A Hybrid Optimized Deep Learning Techniques for Analyzing Mammograms

Satish Babu Bandaru[1], Natarajasivan Deivarajan[2], Rama Mohan Babu Gatram[3]

[1]Research Scholar, Department of Computer Science and Engineering, Annamalai University, Annamalinagar, Tamil Nadu 608002, India[2]Assistant Professor, Department of Computer Science and Engineering, Faculty of Computer Science and Engineering, Annamalai University, Annamalinagar, Tamil Nadu 608002, India
[3]Professor, Department of Computer Science and Engineering (AI & ML), RVR & JC College of Engineering, Guntur, A.P. 522019, India

Abstract
Early detection continues to be the mainstay of breast cancer control as well as the improvement of its treatment. Even so, the absence of cancer symptoms at the onset has early detection quite challenging. Therefore, various researchers continue to focus on cancer as a topic of health to try and make improvements from the perspectives of diagnosis, prevention, and treatment. This research's chief goal is development of a system with deep learning for classification of the breast cancer as non-malignant and malignant using mammogram images. The following two distinct approaches: the first one with the utilization of patches of the Region of Interest (ROI), and the second one with the utilization of the overall images is used. The proposed system is composed of the following two distinct stages: the pre-processing stage and the Convolution Neural Network (CNN) building stage. Of late, the use of meta-heuristic optimization algorithms has accomplished a lot of progress in resolving these problems. Teaching-Learning Based Optimization algorithm (TIBO) meta-heuristic was originally employed for resolving problems of continuous optimization. This work has offered the proposals of novel methods for training the Residual Network (ResNet) as well as the CNN based on the TLBO and the Genetic Algorithm (GA). The classification of breast cancer can be enhanced with direct application of the hybrid TLBO- GA. For this hybrid algorithm, the TLBO, i.e., a core component, will combine the following three distinct operators of the GA: coding, crossover, and mutation. In the TLBO, there is a representation of the optimization solutions as students. On the other hand, the hybrid TLBO-GA will have further division of the students as follows: the top students, the ordinary students, and the poor students. The experiments demonstrated that the proposed hybrid TLBO-GA is more effective than TLBO and GA.

Keywords:
Deep Learning, Convolutional Neural Networks (CNNs), Residual Network (ResNet), Teaching Learning Based Optimization Algorithm (TLBO)

Manuscript received October 5, 2022
Manuscript revised October 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.10.11

1. Introduction

Being one of the most invasive malignant tumors, breast cancer is common among women albeit rare among men. Due to its high death rate in women, this cancer type is accounted for as the worst cancer after the lung cancer. Coincidentally, it was noted from the previous studies that early-stage detection of the breast cancer was able to achieve an above 90% rate of survival. In contrast, late diagnosis would result in the spreading of cancer to the lymph node as well as other parts of the body, and thus, drastically reduce the rate of survival by 27%. Hence, to save the patient's life, early-stage breast cancer detection is essential. The breast's complex structure may result in the specialist be incapable of identifying the local lesions, and other times, the misdiagnosis may also occur due to the involvement of tremendous reading procedures. Medical screenings are an essential feature of the breast cancer's diagnostic procedure [1].

At present, mammography is accounted for as the most effective approach for early breast cancer detection due to its high resolution of the breast's internal anatomy. Mammography's key objective is detection of early cancer symptoms, and also to diagnose the breast masses. For doctors, mammograms are not only a critical screening tool for breast cancer, but it is also essential for the diagnosis, assessment as well as continuous monitoring of people with breast cancer. Mammography screening will involve the following two distinct steps: At first, there is compression of the breast between two small flat plates. Later, there is direct application of a low X-ray dose through the breast, and utilization of a two-dimensional (2D) panel detector for image acquisition [2].

Mammography screening does suffer from certain limitations because of complexities like irregularity in the form of abnormalities, and abnormal tissues hidden within the dense breasts. Hence, during the mammogram examination procedure, there may be either the occurrence of missed cancer cases (i.e., false-negative value) or the misinterpretation of non-cancerous lesions as cancerous (i.e., false-positive value). Since expertise is absolutely necessary for the understanding of mammogram images, it is an operator-dependent modality. Furthermore, the procedure of mammogram reading is expensive, error-prone, tiresome as well as time-consuming. Therefore, Computer-Aided Diagnosis (CAD) systems are able to help the physicians in abnormality detection via computerized procedures for feature extraction as well as classification. These systems will aim to mitigate the effort which is needed for distinguishing the malignant lesions from the benign ones. In addition, these systems will be employed for minimization of the interpretation errors via reduction of the number of false-positives that result in expensive as well as uncomfortable biopsies and also reduction of the number of false-negatives that result in neglect in treating patients, and in turn, diminish the rate of survivability [3].

CAD systems based on standard techniques of machine learning like the Support Vector Machines (SVM), decision tree, in turn, are based on the extraction of hand-crafted features, the feature selection as well as the classification of lesions. Whereas the deep learning is able to automatically distinguish features during training. The Convolution Neural Networks (CNNs) are variations of the Multi-Layer Perceptron (MLP) are one of the most popular deep learners. These networks are capable of directly recognizing visual patterns from the input image pixels. Thus, they are quite well-known in the medical image analysis as well as vision system fields. In CNN, the feature maps are created using filters to obtain a convolution operation with the addition of a bias on sub-regions of the whole image [4]. Earlier studies have employed a hybrid CNN for the accomplishment of high mass as well as pathological classification. Moreover, there was utilization of a convolutional sparse auto-encoder for CNN formation that in turn had acquired feasible outcomes in the breast density segmentation for mammography risk scoring [5].

Transfer learning is a technique employed to leverage a model already trained for a task to be used for another task that is related to it. However, it is not possible to ensure successful results by training the model with random weight on a scarce dataset. Hence, whenever the researchers are unable to acquire sufficient datasets, they will often use other sufficient datasets that have been collected for similar tasks. With utilization of weights pre-trained on other similar datasets, the researchers are able to fine-tune the model. For the domain of medical image analysis, transfer learning is employed for acquisition of accurate annotations of the lesions because it is tedious to collect a sufficient number of well-refined images as a result of issues of privacy. Thus, this work will assess the transfer learning's effectivity by means of experiments with public as well as in-house mammogram datasets, and also will grant public-availability of the weights pre-trained on the datasets such so that it can be leveraged by other researchers within the mammography community [6]. However, the weights and the architecture of the deep learner needs to be optimal for the model to perform well. The metaheuristic algorithms are utilized for optimizing the deep learners.

In this work, the hybrid TLBO-GA with ResNet and CNN-18 is proposed for improving the detection of cancer in mammography. The rest of the paper is organized into four sections. Section two reviews some of the related works available in the literature. Section three details all the techniques used in this investigation. Section four presents experimental results and section five concludes the work.

2. Related works

Agnes et al., [7] had developed Multiscale All CNN (MA-CNN) for effective breast cancer diagnosis. As a CNN-based approach, the MA-CNN was used for accurate classification of the mammogram images. The CNNs were excellent in the extraction of task-specific features due to close association of the feature learning with the classification task for the accomplishment of enhanced accuracy. The mammogram images on the mini- Mammographic Image Analysis Society (MIAS) dataset into the three distinct classes as follows: normal, malignant, and benign. Moreover, the model had enhanced the classification system's accuracy via fusion of the wider context of information by means of multiscale filters with no negotiation of the computation speed. It was evident from the experimental outcomes that the MA-CNN classified the mammogram images with 96% overall sensitivity as well as 0.99 AUC.

Chouhan et al., [8] presented Diverse Features based Breast Cancer Detection (DFeBCD) system for classification of a mammogram as either normal or abnormal. There was utilization of three sets of static feature and the fourth set of features was extracted using the proposed DFeBCD. These distinct features were trained with the following two classifiers: SVM, and the Emotional Learning inspired Ensemble Classifier (ELiEC). The application of 5-folds cross-validation was able to ensure the system performance's reliability. The experimentations had showed that the DFeBCD system's performance on dynamically generated features was superior to that of the static features. In addition, the system performance had improvements of almost 2-3% through the hybridization of all four feature types. The ELiEC classifier had better performance than that of the SVM with utilization of hybrid as well as dynamic features.

Touami & Benamrane [9] had presented an intelligent system for detection as well as analysis of the micro-calcifications in mammography with utilization of the Particle Swarm Optimization (PSO), region growing technique as well as the Probabilistic Neural Network (PNN) for the early detection of breast cancer, and also for avoidance of the breast's ablation.

Sangeetha & Prakash [10] had proposed the following four distinct stages for the Hybrid Inception Recurrent Residual Convolutional Neural Network (HIRResCNN) framework: pre-processing, dimensionality reduction, segmentation, and classification. The pre-processing stage had involved noise removal from the images by means of two distinct filtering algorithms: Median, and mean filtering. Later, the edge detection was done with the canny edge detector. The image smoothening was done with the canny edge detector's Gaussian filtering. For the subsequent dimensionality reduction stage, correlation of the attributes was executed with the Principal Component Analysis (PCA) which was inclusive of the associated features. Thus, there would be minimization of the huge dataset, and its expression would be done with just a variable. The third stage, i.e., the segmentation stage, would perform foreground as well as background subtraction for accurate breast cancer detection. In the final stage, for breast cancer detection as well as classification, there was introduction of a HIRResCNN, that would integrate Harmony Search Optimization (HSO) to tune the bias as well as the weight parameters, and also boost the classification accuracy with the HIRResCNN-HSO model. As a powerful Deep CNN (DCNN) model, the HIRResCNN was a combination of the Strength of Recurrent CNN (RCNN), the Residual Network (ResNet) as well as the Inception Network (Inception-v4). The MIAS dataset was employed for the execution of multiple experimentations as well as the result comparison with other available techniques. The proposed HIRResCNN classifier had yielded about 92.6% rate of accuracy for the breast cancer detection.

Cai et al., [11] had analyzed the mammographic images with techniques of image processing as well as a pipeline structure. The first step had involved improvements done on both the mammogram image quality as well as on the contrast of the image's abnormal areas via image contrast improvement as well as a noise decline. Later, there was utilization of a color space-based method for the image segmentation followed by the mathematical morphology. A combination of the Gray-Level Co-occurrence Matrix (GLCM) as well as the Discrete Wavelet Transform (DWT) method were employed for feature image extraction. Eventually, the feature classification was done by the Advanced Thermal Exchange Optimizer, a novel optimized version of the CNN as well as a novel improved metaheuristic. The experimental outcomes had showed that the cancer case diagnosis for the proposed method as well as its application on the MIAS database had achieved 93.79% of accuracy, 96.89% of sensitivity as well as 67.7% of specificity.

Ittannavar & Havaldar [12] had offered the proposal of a new multi-objective optimization for segmentation of the breast masses in mammograms. Two benchmark datasets such as the Digital Database for Screening Mammography (DDSM), and MIAS, were used for collection of the mammographic images. This was followed by the utilization of image normalization as well as Contrast-Limited Adaptive Histogram Equalization (CLAHE) methods for enhancement of the mammographic images' visual ability as well as contrast. After denoising the image, there was utilization of electromagnetism-like (EML) for segmentation of the mammogram image into the cancerous and non-cancerous portions. The proposed EML technique had benefits such as enhanced robustness to hold the image details as well as adaptivity to the local context. In the end, there was execution of template matching for detection of the cancer regions, and this was followed by analysis of the proposed model's effectivity with regards to Jaccard coefficient, dice coefficient, specificity, sensitivity as well as accuracy. For the DDSM dataset, the proposed model, on an average, had accomplished a sensitivity of 92.3%, a specificity of 99.21% as well as an accuracy of 98.68%. Moreover, for the MIAS dataset, the proposed model, on an average, had accomplished a sensitivity of 92.11%, a specificity of 99.45% as well as an accuracy of 98.93%.

Kanya Kumari & Naga Jagadesh [13] had utilized the following four-step procedure: pre-processing, feature extraction, feature selection, and classification. At first, there was acquisition of the medical images followed by its pre-processing with CLAHE, and later, the feature retrieval was done with the GLCM for extraction of the texture, shape as well as intensity-based features. Later, there was application of the feature selection for acquisition of better features. In this work, the authors had put forward a new feature selection technique known as the Weighted Adaptive Binary TLBO (WA-BTLBO) and had employed the classification accuracy as its fitness function. The XGBoost classifier was used for training as well as testing of the selected features while result comparison was done with other classifiers such as the K-Nearest Neighbor (KNN), random forest, ANN, and SVM. The experiments were conducted with publicly available mammogram medical images from the MIAS. It was evident from the results that the WA-BTLBO with XGBoost classifier had superior performance in comparison to other techniques of feature selection such as the PSO, and the Binary TLBO (BTLBO) as well as other advanced methods with regards to classification of the MIAS mammogram images as either normal or abnormal. Since this work had aided radiologists as well as physiologists in breast cancer detection in women, it was able to extend the patient's life span.

3. Methodology

A CNN is a deep artificial neural network that is chiefly employed for image classification. Deep learning's effectivity has experienced much increase due to the usefulness of convolutional nets (CNNs) in image processing. In particular, deep learning has a crucial role in the domain of Medical Image Processing [14]. In CNN, each neuron will take inputs from a local receptive field in the earlier layer and the local features are extracted. Each convolution layer will constitute multiple feature maps wherein each one is in the form of a plane within constraints are placed on the individual neurons This work's aim is to present a CAD system methodology that will aid the radiologists in classifying mass lesions on the basis of a CNN with a limited number of layers without any loss of features to obtain a good prediction as well as be scalable to massive datasets.

This section will discuss about the VGGNet, the ResNet, the Inception network, GA and the TLBO with ResNet approaches. All these approaches will have its input be the mammogram, and its output be the mammogram's classification as either Malignant or Benign.

3.1 Deep Learning Methods

3.1.1 Visual Geometry Group Neural Network (VGGNet):

The key premise of the VGGNet architecture is deeper networks with smaller filters. Cf For the ILSVRC 2014 challenge, this network had performed exceedingly well. It had the best performance in the image localization task, and for image classification task having a top five error rate of 7.3%. In the VGGNet a 3 x 3 convolutional filter size is used and there is an increase in the number of layers from 8 in the AlexNet to either 16 or 19 based on the model and are referred to as VGG16, and VGG19, respectively. Although the VGG19 had slightly better performance, but it had more memory consumption [15]

3.1.2 GoogLeNet:

GoogLeNet is a deeper network with 22 layers and had won the ILSVRC 2014 challenge with a top five error rate of 6.7%. This architecture's uniqueness is its inception modules which makes it very efficient with regards to the computational cost. Moreover, due to the absence of fully connected layers, this architecture has been able to save many parameters. Instead of a sequential manner, this architecture will employ diverse types of the filter operations in a parallel manner. With reduction of the GoogLeNet's feature depth of GoogLeNet with 1 x 1 convolutional bottleneck layers, the computational complexity became easier to handle.

3.1.3 Residual Neural Network (ResNet):

In 2015, the ResNet was the winner of the ILSVRC competition with a top five error rate of 3.6%. When compared with all the earlier architectures, the ResNet had superior performance in almost all the following tracks: image classification, detection, and localization. Due to its 152 layers, the ResNet is a much deeper architecture. This ultra-deep architecture will employ a stack of residual blocks in which each block is composed of two 3 x 3 convolutional layers. With the residual mapping model, there is resolution of the problem of optimization in the deeper models. In addition, the ResNet's residual blocks or skip connections will aid the architecture in avoiding the problem of gradient diminishing.

3.2 Teaching-Learning-Based Optimization (TLBO) Algorithm

Teaching-learning is a critical procedure in which every individual will try to learn something from other individuals for their self-improvement. This algorithm will simulate a classroom's traditional phenomenon of teaching learning with the following two fundamental learning modes: (i) learning via the teacher (termed the teacher phase), and (ii) learning via interactions with other learners (termed the learner phase). The TLBO technique starts with a population of students (that is, learner), and there is a similarity between the various subjects offered to the learners as well as the various design variables of the problem of optimization. Furthermore, the learner's results will be akin to the optimization problem's fitness value. The overall population's best solution will be treated as the teacher. Detailed description of the the teacher phase and the learner phase in TLBO is given below [16]:

Teacher phase: The learning of the students/learners from the teacher is simulated in this phase. A teacher will convey knowledge amongst the students, and also will make an effort to raise the class's mean result. Assume that 'n' number of learners are offered with 'm' number of subjects, at any sequential teaching-learning iteration i, Mj, i will indicate the mean result of the learners for a particular subject, 'j'. As a teacher is the most experienced as well as knowledgeable person on a particular subject, the overall population's best learner is accounted for as the algorithm's teacher. Suppose that Xtotal − kbest, i will be the best learner's result with due consideration of all the subjects. It will be accounted for as that particular cycle's teacher. Eq. (1) will express the difference between the teacher's result as well as the learners' mean result each subject as follows:

$$Difference\_Mean_{j,i} = r_i(X_{j,kbest,i} - T_F M_{j,i}) \qquad (1)$$

For the above equation, Xj, kbest, i will indicate the result of the teacher in subject j, TF will indicate the teaching factor, that will determine the value of mean to be changed, and ri will indicate the random number which exists in the range [0, 1]. TF can have a value of either 1 or 2. This value is randomly chosen with equivalent probability as below:

$$T_F = round[1 + rand(0,1)\{2-1\}] \qquad (2)$$

Where in, rand will indicate the random number in the [0, 1] range. However, the TF is not a TLBO algorithm parameter. Moreover, its value is not provided an input to the algorithm, and also the algorithm will employ Eq. (2) to randomly determine its value.

Based on the Difference_Meanj, i, the below Eq. (3) will update the existing solution in the teacher phase as below:

$$X'_{j,k,i} = X_{j,k,i} + Difference\_Mean_{j,i} \tag{3}$$

Here, $X'_{j,k,i}$ will indicate the updated value of $X_{j,k,i}$. $X'_{j,k,i}$ will be accepted if it will provide a better function value. At the end of the teacher phase, all the accepted function values will become the learner phase's inputs.

The TLBO algorithm's performance does get affected by the values of ri as well as TF. While ri will indicate a random number within the [0, 1] range. TF will indicate the teaching factor. However, the algorithm will randomly generate the ri and TF values, and these parameters are not offered as the algorithm's inputs (in contrast, the crossover as well as the mutation probabilities in the GA are offered as inputs). Therefore, the TLBO algorithm does not require the tuning of ri as well as TF (which is contrary to the tuning of the GA's crossover as well as the mutation probabilities) [17].

Learner phase: Learning of the students/learner via their interactions with one another is simulated in learner phase. Also, the learners will gain knowledge via discussions as well as interactions with other learners. Below is an expression of this phase's learning phenomenon.

There is random selection of two distinct learners, P and Q, so that $X'_{total-P,i} \neq X'_{total-Q,i}$, in which $X'_{total-P,i}$ will indicate the updated values of $X_{total-P,i}$, and $X'_{total-Q,i}$ will indicate the updated values of $X_{total-Q,i}$, towards the end of the teacher phase in the below Eq. (4) and Eq. (5):

$$X''_{j,P,i} = X'_{j,P,i} + r_i(X'_{j,P,i} - X'_{j,Q,i}), \text{ If } X'_{total-P,i} > X'_{total-Q,i} \tag{4}$$

$$X''_{j,P,i} = X'_{j,P,i} + r_i(X'_{j,Q,i} - X'_{j,Q,i}), \text{ If } X'_{total-Q,i} > X'_{total-p,i} \tag{5}$$

(The aforementioned equations are for the problems of maximization while the reverse is true for the problems of minimization.)

$X''_{j,P,i}$ will be accepted if it is able to offer a better function value.

3.2.1 Genetic Algorithm (GA)

The GA's basic premise is to replicate the 'survival of the fittest' concept; it will simulate a natural system's observed processes wherein the strong will have the tendency to adapt as well as survive while the weak will have the tendency to perish. Being a population-based technique, the GA will rank the population members on the basis of their solutions' fitness. Specific genetic operators like crossover, reproduction, and mutation are employed to form a new population in the GA. A set of strings (termed chromosomes) is used to represent the population. Each generation will create a new chromosome (i.e., a population member) by means of the information derived from the previous population's fittest chromosomes. The GA will produce population of feasible solutions, and later on, it will recombine these solutions for guiding their search towards the search space's more promising areas [18].

A chromosome's fitness function value will determine its ability to endure as well as yield offspring. While a high value of fitness will indicate the better solution for problems of maximization, a low value of fitness will indicate the better solution for problems of minimization. The reproduction procedure will pick out the population's fittest candidates, the crossover procedure will combine the fittest chromosomes, and then, will pass on the superior genes to the succeeding generation. On the hand, the mutation operation will alter certain genes within a chromosome.

The GA's operation will commence with the determination of an initial population either in a random manner or by means of certain heuristics. There is utilization of the fitness function to assess the population members, and this is followed by ranking these members on basis of their performances. Upon completion of assessment for all the population members, a popularly employed approach in the GA is to discard the low-ranked chromosomes while using the remaining population members for reproduction.

The roulette wheel technique [19] will form the basis of the selection method as detailed below:

With utilization of Eq. (6), there is evaluation of each chromosome's probability of selection:

$$P_i = \frac{f_i}{\sum_i f_i} \tag{6}$$

For the above equation, fi will indicate the value of fitness for chromosome i.

Evaluation of the cumulative probability will be according to the below Eq. (7):

$$CP_i = \sum P_i + CP_{i-1} \quad (7)$$

Here, Pi will indicate the probability of picking chromosome i while CP_{i-1} will indicate the cumulative probability of chromosome i − 1.

Let us generate a random number r ∈ [0, 1]. Eq. (8) will express the chromosome selection as below:

$$\begin{cases} if\ r < CP_i\ select\ the\ first\ chromosome \\ if\ CP < r < CP_{i+1}\ select\ the\ chromosome\ i+1 \end{cases} \quad (8)$$

The mutation is the GA's final step wherein the mutation operator will randomly mutate on a chromosome's gene. This is a vital step in the GA as it will ensure that all regions of the problem space will be reached. There is utilization of elitism to prevent destruction of the population's best solution during the operations of crossover as well as mutation. Elitism is able to ensure that the new generation's fitness will be at least as good as the existing generation. There will be continuation of the new populations' evaluation as well as generation till either the arrival at a maximum number of generations or the detection of an optimum solution.

The GA's key benefits are inclusive of the requisites' for limited parameter settings as well as its self-initialization from potential solutions in lieu of from a single solution. Nevertheless, due to the random nature of both crossover as well as mutation procedures, the GA's chief shortcoming is its lack of rapid convergence towards the optimal values. The GA does encompass a myriad of applications such as scheduling, machine learning, signal processing, routing, robotics, manufacturing, mathematics, business, etc.

3.3 Proposed Hybrid Teaching Learning-Based Optimization –Genetic Algorithm (TLBO-GA) with ResNet (18 & 34) and CNN-18

For the proposed hybrid TLBO-GA Resnet18 & 34 as well as the TLBO-GA CNN 18, there is optimization of the deep learners' architecture so as to boost the mammogram's classification as either malignant or benign. Moreover, there is optimization of the batch size, the rate of learning, the activation as well as the number of epochs of the ResNet 18 & 34 and the CNN-18.

The developed hybrid TLBO-GA [20] will combine the TLBO's calculation procedure as well as the GA's three operators. When compared with the TLBO, the hybrid TLBO-GA's optimization solutions can undergo further classification, and improvements on the calculation phases can also be done. Figure 1 will illustrate the hybrid TLBO-GA's flow chart. The aforementioned improvements are described as below.

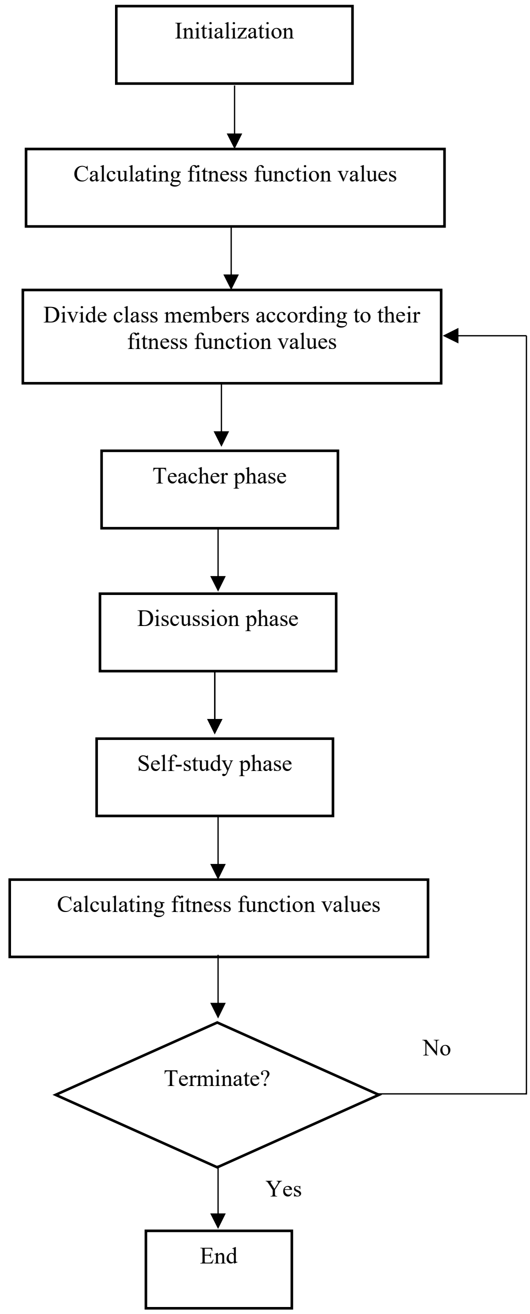

Figure 1 Flow chart of Hybrid TLBO-GA

3.3.1. Operators

The hybrid TLBO-GA will employ the GA operators of encoding, crossover as well as mutation. Direct resolution of the breast cancer problems can be done by the hybrid TLBO-GA with these operators. The encoding operator will have the configuration of the deep learners. The crossover operator will exchange certain fragments between two chromosomes. The mutation operator will involve exchanging its position with another. This operator is essential to increase the search space as well as to prevent the premature convergence.

3.3.2. Hybridizing TLBO-GA

In contrast to the TLBO, every optimization solution in the hybrid TLBO-GA will be split into the following four distinct types: the teacher, the top learners, the ordinary learners, and the poor learners. The student re-classification procedure draws its inspiration from the ranking process in schools. m will indicate the sum of all the students as well as the teachers.

Akin to the TLBO, the teacher will indicate the optimization solution whose values of fitness function are better than that of the entire class's other students. The number of teachers in the hybrid TLBO-GA will be equivalent to 1. The top learners are solutions whose value of fitness function are lower than that of the teacher, but higher than that of other learners. For the hybrid TLBO-GA, 20% of m will be the number of top students. Poor learners are solutions with values of fitness function that are much lower than that of the teacher, the top students as well as the ordinary students. Poor students will neither learn from the teacher nor the other students. Self-study is the only approach used by these students to update their own coding. In the hybrid TLBO-GA, 10% of m will be the number of poor students. Additionally, the objective of setting these students in this algorithm is for prevention of the premature convergence as well as for increasing the search space. About 70% of m will be the number of ordinary students in the hybrid TLBO-GA. Although the ordinary students have higher values of fitness function values of ordinary students are higher than that of the poor students, their values of fitness function are still much lower than that of the teacher as well as the top students.

The hybrid TLBO-GA's "Teacher phase" will simulate a class's teaching procedure. This phase will involve all the top students as well as the ordinary students learning from the teacher. In the hybrid TLBO-GA, the "Discussion phase" will simulate a class's discussion procedure. Even so, this phase will not involve the participation of either the teacher or the poor students. The poor students will not carry out any calculation procedure during the "Teacher phase" as well as the "Discussion phase." However, they will lead the self-learning.

Self-learning procedure will involve all the students carrying out the mutation process. The mutation probability of the poor students would be greater than that of the top students as well as the ordinary students due to the non-participation of the poor students in the first two calculation phases. This procedure is critical to increase the search space, and also to overcome the premature convergence.

Upon utilization of the hybrid TLBO-GA [21] for a deep CNN's training, the batch size, learning rate, activation, number of epoch are encoded as solutions. It is evident from the experimental outcomes that the application of this encoding will result in the evolution procedure's more rapid rate of convergence as well as a consequent higher accuracy threshold. The steps can be generally described as follows:

1. Initialization of a population (TLBO and GA) of individual chromosomes. Each chromosome will indicate the batch size, learning rate, activation, number of epoch of the deep learner.
2. Assignation of a value of fitness (hybrid TLBO-GA) to every population member on the basis of evaluation of its representative network.
3. Picking individuals on the basis of their values of fitness as parents for the reproduction of a new generation of individuals.
4. Reproduction of new children is done via execution of the crossover as well as the mutation operations between the chosen parents.
5. Repetition from Step 2 till fulfilment.

4. Results and Discussion

In this section, the Resnet18, Resnet34, CNN-18, TLBO-Resnet18, TLBO-Resnet34, CNN-18 layer-TLBO, TLBO-GA Resnet 18, TLBO-GA Resnet 34 and CNN-18-layer TLBO-GA methods are evaluated for classifying mammograms. The Curated Breast Imaging Subset of Digital Database for Screening Mammography (CBIS-DDSM) is used for evaluating the various algorithms. The DDSM is a database of 2,620 scanned film mammography studies. It contains normal, benign, and malignant cases with verified pathology information. In this work, 550 Benign and 625 Malignant mammogram images are used for evaluation. Python, open CV, tensor flow, keras are used for the implementation of the algorithms. Table 1 shows the summary of results. The classification accuracy, recall, precision and f measure as shown in figures 3 to 6.

Table 1 Summary of Results

Techniques	Resnet18	Resnet 34	CNN-18 layer	TLBO Resnet18	TLBO Resnet34	CNN-18 layer - TLBO	TLBO-GA Resnet 18	TLBO-GA Resnet 34	CNN-18 layer – TLBO GA
Accuracy	89.53	90.21	92.77	93.28	94.47	95.74	95.23	95.57	97.19
Recall for Benign	0.9055	0.9109	0.9327	0.94	0.9491	0.9564	0.9509	0.9564	0.9727
Recall for Malignant	0.8864	0.8944	0.9232	0.9264	0.9408	0.9584	0.9536	0.9552	0.9712
Precision for Benign	0.8752	0.8836	0.9144	0.9183	0.9338	0.9529	0.9475	0.9495	0.9675
Precision for Malignant	0.9142	0.9194	0.9397	0.9461	0.9545	0.9615	0.9567	0.9614	0.9759
F Measure for Benign	0.8901	0.897	0.9235	0.929	0.9414	0.9546	0.9492	0.9529	0.9701
F Measure for Malignant	0.9001	0.9067	0.9314	0.9361	0.9476	0.9599	0.9551	0.9583	0.9735

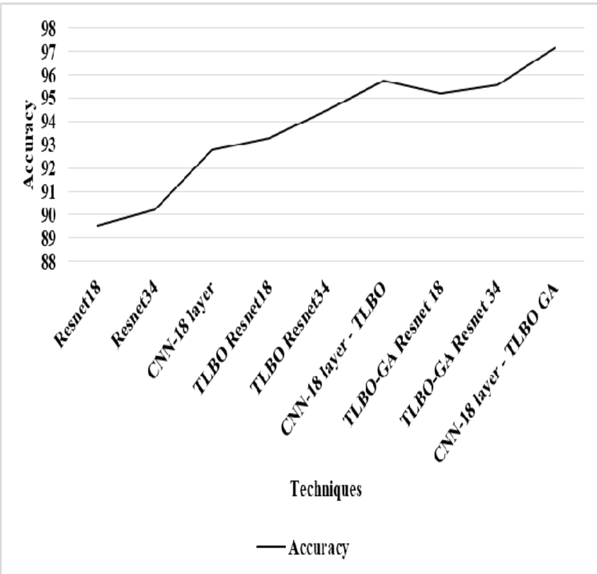

Figure 2 Accuracy for CNN-18 layer-TLBO GA

From the figure 2, it can be observed that the CNN-18 layer - TLBO GA has higher accuracy by for 8.2% Resnet 18, by 7.45% for Resnet 34, by 4.65% for CNN-18 layer, by 4.1% for TLBO Resnet 18, by 2.84% for TLBO Resnet 34, by 1.5% for CNN-18 layer - TLBO, by 2.04% for TLBO-GA Resnet 18 and by 1.68% for TLBO-GA Resnet 34 respectively.

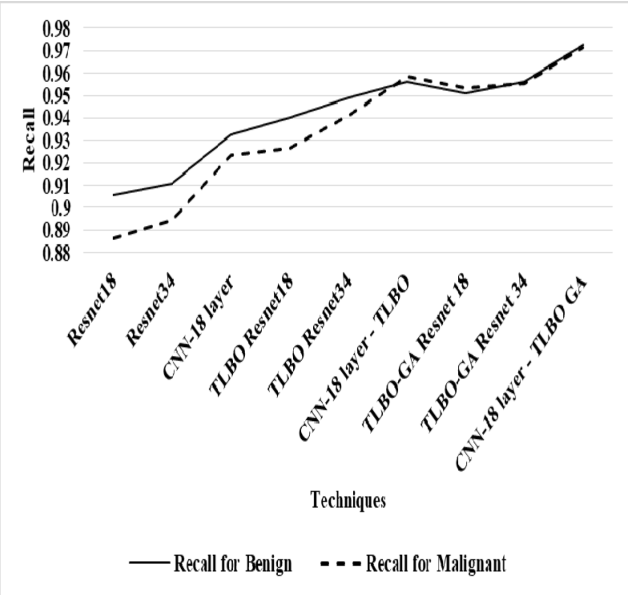

Figure 3 Recall for CNN-18 layer-TLBO GA

From the figure 3, it can be observed that the CNN-18 layer - TLBO GA has higher average recall by for 8.14% Resnet 18, by 7.39% for Resnet 34, by 4.63% for CNN-18 layer, by 4.06% for TLBO Resnet 18, by 2.82% for TLBO Resnet 34, by 1.51% for CNN-18 layer - TLBO, by 2.05% for TLBO-GA Resnet 18 and by 1.67% for TLBO-GA Resnet 34 respectively.

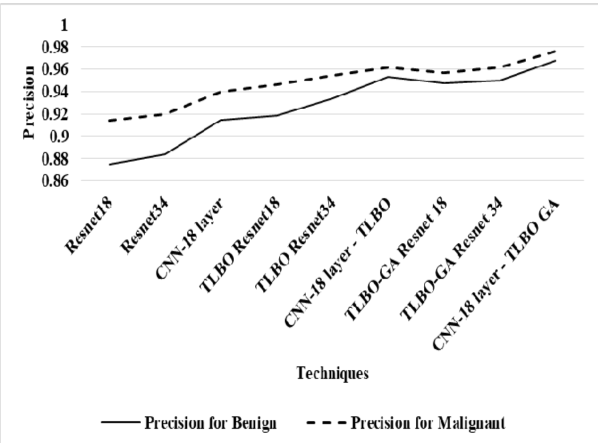

Figure 4 Precision for CNN-18 layer-TLBO GA

From the figure 4, it can be observed that the CNN-18 layer - TLBO GA has higher average precision by for 8.25% Resnet 18, by 7.49% for Resnet 34, by 4.7% for CNN-18 layer, by 4.14% for TLBO Resnet 18, by 2.87% for TLBO Resnet 34, by 1.5% for CNN-18 layer - TLBO, by 2.04% for TLBO-GA Resnet 18 and by 1.68% for TLBO-GA Resnet 34 respectively.

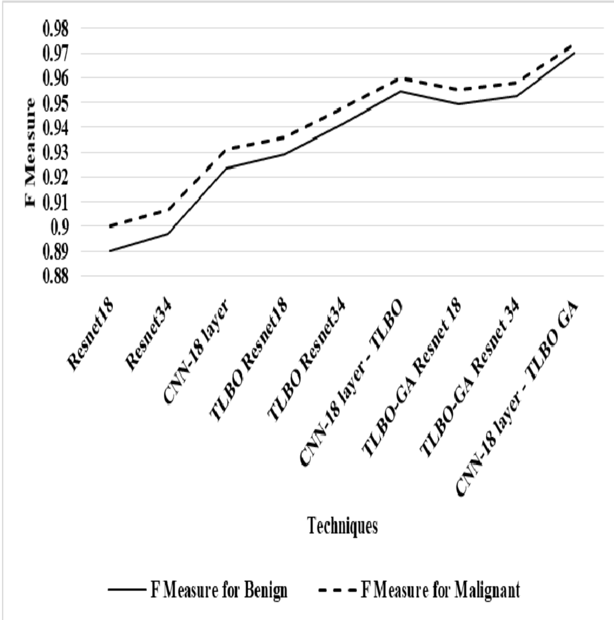

Figure 5 F Measure for CNN-18 layer-TLBO GA

From the figure 5, it can be observed that the CNN-18 layer - TLBO GA has higher average f measure by for 8.22% Resnet 18, by 7.46% for Resnet 34, by 4.67% for CNN-18 layer, by 4.12% for TLBO Resnet 18, by 2.84% for TLBO Resnet 34, by 1.51% for CNN-18 layer - TLBO, by 2.04% for TLBO-GA Resnet 18 and by 1.68% for TLBO-GA Resnet 34 respectively.

5. Conclusion

A mammogram is defined as an image of the breast which is employed for the detection as well as diagnosis of breast cancer. This work's focus is on a CNN-based CAD system that will employ the deep learning concept for classification of the mammogram images into the following three distinct groups: benign, malignant, and normal. Of late, CNNs have garners much interest as a result of their ever-increasing real-world applications. The CNN performance has high dependence on the network structure as well as the chosen method of optimization for network parameter tuning. Nivel as well as effective training methods for the ResNet as well as the CNNs have been submitted in this work. At present, majority of the existing advanced CNN learning approaches are based on the Gradient Descent concept. Unlike the standard methods of CNN training, proposal for the Hybrid TLBO-GA-based methods have been offered for optimization of the ResNet as well as the CNNs. Successful application of the hybrid TLBO-

GA has been done on problems in which the candidate representations have been done. Results show that the CNN-18 layer - TLBO GA has higher accuracy by for 8.2% Resnet 18, by 7.45% for Resnet 34, by 4.65% for CNN-18 layer, by 4.1% for TLBO Resnet 18, by 2.84% for TLBO Resnet 34, by 1.5% for CNN-18 layer - TLBO, by 2.04% for TLBO-GA Resnet 18 and by 1.68% for TLBO-GA Resnet 34 respectively.

References

[1] Malebary, S. J., & Hashmi, A. (2021). Automated breast mass classification system using deep learning and ensemble learning in digital mammogram. IEEE Access, 9, 55312-55328.

[2] Hamed, G., Marey, M. A. E. R., Amin, S. E. S., & Tolba, M. F. (2020, April). Deep learning in breast cancer detection and classification. In The International Conference on Artificial Intelligence and Computer Vision (pp. 322-333). Springer, Cham.

[3] Al-Antari, M. A., Al-Masni, M. A., & Kim, T. S. (2020). Deep learning computer-aided diagnosis for breast lesion in digital mammogram. Deep Learning in Medical Image Analysis, 59-72.

[4] El Houby, E. M., & Yassin, N. I. (2021). Malignant and non-malignant classification of breast lesions in mammograms using convolutional neural networks. Biomedical Signal Processing and Control, 70, 102954.

[5] Cai, H., Huang, Q., Rong, W., Song, Y., Li, J., Wang, J., ... & Li, L. (2019). Breast microcalcification diagnosis using deep convolutional neural network from digital mammograms. Computational and mathematical methods in medicine, 2019.

[6] Jung, H., Kim, B., Lee, I., Yoo, M., Lee, J., Ham, S., ... & Kang, J. (2018). Detection of masses in mammograms using a one-stage object detector based on a deep convolutional neural network. PloS one, 13(9), e0203355.

[7] Agnes, S. A., Anitha, J., Pandian, S., & Peter, J. D. (2020). Classification of mammogram images using multiscale all convolutional neural network (MA-CNN). Journal of medical systems, 44(1), 1-9.

[8] Chouhan, N., Khan, A., Shah, J. Z., Hussnain, M., & Khan, M. W. (2021). Deep convolutional neural network and emotional learning based breast cancer detection using digital mammography. Computers in Biology and Medicine, 132, 104318.

[9] Touami, R., & Benamrane, N. (2021). Microcalcification Detection in Mammograms Using Particle Swarm Optimization and Probabilistic Neural Network. Computación y Sistemas, 25(2), 369-379.

[10] Sangeetha, K., & Prakash, S. (2021). Hybrid Inception Recurrent Residual Convolutional Neural Network (HIRResCNN) with Harmony Search Optimization (HSO) for Early Breast Cancer Detection System. NeuroQuantology, 19(7), 1.

[11] Cai, X., Li, X., Razmjooy, N., & Ghadimi, N. (2021). Breast Cancer Diagnosis by Convolutional Neural Network and Advanced Thermal Exchange Optimization Algorithm. Computational and Mathematical Methods in Medicine, 2021.

[12] Ittannavar, S. S., & Havaldar, R. H. (2022). Segmentation of Breast Masses in Mammogram Image Using Multilevel Multiobjective Electromagnetism-Like Optimization Algorithm. BioMed Research International, 2022.

[13] Kanya Kumari, L., & Naga Jagadesh, B. (2022). An adaptive teaching learning based optimization technique for feature selection to classify mammogram medical images in breast cancer detection. International Journal of System Assurance Engineering and Management, 1-14.

[14] Gnanasekaran, V. S., Joypaul, S., Sundaram, P. M., & Chairman, D. D. (2020). Deep learning algorithm for breast masses classification in mammograms. IET Image Processing, 14(12), 2860-2868.

[15] Khan, H. N., Shahid, A. R., Raza, B., Dar, A. H., & Alquhayz, H. (2019). Multi-view feature fusion based four views model for mammogram classification using convolutional neural network. IEEE Access, 7, 165724-165733.

[16] Chen, X., Xu, B., Yu, K., & Du, W. (2018). Teaching-Learning-Based Optimization with Learning Enthusiasm Mechanism and Its Application in Chemical Engineering. Journal of Applied Mathematics, 2018.

[17] Rao, R. V., & Patel, V. (2013). An improved teaching-learning-based optimization algorithm for solving unconstrained optimization problems. Scientia Iranica, 20(3), 710-720.

[18] Ab Wahab, M. N., Nefti-Meziani, S., & Atyabi, A. (2015). A comprehensive review of swarm optimization algorithms. PLoS One, 10(5), e0122827.

[19] Belhaj Soulami, K., Kaabouch, N., Saidi, M. N., & Tamtaoui, A. (2020). An evaluation and ranking of evolutionary algorithms in segmenting abnormal masses in digital mammograms. Multimedia Tools and Applications, 79(27), 18941-18979.

[20] Li, Z., Huang, J., Wang, J., & Ding, M. (2021). Development and application of hybrid teaching-learning genetic algorithm in fuel reloading optimization. Progress in Nuclear Energy, 139, 103856.

[21] Esfahanian, P., & Akhavan, M. (2019). Gacnn: Training deep convolutional neural networks with genetic algorithm. arXiv preprint arXiv:1909.13354.