

Elsevier required licence: © <2022>. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

The definitive publisher version is available online at

<https://www.sciencedirect.com/science/article/pii/S0011916422001886?via%3Dihub>

# A Novel Breast Cancer Diagnosis Scheme With Intelligent Feature and Parameter Selections

S.Punitha<sup>a</sup>, Thompson Stephan<sup>b</sup>, Amir H Gandomi<sup>c,\*</sup>

<sup>a</sup>*Department of Computer Science Engineering, Karunya Institute of Technology and Sciences, Tamilnadu, India*

<sup>b</sup>*Department of Computer Science Engineering, Faculty of Engineering and Technology, M. S. Ramaiah University of Applied Sciences, Bengaluru, India*

<sup>c</sup>*Faculty of Engineering and Information Technology, University of Technology Sydney, Australia*

## Abstract:

**Background and Objective:** Breast cancer is the most commonly occurring cancer among women, which contributes to the global death rate. The key to increasing the survival rate of affected patients is early diagnosis along with appropriate treatments. Manual methods for breast cancer diagnosis fail due to human errors, inaccurate diagnoses, and are time-consuming when demands are high. Intelligent systems based on Artificial Neural Network (ANN) for automated breast cancer diagnosis are powerful due to their strong decision-making capabilities in complicated cases. Artificial Bee Colony, Artificial Immune System, and Bacterial Foraging Optimization are swarm intelligence algorithms that solve combinatorial optimization problems. This paper proposes two novel hybrid Artificial Bee Colony (ABC) optimization algorithms that overcome the demerits of standard ABC algorithms. First, this paper proposes a hybrid ABC approach called HABC, in which the standard ABC optimization is hybridized with a modified clonal selection algorithm of the Artificial Immune System that eliminates the poor exploration capabilities of standard ABC optimization. Further, this paper proposes a novel hybrid Artificial Bee Colony (Hybrid ABC) optimization where the strong explorative capabilities of the chemotaxis phase of the bacterial foraging optimization are integrated with a spiral model-based exploitative phase of the ABC by which the proposed Hybrid ABC overcomes the demerits of poor exploration and exploitation of the standard ABC algorithm.

**Methods:** In this work, the two proposed hybrid approaches were used in concurrent feature selection and parameter optimization of an ANN model. The proposed algorithm is implemented using various back-propagation algorithms, including resilient back-propagation (HABC-RP and Hybrid ABC-RP), Levenberg Marquart (HABC-LM and Hybrid ABC-LM), and momentum-based gradient descent (HABC-MGD and Hybrid ABC-GD) for parameter tuning of ANN. The Wisconsin breast cancer dataset was used to evaluate the performance of the proposed algorithms in terms of accuracy, complexity, and computational time.

**Results:** The mean accuracy of the proposed HABC-RP was 99.14% and 99.54% for Hybrid ABC which is better than the results found in the existing literature. HABC-RP attained a sensitivity of 98.32%, a specificity of 99.63%, and a precision of 99.38% whereas Hybrid ABC attained sensitivity of 99.08% and Specificity of 99.81%.

**Conclusions:** HABC-RP and Hybrid ABC-RP yielded high accuracy with a low complexity ANN structure compared to other variants. After evaluation, interestingly it is found that the Hybrid ABC-RP has achieved the highest mean accuracy of 99.54% with low complexity of 10.25 mean connections when compared to other variants proposed in this paper. It can be concluded that the concurrent selection of input features and tuning of parameters of ANN plays a vital role in increasing the accuracy of a breast cancer diagnosis. The proposed HABC-RP and Hybrid ABC-RP showed better results when compared to the existing breast cancer diagnosis systems taken for comparison. In the future, the proposed two-hybrid approaches can be used to generate optimal thresholds for the segmentation of tumors in abnormal images. HABC and Hybrid ABC can be used for tuning the parameters of various classifiers.

**Keywords:** Neural networks, Bacterial Foraging Optimization, Artificial Bee Colony, Artificial Immune Systems, Levenberg Marquardt backpropagation, Resilient backpropagation.

## 1. Introduction

Breast Cancer at its early stages has to be diagnosed to have control over the increased mortality rate of women. Annually, millions of invasive breast cancers cases are occurring every year in which 400,000 women expire because of breast cancers according to the World Health Organization (WHO) [1]. According to 2019 statistics, 12.4% of women population were affected in the US because of breast cancers. In 2020, 276,480 invasive breast cancer patients are expected and 48,530 new non-invasive cases are expected [2]. India registers 28 to 35 % of affected women cities like Chennai, Bhopal, Mumbai, Delhi, Bangalore Ahmadabad [3].

Early detection plays a major impact on the prognosis of breast cancer [4]. Screening methods for breast cancer diagnosis are Magnetic Resonance Imaging (MRI), clinical check-ups, self-checks, ultrasound scanning, mammography and biopsies for suspected cases [5]. Traditional methods for breast cancer detection fails because of the factors such as high time consumption and inaccuracy in diagnosing complicated cases. Hence, automated methods eliminate the unwanted biopsies, unwanted treatments and reduce the cost incurred [6]. ANN-based intelligent systems are used for effective and automatic breast cancer diagnosis [7]. Its

powerful decision-making capabilities increase the prediction rate so that it is used for diagnosing various diseases [8]. ANN has predominantly outperformed statistical techniques in classifying patterns for real-time applications [9].

The algorithms that are inspired by the natural evolution process is the Evolutionary Algorithms (EA). The commonly used EA that works under the principle of survival to fittest is the Genetic algorithm (GA) [10], Differential Evolution (DE) [11], Evolution Strategy (ES) [12]. Enhancements have been made to the standard EA by many researchers. An integrated algorithm called ES-DE where DE is coupled with eagle strategy is proposed to have an improved search process for constrained problems [13]. A biography-based EA is proposed by Simon, that can be used for recombination and crossover of GA [14]. Another modified variant of GA is proposed for called an Adaptive Network-based Fuzzy Inference Systems [15].

Swarm intelligence is a sub class of metaheuristics that can be used in solving various optimization problems of real-time applications as demonstrated in [16, 17]. The common swarm intelligence approaches are the Ant colony Optimization (ACO) inspired by the foraging nature of ants [18], Particle Swarm Optimization (PSO) inspired by the social behaviour of birds or fishes [19] and can be used to solve various optimization problems as demonstrated in [20]. Another technique that is proposed based on the micro bat's echolocation capability is the Bat Algorithm (BA) [21]. A metaheuristic approach inspired by the foraging behaviour of honey bees is introduced called the ABC optimization[22]. A swarm approach based on the behaviour of dragonflies called Dragonfly Algorithm (DA) is proposed [23]. Krill Herd (KH) algorithm inspired by the herding of a krill is proposed [24]. A swarm optimization called Monarch Butterfly Optimization (MBO) inspired by the migration behaviour of monarch butterflies is proposed [25]. A bio-inspired technique called Bacterial Foraging Optimization (BFO) inspired by the foraging process of E. coli bacteria is proposed by [ 26]. An Artificial Immune System (AIS) for solving various engineering optimization problems is proposed [27]. Another optimization algorithm called Salp Swarm Algorithm (SSA) inspired by the swarming behaviour of salps in the ocean is proposed [28]. A Marine Predictor Algorithm (MPA) is inspired by the movements that happen in the foraging process of predators [29]. A metaheuristic hybrid approach for cancer diagnosis is proposed by integrating ABC with MBO and is used for cancer diagnosis [30]. This paper has focussed on improving the exploration and exploitation of the standard ABC by hybridizing ABC with BFO. The proposed algorithm is called the Hybrid ABC algorithm. This proposed Hybrid ABC has been used for optimal feature subset selection and

parameter optimization of an ANN and the resulting optimized ANN classifier is utilized for improved breast cancer diagnosis.

ANN performance mainly depends on generating an optimal ANN topology [31]. ANN topology design has been focussed on by many pieces of research in the existing literature. Choosing optimal hidden layer and nodes directly influence the predictive capability of the ANN classifier. Too many hidden nodes increase ANN complexity which causes problems of overfitting where the classifier is over-trained producing accurate results in the training phase but showing decreased generalization in the testing phase. Fewer nodes in the hidden layer results in inadequate ANN learning of input and output patterns which will affect the generalization power of the ANN classifier. One of the commonly used traditional methods is called the constructive method in which the ANN is constructed using minimum neurons and gradually increased till maximum accuracy is achieved. Further, destructive or pruning methods starts with maximum neuron size and are then eliminated gradually till accuracy starts decreasing or till it has gained its maximum accuracy. These methods failed to generate optimal hidden nodes for large datasets decreasing the accuracy because of which it cannot be applied for medical diagnosis purposes. Further, manual selection of hidden nodes failed to generate optimal hidden nodes size making the ANN classifier either undertrained or overtrained. Considering the demerits of the traditional methods, our motivation is to optimize the ANN design process with the help of Evolutionary Algorithms (EA) where EA will be used to generate the optimal hidden node sizes for an ANN model. Due to the strong diversification and intensification characteristic of ABC optimization, our work has utilized ABC to search for the optimal hidden nodes for ANN and the generalization capability of the resulting optimized ANN model is tested using WBCD.

Successful learning increases the generalization capability of an ANN. Backpropagation is the widely used ANN learning method in which the weights are optimized by calculating error function gradient and propagating back in the ANN [32]. The merits of backpropagation are its simplicity, fast execution and implementation ease. Learning with the help of backpropagation is dependent on the appropriate initialization of weights [33]. Improper initial weights make the training process stuck in the local optimum resulting in inadequate learning. Generation of Random weights is the traditional way of the weight initialization process. The trial and error method has been used by researchers for weights initialization. These methods generate improper weights making them too low or high resulting in a longer learning process. Further, the zero-initialization method does not break the symmetry of the ANN classifier. Considering the demerits of the traditional initialization methods, our motivation is to hybridize the ANN design

with an efficient meta-heuristic approach such as ABC optimization and we have used it for searching the optimal initial weights of a backpropagation training method and the resulting ANN classifier has been tested using WBCD.

Further, ANN performance is dependent on the appropriate selection of features that acts as input of an ANN classifier [34,35]. ANN neurons size is dependent on a number of input features and hence, this paper has coupled the feature selection process with the ANN design process. Too many input features make ANN complex requiring more computational time with a large set of input neurons and connections. Inappropriate input features also make ANN deal with the problems of overfitting and underfitting. Further, Feature selection can also eliminate the curse of dimensionality in which memory storage and other requirements increase when full feature subsets are used. Traditional methods like filter methods work by calculating the correlation between input features and output variables that are tested using scores generated by a statistical test. In the case of the wrapper method, the usefulness of a feature subset is evaluated by training classifier using that particular set. Researches have proved that filter methods are faster consuming less time but fail to produce optimal features when tested on different classifiers whereas the wrapper method has produced the best feature subsets. Based on this, our paper has used a wrapper-based approach without the use of any primary statistical filtering.

Many researchers have utilized swarm-based approaches such as ABC optimization for generating an optimized ANN by selecting optimal initial weights, optimal learning rate, optimal hidden layer and optimal hidden node size [36]. Several studies have not concentrated on the selection of the best training method to be integrated with ANN classification. Our work has concentrated on choosing the best training algorithm which can be induced with ANN classification for successful training of WBCD datasets.

The standard ABC used in existing works for feature selection and ANN parameter optimization lacks local search capability due to inefficient exploitation. Loss of diversified solutions in the initialization phase weakens the explorative capability of ABC optimization. The standard ABC suffers from problems of unbalanced exploration and exploitation making the search process prematurely converge at the local solutions. Because of the above-mentioned reasons, the standard ABC fails to bring the global optimum input subsets and ANN design parameters and it brings only localized solutions. Hence, the motivation is to hybridize the standard ABC optimization algorithm to enhance the exploration and exploitation processes such that the developed hybrid ABC optimization is capable of generating the optimal input features

and design parameters for an ANN classifier and optimized ANN model performance is evaluated for its prediction accuracy using WBCD.

## 1.1 Artificial Bee Colony (ABC) Optimization

ABC is a population-based metaheuristic search technique. It lies on swarm intelligence based on the foraging process of the honey bees. This algorithm is introduced by Karaboga in 2005 and it is capable of solving multidimensional problems. This intelligent algorithm has stochastic process and they are flexible and simple. Algorithm (1) shows the ABC optimization process.

### Algorithm 1: ABC Optimization

**Initialization:** Random generation of food sources using Equation (1)

$$B_m^l = B_m^l + \text{random}(0,1) * (B_{max}^l - B_{min}^l) \quad (1)$$

$B_m^l$  represents  $m^{\text{th}}$  food source with  $l^{\text{th}}$  parameter and  $m=1, 2, \dots, X$  and  $X$  represents food sources size. Where  $l=1, 2, \dots, Y$  and ‘ $Y$ ’ is the dimension of the considered optimization problem.  $\text{random}(0,1)$  is step size based on a random number generated in between 0 and 1.  $B_{max}^l$  and  $B_{min}^l$  is maximum and minimum bound of the  $l^{\text{th}}$  parameter of the considered optimization problem.

**Evaluating food source quality:** Fitness is identified for an individual food source  $B_m$ .

**Employed bee Phase:** Each  $B_m^l$  is given to employee bees and they look for  $F_m^l$  which is the neighbourhood of  $B_m^l$  using Equation (2).

$$F_m^l = B_m^l + \text{random}[-1,1](B_m^l - B_n^l) \quad (2)$$

$B_n^l$  is a random food source where  $n \in \{1, 2, \dots, X\}$ ,  $X$  represents food sources size. ‘ $l$ ’ is random and  $l = \{1, 2, \dots, Y\}$  and ‘ $n$ ’ not be equal to ‘ $m$ ’ for exploitation. If fitness of  $F_m^l$  is more than  $B_m^l$ , employee bee rejects  $B_m^l$  and saves  $F_m^l$ .

**Onlooker Bee Phase:** Information of selected food source are shared to onlookers where a probability  $X_i$  is evaluated using  $\text{fitness}(B_m)$  by Equation (3).

$$X_i = \frac{\text{fitness}(B_m)}{\sum_{m=1}^X \text{fitness}(B_m)} \quad (3)$$

The value  $X_i$  each food source is compared with a random number between 0 and 1. Food sources that have greater  $X_i$  than the random number are assigned to onlooker bees and they exploit those food sources using Equation (2).

**Food source selection:** Good quality food source selected and saved.

**Scout Bee Phase:** New food sources are derived by scouts using Equation (1). Unimproved food sources for certain iterations are identified using counter checks are identified, eliminated and replaced by scouts.

## **1.2 Artificial Immune System (AIS)**

AIS is an intelligent system inspired by Biological Immune System (BIS). Intelligent theories of AIS are negative selection, danger theory, immune network model, clonal selection, and population compression [74]. Negative selection, positive selection, and population compression techniques are used in our proposed work due to their strong diversification process. Negative selection is based on the way BIS filters out T-cells population. Each T-cell is evaluated for its tolerance level to self-cells; T-cells that respond to self-cells are eliminated. Thus, T-cells that do not recognise self-cells are maintained. Clonal selection is inspired by the way BIS eliminates antigens from the body via cloning and maturation. [75, 76] used clonal selection to identify optimum solutions. Positive selection selects T-cells that only recognise non-self-cells and the rest of the T-cells are eliminated.

## **1.3 Investigation of ABC and AIS for exploration and exploitation**

. During exploration, population compression, positive selection, and negative selection force the AIS to search for new promising regions, which ensures strong exploration and avoidance of premature convergence and interference between individuals. In contrast, ABC uses scout bees for exploration only with a random search process. This forces the algorithm to generate solutions that are concentrated in a local area, thereby losing its exploration capability at initialization. As a result, the algorithm prematurely converges and returns sub-optimal solutions. Hence AIS is better at exploration than the ABC algorithm.

In the context of exploitation, the clonal selection of AIS is weak since it performs local search only by producing clones of the best antibodies. The amount of maturation depends on the affinity of antibodies and hence some of the antibodies may not be properly exploited. Thus, the AIS always lacks in the local search process. Comparatively, ABC has good exploitation, since exploitation is carried by both employee and onlooker bees. Hence, ABC has better exploitation as compared to AIS.

## **1.4 Bacterial Foraging Optimization**

BFO is swarm-based which is inspired by the intelligent foraging process of Escherichia Coli bacteria [37]. E. coli searches for nutrient regions for their survival. Bacterium represents possible solutions to the optimization process. It randomly moves to find nutrient regions in presence of attractants and repellents called chemotaxis of the bacteria. Four phases are in the



bacterial search process. They are (i) chemotaxis phase, (ii) swarming phase, (iii) reproduction phase, (iv) elimination and dispersal phase. The chemotaxis is considered an important phase of the forging process. The bacterium tumbles randomly in the direction specified by Equation (4) then swims in the direction of the tumble represented by Equation (5).

$$\phi(k) = \frac{\Delta(k)}{\sqrt{\Delta^T(k)\Delta(k)}} \quad (4)$$

$\phi(k)$  is a unit vector representing the tumble direction which is based on vector  $\Delta(k)$  between  $[-1, 1]$ .

$$\theta^k(k+1, l, n) = \theta^k(k, l, n) + X(k)\phi(k) \quad (5)$$

$\theta^k(k+1, l, n)$  is new bacterial position of  $k^{\text{th}}$  bacterium in  $k^{\text{th}}$  chemotaxis,  $l^{\text{th}}$  reproduction and  $n^{\text{th}}$  elimination dispersal step.  $\theta^k(k, l, n)$  is old position of bacterium.  $X(k)$  is the step size for  $k^{\text{th}}$  bacterium taken in problem space. If  $\theta^k(k+1, l, n)$  is better than  $\theta^k(k, l, n)$ , it swims in the same direction as that of tumble otherwise tumbles in a different direction. This continues till maximum swim step size  $M_s$ . The algorithm for the chemotaxis phase of BFO is in Algorithm (2).

### Algorithm 2: Chemotaxis phase of BFO

Initialize  $B, C_{cl}, M_s$  and place bacteria randomly

Elimination dispersal loop:  $n=n+1$

Reproduction loop:  $l= l + 1$

Chemotaxis loop:  $k = k+1$

WHILE  $k < C_{cl}$

For  $k=1$  to  $B$ , perform Chemotaxis for  $k^{\text{th}}$  bacterium

Compute fitness ( $\theta^k(k, l, n)$ )

Save  $\theta^{best} = \theta^k(k, l, n)$

Tumble using Equation (4)

Swim using equation (5)

Compute fitness ( $\theta^k(k+1, l, n)$ )

Initialize counter  $y = 0$

WHILE  $y < M_s$

$y = y + 1$

IF  $fitness(\theta^k(k+1, l, n)) > fitness(\theta^{best})$

Save  $\theta^{best} = \theta^k(k+1, l, n)$

Swim using equation (5)

ELSE

$$y = M_s$$

ENDIF

ENDWHILE

Proceed with next bacterium if  $k \neq B$

ENDWHILE

### 1.5 Comparative Investigation of the exploitation and exploration of ABC and BFO

The performance of BFO depends on step size  $X(k)$  taken during the foraging process. Fixed step sizes create an imbalance between exploration and exploitation. Larger  $X(k)$  makes BFO converge faster making few movements reaching global optimum point finding out promising areas [38]. Proper step sizes should be used to create balance in exploration and exploitation. A larger step size introduces oscillation making back and forth movement around the optimum locations. Larger step sizes may not make bacterium search space completely without reaching remotely located global optimum points. For BFO, Proper exploration is guaranteed in the tumble direction when step sizes are large whereas ABC uses scout bees for exploration only by random search. Hence, Loss of diversified solutions exists in initialization making solutions concentrated only around a local area returning suboptimal solutions. Hence, BFO is better in exploration than standard ABC.

Smaller  $X(k)$  makes bacterium take small steps which makes it reach various locations in the fitness landscape. However, the algorithm converges slower requiring high computational time to search complete search space. Comparatively, ABC has good exploitation since the intensification process is done by both employee and scout bee phases. Hence, ABC has better exploitation when compared to BFO.

### 1.6 Issues with standard ABC Optimization addressed by the proposed Hybrid algorithms

Even still the standard ABC is widely used by the researchers in the feature selection process and fine-tuning of parameters of an ANN classifier for different medical diagnoses, they suffer from the below-listed issues which the proposed Hybrid ABC resolves.

- (i) The new food source is generated by modifying a single parameter of the current food source, which leads to the same food sources converging at similar optimum locations.
- (ii) Local search is carried out with the help of worker bees and the onlooker bees, the local search process is conducted only by the scout bees, leading to an imbalance in local and global search capabilities.
- (iii) The exploitative selection based on greedy selection by employee bees generates uphill movements and downhill movements are neglected completely which rejects the worst solutions producing suboptimal solutions.
- (iv) Local search of ABC has the tendency of revisiting solutions that causes looping that leads to premature converge.
- (v) Loss of diversification at initialization due to the inefficient random search process.

Based on the comparative investigation and to overcome the aforementioned problems, this paper focuses on the following research objectives:

- (i) Development of a Hybrid ABC Optimization (HABC) that combines the strong explorative capabilities of AIS with a Tabu-based employee bee phase of ABC to enhance the local and global search capabilities of the standard ABC optimization.
- (ii) To develop a Hybrid ABC Optimization by integrating the explorative phase of BFO using an adaptive step size with ABC Optimization such that the ABC overcomes the problem of poor diversification at the initialisation phase. Further, the exploitation of ABC gets improved by using the spiral model-based employee bee phase.
- (iii) To evaluate the proposed hybrid algorithms in optimal input feature selection and tuning of ANN parameters such as the initial weights and hidden node size using the Wisconsin breast cancer dataset. The derived optimized ANN is evaluated with respect to accuracy, complexity and computational time.

## **2. Related works**

ABC and ACO are hybridized and proposed for finding out optimal feature subsets [39]. The time-consuming global search of ABC is improved by using generated features produced by ACO that are given to the employee phase of ABC. The best ants are generated by the bee colony exploitation phase which eliminates ant's stagnation behaviour. The proposed system has focussed on feature selection and there is no tuning of parameters of the classifier used. The

exploitative greed selection of ABC has been used which may allow only a better solution leading to premature convergence. Another Hybrid algorithm based on DE and ABC is proposed for medical diagnosis where the exploration of the DE algorithm is integrated with the ABC onlooker bee phase which is improved using the DE operators and this system has tried to achieve the exploration and exploitative balance[40]. The proposed algorithm is tested for using UCI datasets where the proposed achieved F-measure of 97.6, 92.2, 96.4, for RBF networks, decision tree and Naive Bayes classifier respectively using WBCD. The proposed system has focussed on feature selection and there is no tuning of parameters of the classifier used.

A breast cancer diagnosis system with the help of ANN and ABC is proposed [41]. The ABC finds out optimal hidden layers and hidden nodes of ANN trained using Levenberg–Marquardt under WBCD. The accuracy is 95.9% using the testing set. The system uses standard ABC and hence lacks in the exploration at initialization and proposed does not optimize parameters of ANN. ABC based system is proposed using FFNN [42] for medical diagnosis. The ABC is utilized to produce optimal initial weights which are adjusted using backpropagation for certain iterations. The proposed system uses a breast cancer dataset and achieved an accuracy of 97.94%. No feature selection is used and standard ABC used may produce early convergence due to lack of diversification. A hybrid approach using Dragon Fly Optimization with ABC is introduced for optimizing weights and bias of ANN [43]. The proposed algorithm integrates levy flight operations of Dragon Fly Optimization with the local search of ABC. Scout bee phase is modified using an improved search equation. The proposed achieves balanced exploration and exploitation and avoids the local minima with good convergence speed. The proposed is tested using a breast cancer dataset and gained an accuracy of 96%. The system does not concentrate on feature selection and optimal selection of hidden nodes which may affect convergence.

A system is proposed for medical diagnosis using ABC and decision tree-based gradient boosting [44]. WBCD and Haberman's survival dataset is used. ABC selects feature sets from the two datasets. A regression tree is used as a classifier and Gradient descent is used to find the direction of the gradient. The proposed has not been evaluated for complexity. Two-hybrid approaches using ABC and PSO is proposed [45]. The employee bee phase is integrated with PSO finding new velocity with updating position. Another approach where onlooker bee, scout bee is enhanced using mutation of genetic algorithm. The UCI repository is used where WBCD is utilized for testing where two algorithms achieved 99.14%. It hasn't focussed on tuning parameters of the classifier.

An efficient ABC by integrating ABC with Broyden–Fletcher–Goldfarb–Shannon (BFGS) using limited memory is proposed [46]. This proposed is used to tune parameters of DNN with cascaded auto encoder forming softmax classification. The candidate solution is produced using a neighbourhood process executed using BFGS with limited memory. The proposed algorithm is tested under 15 datasets. The classification accuracy of WBCD is 73.03%. The step size of the neighbourhood search is not made adaptive and does not focus on feature selection. An optimized approach utilizing standard ABC is proposed for classifying DNA microarrays [47]. ABC is used for the feature selection process. The approach is simple but hasn't focussed on tuning parameters of the classifier. ABC coupled feature selection is proposed for medical diagnosis [48]. It identifies feature subsets using employee and onlooker phases from total feature sets. The proposed system selected 2 features from a set of 9 in WBCD and achieved an accuracy of 96.69%. The proposed system is simple but feature selection may strike at local optimum due to imbalanced exploration and exploitation of ABC.

A system using standard ABC evaluated using the UCI repository is proposed [49]. The proposed achieves good permutation when compared to the standard employee bee phase. The proposed has an accuracy of 75.87% and it focussed on improving the exploitation capability of artificial bee colony but did not focus on parameter optimization of the classifier used. A wrapper approach is proposed for feature selection [50]. Two sets of features are used where the first set contains similar features and the second set contains dissimilar features. A three-layered architecture chooses a feature from each set and undergoes training. This is done till certain iterations until the error increases. During the evaluation, WBCD had a mean accuracy of 98.76% with a number of connections of 14.36. The computational time was 17 minutes 24 seconds for 30 runs. The proposed has not focused on parameter tuning of ANN.

A hybrid algorithm using branch & bound and ABC is proposed for feature selection [51]. The algorithm applies branch and bound on total features. Then, the proposed system applies ABC on the total subset to gather a second set of features. A union operation was then performed on two sets. The system is checked using various medical datasets based on relevance feedback and precision/recall values. The proposed is tested for classification. ABC based feature selection is proposed [52]. Optimal features for classification are selected by ABC where the J48 decision tree is used as a classification model. This predictive model is utilized for discriminating features of datasets from the University of California in which the accuracy of the proposed model using WBCD is 96.99% where 4 were selected from nine features. Totally 699 instances are utilized for checking performance using ' $k$ ' fold cross-validation method.

The proposed approach in [78] which is presented for breast cancer diagnosis is based on integrating the employee bee phase with simulated annealing improved local search process. First, the construction of feasible and random solutions was done and the estimation of the probabilities of the best solution using roulette selection was implemented. AIS techniques such as negative selection and population compression were incorporated in the ABC framework. The proposed algorithm was evaluated using the WBCD dataset. The approach proposed in [79] for breast cancer diagnosis is based on integrating the employee bee phase of the ABC algorithm with the bubble net attacking phase of the whale algorithm. The employee bees in the exploitation phase of the proposed algorithm in [79] follow the local search process of the humpback whales and find better food sources. The exploration process of finding the optimal regions is accomplished using various mutation strategies. The proposed algorithm was evaluated using various breast cancer datasets available in the UCI repository and a few breast cancer databases. Compared to the approaches presented in [78 & 79], the hybrid approaches in this paper proposed for breast cancer diagnosis integrates standard ABC algorithm with modified clonal selection algorithm of the Artificial Immune System and the chemotaxis phase of the bacterial foraging optimization with a spiral model-based exploitative phase of the ABC. Both the algorithms are evaluated using the WBCD dataset.

### **3. Materials and Proposed Methodology**

The proposed algorithm is implemented using a wrapper-based approach which does not use any statistical processing methods. The wrapper method is used to simultaneously select the feature subsets and other parameters of an ANN for effective breast cancer diagnosis. The proposed Hybrid ABC that hybridizes the spiral model-based exploitation phase of standard ABC with BFO is novel when compared to previous similar works. The Wisconsin breast cancer dataset is taken as the input and divided into three subsets called training, testing and validation set. The feature subsets that are optimally generated by our Hybrid ABC are selected from training, testing and validation subsets where the other features are eliminated. These features are used to train ANN. The initial weights and hidden node size are optimally generated by Hybrid ABC. The ANN error using the validation set is checked continuously for six iterations and if the ANN error increases continuously then the training process is stopped. The fitness of the resulting ANN is evaluated using Equations (7 and 8). High fitness ANN is selected and tested by testing set and accuracy and complexity of final network are calculated using Equation (6)

$$Connections = R * S + S * T + S + T \quad (6)$$

'R' is the number of input features, 'S' is hidden node size and 'T' is a number of output nodes. The network with the least connections confirms that the ANN is less complex. The fitness is calculated using Equation (6). Higher ANN error resembles the ANN with low fitness.

$$ANN\ Error = \left( \frac{I_{max} - I_{min}}{m * n} \right) \sum_{i=1}^m \sum_{j=1}^n (I_j^i - K_j^i)^2 \quad (7)$$

$$Fitness = \frac{1}{ANN\ Error} \quad (8)$$

Where 'm' is output nodes size, 'n' is the validation sample size,  $I_{max}$ ,  $I_{min}$  is the maximum and minimum of actual output.  $K_j^i$  is target output and  $I_j^i$  is the actual output. The initial solution representation has X bits that give random initial weights where  $2^X$  different combinations can be explored. Y bits gives the size of the hidden node where  $2^Y$  sizes could be explored. Z bits represent feature bits which give the total feature size. If the feature is selected, 'Z' shows one otherwise it shows zero.

### 3.1 Wisconsin Breast Cancer dataset (WBCD)

The WBCD is a publicly available benchmark database that contains breast cancer datasets. It is a database from the UCI repository [77] for breast cancer diagnosis. It contains data of breast lumps taken from breast tissues with the help of fine needles. It contains data of sample size 699 and this contains 458 benign instances. Each sample has features listed in Table 1. Each feature takes a value between one to ten where one shows a benign lump and ten shows a malignant lump.

**Table 1: Features of WBCD**

Label	Attributes
Feature1	Clump thickness
Feature 2	Uniformity of cell size
Feature 3	Uniformity of cell shape
Feature 4	Marginal adhesion
Feature 5	Single epithelial cell size
Feature 6	Bare nuclei
Feature 7	Bland chromatin
Feature 8	Normal nucleoli
Feature 9	Mitoses

### 3.2 Detailed Description of the proposed HABC algorithm

The proposed HABC is framed in such a way that it performs proper exploration using the techniques of AIS, in addition to its random initialisation process. HABC uses modified clonal selection, positive selection, and population compression techniques. The initial positions of the food source of the employee bee phase are the solutions obtained from the first stage. The second stage uses a tabu-based employee bee phase. HABC algorithm is shown in Figure (3). The proposed HABC algorithm is as follows:

- (i) Modified antibody detection is proposed to improve the search speed during initialization.
- (ii) Positive selection and cell suppression techniques of AIS are used in the initialization phase to ensure proper diversification in the exploration stage.
- (iii) Tabu-based employee bee phase is proposed to quit looping problems.
- (iv) During the employee bee phase, few solutions converge at the same local optimum. Such solutions share some similarities and are eliminated by negative selection.
- (v) The scout bee phase with population compression is used to maintain the food source size.

### 3.2.1 Initialization and fitness calculation

Each  $i^{th}$  antibody  $X_i$  represents a solution (antibody) for the optimization problem and it is a vector  $X^j = (X^1, X^2, X^3 \dots \dots X^{dim})$  in which the HABC algorithm generates ‘ $i$ ’ number of solutions  $X_i$  ( $i=1, 2, \dots, n$ ) and ‘ $i$ ’ represents the antibody size. Affinity evaluation is performed and the antibodies are further ranked in the decreasing order of affinity values. Antibodies with the least amount of network error for ANN are considered to be better antibodies. The lower and upper bound of parameter ‘ $j$ ’ is represented as  $X_{min}^j$  and  $X_{max}^j$ , respectively. The initialization step is explained by Algorithm (3).

#### Algorithm 3: HABC Initialization

##### Initial population

For  $i = 1$  to  $p$  do

    For  $j = 1$  to  $D$  do

$$X_i^j = X_{min}^j + rand(0,1)(X_{max}^j - X_{min}^j)$$

    End for

End for

##### Evaluation and ranking

For  $i = 1$  to  $p$  do

    fitness ( $X_i$ )



End for

### 3.2.2 Modified antibody detection phase

The modified antibody detection phase detects multiple antibodies based on the quality of each solution. The helper antibodies that are separated from the total antibody population are classified into three subsets based on the affinity distance between each antibody and the best antibody in the population, based on the threshold values such as  $ctrl1$ ,  $ctrl2$ , and  $ctrl3$ . Helper antibodies with better affinity values are grouped as the  $X_{h1}$  antibodies whose affinity values are close to the affinity value of the best antibody in the population. These antibodies are cloned, and the amount of cloning is directly proportional to antibody affinity. The  $X_{h1}$  antibodies are better; hence, a large number of new copies of the  $X_{h1}$  parent antibodies are generated. Equation (9) mutates using Gaussian mutation in order to facilitate a local search process around the better antibodies,  $X_{h1}$  antibodies and their clones.

$$X'_{h1} = X_{h1} + \emptyset_g \cdot G(0,1) \quad (9)$$

$X'_{h1}$ , the mutated antibody, is produced after Gaussian mutation;  $\emptyset_g$  indicates the strength of the added Gaussian noise;  $X_{h1}$  is the immature antibody; and  $G(0,1)$  is the Gaussian distribution random number. Intermediate antibodies with intermediate affinity values are grouped as the  $X_{h2}$  antibodies whose affinity values are at an intermediate distance from the affinity value of the best antibody in the population. Hence,  $X_{h2}$  antibodies are worse than  $X_{h1}$  and the number of clones of the parent  $X_{h2}$  that are generated are less as compared to  $X_{h1}$ . In order to facilitate a uniform search process around the intermediate antibodies, Equation (10) is used to uniformly mutate the  $X_{h2}$  antibodies.

$$X'_{h2} = X_{h2} \quad (10)$$

$X'_{h2}$  is the mutated antibody,  $X_{h2}$  is the immature antibody and a random cell is selected and replaced by a uniform random value chosen between the user-defined upper bound ( $Ub$ ) and lower bound ( $Lb$ ) values of that cell. The worst antibodies with lower affinity values are grouped as the  $X_{h3}$  antibodies whose affinity values are a larger distance from the affinity value of the best antibody in the population. The  $X_{h3}$  antibodies are worse than  $X_{h1}$  and  $X_{h2}$ ; hence, the number of clones of the parent  $X_{h3}$  that are generated are less as compared to  $X_{h1}$  and  $X_{h2}$ . Hence, in order to facilitate a global search process around the worst antibodies,  $X_{h3}$  antibodies are mutated using Levy Mutations (LM) according to Equation (11).

$$X'_{h3} = X_{h3} + \phi_c \cdot levy(X_{h3}) \quad (11)$$

$X'_{h3}$  is the mutated antibody that is produced after Cauchy mutation and  $\phi_c$  is the mutation rate. The amount of mutation depends on the antibody affinity value. The antibodies with higher affinity values are subjected to less mutation, and the antibodies with the worst affinity values are subjected to more mutation. Then, the mutated antibodies enter into the pool of antibodies along with the original antibodies. Next, the affinity of these antibodies is evaluated, and they are ranked in a descending pattern. The worst antibodies are eradicated from the total population, mutated antibodies are removed, and only the better antibodies are selected, which maintains the population size of the helper antibody. Algorithm (4) shows the Modified antibody detection phase.

#### **Algorithm 4: Modified antibody detection phase**

Replace the worst ' $n$ ' antibodies by randomly generated antibodies  $X_s$

For each helper antibody ( $X_h$ )

Select  $X_{h1}, X_{h2}$ , and  $X_{h3}$  antibodies from parent antibodies and clone them.

$X_{h1}$  if  $\{|\text{Affinity}(X_h) - \text{Affinity}(X_{\text{best}})|\} \leq \text{ctrl1}$

$X_{h2}$  if  $\{|\text{Affinity}(X_h) - \text{Affinity}(X_{\text{best}})|\} \leq \text{ctrl2}$  and if  $\{|\text{Affinity}(X_h) - \text{Affinity}(X_{\text{best}})|\} \geq \text{ctrl1}$

$X_{h3}$  if  $\{|\text{Affinity}(X_h) - \text{Affinity}(X_{\text{best}})|\} \geq \text{ctrl2}$

$C = 0$ ,  $C_l$  the number of clones derived for parent antibody

While ( $C < C_l$ )

If  $X_{h1}$  clones

Mutate the clones using equation (7)

$C = C + 1$

Else If  $X_{h2}$  clones

Mutate the clones using equation (8)

$C = C + 1$

Else

Maturate the clones using equation (9)

$C = C + 1$

Set of mutated antibodies are generated  $\overline{X_h}$

Calculate the fitness ( $X_h, \overline{X_h}$ )

Rank according to fitness and best antibodies based on fitness are selected

End for

Generate  $X_{new} = X_{new} \cup X_s$

### 3.2.3 Tabu-based employee bee phase

The positions of the initial food source in the employee bee phase are the solutions obtained from the modified antibody detection phase. Worker bees are assigned to a food source in which they search for neighbouring food sources around current food sources according to Equation (12):

$$Y_m^n = X_m^n + random[-1,1](X_m^n - X_k^n) \quad (12)$$

where  $X_k$  is a randomly selected food source in which  $k \in \{1, 2, \dots, F\}$  and  $F$  represents the number of food sources. ' $n$ ' is the random integer in which  $n = \{1, 2, \dots, d\}$ , and ' $m$ ' should not be equal to ' $k$ ' for proper exploitation.  $random[-1,1]$  is a random integer between -1 to 1 that represents the step size of the exploitation process. The employee bee phase accepts only the better food source and makes only uphill movements; downhill movements are neglected completely. Because of this, the local search process tends to cyclically revisit the same solutions repeatedly making an algorithm to return suboptimal solutions.

To avoid these problems, the tabu-based selection mechanism accepts even the worst solutions. This phase uses a short-term memory called Tabu List (TL) in which the recent ' $n$ ' moves are stored, dynamically maintained, updated by the worker bees, and used in the exploitative selection. The ' $n$ ' represents the tabu tenure. According to HABC,  $Y_i^j$  is accepted even if it is worse than  $X_i^j$  if  $Y_i^j$  not in TL. If  $Y_i^j$  is in TL, then it should be checked for an aspiration condition. According to the aspiration condition, if  $Y_i^j$  is in TL but  $Y_i^j$  is better than the best solution that was found until the current iteration, then  $Y_i^j$  is accepted, otherwise, it is rejected.  $Y_i^j$  is accepted directly if it is better than  $X_i^j$  and is not in the TL. Thus, Tabu based employee bee phase avoids the problems of looping by ensuring that every solution is visited only once by the worker bees. Algorithm(5) shows tabu based employee bee phase.

#### Algorithm 5: Tabu based employee bee phase

Assign employee bees to  $X_{new}$

$$X_i^j = \{X_1^1, X_2^2, X_3^3 \dots \dots \dots X_n^D\}$$

```

For  $i = 1$  to  $p$  do
  For  $j = 1$  to  $D$  do
    For  $X = 1$  to  $n$ 
       $Y_i^j = X_i^j + \text{random}[-1,1](X_i^j - X_k^j)$ 
      If  $\text{fitness}(Y_i^j) \geq \text{fitness}(X_i^j)$ 
        If  $Y_i^j$  is not in TL
          Accept  $Y_i^j$  and Reject  $X_i^j$ 
        Else If
           $Y_i^j$  holds aspiration condition
            Accept  $Y_i^j$  reject  $X_i^j$ 
          Else
            Accept  $X_i^j$  and Reject  $Y_i^j$ 
            Update TL
        End for
      End for
    End for
  End for
End for

```

### 3.2.4 Negative Selection Phase

To accelerate good convergence in the search process, similar food sources must be eliminated in order to maintain proper diversification. Euclidian distance is measured between food sources that is represented as  $D(X_i^j, X_k^j)$ . If  $D(X_i^j, X_k^j)$  less than the threshold  $T_h$  then the food source with less fitness is eliminated, either  $X_i^j$  or  $X_k^j$ , and replaced by a newly created random food source; otherwise, the food source is retained as either  $X_i^j$  and  $X_k^j$ . The negative selection phase is explained by Algorithm (6).

#### Algorithm 6: Negative selection phase

```

For  $i = 1$  to  $p$  do
  For  $j = 1$  to  $D$  do
    For  $X = 1$  to  $n$ 

```

$$\text{Calculate } D(X_i^j, X_k^j) = \sqrt{\sum_{j=1}^D (X_i^j - X_k^j)^2}$$

$$\text{If } D(X_i^j, X_k^j) < T_h$$

If fitness ( $X_k^j$ ) > fitness ( $X_i^j$ )

Eliminate  $X_i^j$

Replace  $X_i^j$  with new randomly generated food source

Else

Retain  $X_i^j$  and Eliminate  $X_k^j$

Replace  $X_k^j$  with new randomly generated food source

Else

Retain  $X_i^j$  and Retain  $X_k^j$

End for

End for

End for

### 3.2.5 Onlooker Bee Phase

The fitness value is determined by Equation (13). The description of the onlooker bee process is shown in Algorithm (7).

$$E_i = \frac{fitness(X_i)}{\sum_{i=1}^n fitness(X_i)} \quad (13)$$

#### Algorithm 7: Onlooker Bee Phase

For i = 1 to n do

Calculate the probability using Equation (11)

If  $rand(0,1) < E_i$

Select  $X_i$

Else

Reject  $X_i$

End for

Memorize the best food source  $X_{best}$

### 3.2.6 Scout Bee Phase with population compression

New food sources are discovered by scout bees. If the fitness of the food source is not improved during certain iterations and if the counter associated with the food source crosses the counter value  $\text{count}_{\max}$ , then those food sources are replaced by scout bees. 's' new random food sources are generated and the lower quality food sources are eliminated until the population size is maintained. The detailed description is shown in Algorithm (8).

**Algorithm 8: Scout Bee Phase with population compression**

While (  $X_i$  does not improve)

$\text{count} = \text{count} + 1$

If  $\text{count} > \text{count}_{\max}$

Replace  $X_i$  with a newly generated food source

Else

$\text{count} = 0$

Generate 's' number of new random food sources

Compress low fitness food sources until population size

Iteration = Iteration +1

**3.3 Proposed Hybrid ABC Optimization**

Proposed Hybrid ABC is a hybrid version of standard ABC that utilizes the optimal locations identified by the chemotaxis phase of BFO as the initial food source. The standard ABC is weak in exploration due to localized initial food sources and poor random search. Hence, the proposed hybrid ABC in its explorative phase adapts the chemotaxis phase with adaptive step sizes so that it can explore the entire search region with larger steps in proper tumble direction discovering highly nutrient and promising areas. Exploitation is done such that optimum solutions found in the exploration phase are used as the initial food sources in a spiral dynamic model [54]. In this phase, employee bees are assigned to the food sources where step sizes are varied during the search process so that the bees search all around the problem space efficiently and find out better food sources. To make the step sizes adaptive, they are calculated based on the fitness using the dynamic spiral radius and angular displacement of the food source which is inversely proportional to the fitness of the food source. At initial iterations in the spiral, where larger steps are taken the bees performs global search moving towards global optimum and as the iteration proceeds bees focus on the local search process taking smaller steps converging towards the global optimum. The centre of the spiral is updated using the best food source in that particular iteration. Using adaptive step sizes calculated based on the fitness of food source, employee bees

on spiral path search the search space covering remote locations that are not visited in the exploration phase. Thus, proposed Hybrid ABC using adaptive step size can also overcome problems of oscillations that happen with larger step sizes. The proposed Hybrid ABC is given in Figure (1). The methodology of the proposed Hybrid ABC is summarized as follows:

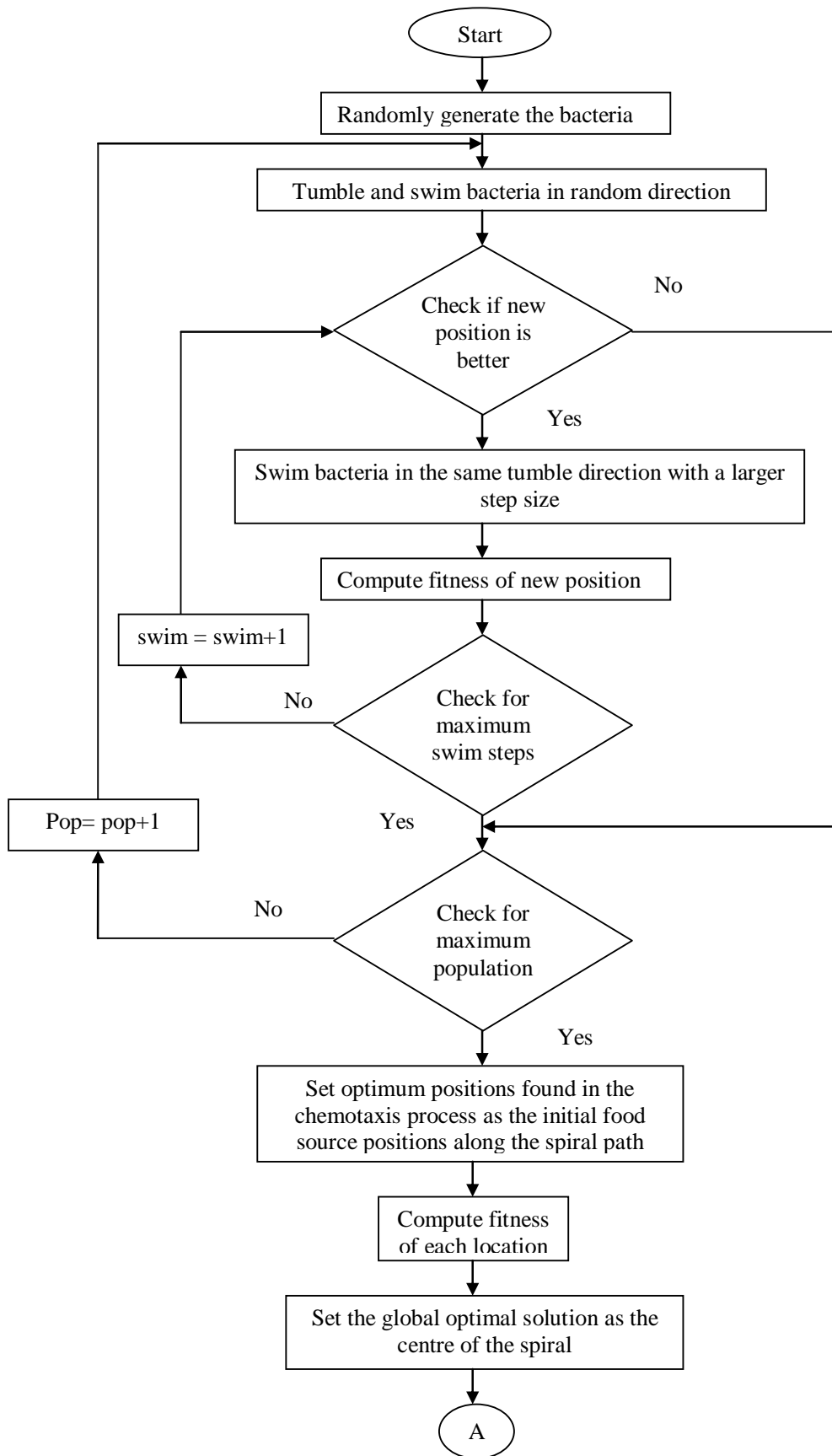
- (i) The chemotaxis process of BFO with an adaptive step size is performed in the initialisation phase to ensure proper diversification in the initial stages.
- (ii) Employee bee phase is incorporated using spiral dynamics model in which search is guided by global optimum locations to ensure efficient exploitation.
- (iii) Step size is adaptively varied in which at initial stages where the fitness of food sources is less, rotation angles and radius on the spiral are kept high to enhance a global search process and as the fitness improves, they are reduced gradually to enhance the local search process.

### 3.3.1 Initialization based on chemotaxis process

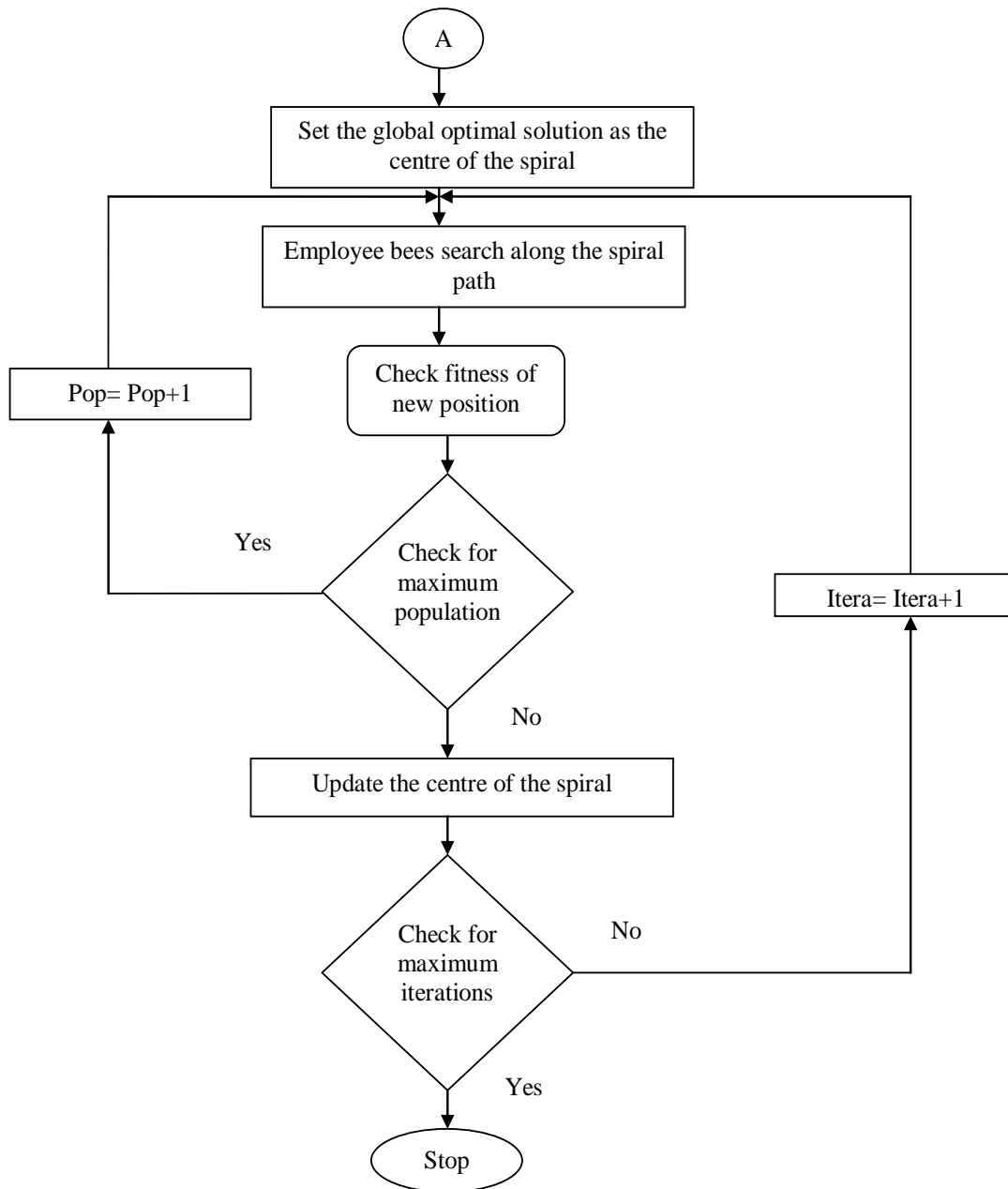
A bacterium represents a solution of optimization considered. A pre-defined bacterium set ' $x$ ' is generated randomly using a uniform distribution. Each bacterium is represented by  $Z(j) = \theta^j(l, m, n)$  such that  $\forall_j = 1, 2, \dots, x$ .  $\theta^i$  represents  $i^{\text{th}}$  bacterium. ' $l$ ' is chemotactic loop number, ' $m$ ' gives reproduction loop number and ' $n$ ' gives elimination-dispersal loop number. The chemotaxis is executed till  $C_{cl}$  times. The bacterium tumbles in a random direction and swims in the same direction of tumble using Equation (14).

$$Z(j + 1) = \theta^j(l + 1, m, n) = Z(j) + \text{step size} * \text{ran direction} \quad (14)$$

The new position of  $j^{\text{th}}$  bacterium is given by  $Z(j + 1) = \theta^j(l + 1, m, n)$ . Then if  $Z(j + 1)$  is better than  $Z(j)$  then bacteria swim in the direction as tumble with a larger step size. This continues till maximum swim steps  $N_{max}$  are attained. The algorithm steps of initialization are given in Algorithm (9).







**Figure 1: Flowchart of the proposed Hybrid ABC Optimization**

**Algorithm 9: Generation of Initial population**

Randomly place bacteria  $Z(j + 1)$

**FOR**  $l = 1$  to  $C_{cl}$

**FOR**  $x = 1$  to  $x_{max}$

Bacteria tumbles randomly

$$Z(j + 1) = Z(j) + \textit{step size} * \textit{ran direction}$$

**IF** ( $\textit{swimstep} < N_{max}$ )

**IF**  $fitness(Z(j + 1)) > fitness(Z(j))$

Swim with larger step size in the same direction as tumble

Swim  $Z(j + 1) = Z(j) + step\ size * ran\ direction$

Compute  $fitness(Z(j + 1))$

$swimstep = swimstep + 1$

**ELSE**

$l = l + 1$

$x = x + 1$

**ELSE**

$l = l + 1$

**END FOR**

**END FOR**

The set of solutions discovered in the chemotaxis process is used as initial food sources of the spiral based employee bee phase.

### 3.3.2 Spiral Dynamics Model-based Employee Bee Phase

The spiral dynamics model is inspired by natural spiral models like shells, fingerprints, etc. It contains an efficient spiral path that enhances intensification and diversification. Initial phases allow search agents to intensify using small step sizes and the final phase allows to diversify using large step sizes while searching. Using the spiral dynamic model, the step size is varied dynamically the search process is guided by the global optimum location identified which is updated at each iteration. Thus, it improves the convergence of search operation with less computational time attaining more accuracy. This model uses two parameters called the spiral radius  $I_d$  and spiral angular displacement  $\theta_d$  for an efficient search process. The optimal selection of  $I_d$  and  $\theta_d$  plays a vital role in finding out the global optimum locations since fixed values  $I_d$  and  $\theta_d$  may lead to an imbalance of exploration and exploitation allowing suboptimal locations. Smaller values may make the search process converge fast but make the algorithm stuck in its local optimum whereas larger values introduce oscillations in the search process.

Based on the above discussion, the proposed Hybrid ABC utilizes the benefits of the spiral dynamic model by using it as its search path in the employee bee phase to search for better food sources. The optimum locations discovered during exploration is set as initial food source

positions of improved employee bee phase-based on the spiral dynamics model. The best food source saved at every iteration is updated as the centre of the spiral trajectory path as indicated using Equation (15)

$$Z^* = Z_g(0) \quad (15)$$

$Z^*$  represents the centre of the spiral path which will be equal to  $Z_g(0)$  that indicates the best fitness food source found during any particular iteration. The objective of our work considered is to minimize ANN network error with maximum fitness value according to Equation (8) and hence we take  $maximize(fitness(Z_x(j)))$   $x = 1, 2, \dots \dots x_{max}$  where 'x' represents a number of search points that is equal to the bacterium set. All employee bees are assigned to a food source on the spiral path and the bees searches over the spiral path to find a neighbouring food source that is better than the old food source given by Equation (16).

$$Z(j + 1) = S(I_d, \theta_d)Z(j) - (S(I_d, \theta_d) - M_k)Z^* \quad (16)$$

Where  $Z(j + 1)$  is the new food source,  $Z(k)$  is the old food source,  $S(I_d, \theta_d)$  is adaptive spiral radius and angular displacement.  $M_k$  is the identity matrix with  $k*k$  dimensions. Employee bee search process is made more efficient using  $(I_d, \theta_d)$  which is adaptively calculated using the quality of the food source. This dynamic  $(I_d, \theta_d)$  which is responsible for smooth convergence of the search process is represented by Equations (17&18)

$$I_d = \frac{I_{max} - I_{min}}{L_s |fitness(Z(j))|} \quad (17)$$

$$\theta_d = \frac{\theta_{max} - \theta_{min}}{L_s |fitness(Z(j))|} \quad (18)$$

Where  $I_d$  and  $\theta_d$  represents adaptive radius and adaptive angular displacement that is inversely proportional to the fitness of the food source.  $L_s$  is the positive constant.  $fitness(Z(j))$  is the fitness of the food source considered.  $I_{max}$  and  $\theta_{max}$  is considered as maximum radius and maximum angular displacement for the spiral path.  $I_{min}$  and  $\theta_{min}$  is minimum radius and minimum angular displacement for the spiral path. The minimum and maximum radius are set in between 0 to 1. The minimum and maximum angular displacement are set between 0 and  $2\pi$ . The adaptive radius and angular displacement are set minimum when the fitness of the food source is high to enhance the exploitative process around better food sources. The adaptive radius and angular displacement are set high when the fitness of the food source is low to enhance the exploration of

new promising global optimal regions. Using  $I_d$  and  $\theta_d$ , employee bees search any location placed remotely that was not explored during the exploration process with the search process guided by the best optimal food source and hence the centre of the spiral path used by employee bees has the best food source instead of a random food source in the spiral based employee bee phase. The detailed algorithmic steps of this phase are given in Algorithm (10)

**Algorithm 10: Spiral Dynamics Model-based Employee Bee Phase**

Set optimum locations discovered in chemotaxis as initial food sources of ABC on spiral

$$Z_i(0) = Z_{best}$$

Compute  $fitness(Z(j))$

Set  $Z^* = Z_g(0)$  center of spiral

**FOR**  $j = 1$  to  $m$

**FOR**  $i$  to  $i_{max}$

The bee moves spirally to find a neighbouring food source position

$$Z_i(j + 1) = S_i(I_d, \theta_d)F_i(k) - (S_i(I_d, \theta_d) - M_k)Z^*$$

**IF**  $fitness(Z(j + 1)) < fitness(Z(j))$

Reject  $Z(j + 1)$

**ELSE**

Reject  $Z(j)$

**END IF**

$$i = i + 1$$

**END FOR**

Set  $Z^* = Z_g(j + 1)$  as center of spiral

**END FOR k**

**END**

**3.3.3 Onlooker Bee Phase**

The information about selected food sources in the spiral dynamic model-based employee bee phase is shared with the onlooker bees. Further, the onlooker bees calculate the probability  $X_i$  for the food source, it receives based on quality given by its fitness value as represented in Equation (19). The onlooker bee phase is given in algorithm (11)

$$X_i = \frac{fitness(Z_i)}{\sum_{i=1}^m fitness(Z_i)} \quad (19)$$

**Algorithm 11: Onlooker Bee Phase**

For i= 1 to n do

Calculate probability using equation (19)

    If  $random(0,1) < X_i$

        Select  $Z_i$

    Else

        Reject  $Z_i$

End for

Memorize the best food source  $Z_{best}$

**3.3.4 Scout Bee Phase**

When a food source  $Z_i$  does not improve, the proposed algorithm uses the scout bee phase of the standard ABC process. A detailed description of the scout bee phase is given in the algorithm (12).

**Algorithm 12: Scout Bee Phase**

While (  $Z_i$  does not improve)

    limit = limit + 1

    If limit > limit<sub>max</sub>

        Replace  $Z_i$  with a newly generated food source

    Else

        limit = 0

Iteration= Iteration +1

**4 Performance Evaluation of proposed Hybrid ABC using WBCD dataset**

The proposed hybrid algorithms are implemented using the proposed wrapper architecture. The section focuses on the following factors: (i) The WBCD dataset is used by the proposed algorithms for evaluating the performance in terms of accuracy, complexity and computational time. (ii) The proposed algorithms are evaluated using an ANN classifier where three different back propagation variations called the RP, LM and GD are used for training purposes. The proposed hybrid algorithms derive the optimal

initial weights, hidden nodes and optimal feature subsets for an ANN classifier. (iii) The proposed algorithms find out the best backpropagation variant that performs with less network error, fewer connections and computational time.

#### 4.1 Parameter Settings and Experimental Setup

The algorithm was implemented using MATLAB 8.5 software. The neural network toolbox was used to implement the backpropagation training algorithm. Default training parameters were used for backpropagation training during implementation. The winner take all approach was used in output nodes for classification and the output nodes represented classes such as benign/malignant. HABC uses [77] that consists of WBCD datasets. The values between 0 to 1 were rescaled and missing attribute values were filled using mean values of non-missing attributes. The training dataset was formed using the first 349 samples, which was 50% of the total dataset, the second 175 samples that comprised 25% of the total dataset was used as the validation dataset, and the last 175 samples that comprised the remaining 25% was used as the testing dataset. The Input nodes size represents the size of the total feature set and hence it is considered as 9., The output node size represents the number of classifications and hence it is considered as 2 (benign, malignant). According to the initial solution representation, The size of L has been considered as 15 where  $2^L$  different bit combinations can be explored for initial weights. The value of M has been set as 2 where four different combinations of hidden node sizes can be explored. This can be set considering the size of the dataset and the number of input features. The activation functions at hidden and output nodes are hyperbolic tangent and pure linear, respectively. The parameter settings of the ANN classifier are shown in Table (2). The parameter settings of the HABC are given in Table (3). The values of thresholds ( $Limit_1$ ) and ( $Limit_1$ ) are set using a trial and error method. The count limit is set as 10 to enhance global exploration. The employee bees size and the onlooker bees size are set according to the problem space and they are set equal which are again equal to antibodies size (s). The mutation rates  $\phi_g$  and ( $\phi_c$ ) were set using the trial and error method to enhance the local and global search process.

**Table 2: Parameter settings of ANN classifier**

Parameter	Value
Training	Back propagation
Input nodes size	9

Output nodes size	2
Number of bits for initial weights	15
Number of bits for hidden node size	2
Number of bits for input features	9
Activation function for hidden node	Hyperbolic Tangent
Activation function for output node	Pure Linear
Training set samples	349 (50%)
Validation set samples	175 (25%)
Testing set samples	175 (25%)

**Table 3: Parameter settings of HABC**

Parameter	Value
Employee Bees size	30
Onlooker Bees size	30
Scout bees size	1
Total colony size	60
Count limit ( $Count_{max}$ )	10
Number of antibodies (s)	30
Helper antibody selection ( $A_h$ )	80%
Suppressor antibody selection ( $A_s$ )	20%
Control Threshold ( $Limit_1$ )	0.0001
Control Thershold ( $Limit_2$ )	0.5
Mutation rate ( $\emptyset_g$ )	0.4
Mutation rate ( $\emptyset_c$ )	0.6

The parameter settings of the Hybrid ABC are given in Table (4). The number of employee bees and onlooker bees are equally set. The bacterial step size, maximum and minimum spiral radius and angular displacement are chosen based on the trial and error method. The Hybrid ABC is implemented for various generation sizes as given in Table (4)

**Table 3: Parameter settings of Hybrid ABC**

Parameter	Value
Number of Employee Bees	50
Number of Onlooker Bees	50
Number of scout bees	1
Total colony size	100
Bacteria population size	50
Number of chemotaxis( $N_s$ )	30
Bacterial step size ( $C$ )	0.04
Maximum Spiral Radius( $R_{max}$ )	1
Minimum Spiral Radius( $R_{min}$ )	0.65
Maximum Spiral Angular displacement( $\theta_{max}$ )	1.8
Maximum Spiral Angular displacement( $\theta_{min}$ )	1.03
Generation sizes	10,20,30

#### 4.2 Proposed HABC in terms of accuracy, complexity and computational time

The proposed HABC was evaluated for the number of connections and classification accuracy using ten runs for different generation sizes, as shown in Table (5). HABC-RP achieved an average accuracy of 99.14% at a generation size of 20. The average number of connections of HABC-RP was 12.40, which was less as compared to HABC-LM and HABC-MGD. HABC produced a less complex ANN when trained using RP. The convergence of the validation error of the proposed HABC over different generations is shown in Figure (2).

**Table 5: Performance evaluation of the proposed HABC algorithm**

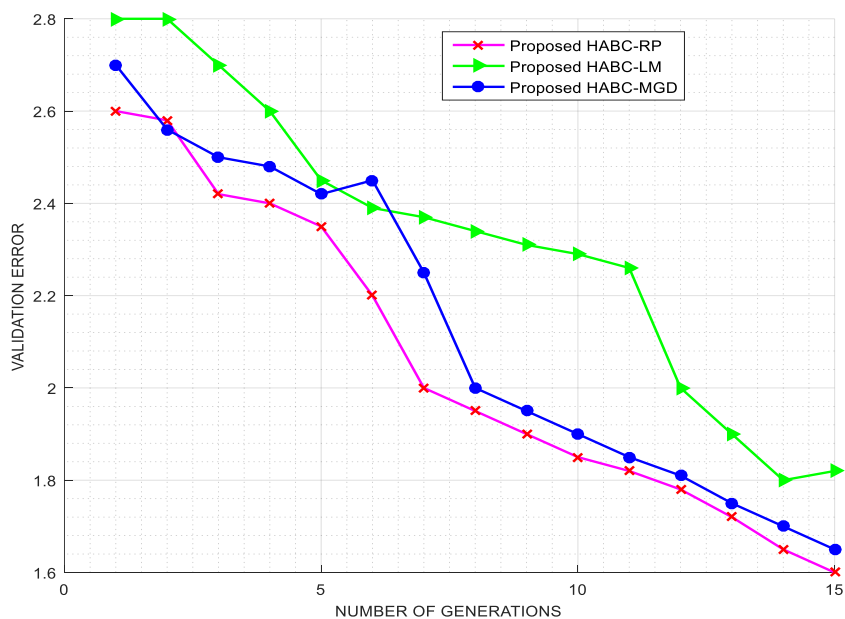
Max Generation Size	Proposed HABC-RP		Proposed HABC-LM		Proposed HABC-MGD	
	Classification accuracy (%)	Number of connections	Classification accuracy (%)	Number of connections	Classification accuracy (%)	Number of connections



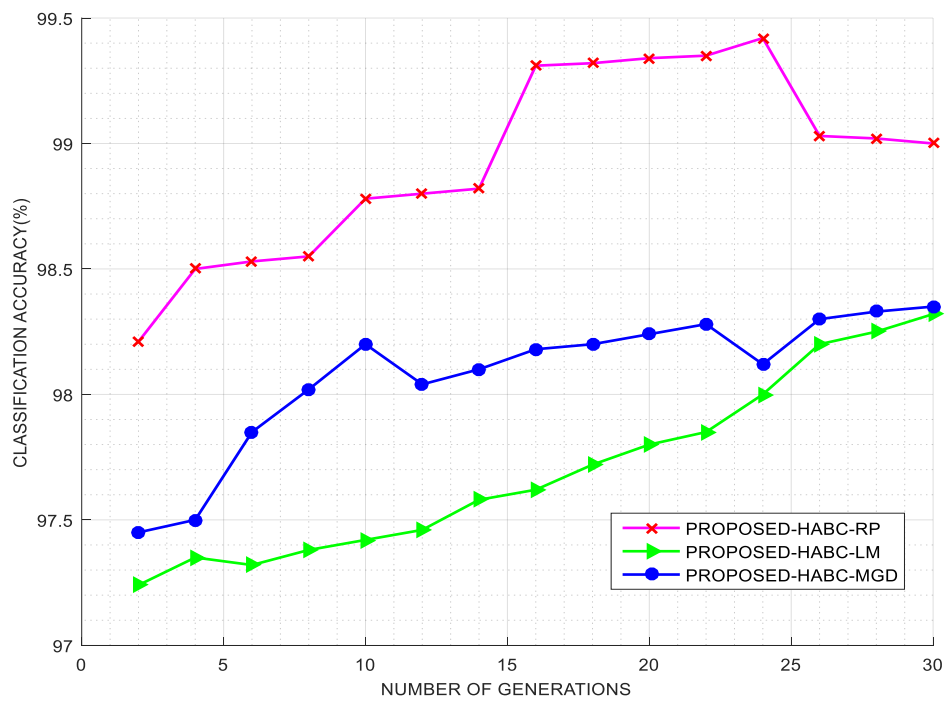
	Best	Mean	Best	Mean	Best	Mean	Best	Mean	Best	Mean	Best	Mean
10	98.42	97.81	12	13.56	97.12	96.53	12	17.42	97.23	96.51	12	15.65
20	99.42	99.14	10	12.40	97.82	96.71	13	16.34	98.12	97.28	12	15.82
30	99.02	98.58	11	13.23	98.35	98.05	14	15.46	98.71	98.34	13	15.21

Followed by HABC-RP, HABC-MGD was accurate, which was 98.34%, with mean connections of 15.21 at a generation size of 30. Next to HABC-MGD, HABC-LM had the highest average accuracy, which was 98.05% at a generation size of 30 with mean connections of 15.46. The classification accuracy for different generation sizes of the HABC algorithm is shown in Figure (3). The accuracy of HABC-RP was 1.11% higher than HABC-LM and 0.81% higher than HABC-MGD.

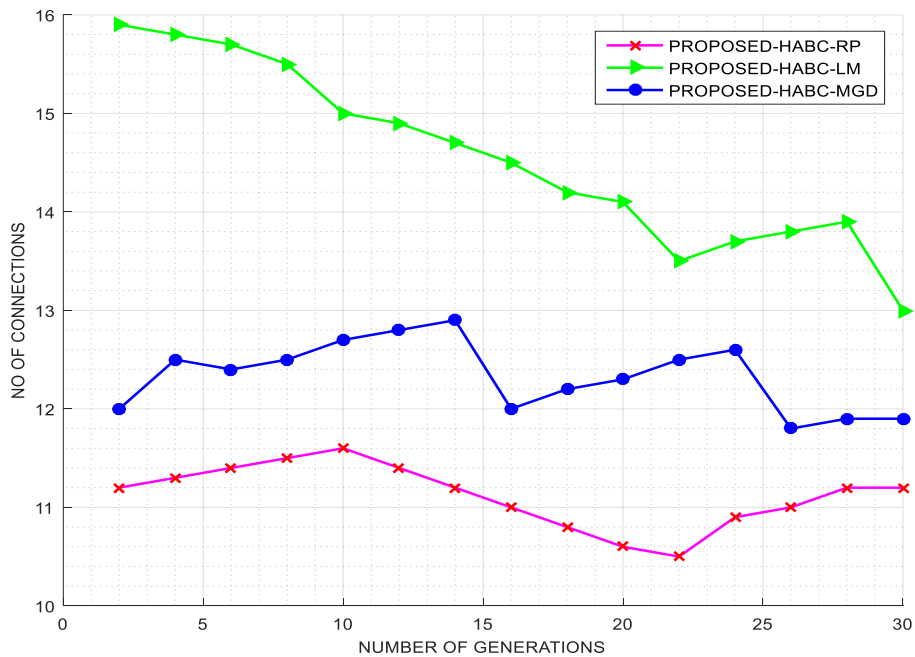
The confusion matrix with classification performance of HABC is shown in Table (6). Figure (4) shows the number of connections for different generation sizes of the HABC algorithm. HABC-RP utilized a lower number of connections, followed by HABC-MGD and HABC-LM. The average hidden node count of HABC-RP was 32% lower than HABS-AIS-LM and 27.78% lower than HABC-MGD. HABC-RP utilized 19.79% fewer connections than HABC-LM and 18.47% fewer connections than HABC-MGD.



**Figure 2: Convergence of validation error for HABC**



**Figure 3: Performance of HABC with respect to classification accuracy**



**Figure 4: Performance of HABC with respect to number of connections**

**Table 6: Confusion Matrix of HABC algorithm across ten runs**

Methods of comparison	Actual	Number of cases	Test outcome-Predicted	
			Malignant	Benign
HABC-RP	Malignant	650	646(TP)	11 (FN)
	Benign	1100	4 (FP)	1089(TN)
HABC-LM	Malignant	650	640(TP)	24FN)
	Benign	1100	10 (FP)	1076(TN)
HABC-MGD	Malignant	650	642(TP)	21 (FN)
	Benign	1100	10 (FP)	1079(TN)

The performance metrics of the proposed HABC for different back propagation variants are listed in Table (7). HABC-RP showed sensitivity and specificity, followed by HABC-MGD and HABC-LM. Figure (5) shows the hidden node count for HABC-RP, HABC-LM, and HABC-MGD for different generation sizes.

**Table 7: Performance of HABC algorithm based on different metrics**

Metrics	Proposed HABC-RP	Proposed HABC-LM	Proposed HABC-MGD
Sensitivity (%)	98.32	96.38	96.83
Specificity (%)	99.63	99.07	99.08
Accuracy (%)	99.14	96.38	97.32
Precision (%)	99.38	98.46	98.76
Negative predictive Value (NPV) (%)	99	97.81	98.09
F-measure	0.98	0.97	0.97

In Figure 5, the performance of HABC with respect to classification accuracy for individual generations until 30 has been shown. According to figure 5, the classification accuracy of the HABC-ANN is increasing as the generation size increases. The lowest performance of the proposed ANN is obtained at the initial generation sizes where the search space has not been explored and as the iteration proceeds towards the global optimal regions optimal parameters has

been chosen and hence the accuracy has been increased gradually. According to table 4, the lowest mean classification accuracy of 96.51% is attained for HABC-MGD for a generation size of 10.

Based on Table (8), the HABC-ANN with the feature selection process enhanced classification accuracy with less complexity as compared to without the feature selection process.

**Table 8: Influence of feature selection on HABC-RP**

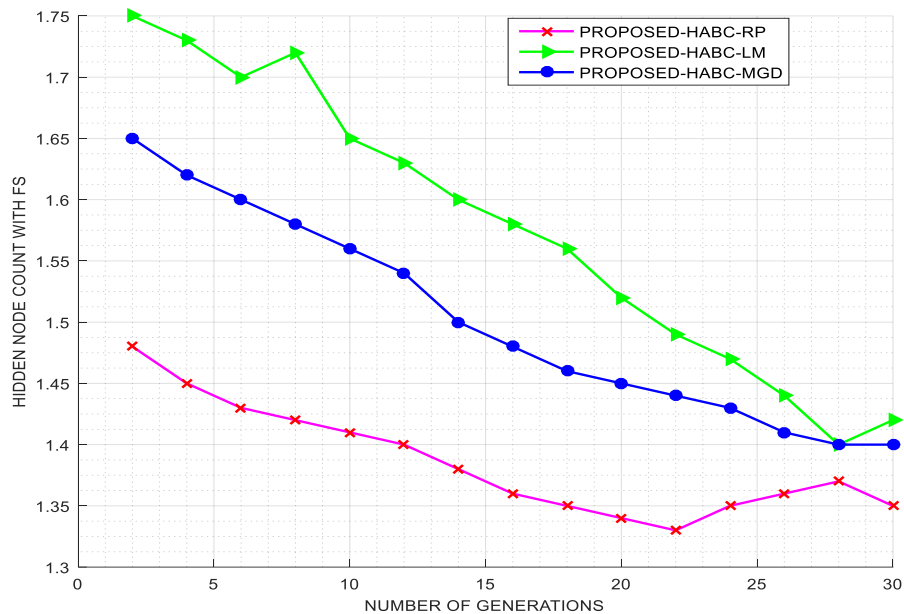
Max Generation Size	Feature Selection	Hidden Node Count (Avg)	No of Selected Features (Avg)	No of connections (Avg)	Accuracy (Avg) (%)
10	With Feature Selection	1.5	5.8	13.6	97.81
	Without Feature Selection	2.1	9	27.2	89.34
20	With Feature Selection	1.3	5	12.4	99.14
	Without Feature Selection	1.8	9	23.6	91.42
30	With Feature Selection	1.4	5.1	13.2	98.58
	Without Feature Selection	1.9	9	24.8	90.05

Table (9) depicts the best network of HABC-RP, with selected features shown as F<sub>2</sub>, F<sub>3</sub>, F<sub>5</sub>, F<sub>6</sub> respectively. The details of the feature set are shown in Table (1).

**Table 9: Confusion Matrix of HABC-RP for the best network with selected features**

Feature Selection(FS)	Cases(actual)		Cases(Predicted)		Selected Feature(SF)
			Benign	Malignant	
With Feature Selection	Benign	110	110	1	F <sub>2</sub> , F <sub>3</sub> , F <sub>5</sub> , F <sub>6</sub>
	Malignant	65	0	64	

Without Feature Selection	Benign	110	102	7	F <sub>1</sub> , F <sub>2</sub> , F <sub>3</sub> , F <sub>4</sub> , F <sub>5</sub> , F <sub>6</sub> , F <sub>7</sub> , F <sub>8</sub> , F <sub>9</sub>
	Malignant	65	8	58	



**Figure 5: Hidden node count for HABC**

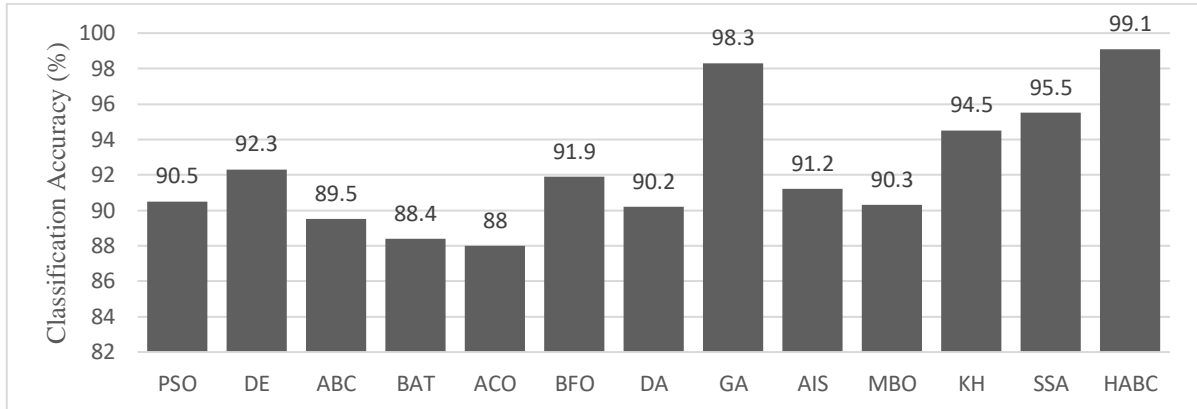
The performance of the HABC optimized ANN classifier with respect to the average computational time is shown in Table (9). The average computational time for ten independent runs was calculated for generation sizes of 10, 20, and 30. The computational time of HABC-LM was less than the computational time for HABC-RP and HABC-MGD.

**Table 9: computational time for HABC-RP**

Max Generation	Average CPU Time(s)		
	Proposed HABC-RP	Proposed HABC-LM	Proposed HABC-MGD
10	327.5	201.5	402.8
20	720.2	689.6	745.9
30	1005.9	936.5	1154.1

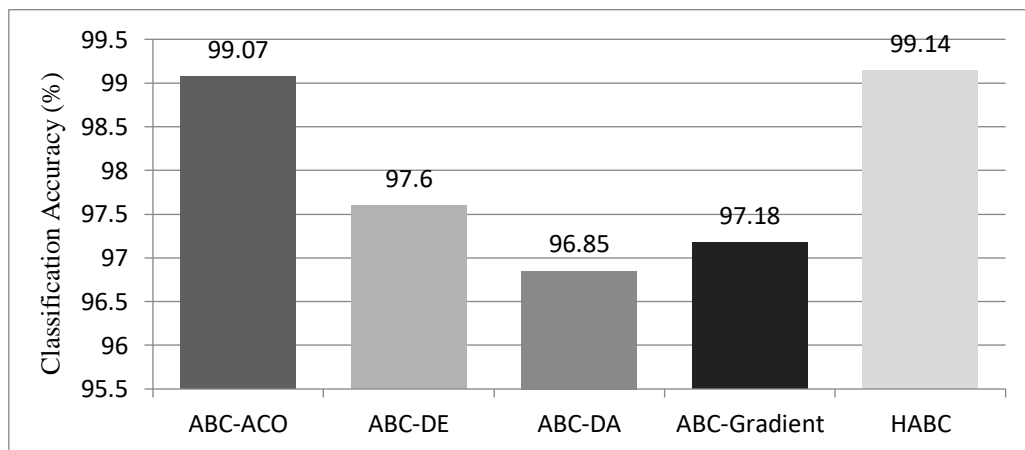
Figure (6) compares HABC with Particle Swarm Optimization (PSO) Differential Evolution (DE), Artificial Bee Colony Optimization (ABC), Bat Algorithm (BA), Ant Colony Optimization (ACO), Bacterial Foraging Optimization (BFO), Dragonfly Algorithm (DA),

Genetic Algorithm (GA), Artificial Immune System (AIS), Monarch Butterfly Optimization (MBO), Krill Herd (KH) Algorithm, and Salp Swarm Algorithm (SSA) using the WBCD dataset. HABC achieved the highest accuracy.



**Figure 6: Performance comparison of HABC and evolutionary methods**

Figure (7) shows a comparison between HABC and existing ABC-based hybrid algorithms using WBCD, including the ABC-ACO [45], ABC-DE [46], ABC-DA [43], and ABC-Gradient Decision Tree [66].



**Figure 7: Comparison of ABC-based hybrid algorithms**

#### 4.3 Performance of the proposed Hybrid ABC in terms of accuracy, complexity and computational time

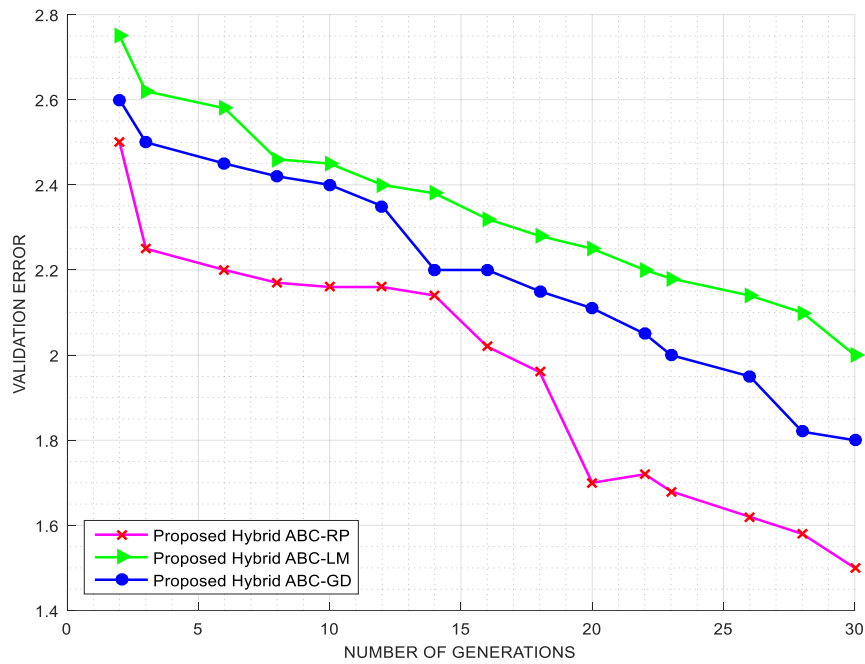
The proposed algorithm is implemented for different generation sizes 10, 20 and 30 where the average number of connections and classification accuracy for ten independent runs is

calculated. The different back propagation approaches give significant variations in the accuracy and complexity of an ANN even when evaluated using the same dataset. Hence, the proposed algorithm is evaluated using different back propagations such as resilient back propagation, Levenberg Marquardt and momentum-based gradient descent represented as Hybrid ABC-RP, Hybrid ABC-LM and Hybrid ABC-GD respectively as given in Table (10). The objective is to produce an ANN network with optimal input features, initial weights and hidden node size with the least network error, complexity and computational time. The convergence of validation error for the proposed HybridABC for different generations is shown in Figure (8). The figure shows that the proposed Hybrid ABC has produced the least network errors during validation when compared to the other two variants namely Hybrid ABC-LM and Hybrid ABC-GD.

**Table 10: Performance evaluation of the proposed Hybrid ABC algorithm**

Max Generati on Size	Proposed Hybrid ABC-RP				Proposed Hybrid ABC -LM				Proposed Hybrid ABC-GD			
	Classification Acc (%)		No of connections		Classification Acc (%)		No of connections		Classification Acc (%)		No of connections	
	Best	Mean	Best	Mean	Best	Mean	Best	Mean	Best	Mean	Best	Mean
10	98.35	98.25	8	11.65	98.21	97.22	10	13.56	98.95	98.32	9	12.65
20	100.00	99.54	7	10.25	98.33	98.58	9	12.34	98.60	98.65	7	11.03
30	99.11	98.92	7	11.26	99.03	98.97	8	11.21	99.12	99.08	7	11.03

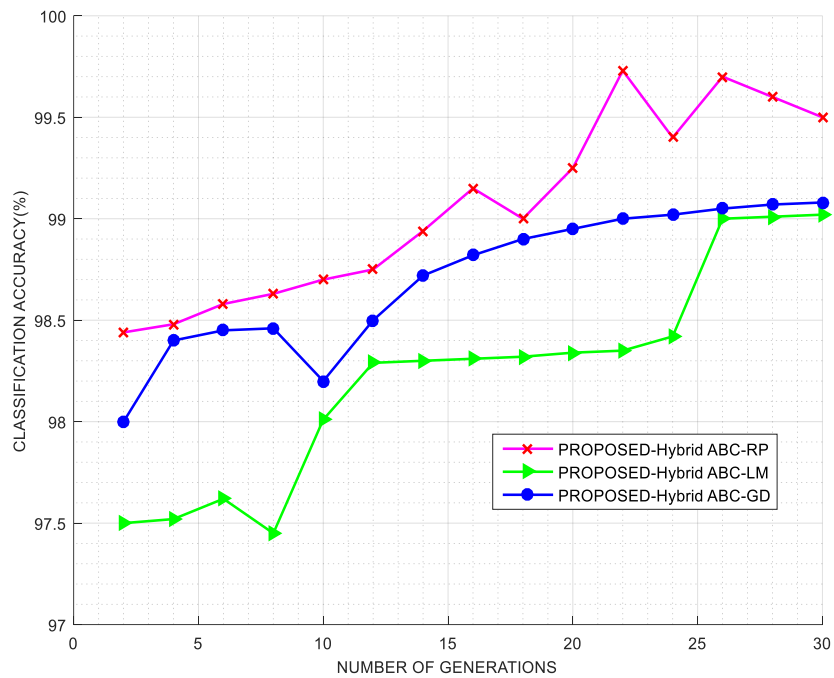
Hybrid ABC-RP achieved the highest mean ANN accuracy for ten runs which is 99.54% under the generation size of twenty. Followed by Hybrid ABC-RP, the Hybrid ABC-GD achieved the highest average accuracy which is 99.08% under generation size of thirty. Next to Hybrid ABC-GD, Hybrid ABC-LM achieved 98.97% under generation size of thirty. The accuracy of Hybrid ABC-RP is 0.46% more than Hybrid ABC-GD and 0.58% more than Hybrid ABC-LM.



**Figure 8: Convergence of Validation error for Hybrid ABC**

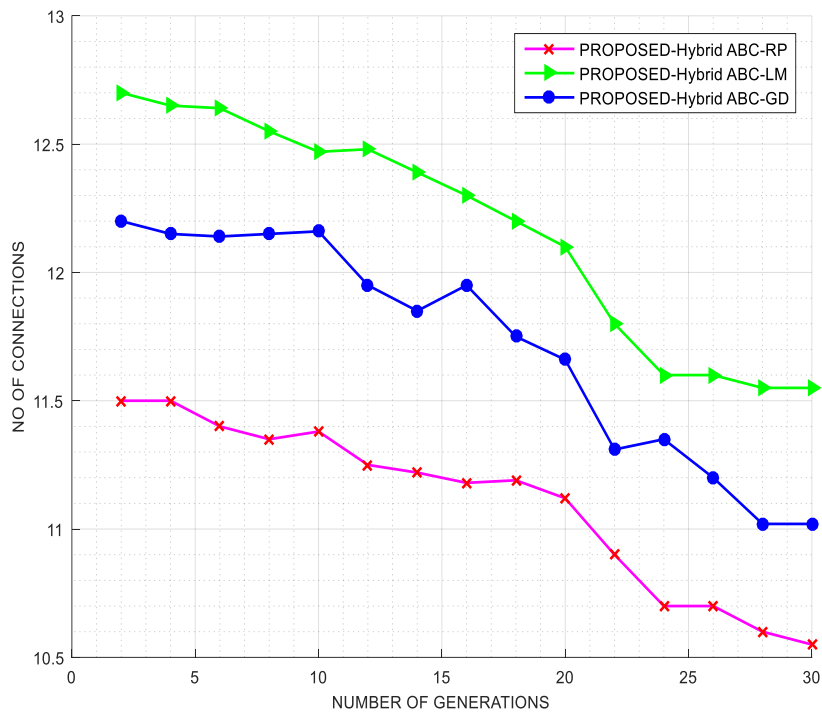
The average number of connections for Hybrid ABC-RP is 10.25 under a generation size of 20. Followed by Hybrid ABC-RP, Hybrid ABC-GD produced 11.03 average connections under generation size of thirty. Next to Hybrid ABC-GD, the Hybrid ABC-LM produced 11.21 average numbers of connections under generation size of thirty. Hence it is concluded that the Hybrid ABC-RP produced the least complex ANN when compared to the other two variants. Hybrid ABC-RP achieved low complexity of 7.07% less than Hybrid ABC-GD and 8.56% less complexity than Hybrid ABC-LM. Figure (9) shows the evolution of classification accuracy for different generation sizes for a single independent run of the Hybrid ABC algorithm. Hybrid ABC-RP achieved the best accuracy of 0.58% more than Hybrid ABC-LM and 0.48% more than Hybrid ABC-GD.



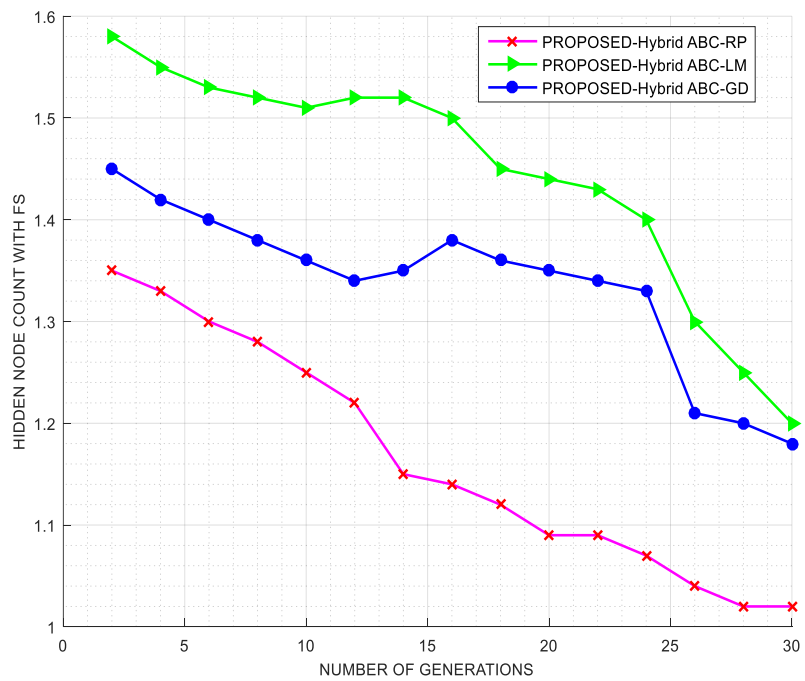


**Figure 9: Performance of Hybrid ABC in terms of classification accuracy**

Figure (10) shows the evolution of a number of connections for different generation sizes. According to Figure (10), Hybrid ABC-RP has used the least connections producing low complexity network followed by Hybrid ABC-GD and Hybrid ABC-LM. The least number of connections is by Hybrid ABC-RP and Hybrid ABC-GD which is 7 connections that are 12.5% less than Hybrid ABC-LM. Figure (11) shows the evolution of the hidden node count over different generation sizes.



**Figure 10: Evolution of number of connections of Hybrid ABC for different generations**



**Figure 11: Hidden node count for Hybrid ABC**

The confusion matrix showing the classification performance of Hybrid ABC is presented in Table (11). A total number of 650 malignant cases and 1100 benign cases are investigated and True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) are calculated.

**Table 11: Confusion Matrix of the proposed Hybrid ABC algorithm for ten runs**

Comparative methods	Actual	No of cases	Test outcome-Predicted	
			Malignant	Benign
Hybrid ABC-RP	Malignant	650	648(TP)	6(FN)
	Benign	1100	2(FP)	1094(TN)
Hybrid ABC-LM	Malignant	650	645 (TP)	13(FN)
	Benign	1100	5(FP)	1087 (TN)
Hybrid ABC-GD	Malignant	650	646(TP)	12(FN)
	Benign	1100	4(FP)	1088(TN)

**Table 12: Hybrid ABC algorithm based on different metrics**

Metrics	Proposed Hybrid ABC-RP	Proposed Hybrid ABC-LM	Proposed Hybrid ABC-GD
Sensitivity (%)	99.08	98.02	98.17
Specificity (%)	99.81	99.54	99.63
Accuracy (%)	99.54	98.97	99.08
Precision (%)	99.69	99.23	99.38
Negative predictive Value (%)	99.45	98.81	98.90
F-measure	0.9938	0.9243	0.9877

**Table 13: Performance of Hybrid ABC-RP based on feature selection**

Max Generation Size	Feature Selection (FS)	Average Hidden Node Count	Average Number of Selected Features	Average Number of connections	Average Accuracy
10	With FS	1.4	4.8	11.7	98.25
	Without FS	2.0	9	26.0	92.39
20	With FS	1.1	4.1	10.3	99.54
	Without FS	1.6	9	21.2	93.71
30	With FS	1.3	4.6	11.3	98.92
	Without FS	1.8	9	23.6	92.12

The performance metrics such as the sensitivity, specificity and other measures of the proposed Hybrid ABC are listed in Table (12). The sensitivity and specificity are high for Hybrid

ABC-RP followed by Hybrid ABC-GD and Hybrid ABC-LM. Table (13) shows the performance of the Hybrid ABC-RP optimized ANN classifier in terms of the feature selection process since it has achieved the highest accuracy and less complexity network. The Hybrid ABC-RP optimized ANN classifier with feature selection enhances the classification accuracy with less complexity when compared without feature selection. Table (14) shows the confusion matrix consists of actual and predicted cases that contain two types of classification called the benign and malignant for best network obtained using RP with optimally selected feature subsets.

**Table 14: Confusion Matrix of Hybrid ABC-RP for the best network with selected features**

Feature Selection	Actual Cases		Predicted Cases		Selected Feature Set
			Benign	Malignant	
With Feature Selection	Benign	110	110	0	uniformity of cell shape, uniformity of cell size, Bare nuclei
	Malignant	65	0	65	
Without Feature Selection	Benign	110	104	5	Total dataset from Table 1
	Malignant	65	6	60	

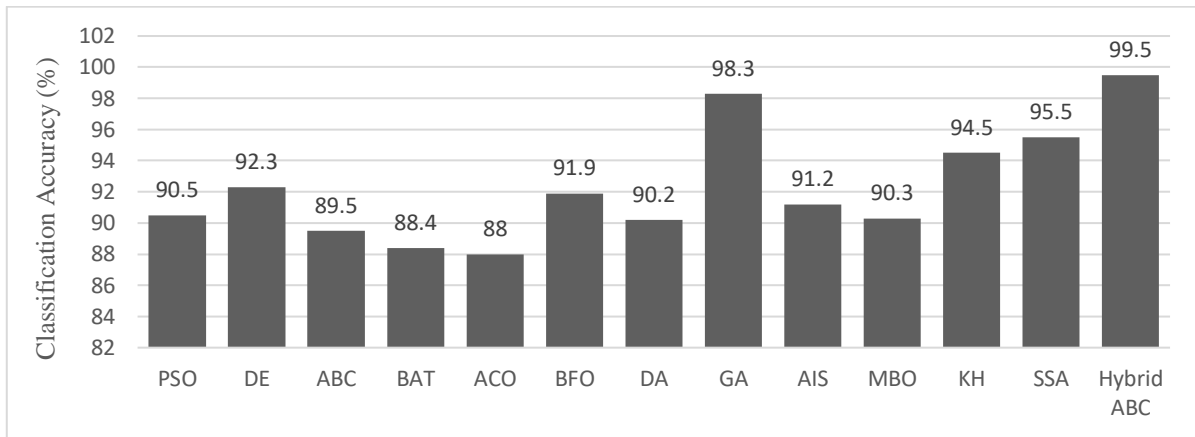
The proposed Hybrid ABC optimized ANN classifier performance in terms of the average computational time is shown in Table (15). The average computational time is calculated for ten independent runs for different generation sizes such as 10, 20 and 30. The Hybrid ABC-LM requires less computational time followed by Hybrid ABC-RP and Hybrid ABC-GD.

**Table 15: Performance-based on computational time**

Maximum Generation Size	Average CPU Time(s)		
	Proposed Hybrid ABC-RP	Proposed Hybrid ABC-LM	Proposed Hybrid ABC-GD
10	240.2	132.5	355.3
20	605.9	520.6	689.6
30	876.1	754.8	1100.6

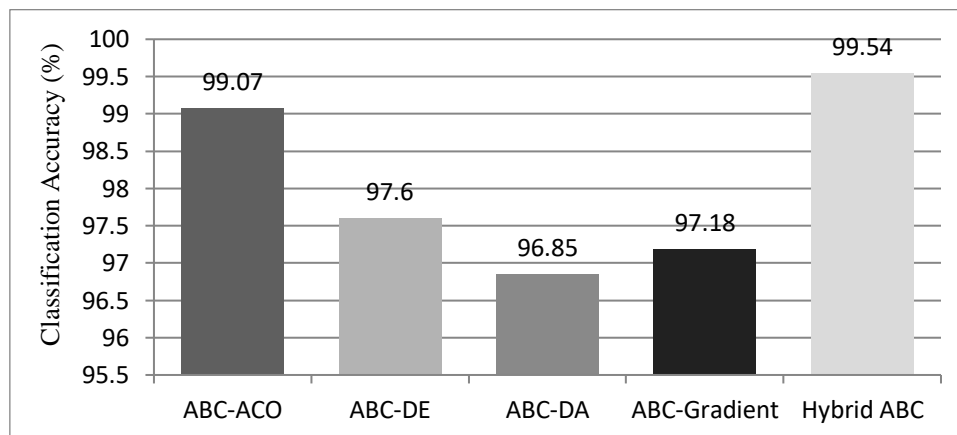
Figure (12) shows a comparison between Hybrid ABC with other evolutionary algorithms like particle Swarm Optimization (PSO) Differential Evolution (DE), Artificial Bee Colony Optimization (ABC), Bat Algorithm (BA), Ant Colony Optimization (ACO), Bacterial Foraging Optimization (BFO), Dragonfly Algorithm (DA), Genetic Algorithm (GA), Artificial Immune System (AIS), Monarch Butterfly Optimization (MBO), Krill Herd (KH) Algorithm, and Salp

Swarm Algorithm (SSA) using the WBCD dataset. The aforementioned algorithms are executed using ANN with RP for classifying the datasets of WBCD. Each is executed for 10 runs for 10, 20 and 30 generations and the best are taken and compared with the proposed Hybrid ABC algorithm.



**Figure 12: Comparison with evolutionary methods**

Figure (13) shows a comparison of existing ABC based hybrid approaches that used the WBCD dataset namely ABC-ACO [39], ABC-DE [40], ABC-DA [43], and ABC-Gradient [44].



**Figure 13: Comparison with ABC based Hybrid algorithms**

Table (16) shows the comparison of the existing breast cancer diagnosis schemes tested under with proposed Hybrid ABC using WBCD. From table (10), it can be concluded that Hybrid ABC outperforms other existing schemes.

**Table 16: Existing breast cancer diagnosis schemes comparison**

First author, Year	Method	Accuracy (%)
Quinlan,1996 [56]	C4.5	94.74

Hamilton,1996 [57]	RAIC	95.00
Nauck,1999 [58]	NEFCLASS	95.06
Pena-Reyes,1999 [59]	FUZZY-GA	97.36
Setino, 2000 [60]	Neuro-rule	98.10
Albrecht, 2002 [61]	LSA Machine	98.80
Fogel, 1995 [62]	ENN	98.05
Abonyi, 2003 [63]	SFC	95.57
Polat, 2007 [64]	LS-SVM	98.54
Gujaro-Berdinas, 2007 [65]	LIS	96.00
Karabatak, 2009 [66]	AR+NN	97.40
Stoean, 2013 [67]	SVM+EA	97.07
Fadzil Ahmad, 2014 [68]	GANN-MLP	98.29
Shunmugapriya, 2017 [3R9]	ABC-ACO-J48	99.07
H. Rao, 2018 [44]	ABC+ GD	97.18
Ghanem, 2018 [43]	ABC-DA+MLP	96.85
Karthik, 2018 [69]	DNN	98.62
Reyhaneh,2020[70]	GO-FS	97.85
Nayak ,2020[71]	EHO-NN	98.37
Dalwinder,2020[72]	Ant Lion-MLP	98.61
Srinivasa,2021[73]	BCRO-Naïve Bayes	98.82
Proposed Work	HABC-RP	99.14
Proposed Work	Hybrid ABC-RP	99.54

## 5 Conclusion

The paper introduced two hybrid ABC versions for concurrent feature selection and parameter tuning of ANN for diagnosing breast cancer. The strength of the proposed algorithms lies in deriving an optimal ANN classifier using an automatic and optimal selection of parameters. It uses a simple and wrapper-based approach. The proposed algorithms were investigated with the help of various backpropagation algorithms such as RP, LM and GD. It has been concluded that RP achieved higher accuracy and low complexity ANN network. The result shows that it is vital to do feature selection along with an optimal selection of hidden node size and initial weights. During comparison with existing works, the proposed HABC-RP and Hybrid ABC-RP showed promising results. The comparison of the proposed HABC-RP and Hybrid ABC-RP with existing evolutionary approaches proved better results.

## REFERENCES

- [1] N. R. Pal, B. Bhowmick, S. K. Patel, S. Pal, and J. Das, "A multi-stage neural network aided system for detection of microcalcifications in digitized mammograms," *Neurocomputing*, vol. 71, no. 13-15, pp. 2625–2634, 2008.
- [2] [https://www.breastcancer.org/symptoms/understand\\_bc/statistics](https://www.breastcancer.org/symptoms/understand_bc/statistics)
- [3] Population-Based Cancer report. [http://ncrpindia.org/annual\\_reports.aspx](http://ncrpindia.org/annual_reports.aspx) [Online]
- [4] M. M. Eltoukhy, I. Faye, and B. B. Samir, "A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram," *Computers in Biology and Medicine*, vol. 40, no. 4, pp. 384–391, 2010.
- [5] M. Zhao, C. Fu, L. Ji, K. Tang, and M. Zhou, "Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5197–5204, 2011.
- [6] M. F. Akay, "Support vector machines combined with feature selection for breast cancer diagnosis," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3240–3247, 2009.
- [7] L. C. Jain, "Advances in design and application of neural networks," *Neural Computing and Applications*, vol. 19, no. 2, pp. 167–168, 2010.
- [8] M. Karabatak and M. C. Ince, "An expert system for detection of breast cancer based on association rules and neural network," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3465–3469, 2009.
- [9] E. D. Übeyli, "Implementing automated diagnostic systems for breast cancer detection," *Expert Systems with Applications*, vol. 33, no. 4, pp. 1054–1062, 2007.
- [10] Goldberg DE, Holland JH, "Genetic algorithms and Machine Learning," *Machine Learning*, vol. 3, no. 2-3, pp. 95–99, 1988.
- [11] Storn R, Price K. Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces. *J Global Optim* 1997; 11:341-59.
- [12] Rechenberg I. *Evolution strategy: optimization of technical systems by means of biological evolution*, 104. Stuttgart: Fromman-Holzboog; 1973.

- [13] A. H. Gandomi, X.-S. Yang, S. Talatahari, and S. Deb, "Coupled eagle strategy and differential evolution for unconstrained and constrained global optimization," *Computers & Mathematics with Applications*, vol. 63, no. 1, pp. 191–200, 2012.
- [14] Simon D. Biogeography-based optimization. *Evol Comput IEEE Trans* 2008; 12:702-13.
- [15] A. Sarkheyli, A. M. Zain, and S. Sharif, "Robust optimization of ANFIS based on a new modified GA," *Neurocomputing*, vol. 166, pp. 357–366, 2015.
- [16] Thompson and K. Suresh Joseph, "Particle Swarm Optimization-Based Energy Efficient Channel Assignment Technique for Clustered Cognitive Radio Sensor Networks," *The Computer Journal*, Oxford University Press, vol. 61, no. 6, pages 926-936, Jun. 2018.
- [17] S. Thompson and K. Suresh Joseph, "PSO assisted OLSR Routing for Cognitive Radio Vehicular Sensor Networks," in *ACM International Conference on Informatics and Analytics*, pages 1–8, Aug. 2016. DOI 10.1145/2980258.2980457.
- [18] Colorni A, Dorigo M, Maniezzo V. Distributed optimization by ant colonies. In: *Proceedings of the first European conference on artificial life*; 1991. p. 134-42.
- [19] Eberhart RC, Kennedy J. A new optimizer using particle swarm theory. In: *Proceedings of the sixth international symposium on micro machine and human science*; 1995. p. 39-43.
- [20] S. Punitha, S. Ravi, M. A. Devi, and J. Vaishnavi, "Particle swarm optimized computer aided diagnosis system for classification of breast masses," *Journal of Intelligent & Fuzzy Systems*, vol. 32, no. 4, pp. 2819–2828, 2017.
- [21] Yang X-S. A new metaheuristic bat-inspired algorithm. In: *Nature inspired cooperative strategies for optimization (NICSO 2010)*. Springer; 2010. p. 65-74
- [22] Karaboga D, Basturk B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *J Global Optim* 2007; 39:459-71
- [23] Mirjalili S, Mirjalili S, Hatamlou A. Multi-verse optimizer: a nature-inspired algorithm for global optimization. *Neural Comput Appl* 2015:1-19 /03/17 2015.



- [24] A. H. Gandomi and A. H. Alavi, "Krill herd: A new bio-inspired optimization algorithm," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 12, pp. 4831–4845, 2012.
- [25] G.-G. Wang, S. Deb, and Z. Cui, "Monarch butterfly optimization," *Neural Computing and Applications*, vol. 28, no. 3, pp. 1–20, 2015.
- [26] P. K. M, "Biomimicry of bacterial foraging for distributed optimization and control," *IEEE Control Systems*, vol. 22, no. 3, pp. 52–67, 2002.
- [27] Castro, L. & Timmis, "Artificial immune systems as a novel soft computing paradigm", *Journal of Soft Computing*7: 526,2003.
- [28] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris, and S. M. Mirjalili, "Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems," *Advances in Engineering Software*, vol. 114, pp. 163–191, 2017.
- [29] A. Faramarzi, M. Heidarinejad, S. Mirjalili, and A. H. Gandomi, "Marine Predators Algorithm: A Nature-inspired Metaheuristic," *Expert Systems with Applications*, p. 113377, 2020.
- [30] S. Punitha, A. Amuthan, and K. Suresh Joseph, "Enhanced Monarchy Butterfly Optimization Technique for effective breast cancer diagnosis," *Journal of Medical Systems*, vol. 43, no. 7, 2019
- [31] D. Furundzic, M. Djordjevic, and A. J. Bekic, "Neural networks approach to early breast cancer detection," *Journal of Systems Architecture*, vol. 44, no. 8, pp. 617–633, 1998.
- [32] J.F. Kolen, J.B. Pollack, "Back propagation is sensitive to initial conditions," *Adv Neural Information Process System*, vol 3 , pp. 860–867, 1991.
- [33] K. P. Ferentinos, "Biological engineering applications of feedforward neural networks designed and parameterized by genetic algorithms," *Neural Networks*, vol. 18, no. 7, pp. 934–950, 2005.
- [34] R. Setiono and H. Liu, "Neural-network feature selector," *IEEE Transactions on Neural Networks*, vol. 8, no. 3, pp. 654–662, 1997.
- [35] A. Verikas and M. Bacauskiene, "Feature selection with neural networks," *Pattern Recognition Letters*, vol. 23, no. 11, pp. 1323–1335, 2002.
- [36] M. M. Kabir, M. M. Islam, and K. Murase, "A new wrapper feature selection approach using neural network," *Neurocomputing*, vol. 73, no. 16-18, pp. 3273–3283, 2010.
- [37] P. K. M, "Biomimicry of bacterial foraging for distributed optimization and control," *IEEE Control Systems*, vol. 22, no. 3, pp. 52–67, 2002.

- [38] I. H. Osman and G. Laporte, "Metaheuristics: A bibliography," *Annals of Operations Research*, vol. 63, no. 5, pp. 511–623, 1996.
- [39] P. Shunmugapriya and S. Kanmani, "A hybrid algorithm using ant and bee colony optimization for feature selection and classification (AC-ABC Hybrid)," *Swarm and Evolutionary Computation*, vol. 36, pp. 27–36, 2017.
- [40] E. Zorarpacı and S. A. Özel, "A hybrid approach of differential evolution and artificial bee colony for feature selection," *Expert Systems with Applications*, vol. 62, pp. 91–103, 2016.
- [41] D. J. Kumari, "Structural Redesign of Artificial Neural Network for Predicting Breast Cancer with the Aid of Artificial Bee Colony," *Indian Journal of Science and Technology*, vol. 10, no. 15, pp. 1–8, 2017.
- [42] P. P. Sarangi, A. Sahu, and M. Panda, "Training a Feed-Forward Neural Network Using Artificial Bee Colony with Back-Propagation Algorithm," *Intelligent Computing, Networking, and Informatics Advances in Intelligent Systems and Computing*, pp. 511–519, 2013.
- [43] W. A. H. M. Ghanem and A. Jantan, "A Cognitively Inspired Hybridization of Artificial Bee Colony and Dragonfly Algorithms for Training Multi-layer Perceptrons," *Cognitive Computation*, vol. 10, no. 6, pp. 1096–1134, 2018.
- [44] H. Rao, X. Shi, A. K. Rodrigue, J. Feng, Y. Xia, M. Elhoseny, X. Yuan, and L. Gu, "Feature selection based on artificial bee colony and gradient boosting decision tree," *Applied Soft Computing*, vol. 74, pp. 634–642, 2019.
- [45] H. Djellali, A. Djebbar, N. G. Zine, and N. Azizi, "Hybrid Artificial Bees Colony and Particle Swarm on Feature Selection," *Computational Intelligence and Its Applications IFIP Advances in Information and Communication Technology*, pp. 93–105, 2018.
- [46] H. Badem, A. Basturk, A. Caliskan, and M. E. Yuksel, "A new efficient training strategy for deep neural networks by hybridization of artificial bee colony and limited-memory BFGS optimization algorithms," *Neurocomputing*, vol. 266, pp. 506–526, 2017.
- [47] B. A. Garro, K. Rodríguez, and R. A. Vázquez, "Classification of DNA microarrays using artificial neural networks and ABC algorithm," *Applied Soft Computing*, vol. 38, pp. 548–560, 2016.
- [48] S. Palanisamy, S. Kanmani, "Artificial Bee Colony Approach for Optimizing Feature Selection," *International Journal of Computer Science Issues*, vol.9, no.3, pp.432-438, 2012.

- [49] M. Schiezero and H. Pedrini, "Data feature selection based on Artificial Bee Colony algorithm," *EURASIP Journal on Image and Video Processing*, vol. 2013, no. 1, 2013.
- [50] M. M. Kabir, M. M. Islam, and K. Murase, "A new wrapper feature selection approach using neural network," *Neurocomputing*, vol. 73, no. 16-18, pp. 3273–3283, 2010.
- [51] G. Nagarajan, R. Minu, B. Muthukumar, V. Vedanarayanan, and S. Sundarsingh, "Hybrid Genetic Algorithm for Medical Image Feature Extraction and Selection," *Procedia Computer Science*, vol. 85, pp. 455–462, 2016.
- [52] Shunmugapriya Palanisamy and Kanmani , "Artificial Bee Colony Approach for Optimizing Feature Selection," *IJCSI International Journal of Computer Science Issues*, Vol. 9, Issue 3, No 3, May 2012
- [53] M. S. Uzer, O. Inan, and N. Yılmaz, "A hybrid breast cancer detection system via neural network and feature selection based on SBS, SFS and PCA," *Neural Computing and Applications*, vol. 23, no. 3-4, pp. 719–728, 2012.
- [54] K. Tamura and K. Yasuda, "Primary study of spiral dynamics inspired optimization," *IEEEJ Transactions on Electrical and Electronic Engineering*, vol. 6, no. S1, pp. 98–100, 2010.
- [55] L. Prechelt Proben1: a set of neural network benchmark problems and benchmarking rules. *Technical Report*, University of Karlsruhe, Karlsruhe, Germany, 1994
- [56] J. R. Quinlan, "Improved Use of Continuous Attributes in C4.5," *Journal of Artificial Intelligence Research*, vol. 4, pp. 77–90, 1996.
- [57] Hamilton HJ, Shan N, Cercone N, RIAC: a rule induction algorithm based on approximate classification. "International conference on engineering applications of neural networks", University of Regina, 1996.
- [58] D. Nauck and R. Kruse, "Obtaining interpretable fuzzy classification rules from medical data," *Artificial Intelligence in Medicine*, vol. 16, no. 2, pp. 149–169, 1999.
- [59] C. A. Peña-Reyes and M. Sipper, "A fuzzy-genetic approach to breast cancer diagnosis," *Artificial Intelligence in Medicine*, vol. 17, no. 2, pp. 131–155, 1999.
- [60] R. Setiono, "Generating concise and accurate classification rules for breast cancer diagnosis," *Artificial Intelligence in Medicine*, vol. 18, no. 3, pp. 205–219, 2000.
- [61] A. Albrecht, G. Lappas, S. Vinterbo, C. Wong, and L. Ohno-Machado, "Two applications of the LSA machine," *Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP 02.*, pp. 184–189.
- [62] D. B. Fogel, E. C. Wasson, and E. M. Boughton, "Evolving neural networks for detecting breast cancer," *Cancer Letters*, vol. 96, no. 1, pp. 49–53, 1995.
- [63] J. Abonyi and F. Szeifert, "Supervised fuzzy clustering for the identification of fuzzy classifiers," *Pattern Recognition Letters*, vol. 24, no. 14, pp. 2195–2207, 2003.

- [64] K. Polat and S. Güneş, “Breast cancer diagnosis using least square support vector machine,” *Digital Signal Processing*, vol. 17, no. 4, pp. 694–701, 2007.
- [65] B. Guijarro-Berdiñas, O. Fontenla-Romero, B. Pérez-Sánchez, and P. Fraguera, “A Linear Learning Method for Multilayer Perceptrons Using Least-Squares,” *Intelligent Data Engineering and Automated Learning - IDEAL 2007 Lecture Notes in Computer Science*, vol. 4881, pp. 365–374.
- [66] M. Karabatak and M. C. Ince, “An expert system for detection of breast cancer based on association rules and neural network,” *Expert Systems with Applications*, vol. 36, no. 2, pp. 3465–3469, 2009.
- [67] R. Stoean and C. Stoean, “Modeling medical decision making by support vector machines, explaining by rules of evolutionary algorithms with feature selection,” *Expert Systems with Applications*, vol. 40, no. 7, pp. 2677–2686, 2013.
- [68] F. Ahmad, N. A. M. Isa, Z. Hussain, M. K. Osman, and S. N. Sulaiman, “A GA-based feature selection and parameter optimization of an ANN in diagnosing breast cancer,” *Pattern Analysis and Applications*, vol. 18, no. 4, pp. 861–870, 2014.
- [69] S. Karthik, R. S. Perumal, and P. V. S. S. R. Chandra Mouli, “Breast Cancer Classification Using Deep Neural Networks,” *Knowledge Computing and Its Applications*, pp. 227–241, 2018.
- [70] Reyhaneh Yaghobzadeh, Seyyed Reza Kamel, Mojtaba Asgari, and Hassan Saadatmand, “A Binary Grasshopper Optimization Algorithm for Feature Selection,” *International Journal of Engineering Research and*, vol. V9, no. 03, 2020.
- [71] M. Nayak, S. Das, U. Bhanja, and M. R. Senapati, “Elephant herding optimization technique based neural network for cancer prediction,” *Informatics in Medicine Unlocked*, vol. 21, p. 100445, 2020.
- [72] S. Dalwinder, S. Birmohan, and K. Manpreet, “Simultaneous feature weighting and parameter determination of Neural Networks using Ant Lion Optimization for the classification of breast cancer,” *Biocybernetics and Biomedical Engineering*, vol. 40, no. 1, pp. 337–351, 2020.
- [73] P. C. Srinivasa Rao, A. J. Sravan Kumar, Q. Niyaz, P. Sidike, and V. K. Devabhaktuni, “Binary chemical reaction optimization based feature selection techniques for machine learning classification problems,” *Expert Systems with Applications*, vol. 167, p. 114169, 2021.
- [74] S. Forrest and S. A. Hofmeyr, “Immunology as information processing,” in *Design Principles for the Immune System and Other Distributed Autonomous Systems*, SantaFe

Institute Studies in the Sciences of Complexity, L. A. Segel and I. Cohen, Eds. Oxford, U.K. Oxford Univ. Press, pp. 361–387, 2000.

- [75] L N de Charsto and J. V. Zuben, “Learning and Optimization using Clonal Selection Principle,” *IEEE Trans on Evolutionary Computation*, Special issue on Artificial Immune Systems, vol. 6, issue 3 , pp.239-251,2002.
- [76] L N de Charsto and J. Timmis, “An Artificial Immune Network for MultimodalFunction Optimization”, *IEEE Congress on Evolutionary Computation (CEC‘02)*, vol. 1, pp.699-674, May, Hawaii,2002.
- [77] Wolberg WH (1990) Breast cancer Wisconsin (original) dataset. <http://archive.ics.uci.edu/ml/datasets/Breast?Cancer?Wisconsin?%28Original%29>. Accessed 14 Oct 2012.
- [78] S. Punitha, Thompson Stephan, “An Automated breast cancer diagnosis scheme using feature selection and parameter optimization of an ANN model” *Journal of Computer and Electrical engineering*, Volume 90, March 2021, 106958. <https://doi.org/10.1016/j.compeleceng.2020.106958>
- [79] S. Punitha, Thompson Stephan, Ramani Kannan, Ajith, “A Hybrid Artificial Bee Colony with Whale Optimization algorithm for improved breast cancer diagnosis”, *Neural Computing and applications*, 33(20), 13667-1369. <https://doi.org/10.1007/s00521-021-05997-6>