predetermined output. Further, the fitness function is estimated based on entropy measures by increasing the mutual information of the conditional entropy between the regularized and output features. Crossover operation maximizes the genetic diversity for finding the relevant regularized feature vectors. In addition, the selection operation is used to identify the dominant features for the given output by reducing the redundancy on the basis of fitness function. In the genetic algorithm, population size depends on the maximum dataset length, generation is 100, population type is a bit string, elite count is 2, mutation function is 0.1, crossover function is 0.8, and fitness function is entropy measure. Pseudocode of IFSGA is given below. Table 1 describes the total extracted features and selected features by IFSGA.

A Dataset	Extracted features	Selected features
MIAS	1850	555
DDSM	1850	555

 Table 1
 Total extracted features

 and selected features by IFSGA

Pseudo code of IFSGA

Input: Extracted feature values f_i , population size; maximum dataset length, generation; 100, elite count; 2, mutation function; 0.1, and crossover function; 0.8.

Output: Optimal features X_i

For i = 1 // infinite feature selection

Select appropriate energy score $e_l(i)$ and length l for f_i based on equation (8). Calculate new energy score e(i) for f_i using equation (9).

Regularize e(i) of extracted feature values f_i .

End For

While (*iteration* \leq *dataset length*)

Set initial population based on regularized subset of features // genetic algorithm Calculate fitness function based on entropy measure

Crossover and mutation applied to regularized subset of features

Calculate entropy value for regularized features

Features for next generation is selected on the basis of entropy value

Execute the procedure until the stopping criteria is not met.

End while

3.4 Classification

The feature selection method is carried out in this research to reduce the computational complexity of the DNN [39]. The relevant features are selected based on IFSGA to train the DNN and reduce the computational complexity. After selecting the optimal feature vectors, DNN classifier is employed for classifying the mammograms as malignant and benign classes. In image processing domain, DNN classifier is a powerful tool, which is denoted by a non-linear function $f : X_i \rightarrow y$, where y is indicated as output space and X_i is stated as input space. DNN classifier learns extremely for generating more abstract representation of the mammogram images. Though, DNN classifier is a feed-forward network that comprises of several input, hidden and output layers, where the nodes in the neighborhood layers are fully connected and the nodes in the same layers are independent [40]. If the input property is equal to the hidden layer $g^0 = N$, the active nodes in the hidden layer are defined by the Eqs. (11) and (12).

$$b^{m} = W^{m} X_{i} g^{m-1} + a^{m} (1 \le m \le M + 1)$$
(11)

$$g^m = f(b^m)$$
 with $g_j^m = \frac{1}{1 + e^{-b_j^m}} (1 \le m \le M)$ (12)

where, a^m is indicated as bias vector, W^m is stated as weight function, $f(b^m)$ is stated as sigmoid activation function, and N is represented as number of hidden layers. In the output layer, softmax function is used to model the posterior probability of input properties, which is defined in Eq. (13).

$$y_s = g_s^{M+1} = \Pr(s|N) = softmax_s(b^{M+1})$$
 (13)

where, y_s is indicated as output vector and *s* specified as sparsity parameter. Hence, the parameter settings of DNN classifier is indicated as follows; hidden layer 1 is 500, hidden layer 2 is 250, L2 weight regularization is 0.004, maximum epochs is 500, sparsity regularization is 4 and sparsity proportion is 0.15. The input of the DNN is optimal features X_i and the output of the DNN is represented as y_1 and y_2 . The architecture of DNN is graphically denoted in Fig. 6.

4 Experimental result and discussion

The proposed IFSGA-DNN model is implemented using MATLAB (2019a) software with Intel core i7 processor, 16 GB RAM, 2 TB hard disk and 64 bit operating system. The proposed IFSGA-DNN model performance is related with few benchmark models such as SVM with Hough transform [13], cGAN with CNN [14], CNN [18] and ELM [20] for investigating its efficacy. In this research, the performance of IFSGA-DNN model is evaluated by means of precision, accuracy, recall, FPR, FDR, and FOR. The precision, accuracy, recall, FPR, FDR, and FOR are mathematically indicated in the Eqs. (14–19).

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



Fig. 6 Architecture of DNN classifier

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \times 100$$
(15)

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

$$FPR = \frac{FP}{FP + TN} \times 100 \tag{17}$$

$$FDR = \frac{FP}{FP + TP} \times 100 \tag{18}$$

$$FOR = \frac{FN}{TN + FN} \times 100 \tag{19}$$

where, True Positive is represented as *TP*, False Positive is indicated as *FP*, True Negative is stated as *TN*, and False Negative is denoted as *FN*.

4.1 Quantitative investigation on DDSM dataset

In this section, DDSM dataset is undertaken for investigating the performance with different feature selection algorithms (reliefF, Principal Component Analysis (PCA), infinite and IFSGA) and classifiers (KNN, MSVM, Random Forest (RF) and DNN). The DDSM database consists of 2620 mammogram images, where 80% of the mammogram images (2096) is used for training and 20% of the mammogram images (524) are used for testing. By investigating Table 2, IFSGA with DNN showed high performance in breast cancer detection compared to other combinations. The combination IFSGA-DNN achieved 96.67% of precision, 95.22% of recall, and 95.83% of classification accuracy. The feature importance score is measured in proposed DNN classification and shown in the Fig. 7. The graphical analysis of proposed IFSGA-DNN model on DDSM dataset in terms of precision, accuracy, and recall is represented in Fig. 8.

The feature importance is measured in the proposed method and feature score is shown in the Fig. 7. This shows that Inverse difference moment features has the higher score and Angular second moment has the second higher score in the classification. The entropy and HOG has the lower feature score in the classification.

By inspecting Table 3, the proposed IFSGA-DNN model is analyzed by means of FPR, FDR and FOR. The proposed IFSGA-DNN model attained 14.45% of FPR, 4.17% of FDR, and 14.44% of FOR, which is better in breast cancer detection compared to other combinations. Applying the extracted features directly to the DNN tends to improper learning due to variation in the weight calculation in the network. Therefore, the genetic algorithm is applied in this method to calculate the weight values based on features importance and this tends to improve the learning of the DNN. Instead of directly applying the extracted features, the genetic algorithm defined features with feature score helps to improve the

Feature selection	Classifier	Precision (%)	Recall (%)	Accuracy (%)
ReliefF	KNN	65.12	75.65	63.03
	MSVM	87.92	80.11	86.92
	RF	69.23	81.77	68.52
	DNN	85.34	75.69	91.26
PCA	KNN	64.06	67	65.68
	MSVM	75.81	72.41	81.89
	RF	67.49	80.87	72.42
	DNN	75.24	81.28	86.76
Infinite	KNN	68.07	78.92	64.35
	MSVM	88.32	81.05	92.13
	RF	75.88	82.93	71.41
	DNN	86.05	76.58	94.56
Infinite with genetic algorithm	KNN	70	80	66.87
	MSVM	88.33	83.33	92.58
	RF	76.67	83.33	74.21
	DNN	96.67	95.22	95.83

 Table 2
 Performance investigation with different feature selection and classification techniques on DDSM dataset by means of precision, accuracy and recall



Fig. 7 The feature importance in the classification

learning of the DNN. Compared to other feature selection methods, the genetic algorithm has the advantage of self-evolve in the individual chromosome to find the candidate solution that provide prominent feature selection. DNN classifier able to solve more complex issues by utilizing a wide range of architectures related to other simple feed forward networks and fully connected networks.



Fig. 8 Graphical analysis of proposed IFSGA-DNN model on DDSM dataset in terms of precision, accuracy, and recall

Table 3 Performance investigation with different	Feature selection	Classifier	FPR (%)	FDR (%)	FOR (%)
feature selection and	ReliefF	KNN	30.68	34.98	28.74
DDSM dataset by means of		MSVM	15.16	15.16	19.91
FPR, FDR and FOR		RF	19.31	29.54	20.10
		DNN	26.07	8.85	21.33
	PCA	KNN	39.23	34.40	42.42
		MSVM	17.45	14.25	34.22
		RF	22.83	29.07	21.05
		DNN	27.55	20.65	24.71
	Infinite	KNN	27.62	35.38	27.58
		MSVM	15.36	7.97	16.26
		RF	22.04	25.80	19.31
		DNN	22.08	6.95	19.93
	Infinite with	KNN	23.12	33.13	25.17
	genetic algo-	MSVM	24.15	7.42	15.07
	rithm	RF	23.21	25.79	19.09
		DNN	14.45	4.17	14.44

Additionally, DNN classifier includes three key benefits; computationally effective, able to deliver high quality results and maximum usage of unstructured data. The graphical analysis of proposed IFSGA-DNN model on DDSM dataset by means of FPR, FDR and FOR is denoted in Fig. 9.



Fig. 9 Graphical analysis of proposed IFSGA-DNN model on DDSM dataset in terms of FPR, FDR and FOR

4.2 Quantitative investigation on MIAS dataset

In this section, MIAS dataset is undertaken to assess the performance with dissimilar feature selection algorithms and classifiers. The genetic algorithm learning ability is increased by self-evolve chromosome that provide prominent feature selection. Though, MIAS database consists of 322 mammogram images in that 80% of the mammogram images (258) is used for training and 20% of the images (64) are used for testing. In Table 4, the combination (IFSGA-DNN) achieved better performance in breast cancer classification and showed a maximum of 32.84% and a minimum of 2% improvement in the classification accuracy. Similarly, the proposed IFSGA-DNN model attained better performance in breast cancer detection in light of recall and precision. Graphical analysis of proposed IFSGA-DNN model on MIAS dataset in terms of precision, accuracy, and recall is indicated in Fig. 10.

Similarly in Table 5, the proposed IFSGA-DNN model achieved significant performance in breast cancer detection on mammogram images in terms of FPR, FDR and FOR. Graphical analysis of proposed IFSGA-DNN model on MIAS dataset by means of FPR, FDR and FOR is stated in Fig. 11. In this research, the proposed IFSGA effectively improves the learning performance of DNN and gives better understanding which features are important for breast cancer classification.

4.3 Cross validation

In this segment, the proposed IFSGA-DNN model performance is validated with different training and testing percentages like 60:40, 70:30, and 80:20 on both DDSM and MIAS databases. By investigating Table 6, the proposed IFSGA-DNN model achieved better performance in 80:20 combination on both DDSM and MIAS datasets in light of FPR, FDR, FOR, recall, precision and classification accuracy.

Feature selection	Classifier	Precision (%)	Recall (%)	Accuracy (%)
ReliefF	KNN	68.03	76.97	64.59
	MSVM	88.85	81.68	89.51
	RF	71.10	84.76	68.69
	DNN	86.52	78.51	92.48
PCA	KNN	64.27	68.11	67.72
	MSVM	76.11	74.91	83.32
	RF	67.52	83.50	74.75
	DNN	78.13	83.04	87.56
Infinite	KNN	69.90	81.09	66.20
	MSVM	89.55	83.22	94.08
	RF	77.99	84.41	73.37
	DNN	87.89	77.46	95.42
Infinite with genetic algorithm	KNN	71.10	80.77	69.20
	MSVM	90.52	86.33	95.01
	RF	78.61	85.57	74.81
	DNN	98.93	97.85	97.43

 Table 4
 Performance investigation with different feature selection and classifiers on MIAS dataset in light of precision, accuracy and recall



 $\ensuremath{\textit{Fig.10}}$ Graphical analysis of proposed IFSGA-DNN model on MIAS dataset in terms of precision, accuracy, and recall

4.4 Comparative investigation

Comparative analysis between the proposed IFSGA-DNN model and the existing models; SVM with Hough transform [13], cGAN with CNN [14], CNN [18] and ELM [20] are given in Table 7. R. Vijayarajeswari et al. [13] developed a new system for breast cancer detection based on SVM and Hough transform. In this system, Hough transform was used to isolate the feature vectors of the collected mammogram images. Further, SVM was applied for classifying the mammograms

Table 5 Performance investigation with different feature selection and classifiers on MIAS dataset in light of EPR EPR EDR and EOR	Feature selection	Classifier	FPR (%)	FDR (%)	FOR (%)
	ReliefF	KNN	29.80	32.63	27.75
		MSVM	14.02	14.37	17.98
		RF	19.28	26.75	19.55
		DNN	25.05	7.77	21.30
	PCA	KNN	36.74	33.68	41.24
		MSVM	14.95	11.56	34.15
		RF	19.85	26.94	18.76
		DNN	25.97	18.75	21.85
	Infinite	KNN	25.26	34.81	26.04
		MSVM	13.86	7.59	16.11
		RF	21.66	23.63	17.10
		DNN	20.70	6.04	17.44
	Infinite with	KNN	24.41	32.38	24.37
	genetic algo-	MSVM	13.96	7.40	12.62
	rithm	RF	18	24.84	16.44
		DNN	12.88	4.06	9.40



Fig. 11 Graphical analysis of proposed IFSGA-DNN model on MIAS dataset in terms of FPR, FDR and FOR

as normal, malignant and benign patients. In this paper, MIAS database was used to validate the developed system performance. From the experimental analysis, the developed system attained 94% of accuracy in breast cancer classification. In addition, Singh et al. [14] utilized cGAN for segmenting the breast cancer regions and then CNN was applied to classify four cancer shapes; round, lobular,

Cross validation	Precision (%)	Recall (%)	Accuracy (%)	FPR (%)	FDR (%)	FOR (%)
DDSM dataset						
60:40	91.01	92.12	90.83	25.36	10.23	28.99
70:30	95.33	94.34	93.83	21.67	6.26	22.56
80:20	96.67	95.22	95.83	14.45	4.17	14.44
MIAS dataset						
60:40	91.08	92.11	92.79	28.12	96.66	29.72
70:30	96.89	95.25	94.78	21.04	7.69	23.58
80:20	98.93	97.85	97.43	12.88	4.06	9.40

Table 6 Cross validation of proposed IFSGA-DNN model with different testing and training percentage

Table 7 Comparative investigation in fivefold cross-validation	Methodology	Dataset	Accuracy (%)
	SVM with Hough transform [13]	MIAS	94
	cGAN with CNN [14]	DDSM	80
	CNN [18]	MIAS	90.50
	ELM [20]	DDSM	95.73
		MIAS	96.02
	Proposed model IFSGA-DNN	DDSM	95.83
		MIAS	97.43

Bold values indicate the result received after applying our proposed method

irregular and oval. From the simulation result, the developed model attained 80% of classification accuracy in breast cancer detection on DDSM dataset.

Ting et al. [18] used the CNN model to enhance breast cancer detection performance. In this literature, the developed CNN model attained 89.47% of recall and 90.50% of accuracy on MIAS dataset. In addition, Xie et al. [20] used level set model to eliminate the interference in the mammogram images. Next, feature selection was done using ELM for identifying the optimal features from the multi-dimensional features. At last, SVM was applied to distinguish the benign and malignant classes. From the experimental investigation, the developed model attained 95.73% of accuracy in DDSM dataset and 96.02% of accuracy in MIAS dataset. Compare to these existing models, the proposed IFSGA-DNN model attained better performance in breast cancer detection on both MIAS and DDSM datasets.

5 Conclusion

In this research, a new IFSGA-DNN model is proposed for automatic breast cancer segmentation and classification. After collecting the images from MIAS and DDSM databases, multi-level multi objective EML optimization algorithm is proposed to

segment the cancerous and non-cancerous regions from the mammogram images. Then, hybrid feature extraction is carried out to extract the features from the segmented cancer and non-cancer regions. Further, the discriminative features are selected from the extracted features by using IFSGA and the selected features are fed to DNN classifier to classify the benign and malignant classes. The proposed IFSGA-DNN method has accuracy improvement of 1.41% compared to ELM, 6.93% compared to CNN, and 3.43% compared to SVM with Hough transform in MIAS dataset. The proposed IFSGA-DNN method has accuracy improvement of 0.1% compared to ELM, and 15.83% compared to cGAN with CNN in DDSM dataset. In this study, the proposed IFSGA effectively removes the irrelevant features from the extracted features, so the complexity of the proposed model is linear O(n), where n is represented as input size and O is indicated as order of magnitude. The developed method is supervised method for the breast cancer classification and clustering algorithms is required for unsupervised classification. In the future work, a new clustering based segmentation algorithm can be developed to improve the performance of breast cancer detection. The future work of this method involves in applying the proposed method in the cloud to test the performance in clinical data.

References

- Beura, S., Majhi, B., Dash, R.: Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer. Neurocomputing 154, 1–14 (2015)
- Pawar, M.M., Talbar, S.N.: Genetic fuzzy system (GFS) based wavelet co-occurrence feature selection in mammogram classification for breast cancer diagnosis. Perspect. Sci. 8, 247–250 (2016)
- Verma, B., McLeod, P., Klevansky, A.: Classification of benign and malignant patterns in digital mammograms for the diagnosis of breast cancer. Expert Syst. Appl. 37(4), 3344–3351 (2010)
- Shastri, A.A., Tamrakar, D., Ahuja, K.: Density-wise two stage mammogram classification using texture exploiting descriptors. Expert Syst. Appl. 99, 71–82 (2018)
- Agnes, S.A., Anitha, J., Pandian, S.I.A., Peter, J.D.: Classification of mammogram images using multiscale all convolutional neural network (MA-CNN). J. Med. Syst. 44(1), 30 (2020)
- Wang, X., Liang, G., Zhang, Y., Blanton, H., Bessinger, Z., Jacobs, N.: Inconsistent performance of deep learning models on mammogram classification. J. Am. Coll. Radiol. (2020). https://doi.org/10. 1016/j.jacr.2020.01.006
- Arevalo, J., González, F.A., Ramos-Pollán, R., Oliveira, J.L., Lopez, M.A.G.: Representation learning for mammography mass lesion classification with convolutional neural networks. Comput. Methods Prog. Biomed. 127, 248–257 (2016)
- Chougrad, H., Zouaki, H., Alheyane, O.: Multi-label transfer learning for the early diagnosis of breast cancer. Neurocomputing (2019). https://doi.org/10.1016/j.neucom.2019.01.112
- Patil, R.S., Biradar, N.: Automated mammogram breast cancer detection using the optimized combination of convolutional and recurrent neural network. Evolut. Intell. (2020). https://doi.org/10.1007/ s12065-020-00403-x
- Mohanty, F., Rup, S., Dash, B., Majhi, B., Swamy, M.N.S.: Digital mammogram classification using 2D-BDWT and GLCM features with FOA-based feature selection approach. Neural Comput. Appl. (2019). https://doi.org/10.1007/s00521-019-04186-w
- Mohanty, F., Rup, S., Dash, B., Majhi, B., Swamy, M.N.S.: Mammogram classification using contourlet features with forest optimization-based feature selection approach. Multimed. Tools Appl. 78(10), 12805–12834 (2019)
- Nirmala, G., Suresh Kumar, P.: A novel bat optimized runlength networks (BORN) for an efficient classification of breast cancer. J. Ambient Intell. Humaniz. Comput. (2020). https://doi.org/10.1007/ s12652-020-01890-7

- Vijayarajeswari, R., Parthasarathy, P., Vivekanandan, S., Basha, A.A.: Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform. Measurement 146, 800–805 (2019)
- Singh, V.K., Rashwan, H.A., Romani, S., Akram, F., Pandey, N., Sarker, M.M.K., Saleh, A., Arenas, M., Arquez, M., Puig, D., Torrents-Barrena, J.: Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network. Expert Syst. Appl. 139, 112855 (2020)
- 15. Li, H., Zhuang, S., Li, D.A., Zhao, J., Ma, Y.: Benign and malignant classification of mammogram images based on deep learning. Biomed. Signal Process. Control **51**, 347–354 (2019)
- Mohanty, F., Rup, S., Dash, B., Majhi, B., Swamy, M.N.S.: An improved scheme for digital mammogram classification using weighted chaotic salp swarm algorithm-based kernel extreme learning machine. Appl. Soft Comput. 91, 106266 (2020)
- 17. Sadad, T., Munir, A., Saba, T., Hussain, A.: Fuzzy C-means and region growing based classification of tumor from mammograms using hybrid texture feature. J. Comput. Sci. 29, 34–45 (2018)
- Ting, F.F., Tan, Y.J., Sim, K.S.: Convolutional neural network improvement for breast cancer classification. Expert Syst. Appl. 120, 103–115 (2019)
- Kaur, P., Singh, G., Kaur, P.: Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification. Inform. Med. Unlocked 16, 100151 (2019)
- 20. Xie, W., Li, Y., Ma, Y.: Breast mass classification in digital mammography based on extreme learning machine. Neurocomputing **173**, 930–941 (2016)
- Ahmed, L., Iqbal, M.M., Aldabbas, H., Khalid, S., Saleem, Y., Saeed, S.: Images data practices for semantic segmentation of breast cancer using deep neural network. J. Ambient Intell. Humaniz. Comput. (2020). https://doi.org/10.1007/s12652-020-01680-1
- Shayma'a, A.H., Sayed, M.S., Abdalla, M.I., Rashwan, M.A.: Breast cancer masses classification using deep convolutional neural networks and transfer learning. Multimed. Tools Appl. **79**(41), 30735–30768 (2020)
- Khan, H.N., Shahid, A.R., Raza, B., Dar, A.H., Alquhayz, H.: Multi-view feature fusion based four views model for mammogram classification using convolutional neural network. IEEE Access 7, 165724–165733 (2019)
- Shen, Y., Wu, N., Phang, J., Park, J., Liu, K., Tyagi, S., Heacock, L., Kim, S.G., Moy, L., Cho, K., Geras, K.J.: An interpretable classifier for high-resolution breast cancer screening images utilizing weakly supervised localization. Med. Image Anal. 68, 101908 (2021)
- 25. Suganthi, M., Madheswaran, M.: An improved medical decision support system to identify the breast cancer using mammogram. J. Med. Syst. **36**(1), 79–91 (2012)
- de Oliveira Martins, L., Silva, A.C., De Paiva, A.C., Gattass, M.: Detection of breast masses in mammogram images using growing neural gas algorithm and Ripley's K function. J. Signal Process. Syst. 55(1-3), 77–90 (2009)
- Melloul, M., Joskowicz, L.: Segmentation of microcalcification in X-ray mammograms using entropy thresholding. In: CARS 2002 computer assisted radiology and surgery, pp. 671–676. Springer, Berlin (2002)
- Heath, M., Bowyer, K., Kopans, D., Kegelmeyer, P., Moore, R., Chang, K., Munishkumaran, S.: Current status of the digital database for screening mammography. In: Digital mammography, pp. 457–460. Springer, Dordrecht (1998)
- Koo, K.M., Cha, E.Y.: Image recognition performance enhancements using image normalization. HCIS 7(1), 1–11 (2017)
- Kuo, Y.L., Tai, C.W.: A simple and efficient median filter for removing high-density impulse noise in images. Int. J. Fuzzy Syst. 17(1), 67–75 (2015)
- Reza, A.M.: Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. J. VLSI Signal Process. Syst. Signal Image Video Technol. 38(1), 35–44 (2004)
- 32. Qi, X., Zhang, L., Chen, Y., Pi, Y., Chen, Y., Lv, Q., Yi, Z.: Automated diagnosis of breast ultrasonography images using deep neural networks. Med. Image Anal. **52**, 185–198 (2019)
- Cuevas, E., Zaldívar, D., Perez-Cisneros, M.: Otsu and Kapur segmentation based on harmony search optimization. In: Applications of evolutionary computation in image processing and pattern recognition, pp. 169–202. Springer, Cham (2016)

- Pruthi, J., Gupta, G.: Image segmentation using genetic algorithm and OTSU. In: Proceedings of fifth international conference on soft computing for problem solving, pp. 473–480. Springer, Singapore (2016)
- 35. Avuti, S.K., Bajaj, V., Kumar, A., Singh, G.K.: A novel pectoral muscle segmentation from scanned mammograms using EMO algorithm. Biomed. Eng. Lett. 9(4), 481–496 (2019)
- Anusha, R., Jaidhar, C.D.: Human gait recognition based on histogram of oriented gradients and Haralick texture descriptor. Multimed. Tools Appl. (2020). https://doi.org/10.1007/ s11042-019-08469-1
- 37. Chahi, A., Ruichek, Y., Touahni, R.: Local directional ternary pattern: a new texture descriptor for texture classification. Comput. Vis. Image Underst. **169**, 14–27 (2018)
- Rangappa, V.G., Prasad, S.V.A.V., Agarwal, A.: Infinite feature selection based arrhythmia classification using semantic features and random forest. Int. J. Intell. Eng. Syst. 12(3), 203–213 (2019)
- Nguyen, T.G., Phan, T.V., Hoang, D.T., Nguyen, T.N., So-In, C.: Efficient SDN-based traffic monitoring in iot networks with double deep Q-network. In: International conference on computational data and social networks, pp. 26–38. Springer, Cham (2020)
- Birbil, Şİ, Fang, S.C.: An electromagnetism-like mechanism for global optimization. J. Glob. Optim. 25(3), 263–282 (2003)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

S. S. Ittannavar¹ · R. H. Havaldar²

R. H. Havaldar ravirajhavaldar@klescet.ac.in

- ¹ Department of Electronics and Communication Engineering, Hirasugar Institute of Technology, Nidasoshi, India
- ² Department of Biomedical Engineering, KLE Dr. MS Sheshgiri College of Engineering and Technology, Belagavi, India