Hybrid-based bat optimization with fuzzy C-means algorithm for breast cancer analysis

Chocko Valliappa, Reenadevi Rajendran[1], Sathiyabhama Balasubramaniam[1], Sankar Sennan[1], Sathiya Thanikachalam[1], Yuvarajan Velmurugan[1], Nirmalesh Kumar Sampath Kumar[2]

CEO, Vee Technologies; Vice-Chairman, Sona Group of Institutions, Bengaluru, Karnataka, [1]Department of Computer Science and Engineering, Sona College of Technology, [2]Department of Knowledge Transfer and Valourization, Sona College of Technology, Salem, Tamil Nadu, India

ABSTRACT

Background: Breast cancer is one of the most frequent types of cancer among women and early identification can reduce the mortality rate drastically. Feature selection is one of the significant tasks in the breast cancer analysis process. Several types of feature selection algorithms have been implemented to select the most appropriate feature for breast cancer analysis. However, they have to take a longer time to converge, over-fitting problems and providing less accuracy. Hence, a hybrid bat optimization algorithm combined with chaotic maps and fuzzy C-means clustering algorithm (BSCFC) is proposed for feature selection. **Aims and Objectives:** An integrated optimized bat optimization algorithm combined with chaotic maps and fuzzy C-means clustering algorithm (BSCFC) is proposed to determine the relevant feature. **Materials and Methods:** Breast cancer mini-Mammographic Image Analysis Society database (MIAS) dataset is used for analysis. Further, median filters are used for preprocessing, Region of Interest (ROI) was utilized for segmentation, gray level co-occurrence matrix (GLCM), and texture analysis are utilized in the feature extraction process. A hybrid bat optimization algorithm combined with chaotic maps and fuzzy C-means clustering algorithm (BSCFC) is proposed for feature selection. K nearest neighbor (KNN) classifier is used for classification. **Results:** Performance of the proposed system is evaluated using standard measures and achieved the highest accuracy rate of (98.2%), specificity of (97.3%), and sensitivity of (98.3%) as compared to other relevant methods such as bat, chaotic bat, chaotic crow search, ant lion optimization, and chaotic ant lion optimization algorithm. **Conclusion:** The proposed BSCFC algorithm is designed to improve the performance of convergence speed and control balance between exploration and exploitation rate using five types of chaotic maps namely sinusoidal, sine, gauss, logistic, and tent maps. The results show that the BSCFC with sinusoidal maps can significantly boost the classification performance of the BSCFC algorithm in classifying the breast cancer images with reduced features, which in turn optimizes the radiologists' time for their interpretation.

Keywords: Bat algorithm, breast cancer, chaotic map, feature selection, fuzzy C-means, optimization

Introduction

Breast cancer is the most widely diagnosed cancer worldwide, according to the statistics released by the International Agency for Research on Cancer.[1] In the past two decades, the number of people diagnosed with cancer nearly doubled.[2] Breast cancer can be cured if detected early and treated appropriately.[3] Breast cancer treatment can be highly effective and achieving survival probabilities of 90% or higher, particularly when the disease is identified early.[4] The feature is used to describe the attributes of breast tumors for the classification task. In model construction, feature selection is a crucial technique for determining a subset of appropriate features

Access this article online	
	Quick Response Code
Website: www.ijncd.org	
DOI: 10.4103/2468-8827.330652	

Address for correspondence: Ms. Reenadevi Rajendran, Department of Computer Science and Engineering, Sona College of Technology, Salem - 636 005, Tamil Nadu, India. E-mail: reenadevi@sonatech.ac.in

How to cite this article: Valliappa C, Rajendran R, Balasubramaniam S, Sennan S, Thanikachalam S, Velmurugan Y, *et al*. Hybrid-based bat optimization with fuzzy C-means algorithm for breast cancer analysis. Int J Non-Commun Dis 2021;6:S62-8.

Submitted: 30-Jul-2021 **Revised:** 05-Sep-2021
Accepted: 22-Sep-2021 **Published:** 19-Nov-2021

to diagnose breast cancer.[5] Feature selection algorithms are extremely useful in supervised learning that optimize specific functions by choosing the appropriate features on a target class to maximize predictive accuracy. Various techniques for selecting features are used to optimize the predictive models. In recent decades, several different meta-heuristic algorithms have been developed to solve the optimization problems.[6] They can be used to achieve optimal results in a variety of applications. However, all of these techniques are trapped in local optima and are computationally costly.

A better selection based on the global search method is necessary to overcome this problem. Many meta-heuristic methods were proposed for the selection of features, based on the bio-inspiration of nature.[7] Bat algorithm is one of the contemporary, modern meta-heuristic methods that were inspired by bat social behavior.[8] Bat algorithm is characterized by the fact that its equations are straightforward, which makes it possible to apply in many fields including selecting the feature and classification. For every optimization algorithm, exploration and exploitation are the two important components. Much exploration leads to expanding, time-consuming solutions, and a chance of being optimized and not converged. The stability among exploitation and exploration is, therefore, a difficult task in the optimization problem. Trapping in local optima, as well as slow convergence speed, are the challenges in real complex problems for evolutionary algorithms. This work will investigate and balance exploitation rates and increase BSCFC's effectiveness in local optimization as well as convergence using fuzzy logic and chaotic map. Chaotic strategy recently succeeded in optimizing meta-heuristics in particular for breast cancer diagnosis.[9] Parameter tuning has been accomplished through the use of chaotic sequences. Due to high mixing capabilities, chaos can be predicted to be more mobile and diverse when a stable parameter is replaced by a chaotic map.

As chaotic maps may lead to various behaviors, they are used to substitute the parameters of the bat algorithm. In this work, sinusoidal, sine, gauss, logistic, and tent maps are integrated into the bat algorithm. The fuzzy c-means clustering algorithm (FCM) is a well-known fuzzy clustering algorithm which introduced the fuzzy sets.[10,11] When objects in datasets could not be partitioned into well-separated clusters, fuzzy clustering becomes extremely useful. It is used to maximize the classification accuracy, minimize errors, and reduce misclassify features. The advantages of the proposed work are as follows:

1. An integrated optimized BSCFC-based bat algorithm with chaotic maps and FCM to determine the relevant feature
2. Five different chaotic maps are used to overcome the lack of convergence of the bat algorithm
3. FCM is used and assesses bats in high-dimensional space which deals with uncertainty effectively.

Related work

Numerous techniques are used in investigating digital mammograms used for breast cancer analysis in the past few decades. In this section, various hybrid-based bio-inspired optimization algorithms are reported.

Tawhid and Dsouza implemented a hybrid feature selection algorithm by combining bat algorithm with particle swarm optimization algorithm. They have used a benchmark dataset for evaluation.[12] Al-Betar *et al*. developed a filter-wrapper-based algorithm by utilizing minimum redundancy and maximum relevancy as a filter as well as a modified bat algorithm as a wrapper approach for gene expression data.[13] Yu *et al*. designed a chaotic-based bat algorithm for optimization problems. The authors have used an inertial weight to update the velocity for global exploration.[14] Agrawal *et al*. proposed an improved optimization method by integrating the gray wolf and bat algorithm for updating the weights in a multi-layer perceptron neural network to minimize the classification error.[15] Tripathi *et al*. designed a hybrid model by using binary bat optimization techniques with a new fitness function and by combining them with the Radial Basis Neural Function Network to classify the credit score.[16] Eskandari and Javidi proposed a clustering-based hybridized bat algorithm for the feature selected. The authors have compared the algorithm with the variants of the bat algorithm and showed that the proposed algorithm provides better results.[17]

Anter and Ali proposed a hybrid algorithm by integrating the theoretical chaos with the crow search algorithm and Fuzzy c-mean algorithm for health-care problems. The authors used various chaotic maps to improve the position of the row. FCM clustering algorithm has been used to improve the target area for exploration.[18] The wrapper-based bat algorithm for feature selection was created by Naik *et al*. They have used a fitness function for evaluation and a one-pass generalized neural network for classification.[19] Zhang *et al*. developed a *t*-test-based adaptive variable size bat algorithm for feature reduction. The *t*-test is utilized for the initial population generation. In this, a variable step-size adaptive algorithm is employed to speed up convergence and prevent local optima.[20]

Many nature-inspired optimization algorithms presently available in the literature take a longer time to converge, over-fitting problems and providing less accuracy. Hence, in the proposed research work, the authors have developed an integrated framework that incorporates chaos theory, bat algorithm, and fuzzy logic to obtain the best local optima and quicker convergence with less computational time.

Materials and Methods

Bat algorithm

In the recent decades, numerous nature-inspired optimization algorithms have been implemented and are becoming popular choices for their applications in engineering optimization techniques.[21-23] Xin-She Yang implemented a metaheuristic algorithm based on microbats echolocation, namely the bat algorithm.[24] Conventionally, microbats use echolocation to look for prey. While microbats emit short pulses during roams, their pulse rate increases as potential predator approaches and the rate of frequency is attuned. An increase in frequency, known as frequency-tuning, combined with pulse emission acceleration will reduce the echolocations wavelength and thereby improve detection accuracy. The echolocation features of microbats can be idealized as per the following three rules:

1. Each bat uses echolocation to determine the distance, and they also instinctively know the dissimilarity between food/prey and obstacles
2. Bats fly at position x_i randomly, with varying frequencies of wavelength µµ and loudness Al_0, with velocity vl_i at location x_i. Bats can adjust the pulse emission rates automatically rt ϵ (0, 1) reliant on the closeness of the target and the wavelength of its emissions
3. Although the loudness can change in a variety of ways, ranges are assumed from a high Al_0 to a low Al_{min}.

Position x_i and velocity vl_i for every bat is to be defined in a *d*-dimensional search space, and as the iterations go they must be updated successively. Equations (1), (2), and (3) are used to calculate the updated solutions and velocities at time step *t*.

$$fl_i = (fl_{max} - fl_{min})\beta + fl_{min} \tag{1}$$

$$vl_i^t = (x_i^{t-1} - x^*)fl_i + vl_i^{t-1} \tag{2}$$

$$x_i^t = vl_i^t + x_i^{t-1} \tag{3}$$

Where β is in the range (0, 1), a vector produced at random from a uniform distribution. Therefore, after comparing all the solutions across all of the n bats in the current iteration, x^*, the current global best location is discovered. Because the product $\mu_i fl_i$ represents the velocity increment, either fl_i or μ_i can be used to modify the velocity while keeping the other element constant depending on the nature of the problem. According to the domain size $f_{min} = 0$ and $f_{max} = 2$, the actual range can vary. At first, each bat is assigned a frequency at random that is drawn evenly from a pool of frequencies (f_{min}, f_{max}).

The new solution with each bat can be generated as:

$$x_{new} = x_{old} + \varepsilon Al^t \tag{4}$$

Where ε is the random value varies from (−1, 1), whereas $Al^t = < Al_i^t >$, the average noise of all the bats during this time step. For the sake of simplicity, $Al_0 = 1$ and $Al_{min} = 0$ can be used. When Al_{mi}n = 0, a bat has discovered its target and has momentarily stopped making any noise. Then,

$$Al_i^{t+1} = \alpha A_i^t \qquad rt_i^{t+1} = rt_i^0\,[1 - exp\,exp(-\gamma t)] \tag{5}$$

Where α and γ are both constants. For all, 0< α and γ <1.

$$Al_i^t \to 0, rt_i^t \to rt_i^0, t \to \infty \tag{6}$$

In this work, $\alpha = \gamma = 0.9$ is used in the simulations. However, the objective of this work is to employ chaotic maps to tune these four parameters and see if they can improve the bat algorithm's efficiency.

Chaotic maps

Chaos theory is the branch of mathematics that deals with complex systems whose behavior is highly sensitive to slight changes in conditions so that small alterations can give rise to strikingly great consequences. Because of its inherent search area capacity to optimize, chaos randomness differs significantly from statistically random. A chaotic map is an evolution function that represents chaotic behavior and can be discrete or continuous in time.[25] In this, the following five different types of one-dimensional chaotic maps are used to obtain chaotic sets. They are described as:

1. Sinusoidal map:[26] The sinusoidal map is mathematically defined as:

$$S_{n+1} = \alpha w_t^2 \sin(\pi s_n) \tag{7}$$

where $\alpha = 2.3$.

2. Sine map:[27] The sine map is defined mathematically as:

$$S_{n+1} = \frac{\alpha}{4}(\sin \pi s_n) \tag{8}$$

where 0< $\alpha \leq 4$.

3. Gauss map:[28] The nonlinear Gauss map is written as:

$$S_{n+1} = \beta + \exp(-\alpha s_n^2) \tag{9}$$

where α, β are the actual parameters

4. Logistic map:[29] It describes the composite behavior without the stochastic system's uncertainty, defined as:

$$S_{n+1} = cS_n\,(1 - S_n) \tag{10}$$

where $s_0 \in (0,1), s_0 \notin \{0, 0.25, 0.50, 0.75, 1\}$ with chaotic sequence of $c = 4$.

5. Tent map: [30] Tent map is expressed as:

$$S_{n+1} = \begin{cases} \frac{S_n}{0.07}, S_n & < 0.7 \frac{10(1-S_n)}{3}, S_n < 0.7 \end{cases} \tag{11}$$

Fuzzy C-mean clustering method

FCM is an algorithm for the cluster of data that allows a subset of data to belong to a cluster to a certain level.[31] The input of n vectors Xi, (i = 1, 2, 3,…, n) are divided into c fuzzy groups to decide a center of the cluster in each group to reduce the cost function of Euclidean distance. First of all, the n points (X1, X2,…., Xn) of cluster centers ct_i, and the membership matrix U is computed as:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{ct}\left(\frac{d_{ij}}{d_{kj}}\right)^{\frac{2}{m-1}}} \tag{12}$$

Where $d_{ij} = \|ct_i - x_j\|$ is the Euclidean distance from i^{th} center of the cluster to j^{th} data point, whereas m is the fuzziness index. The cost function is computed as:

$$J(U, c1, \ldots, c2) = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c}\sum_{j=1}^{n} \mu_{ij}^{m} d_{ij}^{2} \tag{13}$$

Calculates a new c fuzzy cluster center ct_i ($i = 1, 2, \ldots, C$) as:

$$ct_i = \frac{\sum_{j=1}^{n} \mu_{ij}^{m} x_j}{\sum_{j=1}^{n} \mu_{ij}^{m}} \tag{14}$$

FCM employs the Gaussian membership feature, as defined:

$$\mu(x) = exponent\left\{-\left[\left(\frac{x - ct_i}{\alpha_i}\right)^2\right]^{b_i}\right\} \tag{15}$$

In this a_i signifies the membership variance, b_i is tuning parameter, and ct_i shows the member's degree center.

Proposed BSCFC optimization algorithm

The chaotic theory, with its topologically blended, ergodic, and intrinsic stochastic nature, was used to optimize the balance between exploration and exploitation rate in the proposed BSCFC algorithm. The chaotic vector sequence updates the random variables of the bat algorithm. Five different chaotic maps, namely sinusoidal, sine, gauss, logistic, and tent maps, are employed in this work to significantly increase the effectiveness and convergence velocity of the BSCFC algorithm. Furthermore, the purpose of using FCM as an objective function is to measure the bats in a repetition that benefits from improving convergence performance, reducing the risk of stuck in local minima, and ensuring the selected features' goodness. The proposed algorithm's process flow is shown in Figure 1.

The chaotic vector is initialized in the form of chaotic maps that are utilized in the iteration process as a random chaotic motion in the BSCFC algorithm. The bat algorithm then generates the crow positions following chaotic motion as well as assigning a fitness value to each bat is done using the FCM objective function. FCM objective function

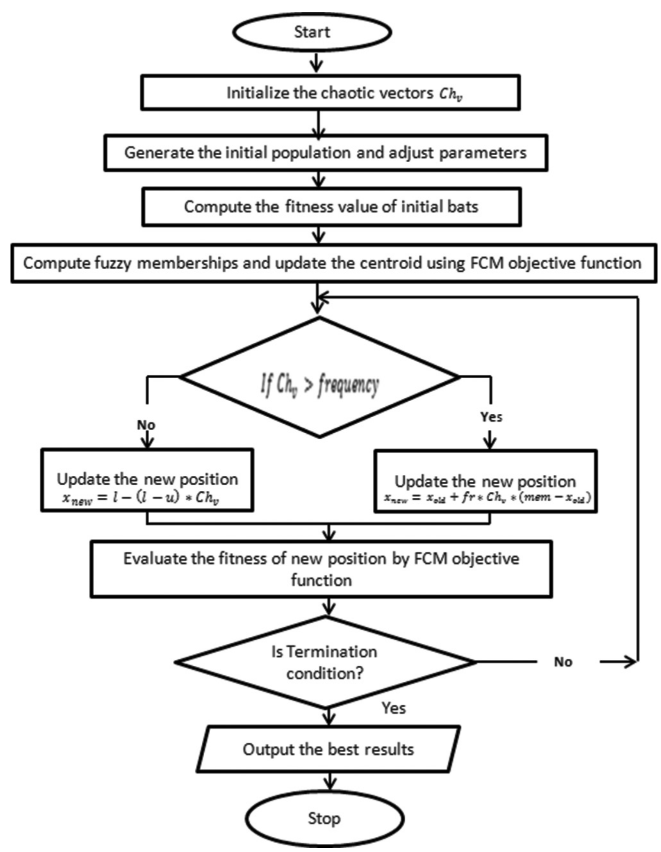

Figure 1: Process flow of the proposed hybrid Bat Optimization algorithm combined with Chaotic maps and Fuzzy C-Means clustering algorithm

serves as a cost function in the proposed BSCFC algorithm. It divides n vectors into c fuzzy groups, computes the clustering center in every group, and minimizes the value function nonsimilarity. To assess the degree of belonging to its group, FCM employs fuzzy partitions to create any value of input data between 0 and 1. Each bat is assigned with the FCM objective function, and these bats are assessed using FCM cost function, and the best fitness of bat being the optimal solution. The fitness function is used as part of the optimizer to evaluate each bat position. To achieve BSCFC's stability and high accuracy, the fitness function is minimized through optimization during the iteration process and it is defined as,

$$Fitness = \text{maxmize}(\text{Accuracy} + Fw_f \times (1 - \frac{L_f}{L_t}) \qquad (16)$$

Where accuracy is the percentage of correctly classified instances divided by the total of instances, L_f denotes the length of the selected feature subset, L_t is the total number of features, and FW_f is the weighted factor which has value in (0, 1) and it is set as 0.8. Here, the problem's dimension depends on the number of attributes in each data set. After the position and fitness are evaluated and updated each time in an iteration that lasts until the criteria are met to achieve the global best fitness function as the global solution to the problem, the good solutions are stored in bat memory.

Results

Dataset description

The proposed BSCFC algorithm was implemented in MATLAB 2019a using an HP Elite 7100 Business PC. Breast cancer mini-Mammographic Image Analysis Society database (MIAS) dataset is used for analysis.[32] Further, median filters are used for preprocessing, region of interest was utilized for segmentation, gray level co-occurrence matrix (GLCM), and texture analysis are utilized in the feature extraction process where the synthesis of all these methods yields a feature vector with 13 values.[33-35] Breast cancer image dataset of a mini-MIAS database with 322 images are used for the breast cancer analysis. In this, 70% of the dataset are used for training and 30% of the datasets are used for testing. The data are partitioned into 70% of training and 30% testing sets using the k-fold cross-validation method, where k = 10 indicates a tenfold data partition and further, K nearest neighbor (KNN) classifier is used for simplifying the data sets based on Euclidean distance measure.[36]

Evaluation criteria

The performance of the proposed BSCFC algorithm is analyzed based on the following evaluation criteria:

a. Standard deviation fitness (SDF) – It depicts the variety of optimal solutions obtained by running stochastic optimizer n times. This measure is used as an indicator of the optimizer's strength and healthiness.[32] It is defined mathematically as,

$$SDF = \sqrt{\frac{1}{n-1}\sum (k_j - Mean)^2} \qquad (17)$$

where Mean is used as the average of solutions.

b. Best Fitness Value (BFTV) – It is the minimum fitness score, which is calculated as

$$BFTV = \frac{k}{\min i = 1} Fitness \qquad (18)$$

c. Worst fitness value (WFV) – It is the maximum fitness score, which is calculated as,

$$WFTV = \frac{k}{\max i = 1} Fitness \qquad (19)$$

d. Accuracy – This accuracy is measured against a subset of features from a given data set using the KNN classifier

$$Average\ Accuracy = \frac{1}{k}\sum_{i=1}^{k} \frac{TP + TN}{TP + FN + FP + TN} \qquad (20)$$

where TP is truly positive, TN is a true negative, FN is a false negative, FP is a false
positive, and k is the number of folds.

e. Sensitivity – Determines the percentage of positives that are correctly identified, which is computed as,

$$Sensitivity = \frac{TP}{TP + FN} \qquad (21)$$

f. Specificity – Determines the percentage of negatives that are correctly identified, which is calculated as,

$$Specificity = \frac{TN}{TN + FP} \qquad (22)$$

Experimental *results*

The training results of five different chaotic maps are all cross-validated by a 10-fold validation for verifying the performance of benign and malignant classification detection in this work. The proposed BSCFC algorithm was compared with the bat, chaotic bat algorithm (CBA), chaotic crow search algorithm (CCS), ant lion optimization (ALO), and chaotic ant lion optimization algorithm (CALO). The performance of the BSCFC algorithm using best, worst, mean, standard derivation of fitness values and accuracy of KNN classifier criteria for the breast cancer data sets can be shown in Table 1.

The proposed BSCFC method only chooses the best six important texture features of contrast, correlation, energy, entropy, homogeneity, and variance from 13 features of

GLCM. These six optimized features are provided as an input to the KNN classifier for benign and malignant classification.

It is observed from the above Table 1 that the proposed BSCFCS method embedded with a chaotic sinusoidal map obtained the higher best fitness value (which is approximately 0.71) for sinusoidal than for gauss, sine, tent, and logistic maps. In terms of worst fitness values, the BSCFCS with gauss, and sinusoidal generates the minimum score fitness values of 0.09 and 0.11 respectively. Moreover, the mean fitness values and standard deviation are also shown in Table 1. Finally, the value of KNN accuracy is the highest for sinusoidal BSCFC algorithm at 0.98 followed by the sine map at 0.96. It is clearly shown that the chaotic map, especially sinusoidal can significantly boost the performance of the BSCFC algorithm.

Table 2 shows the comparative analysis based on the performance of the selected features by different optimizers on the breast cancer data sets in terms of classification accuracy, sensitivity, and specificity. The proposed BSCFCS algorithm is compared with optimization algorithms of a bat, CBA, CCS, ALO, CALO, and the visualization results are shown in Figure 2. In this, the proposed BSCFC achieved the highest accuracy rate of (98.2%), specificity of (97.3%), and sensitivity of (98.3%) as compared to other feature selection optimization algorithms.

The proposed BSCFCS algorithm shows high stability and convergence speed due to less use of parameters and utilization of FCM objective function decreases along with the iterations and escape getting stuck in local minima. This is due to the updated random variables in the bat optimization algorithm by the sequence of chaos vectors. The overall performance of the BSCFCS algorithm is reasonably better than the compared bat, CBA, CCS, ALO, and CALO algorithms

Table 1: Fitness values of bat optimization algorithm combined with chaotic maps and fuzzy c-means clustering algorithm with the bat with an accuracy of K-nearest neighbor classifier

Chaotic maps	Breast cancer data set				Accuracy of KNN
	Best fitness	Worst fitness	Mean	SD	
Sinusoidal	0.71	0.11	0.68	0.01	0.98
Sine	0.65	0.18	0.52	0	0.96
Gauss	0.52	0.09	0.59	0	0.94
Logistic	0.29	0.26	0.29	0.02	0.82
Tent	0.39	0.22	0.45	0	0.76

SD - Standard deviation, KNN - K-nearest neighbor

Table 2: The accuracy of different optimization algorithms

Optimization algorithms	Accuracy (%)	Specificity (%)	Sensitivity (%)
Proposed BSCFC	98.2	97.3	98.3
Bat	86.5	87.4	86.2
CBA	94.1	93.4	94
CCS	92.3	91.2	92
ALO	93.4	92.5	92.7
CALO	79.8	80.2	80

BSCFC - Bat optimization algorithm combined with chaotic maps and fuzzy C-means clustering algorithm, CBA - Chaotic bat algorithm, CCS - Chaotic crow search algorithm, ALO - Ant lion optimization, and CALO - Chaotic ant lion optimization algorithm

Conclusion and Future Work

In this article, a hybrid feature selection BSCFCS method is used, which combines a bat optimization algorithm with chaos theory and an FCM algorithm. The proposed BSCFCS algorithms use the global optimization approach to avoid local minima trapping and chaos theory to overcome the lack of convergence of the bat algorithm while transferring random variables from Gaussian to chaotic behavior. In addition, the FCM is utilized to deal with uncertainty and assess bats in high-dimensional space. The proposed BSCFC algorithm is designed to improve the performance of convergence speed and control balance between exploration and exploitation rate using five types of chaotic maps, namely sinusoidal, sine, gauss, logistic, and tent maps. The BSCFC is tested on one of the challenges of breast cancer diagnosis feature selection. Performance is evaluated on a mini-MIAS dataset of 322 images, which are evaluated using standard measures. The experimental results have shown that BSCFC outperforms as compared to other optimization algorithms of the bat, CBA, CCS, ALO, and CALO in terms of best, mean fitness value, and standard deviation. Moreover, the results show that the BSCFC with sinusoidal maps can significantly boost the classification performance of the BSCFC algorithm in classifying the breast cancer images with reduced features,

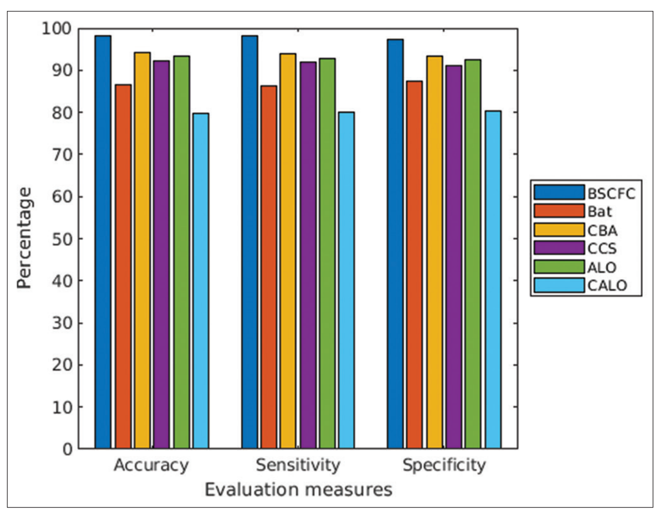

Figure 2: BSCFC – Bat Optimization algorithm combined with Chaotic maps and Fuzzy C-means clustering algorithm, CBA – Chaotic bat algorithm, CCS – Chaotic crow search algorithm, ALO – Ant lion optimization algorithm, and CALO – Chaotic ant lion optimization algorithms

which in turn optimizes the radiologists' time for their interpretation. Further, we develop the proposed algorithm as multiple objective approaches to resolve the problem of optimization and expand the system's efficiency for other forms of cancers such as cervical, lung, skin, and brain.

Ethical approval statement

This article does not contain any studies with human participants or animals performed by any of the authors.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Chugh G, Kumar S, Singh N. Survey on Machine Learning and Deep Learning Applications in Breast Cancer Diagnosis. Cognitive Computation. Springer, Switzerland. 2021:13:1-20.
2. Dibden A, Offman J, Duffy SW, Gabe R. worldwide review and meta-analysis of cohort studies measuring the effect of mammography screening programmes on incidence-based breast cancer mortality. Cancers (Basel) 2020;12:E976.
3. Britt KL, Cuzick J, Phillips KA. Key steps for effective breast cancer prevention. Nat Rev Cancer 2020;20:417-36.
4. Salod Z, Singh Y. A five-year (2015 to 2019) analysis of studies focused on breast cancer prediction using machine learning: A systematic review and bibliometric analysis. J Public Health Res 2020;9:1792.
5. Raj RJ, Shobana SJ, Pustokhina IV, Pustokhin DA, Gupta D, Shankar K. Optimal feature selection-based medical image classification using deep learning model in internet of medical things. IEEE Access 2020;8:58006-17.
6. Agrawal P, Abutarboush HF, Ganesh T, Mohamed AW. Metaheuristic algorithms on feature selection: A survey of one decade of research (2009-2019). IEEE Access 2021 2;9:26766-91.
7. Zanetos A, Dounias G. A Comprehensive Survey on the Applications of Swarm Intelligence and Bio-Inspired Evolutionary Strategies. In: Tsihrintzis G., Jain L. (eds) Machine Learning Paradigms. Learning and Analytics in Intelligent Systems, Springer, Cham. 2020:18.
8. Qasim OS, Algamal ZY. Feature selection using different transfer functions for binary bat algorithm. Int J Math Eng Manag Sci 2020;5:697-706.
9. Reenadevi R, Sathiya T, Sathiyabhama B. Classification of Digital Mammogram Images using Wrapper based Chaotic Crow Search Optimization Algorithm. Annals of the Romanian Society for Cell Biology. 2021:25:2970-9.
10. Bulut H, Onan A, Korukoğlu S. An improved ant-based algorithm based on heaps merging and fuzzy c-means for clustering cancer gene expression data. Sādhanā 2020;45:1-7.
11. Senapati T, Yager RR. Fermatean fuzzy sets. J Ambient Intell Humaniz Comput 2020;11:663-74.
12. Tawhid MA, Dsouza KB. Hybrid binary bat enhanced particle swarm optimization algorithm for solving feature selection problems. Applied Computing and Informatics. 2018:16:117-36.
13. Al-Betar MA, Alomari OA, Abu-Romman SM. A TRIZ-inspired bat algorithm for gene selection in cancer classification. Genomics 2020;112:114-26.
14. Yu H, Zhao N, Wang P, Chen H, Li C. Chaos-enhanced synchronized bat optimizer. Appl Math Model 2020;77:1201-15.
15. Agrawal U, Arora J, Singh R, Gupta D, Khanna A, Khamparia A. Hybrid wolf-bat algorithm for optimization of connection weights in multi-layer perceptron. ACM Trans Multimed Comput Commun Appl (TOMM) 2020;16:1-20.
16. Tripathi D, Edla DR, Kuppili V, Dharavath R. Binary BAT algorithm and RBFN based hybrid credit scoring model. Multimed Tools Appl 2020;79:31889-912.
17. Eskandari S, Javidi MM. A novel hybrid bat algorithm with a fast clustering-based hybridization. Evol Intell 2020;13:427-42.
18. Anter AM, Ali M. Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems. Soft Comput 2020;24:1565-84.
19. Naik AK, Kuppili V, Edla DR. Efficient feature selection using one-pass generalized classifier neural network and binary bat algorithm with a novel fitness function. Soft Comput 2020;24:4575-87.
20. Zhang G, Peng Z, Li X, Yan C, Wang J. TABBA: A Novel Feature Selection Method Based on Binary Bat Algorithm and T-Test. In2021 IEEE 6th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA) IEEE. Chengdu, China 2021. P. 157-160.
21. Sathiyabhama B, Kumar SU, Jayanthi J, Sathiya T, Ilavarasi AK, Yuvarajan V, *et al*. A novel feature selection framework based on grey wolf optimizer for mammogram image analysis. Neural Computing and Applications. 2021:33:1-20.
22. Rajeswari C, Sathiyabhama B, Devendiran S, Manivannan K. A gear fault identification using wavelet transform, rough set based GA, ANN and C4. 5 algorithm. Procedia Eng 2014;97:1831-41.
23. Rajeswari C, Sathiyabhama B, Devendiran S, Manivannan K. Bearing fault diagnosis using wavelet packet transform, hybrid PSO and support vector machine. Procedia Eng 2014;97:1772-83.
24. Yang XS, Slowik A. Bat algorithm. In Swarm Intelligence Algorithms (1st ed.). CRC Press. Taylor & Francis Group. 2020. P. 43-53. https://doi.org/10.1201/9780429422614.
25. El-Hasnony IM, Barakat SI, Elhoseny M, Mostafa RR. Improved feature selection model for big data analytics. IEEE Access 2020;8:66989-7004.
26. Tahir M, Tubaishat A, Al-Obeidat F, Shah B, Halim Z, Waqas M. A novel binary chaotic genetic algorithm for feature selection and its utility in affective computing and healthcare. Neural Computing and Applications. 2020:32:1-22.
27. Demir FB, Tuncer T, Kocamaz AF. A chaotic optimization method based on logistic-sine map for numerical function optimization. Neural Comput Appl 2020;32:14227-39.
28. Koyuncu H. GM-CPSO: A new viewpoint to chaotic particle swarm optimization via Gauss map. Neural Process Lett 2020;52:241-66.
29. Shah AA, Parah SA, Rashid M, Elhoseny M. Efficient image encryption scheme based on generalized logistic map for real time image processing. J Real-Time Image Process 2020;17:2139-51.
30. Arora S, Sharma M, Anand P. A novel chaotic interior search algorithm for global optimization and feature selection. Applied Artif Intell 2020;34:292-328.
31. Chowdhary CL, Mittal M, P K, Pattanaik PA, Marszalek Z. An efficient segmentation and classification system in medical images using intuitionist possibilistic fuzzy C-mean clustering and fuzzy SVM algorithm. Sensors (Basel) 2020;20:E3903.
32. Suckling J, Parker J, Dance D, Astley S, Hutt I, Boggis C, *et al*. (2015). Mammographic Image Analysis Society (MIAS) database v1.21 [Dataset]. https://www.repository.cam.ac.uk/handle/1810/250394.
33. Shah A, Bangash JI, Khan AW, Ahmed I, Khan A, *et al*. Comparative analysis of median filter and its variants for removal of impulse noise from gray scale images. Journal of King Saud University-Computer and Information Sciences. 2020. https://doi.org/10.1016/j.jksuci.2020.03.007.
34. Yengec Tasdemir SB, Tasdemir K, Aydin Z. A review of mammographic region of interest classification. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2020;10:e1357.
35. Hazra R, Banerjee M, Badia L. Machine Learning for Breast Cancer Classification With ANN and Decision Tree. In 2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). IEEE. Vancouver, BC, Canada. 2020. P. 0522-0527.
36. Assegie TA. An optimized K-nearest neighbor based breast cancer detection. J Robot Control 2021;2:115-8.