Journal of Classification 35:300-344 (2018)
DOI: 10.1007/s00357-018-9261-2

A New Chaotic Whale Optimization Algorithm for Features Selection

Gehad Ismail Sayed

Faculty of Computers and Information, Cairo University, Egypt

Ashraf Darwish

Faculty of Science, Helwan University, Egypt

Aboul Ella Hassanien

Faculty of Computers and Information, Cairo University, Egypt

Abstract: The whale optimization algorithm (WOA) is a novel evolutionary algorithm inspired by the behavior of whales. Similar to other evolutionary algorithms, entrapment in local optima and slow convergence speed are two probable problems it encounters in solving challenging real applications. This paper presents a novel chaotic whale optimization algorithm (CWOA) to overcome these problems where chaotic search is embedded in the searching iterations of WOA. Ten chaotic maps are considered to improve the performance of WOA. Experiments on ten benchmark datasets show the novel CWOA is effective for selecting relevant features with a high classification performance and a small number of features. Additionally the performance of CWOA is compared with WOA and ten other optimization algorithms. The experimental results show that circle chaotic map is the best chaotic map to significantly boost the performance of WOA. Moreover, chaotic with modifications of exploration operators outperform the highest performance.

Keywords: Whale optimization algorithm; Chaotic maps; Features selection.

Corresponding Author's Address: G. Sayed, Faculty of Computers and Information, Cairo University, 5 Ahmed Zewail, Ad Doqi, Giza Governorate 12613, Egypt, email: gehad.ismail@egyptscience.net.

Published online: 4 August 2018

1. Introduction

In the last decades, evolutionary computations, which are a subfield of computational intelligence that involves optimization problems, have received much attention for different applications. Evolutionary algorithms have been inspired by creating a relation between the power of natural evolutionary mechanisms and the nature of the solving problem. In literature, there are many different variants of evolutionary algorithms such as genetic algorithms (GA) (Holland, 1992) and particle swarm optimization (PSO) algorithms (Sheikhpour, Sarrama, and Sheikhpour, 2016). The common idea behind all these algorithms is that, given a population of objects or individuals, the environmental pressures, causes natural selection, which causes a rise in the fitness of the population. Evolutionary algorithms implement their structures in different ways according to population intialization, candidate's evaluation, termination condition, selection method, recombination and mutation. The main blocks of the evolutionary algorithms family are population initialization method, termination criteria, fitness function, mutation method, selection method and crossover/recombination method. An evolutionary algorithm may implement these blocks in different ways (Özkaynak, 2015). Regardless of their different structures, these algorithms usually create a random population and evolve it over a predefined number of generations. Selection, reproduction, mutation, and recombination operators perform the evolutionary process. Evolutionary algorithms divide the search process into two phases: exploration and exploitation. For the optimization parameters of each evolutionary algorithm, each one has its own defect such as low classification accuracy, weak generalization ability, slow convergence speed, and so on. Moreover, evolutionary algorithms require some randomness to proceed. To overcome all these problems, many approaches in the literature have been used to improve the performance of evolutionary algorithms. One of the most common mathematical methods that recently has been applied to improve both exploration and exploitation is chaos theory. Chaos theory can study the behavior of systems that are highly sensitive to their initial conditions and can generate a more variable range of numbers instead of random numbers. The behavior of chaotic systems appears to be random; therefore, chaotic systems can be used for the needed randomness by evolutionary algorithms. Chaos theory has provided effectiveness in different fields of sciences, such as chaos control (Strogatz, 1994), synchronization (Sprott, 2010), optimization research (Abdullah, Enayatifa, and Lee, 2012; Emary, Zawbaa, and Hassanien, 2016) and so on. In recent years, the applications of chaos in various disciplines, including optimization research problems, have attracted more attention. Based on chaos theory, the chaotic optimization algorithm (COA) (Wang, Liu, and Liu, 2001) utilizes the nature of chaotic sequence, including the quasi-stochastic property and ergodicity. Currently, chaotic systems are an active area of research in the last few years and have been applied in different sciences research and engineering fields (Hosseinpourfard and Javidi, 2015; Wang et al., 2014).

Feature selection methods provide a way for identifying the important features and removing irrelevant (redundant) ones from the datasets. The main purposes are data dimensionality reduction and improving prediction performance. In real world applications, data representation usually uses many features with redundancy of features. This means that important features may be removed, while unnecessary features remain. Moreover, the relevant features have an influence on the output and contain important information that will be obscured if any of them are removed. There are various heuristic techniques that mimic the behavior of biological and physical systems in the nature. These techniques have been proposed as methods for global optimizations of the problems. A new meta-heuristic optimization algorithm called WOA (Hosseinpourfard and Javidi, 2015) which mimic the hunting behavior of humpback whales is applied for features selection of different datasets. Optimization results demonstrated that WOA is a very competitive and promising algorithm compared to the state-of-the-art optimization algorithms (Mirjalili and Lewis, 2016). However, a main problem that still exists when applying WOA for feature selection is determining how to choose the optimal feature subset of the target dataset. In addition, classification problems usually involve a number of features and not all these features are equally important for classification. Some features may be redundant or even irrelevant, and if they are not eliminated from classification process, it could result in increasing computational time complexity or decreasing classification accuracy. In order to achieve better performance, feature selection or reduction is a necessary step for a complex dataset in classification. Feature selection is proposed in Ebrahimi and Khamehehi (2016) to be an NP-hard combinatorial problem and requires efficient solution algorithms. In this paper, we propose a novel optimization algorithm for feature selection based on chaos theory and WOA and is called chaotic whale optimization algorithm (CWOA). Therefore, WOA is employed with ten chaotic maps for achieving an improved performance of feature selection from ten different datasets. The proposed new algorithm chaotic whale optimization algorithm (CWOA) showed better performance over the classical algorithms such as WOA for feature selection. Chaotic maps adapted WOA work better than the traditional WOA in terms of the quality of the solution and obtaining faster convergence. In order to evaluate the proposed new algorithm, some benchmark functions are utilized. The experimental results showed that the proposed algorithm in this paper improved the performance of evolutionary algorithms for feature selection in various datasets.

The outline of the paper is as follows. In Section 2, we demonstrate the problems of chaos based evolutionary algorithms in the literature. Section 3 introduces the basics of WOA and chaotic maps for feature selection. In Section 4, we describe the proposed new hybrid optimization algorithm in details. Section 5 presents experimental results and analysis. Finally, we introduce concluding remarks and future work in Section 6.

2. Related Work

There are many studies in the literature that focus on improving the performance of evolutionary algorithms. Based on the analysis and results of these studies, we found that chaos theory played an effective role in improving the performance of evolutionary algorithms. As mentioned in the previous section, the problem of evolutionary algorithms, such as GA and PSO, is their premature convergence; where the used algorithm perhaps get stuck in the local optimum (minimum or maximum). One of the prominent algorithms based on hybridizing chaotic optimization algorithm (COA) and particle swarm optimization (PSO) algorithm is introduced in Gadat and Younes (2007) to improve the classification accuracy of data. This algorithm (CPSO) solved the time-consuming problem and introduced a small sample learning ability of LS-SVM. The simulation results showed that the proposed CPSO algorithm could effectively optimize the parameters of LS-SVM model. GA is one of the important algorithms in evolutionary algorithms, and it has been applied in many applications in the literature. To improve the performance of this algorithm, Liu and Zhou (2015) presented a new system to prove that the chaos algorithm has the capability to improve the performance of GA.

Chaos theory describes erratic behavior in nonlinear systems and for this purpose, it uses chaotic maps. Chaotic maps are visualized and can travel as particles in a limited range of nonlinear, dynamic, and nonlinear systems with no definite regularity-traveling path of these particles. Chaotic maps have been employed in Yang and Chen (2002) to manipulate the mutation probability to increase the exploitation of GA. Therefore, a new hybrid method was proposed in Abdullah et al. (2012) to combine GA and chaotic theory for image encryption. Chaotic logistic maps are used in this method for the initial image encryption and hence GA is applied to improve the efficiency of the encryption process of the image. An improved logistic map, namely a double bottom map, was used in Santos et al. (2012) with a PSO algorithm for production optimization problems. Eight chaotic maps for parameter adaptation are used in Alatas, Akin, and Ozer (2000). Their experimental results showed that the proposed algorithm improved the solution quality and sometimes improved the global search capacity. Biogeography-based optimization (BBO) is an optimization algorithm that has been applied in the literature in some application such as ecosystems. Chaotic maps are used in Saremi, Mirjalili, and Lewis (2014a) to improve the performance of BBO. Ten chaotic maps are integrated with the BBO algorithm with the objective to improve the exploration and the exploitation of the BBO algorithm. The Firefly Algorithm (FA) is a new bio-inspired algorithm, which simulates the behavior of fireflies. To solve the problems of low accuracy and local convergence in FA, chaos theory is introduced into the evolutionary process of FA. The comparison of firefly algorithm with chaotic maps is performed in Shoubao, Yu, and Mingjuan (2014), showing that convergence quality and accuracy are improved, with FA and chaos. All these hybrid solutions are used to indicate the great impact of embedding chaotic maps. As, evolutionary algorithms have sensitive dependence on their initial condition and parameters, improving these parameters can have a substantial effect. Chaotic systems are widely used to express the optimization parameters and in the creation of the initial population. Thus, replacing random sequences with chaotic sequences during the evolutionary process can improve the performance of evolutionary algorithms.

Feature selection is an important technique that can be extremely useful in reducing the dimensional data to be processed by the classifier, reducing the execution time and enhancing the recognition rate of the classifier. Several researchers have addressed the feature selection problem as an optimization problem where the fitness function is the accuracy of the given classifier that may be maximized by the selected features. Nature inspired meta-heuristic algorithms are now among the most widely used algorithms for solving optimization problems. To overcome the challenges of parameter optimization and feature selection in the classification process, a feature selection and parameter optimization approach based on GA has been proposed (Huang and Wang, 2006). Recently, chaos embedded methods have been applied in parameter optimization of SVM for feature selection (Li et al., 2012a,b; Liu Wang, and Jin, 2005). A new PSO method that uses chaotic mappings for parameter adaptation of wavelet v-support vector machine is proposed in Wu (2015). In Saremi et al. (2014b), a new chaotic differential evolution optimization approach based on the Ikeda map was proposed to optimize kernel function parameters of SVM. In this approach, a chaotic sequence has also been used in the feature selection process. The analysis of the experimental results of the studies in the literature show that chaotic sequences have an effective potential for optimization of input feature subsets. Although SVM classification performance has been improved significantly in the last years, there are still some outstanding problems. Therefore, a hybridized chaotic search and gravitational search algorithm (GSA) with SVM and a new chaos embedded GSA-SVM (CGSASVM) hybrid system are presented in Chaoshun, Xueli, and Ruhai (2015). In addition, two kinds of chaotic maps, namely logistic maps and tent maps were embedded in PSO to handle feature selection problems Chuang, Yang, and Li (2011). To the best of our knowledge, there is no in literature CWOA is used for feature selection and improving the performance of evolutionary algorithms. This paper proposes a new hybrid optimization algorithm CWOA based on chaos theory and whale bio-inspired algorithm.

3. Basics and Background

3.1 Whale Optimization Algorithm

A. Inspiration Analysis

Whales are extravagant animals and they are considered as the biggest mammals among all animals. A grown-up whale can grow up to 30 meters long and weigh 180 tons. There are several types of whale such as killer, humpback, finback, and blue. Whales never sleep because they need to breathe from the surface of seas and oceans. Moreover, only half of the brain can sleep (Rattenborg, Amlaner, and Lima, 2000). Unlike fish, whales do not have gills for extracting oxygen from the water, so they must come to the surface to get the oxygen or they would drown. Thus, when whales rest they remain partially conscious to obtain the necessary air and to react to the danger. The body of whales is designed to allow them to hold their breath for extended periods while minimizing the amount of energy they use when swimming. Whales live alone or in groups. Some types (such as killer whales) can live in a family all their life period. Humpback whales are one of the biggest whales and their favorite prey is krill and small fish species.

As indicated by (Hof and Van, 2007), whales have basic cells in specific regions of their brains like those of humans, which are called spindle cells. These cells are in charge of judgment, feelings and emotions, and the behavior of humans. Whales have twice number of these cells than an adult human. Whales can think, learn, judge, communicate, and become emotional just as a human does, but with a lower level of smartness than humans. Whales are able to develop their own dialect as well. The special hunting technique of humpback whales is considered their primary point of interest. This method is called bubble-net feeding method as described by Watkins and Schevill (1979), where they create distinctive bubbles along a circle or '9' behavior-shaped path as shown in Figure 1.

By the end of 2011, this behavior was discussed and surveyed according to observations from the surface of the sea or ocean. Moreover, authors in Chuang et al. (2011) investigated the behavior of these whales by using tags sensors. In this work, about 300 tags-derived bubble-net feeding events of 9 individual humpback whales have been captured. In addition,

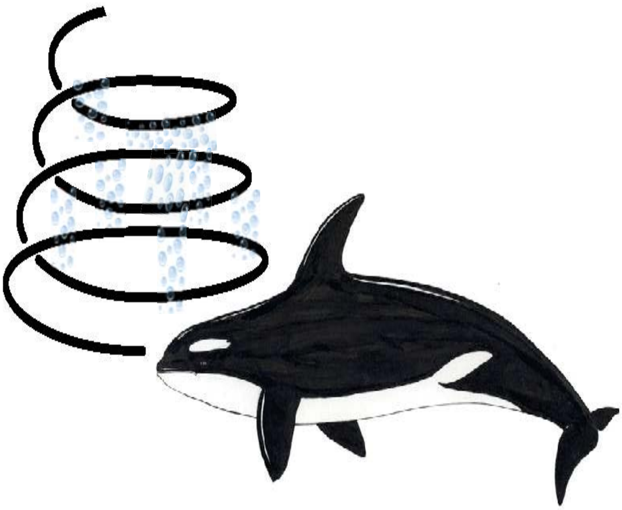

Figure 1. Bubble-net feeding behavior of humpback whales

the authors found that two maneuvers associated with bubble and named them upward-spirals and double- loops. Humpback whales dive around $12m$ down and then start to create bubbles in a spiral shape around the prey and swim up toward the surface of the sea or ocean. The latter maneuver includes three different stages: coral loop, lob tail, and capture loop. Detailed information about these behaviors of, humpback whales and others are discussed in more detail in Goldbogen et al. (2013).

B. Mathematical Model of WOA

This part highlights the mathematical model, including encircling prey, spiral bubble-net feeding maneuver, and search for prey (Hosseinpourfard and Javidi, 2015).

B1. Encircling prey

Humpback whales can perceive the area of prey and enclose it. The position of the optimal design in the hunt or search space is not known from earlier positions, the WOA optimization algorithm supposes that the present best candidate solution is the objective prey or is near to the optimum. In this case, the humpback whales have defined the best search agent; the other search agents then will try to change their positions towards the best agent of search. This behavior can be described by the following equations:

$$D = |C \cdot \vec{X}^*(t) - C\vec{X}_i(t)| \,, \tag{1}$$

$$\vec{X}_i(t+1) = \vec{X}^*(t) - A \cdot D \,, \tag{2}$$

where t indicates the current iteration, A and C are coefficient numbers. $\cdot$ indicates an element-by-element multiplication. $\vec{X}$ is the position matrix of the ith whale with size number of search agents (population size) $\times$ number of dimensions. $\vec{X}^*$ is the position vector of the optimal solution (best search agent position) with size 1 $\times$ number of dimensions which can obtained so far, || is defined as the absolute value. It is important to notice that $\vec{X}^*$ should be updated for each iteration if there exists a better solution. The coefficients A and C can be mathematically formulated according to the following equations:

$$A = 2a \cdot r - a\,, \tag{3}$$

$$C = 2 \cdot r\,, \tag{4}$$

where a is decreased linearly from 2 to 0 over the course of iterations. It is defined in equation (5), where t is the iteration number and Max_{iter} is the maximum number of iterations. r is a random number in [0,1]. The humpback whales can attack the prey with the bubble-net method.

$$a = 2 - t \times \frac{2}{Max_{iter}}\,. \tag{5}$$

B2. Bubble-Net Attacking Method (Exploitation Phase)

Two approaches are designed here to model mathematically the bubble-net behavior of humpback whales:

Shrinking encircling mechanism: This method is achieved by decreasing the value of a in equation 12. A is a random value in the interval where a is decreased from 2 to 0 over the course of iterations. By setting random values for A in [-1, 1], the new position of a search agent can be defined anywhere in between the original position of the agent and the position of the current best agent.

Spiral-updating position : A spiral equation is then created between the position of whale and prey to mimic the helix-shaped movement of humpback whales as follows:

$$\vec{X}_i(t+1) = D \cdot e^{bl} cos(2\Pi l) + \vec{X}^*(t)\,, \tag{6}$$

where $D = |\vec{X}^*(t) - \vec{X}_i(t)|$ indicates the distance of the ith whale to the prey (best solution obtained so far). b is the logarithmic spiral shape constant, l is a random number in [-1, 1], and is an element-by-element multiplication.

During the optimization phase, where humpback whales swim around the prey in a shrinking circle and along a spiral-shaped path simultaneously assuming that the probability of 50 percent to choose between either the shrinking encircling mechanism or the spiral model to update the position of whales. Therefore, the mathematical model of this behavior can be expressed as follows:

$$\vec{X}_i(t+1) = \begin{cases} \vec{X}^*(t) - A \cdot D & \text{if } p < 0.5 \\ D \cdot e^{bl} \cdot cos(2\Pi l) + \vec{X}^*(t) & \text{if } p \geq 0.5\,, \end{cases} \tag{7}$$

where p is a random number in [0,1]. In addition to the bubble-net method, the humpback whales search for prey randomly. The mathematical model of the search is as follows.

B3. Search for Prey (Exploration Phase) :

The same approach based on the variation of the A parameter can be applied to search for prey. Humpback whales search randomly according to the position of each other. We use A with the random values greater than 1 or less than 1 to force search agents to move far away from a reference whale. In contrast to the exploitation phase, we update the position of a search agent in the exploration phase, according to a randomly chosen search agent instead of the best search agent found so far. This mechanism and $|A| > 1$ emphasize exploration and can allow the WOA algorithm to perform a global search. The mathematical model can be described is as follows:

$$D = C \cdot \vec{X_{rand}} - \vec{X}\,, \tag{8}$$

$$\vec{X}_i(t+1) = \vec{X_{rand}} - A \cdot D\,, \tag{9}$$

where $\vec{X_{rand}}$ is a random position vector with size 1 multiplied by the number of dimensions, which has been chosen from the current population. The WOA algorithm starts with a set of random solutions. At each iteration, search agents update their positions with respect to either a randomly chosen search agent or the best solution obtained so far. The a parameter is decreased from 2 to 0 in order to provide exploration and exploitation, respectively. A random search agent is chosen when $|A| > 1$ while the best solution is selected when $|A| < 1$ for updating the position of the search agents. Depending on the value of p, WOA is able to switch between either a spiral or circular movement. Finally, the WOA algorithm is terminated by the satisfaction of a termination criterion.

3.2 Chaotic Maps for Feature Selection

Chaos is unstable dynamic behavior that provides sensitive dependence on the initial conditions and includes infinite unstable periodic motions in nonlinear systems. Chaos has a fine internal structure and three important dynamic characteristics; ergodicity, the sensitive dependence on initial conditions, and the quasi-stochastic property. Ergodic property is the most important one and can search all nodes or states in the search plane by its formulas within certain range.

Chaos strategy is applied to avoid being trapped in local optima and improve the quality of searching global optimum. Therefore, chaos has been employed in numerous optimization applications. Considering that the feature selection problem is an optimization problem with searching range of [0, 1], chaos can be used to optimize this problem. A chaotic map with n dimensions is a discrete-time dynamical system that can be expressed by the following equation:

$$cu_i^{(k+1)} = f(cu_i^{(k)}), i = 1, 2, 3, ..., n\,. \tag{10}$$

By defining the initial state of $cu_i^{(0)}$, a chaotic sequence can be evaluated by running the system function, where chaotic sequences can be defined in the form of $cu_i^{(k)}, k = 0, 1, 2, .$

In this paper, ten chaotic maps (Abdullah et al., 2012) are adopted to represent the selection of features of ten benchmark datasets. These chaotic maps names and its mathematical forms are described in Table 1.

4. Chaos Maps for Whale Optimization Algorithm

In this section, chaotic maps are considered to improve the performance of WOA in terms of avoiding being trapped at the local optima and improving the convergence speed. Ten chaotic maps are used in this paper. These chaotic maps are employed to manipulate the random parameters values of WOA. As the initial values of chaotic maps may have significant effects on the fluctuation pattern, we set the initial point of all chaotic maps to 0.7 while the rest of parameters (e.g. c and d) are initialized as shown in Table 1. These parameters are found to be the best based on trial and error. In WOA, the parameters A, C, P, PI and l are considered the key factors affecting WOA's Convergence behavior. A, C, P and l are affecting bubble-net attacking method (exploitation phase), A and C for shrinking encircling mechanism, while l for the spiral model. The parameter P is the probability to choose between either the spiral model or shrinking encircling mechanism to update the whales' positions during the optimization. Also PI affects the

Table 1. The adapted chaotic maps

No.	Chaotic Map	Mathematical Form	Range/Interval
1	Chebyshev	$u_{i+1} = cos(icos^{-1}(u_i))$	(-1,1)
2	Circle	$u_{i+1} = mod(u_i + d - (\frac{c}{2\pi})sin(2\pi u_i), 1)$, $c = 0.5$ and $d = 0.2$	(0,1)
3	Guass/mouse	$u_{i+1} = \begin{cases} 1, & u_i = 0 \\ \frac{1}{mod(u_i,1)}, & \text{otherwise} \end{cases}$	(0,1)
4	Iterative	$u_{i+1} = sin(\frac{c\pi}{u_i})$, $c = 0.7$	(-1,1)
5	Logistic	$u_{i+1} = cu_i(1 - u_i)$	(0,1)
6	Piecewise	$u_{i+1} = \begin{cases} \frac{u_i}{p}, & 0 \le u_i < p \\ \frac{u_i - p}{0.5 - p}, & p \le u_i < 0.5 \\ \frac{1 - p - u_i}{0.5 - p}, & 0.5 \le u_i < 1 - p \\ \frac{1 - u_i}{p}, & 1 - p \le u_i < 1 \end{cases}$, $P = 0.2$	(0,1)
7	Sine	$u_{i+1} = \frac{c}{4}sin(\pi u_i)$, $c = 4$	(0,1)
8	Singer	$u_{i+1} = \mu(7.86u_i - 23.31u_i^2 + 28.75u_i^3 - 13.302875u_i^4)$, $\mu = 1.07$	(0,1)
9	Sinusoidal	$u_{i+1} = cu_i^2 sin(\pi u_i)$, $c = 2.3$	(0,1)
10	Tent	$u_{i+1} = \begin{cases} \frac{u_i}{0.7}, & u_i < 0.7 \\ \frac{10}{3}(1 - u_i), & u_i \ge 0.7 \end{cases}$	(0,1)

updating position in the exploration phase. We evaluate the performance of each of these parameters along with different chaotic maps singly and in combination of them over WOA. Figure 2 shows the general architecture of the chaotic whale optimization algorithm (CWOA), where the highlighted boxes represent the parts of WOA where chaos maps are applied.

The proposed CWOA is used as a feature selection algorithm that selects the optimal feature subset in a wrapper mode. The CWOA feature selection algorithm starts with randomly initialized search agents with a set of solutions which are feature subsets in our case. Each feature subset has a different combination of features with different size. At each iteration, each search agent updates its position based on a predefined fitness function. Classification performance is used as the fitness function where 5-nearest neighbor is the used classifier. The best feature subset is one which maximizes the classification accuracy and minimize the selected features. The mathematical formula of the fitness function is defined as follows.

$$Fn_t = maximize(Acc + w_f * (1 - \frac{L_f}{L_t}) , \tag{11}$$

where Acc is the classification accuracy calculated as the the number of classified instances divided by total number of instances, w_f is the weighted factor which has value in $[0, 1]$, L_f is the length of selected features subset and L_t is the total number of features. In this paper, a comparative study between different chaotic operators is provided. In addition to, the combination of all of these operators (WOA with chaotic shrinking/spiral/P/PI operators) is provided and benchmarked.

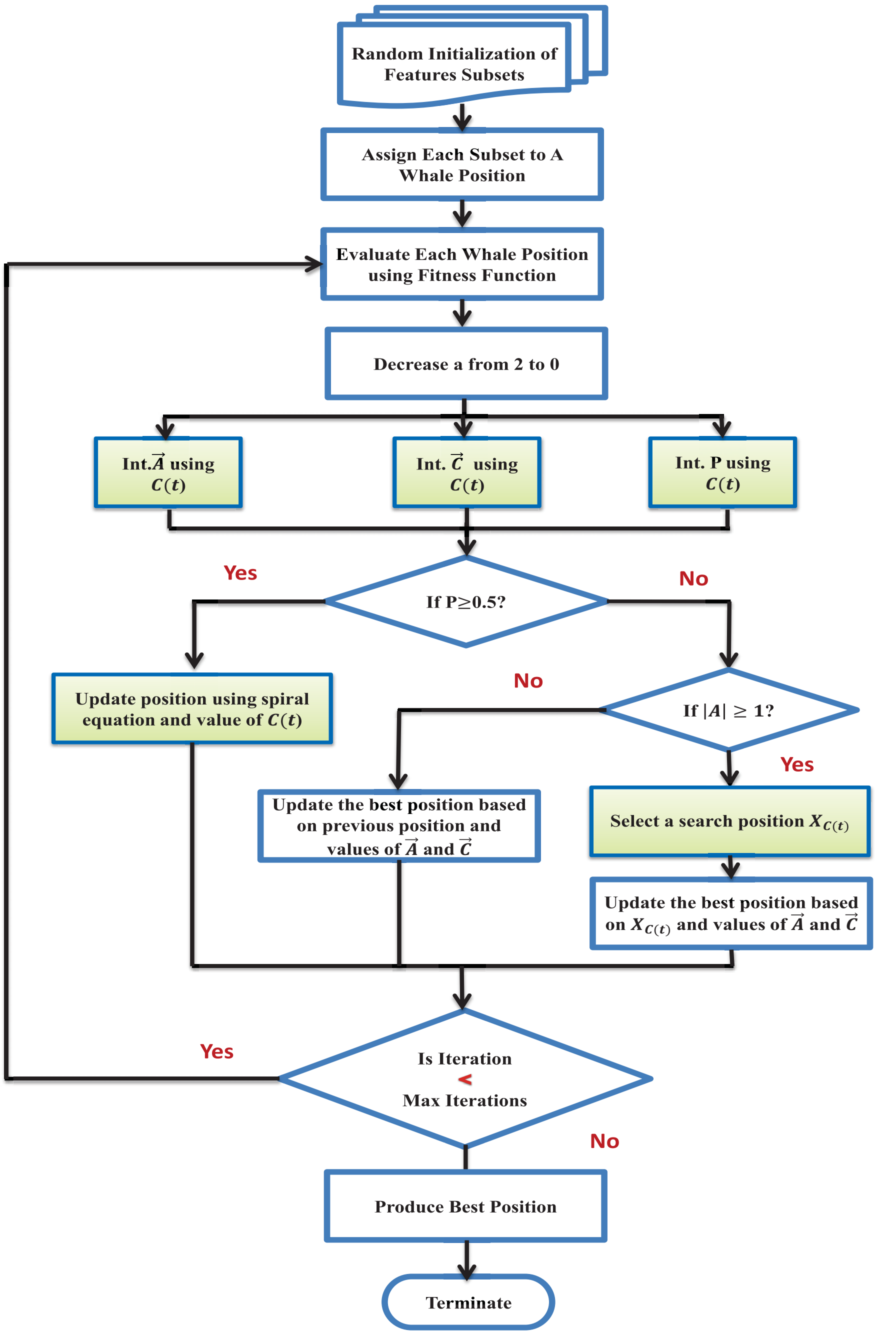

Figure 2. Chaotic whale optimization algorithm architecture (see online version for color)

4.1 Chaotic Maps for Shrinking Circle Mechanism (CWOA-SC)

As mentioned before, A and C are parameters for shrinking the encircling mechanism. In this work, chaotic maps are employed in defining both of them, where the random variable r is substituted by the obtained values from the chaotic map as follows.

$$A = 2a \cdot C(t) - a\,, \tag{12}$$

$$C = 2 \cdot C(t)\,, \tag{13}$$

where $C(t)$ is the obtained value of chaotic map in the $t-th$ iteration.

4.2 Chaotic Maps for Spiral Shaped Mechanism (CWOA-SS)

Parameter l has great influence on spiral updating position of humpbacks whale. Figure 3 shows the behavior of l for 500 iterations. We employed chaotic maps in the spiral equation as follows.

$$\vec{X}_i(t+1) = D \cdot e^{bC(t)} cos(2\Pi C(t)) + \vec{X}^*(t)\,, \tag{14}$$

where $D = |\vec{X}^*(t) - \vec{X}_i(t)|$ indicates the distance of the ith whale to the prey, b is a constant and $C(t)$ is the obtained value of chaotic map in the tth iteration.

4.3 Chaotic Maps for P Parameter (CWOA-P)

In this subsection, the probability of choosing between either the spiral or shrinking mechanism to update the position of whales during optimization is substituted with $C(t)$.

4.4 Chaotic Maps for Choosing Search Agent PI (CWOA-PI)

As $|A| > 1$ has a great impact on updating position of whale during searching for prey (exploration phase), chaotic maps are applied to select the search agent. The mathematical model is defined as follows.

$$D = C \cdot \vec{X_{PI}} - \vec{X}_i \tag{15}$$

$$\vec{X}_i(t+1) = \vec{X_{PI}} - A \cdot D\,, \tag{16}$$

where $\vec{X_{PI}}$ is a chosen position vector from the current population with size 1 multiplied by number of dimensions, $C(t)$ is the chaotic map at the tth iteration.

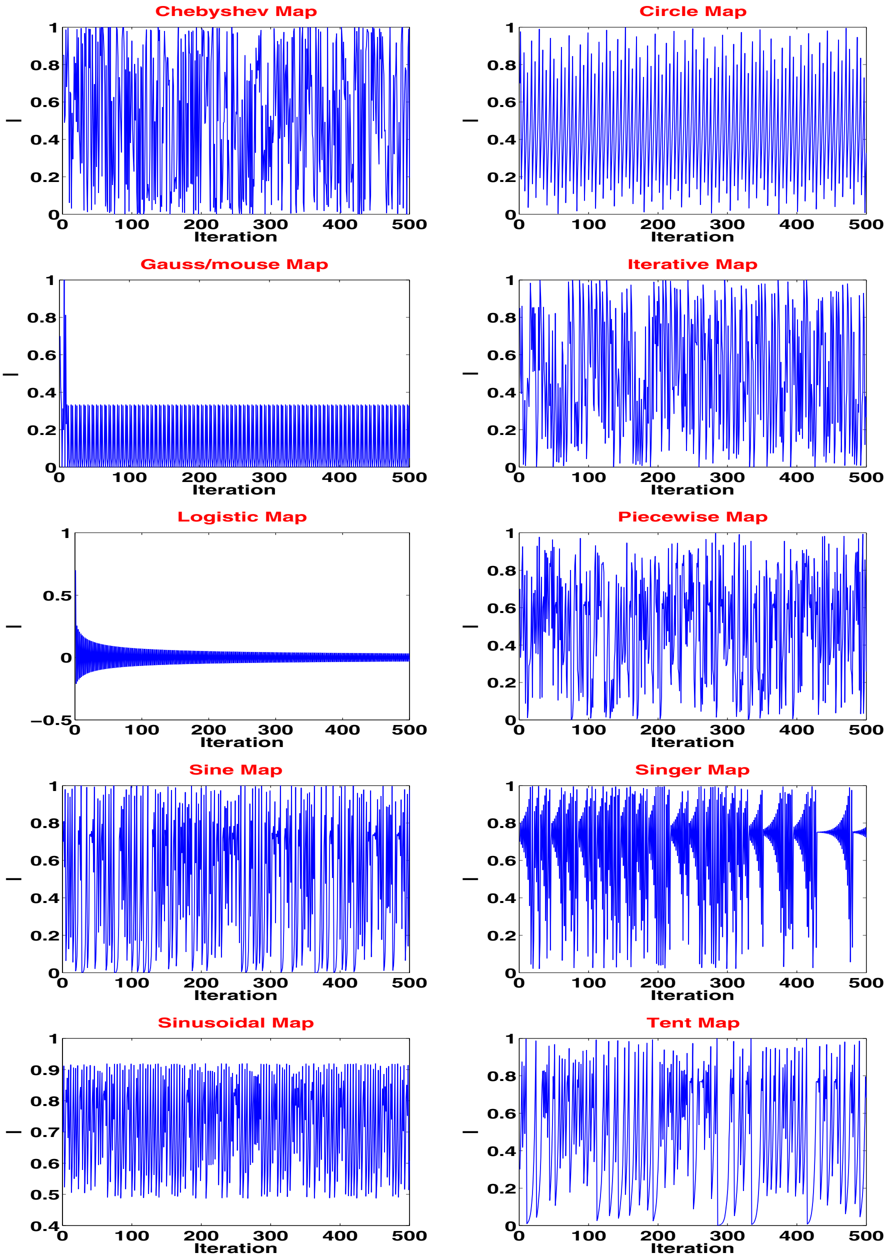

Figure 3. Visualization of Chaotic Maps for l

4.5 Chaotic Maps for Changing All Parameter (CWOA-All)

All the parameters from previous sections are used together in order to improve the performance of WOA. All of these parameters are substituted with the obtained value of chaotic map $C(t)$.

5. Implementation and Results

5.1 Datasets Description

Ten benchmark datasets are used to evaluate the performance of each version of CWOA. The datasets are collected from the UCI dataset repository (Bache and Lichman). A brief description of these datasets are presented in Table 2. Also, we note that some data sets contain missing values (information) in some records. All these missing values for a given feature are replaced by the median value of all known values of this feature in the class. Equation (17) shows the median method for dealing the missing value $x_{i,j}$ for jth feature for a given kth class C.

$$\overline{x}_{i,j} = median_{i:x_{i,j} \in C_k} x_{i,j} \tag{17}$$

5.2 Parameters Initialization and Comparative Setting

The initial parameters settings for WOA versions are presented at Table 3. The performance of different versions of CWOA are compared on ten benchmark datasets with the original WOA and ten other optimization algorithms. The parameter settings for each algorithm used in all experiments are shown in Table 4. All of these algorithms were adopted with their default parameters that proposed in literature, since our main concern is to boost the performance of only whale optimization algorithm.

5.3 Performance Metrics

In this subsection, five different measurements are used to evaluate CWOA algorithms. These measurements are the worst, best, mean fitness value, standard deviation and average features selection size (ASS). They are mathematically defined as following:

$$Std = \sqrt{\frac{\sum_{i=1}^{M} (Q_i - \mu)^2}{M}}\,, \tag{18}$$

$$BestFitness = max_{i=1}^{M} Q_i\,, \tag{19}$$

Table 2. Datasets Description

	No. of Attributes	No. of Instances	No. of Classes	Missing Values
Wisconsin Diagnosis Breast Cancer (WBCD)	32	596	2	No
Mice Protein Expression Dataset (MPED)	82	1080	8	Yes
Parkinson's Disease Detection Dataset (PDD)	23	197	2	No
Cardiotocography	23	2126	3	No
Hepatitis	19	155	2	Yes
Lung Cancer	56	32	3	Yes
Single Proton Emission Computed Tomography (SPECT)	44	267	2	No
Thoracic Surgery	17	470	2	No
Statlog (Heart)	13	270	2	No
Indian Liver Patient Dataset	10	583	2	No

Table 3. Parameters Settings for All WOA Versions

Parameter	Value
Number of Search Agents (Population)	30
Lower Bound	1
Upper Bound	Same as Total Number of Features in The Original Dataset
Number of Iterations (Generation)	50
Dimension	Same as Total Number of Features in The Original Dataset

$$WorstFitness = min_{i=1}^{M} Q_i \,, \tag{20}$$

$$MeanFitness = \frac{1}{M} \sum_{i=1}^{M} Q_i \,, \tag{21}$$

$$ASS = \frac{1}{M} \sum_{i=1}^{M} \frac{length(Q_i)}{L} \,, \tag{22}$$

where M is the maximum number of iterations, Q is the best score obtained so far at each time, and L is number of features in the original dataset.

Table 4. Comparative- based algorithms: A parameters settings, where (PSO) = Particle Swarm Optimization (Kennedy and Eberhart, 1995), ABC = Artificial Bee Colony (Karaboga, 2005), CSO = Chicken Swarm Optimization (Meng et al., 2014), BBO - Biogeography-Based Optimization (Simon and Cleveland, 2008), EHO = Elephant Herding Optimization (Gai-Ge et al., 2015), KH = Krill Herd (Gandomi and Alavi, 2012), BSA = Bird Swarm Algorithm (Meng et al., 2016), FPA=Flower Pollination Algorithm (Yang, 2012), MFO= Moth-Flame Optimization Algorithm (Mirjalili, 2015) and GWO = Grey Wolf Optimizer (Mirjalili, Seyed, and Lewis, 2014).

↓ **Algorithms**	**Parameter Value(s)**
PSO	An inertial weight = 1, A inertia weight damping ratio = 0.9, Personal learning coefficient = 1.5 and Global learning coefficient = 2.0
ABC	A number of colony size = 10, A number of food source = 5 and A number of limit trials = 5
CSO	A number of chicken updated = 10, The percent of roosters population size = 0.15, The percent of hens population size = 0.7and The percent of mother hens population size = 0.05
BBO	A number of keeping best habitats = 2, A lower bound for immigration probability per gene = 0, A upper bound for immigration probability per gene =1, The step size used for numerical integration of probabilities = 1, max immigration rate for each island = 1, max emigration rate, for each island = 1 and max species count for each island = 30
EHO	A number of keeping best elephants = 2, A number of genes in each population member = number of features, A mutation probability = 0.3, A clan operator = 50, and number of elephants in each clan = 10
KH	A number of runs = 3, A number of krill = number of features and Foraging motion operators (vf = 0.02, Dmax = 0.003, Nmax = 0.01, Sr = 0)
BSA	The frequency of birds' flight behaviours = 10, A Cognitive accelerated coefficient = 1.5, A Social accelerated coefficient = 1.5, The two parameters which are related to the indirect and direct effect on the birds' vigilance behaviours = 1
FPA	The probability switch = 0.6
MFO	A number of dimension = number of features, The maximum generation = 30, Search agent number = 50, lower bound = 1 and upper bound = size of features
GWO	A number of dimension = number of features, The maximum generation = 30, Search agent number = 50, lower bound = 1 and upper bound = size of features

In addition, p-Values from Wilcoxon's rank sum test (nonparametric statistical test) with 5% significance level are adopted (Wilcoxon, 1945). The statistical test is needed to indicate that the proposed algorithm provides a significant improvement compared to the other algorithms (Derrac et al., 2011). Wilcoxon's rank sum test is more sensitive than the t-test as it assumes proportionality of differences between two paired samples. Moreover, it is safer than t-tests as it does not assume the normal distribution. Additionally, the outliers affect the Wilcoxon test less than the t-test (Derrac et al., 2011). The best values of p when $p\text{-}value < 0.05$ which can be considered sufficient evidence against the null hypothesis. All the experiments were implemented in MATLAB-R2012 on a computer with Intel Core 2 GHz and 2GB memory.

5.4 Statistical Analysis for the Datasets

Tables 5, 6, 7, 8 and 9 present the statistical results obtained for all used datasets with ten chaos maps in terms of best fitness, worst fitness, mean fitness, standard deviation and average selection size. All of these measurements are calculated on average. Best fitness, worst fitness and mean fitness are used to evaluate the performance of the selected features which maximizing the classification accuracy and minimizing the number of selected features. Average feature selection size is used to evaluate the number of selected feature obtained per time. In order to evaluate the stability, the standard deviation is adopted. In these tables, we show how embedding chaotic maps in searching iterations of the algorithm can significantly boost the performance of WOA. For example, the mice protein expression dataset contains many attributes, which significantly influence the classification performance. The dataset consists of the expression levels of 77 proteins/protein modifications, which produced detectable signals in the nuclear fraction of cortex. 72 mice were included in the experiment, where 38 control mice and 34 trisomic mice (Down syndrome). The eight classes of mice are described based on features such as genotype, behavior and treatment. According to genotypes, mice can be control or trisomic. According to behavior, some mice have been stimulated to learn (context-shock) and others have not (shock-context), and in order to assess the effect of the drug meantime in recovering the ability to learn in trisomic mice, some mice have been injected with the drug and others have not. Feature selection plays an important role in identifying subsets of proteins discriminating between the classes. Thus, it can be used to assess associative learning.

Table 5 shows statistical results for the WDBC and MPED datasets. As it can be seen in this table for WDBC dataset, iterative, circle and piecewise maps outperform the other chaos maps, which owns the highest stability while iterative, sinusoidal, circle, and chebyshev maps do for MPED dataset. Moreover, the highest mean fitness value (best score) is obtained by piecewise and circle map, while circle has the highest stability. These results indicate the performance of circle map is superior in selecting the minimum number of features with good classification performance. Additionally it can be observed that CWOA with modification of PI obtains the best results compared with the original WOA and other version of CWOA, and CWOA with the modification of all parameters obtains the worst results. Furthermore, it can be observed that the circle map used only 8% of the total number of features (81 features) and acquired the highest classification accuracy and highest stability. Thus, it can be concluded that the circle map is superior to the other maps.

Table 6 shows statistical results for the PDD and cardiotocography datasets. As it can be seen in the table for the PDD dataset, the highest stability is obtained by circle, iterative, sinusoidal and singer chaotic maps, while singer, iterative and circle do for the cardiotocography dataset. The highest results are obtained from CWOA with modification of shrinking, spiral and choosing random position PI, while CWOA with modification

Table 5. Statistical Analysis for WDBC and MPED datasets (continued on next page, see online version for color)

	Data sets →	WDBC					MPED				
	Measures →	**Mean**	**Std.**	**Best**	**Worst**	**ASS**	**Mean ⇓**	**Std.**	**Best**	**Worst**	**ASS**
WOA	⇓ Map function	1.64	0.06	1.66	1.28	0.35	1.69	0.07	22	1.34	0.54
CWOA-SC	**Chebyshev**	1.65	0.05	1.68	1.34	0.28	1.67	0.06	1.70	1.33	0.19
	Circle	1.66	0.05	1.71	1.27	0.12	1.70	0.05	1.72	1.33	0.11
	Gauss/mouse	1.64	0.05	1.65	1.32	0.09	1.64	0.05	1.66	1.33	0.20
	Iterative	1.66	0.06	1.68	1.27	0.25	1.68	0.05	1.69	1.34	0.13
	Logistic	1.66	0.05	1.69	1.33	0.28	1.73	0.06	1.75	1.38	0.10
	Piecewise	1.66	0.05	1.69	1.29	0.05	1.70	0.05	1.72	1.35	0.13
	Sine	1.64	0.05	1.68	1.33	0.05	1.72	0.06	1.76	1.36	0.14
	Singer	1.64	0.05	1.66	1.33	0.08	1.68	0.06	1.74	1.35	0.14
	Sinusoidal	1.66	0.05	1.69	1.30	0.12	1.73	0.05	1.75	1.34	0.11
	Tent	1.66	0.06	1.68	1.32	0.04	1.67	0.06	1.69	1.35	0.16
CWOA-SS	**Chebyshev**	1.62	0.06	1.65	1.26	0.06	1.66	0.06	1.70	1.34	0.26
	Circle	1.66	0.06	1.70	1.30	0.13	1.70	0.05	1.73	1.33	0.11
	Gauss/mouse	1.62	0.06	1.64	1.31	0.10	1.69	0.05	1.71	1.30	0.43
	Iterative	**1.64**	**0.04**	**1.65**	**1.37**	**0.10**	**1.68**	**0.04**	**1.71**	**1.30**	**0.66**
	Logistic	1.64	0.06	1.66	1.27	0.91	1.62	0.05	1.63	1.31	0.28
	Piecewise	1.64	0.05	1.66	1.32	0.11	1.70	0.05	1.73	1.29	0.13
	Sine	1.64	0.06	1.67	1.32	0.37	1.64	0.05	1.70	1.32	0.37
	Singer	1.64	0.05	1.66	1.33	0.07	1.65	0.05	1.67	1.33	0.31
	Sinusoidal	1.66	0.05	1.69	1.32	0.30	**1.70**	**0.04**	**1.72**	**1.39**	**0.13**
	Tent	1.64	0.06	1.65	1.30	0.14	1.70	0.06	1.71	1.34	0.62
CWOA-PI	**Chebyshev**	1.66	0.05	1.69	1.32	0.11	**1.75**	**0.04**	**1.76**	**1.31**	**0.09**
	Circle	**1.67**	**0.04**	**1.69**	**1.32**	**0.04**	**1.74**	**0.04**	**1.76**	**1.31**	**0.08**
	Gauss/mouse	1.65	0.06	1.68	1.33	0.07	1.74	0.05	1.76	1.33	0.19
	Iterative	1.66	0.06	1.69	1.29	0.04	1.72	0.06	1.73	1.34	0.10
	Logistic	1.66	0.05	1.68	1.31	0.05	1.72	0.07	1.76	1.32	0.12
	Piecewise	**1.66**	**0.04**	**1.68**	**1.37**	**0.07**	1.73	0.06	1.75	1.30	0.11
	Sine	1.65	0.05	1.67	1.32	0.08	1.72	0.07	1.75	1.32	0.11
	Singer	1.66	0.06	1.68	1.28	0.05	1.69	0.05	1.70	1.33	0.17
	Sinusoidal	1.64	0.05	1.66	1.27	0.11	**1.65**	**0.04**	**1.67**	**1.36**	**0.13**
	Tent	1.64	0.05	1.65	1.26	0.07	1.72	0.06	1.74	1.34	0.15
CWOA-P	**Chebyshev**	1.64	0.06	1.66	1.28	0.08	1.73	0.07	1.76	1.32	0.17
	Circle	1.64	0.05	1.68	1.28	0.22	1.73	0.05	1.75	1.32	0.09
	Gauss/mouse	1.66	0.05	1.68	1.34	0.32	1.72	0.05	1.74	1.35	0.12
	Iterative	1.67	0.06	1.70	1.31	0.09	1.72	0.06	1.75	1.33	0.11
	Logistic	1.60	0.07	1.64	1.33	0.15	1.68	0.06	1.72	1.37	0.14
	Piecewise	1.62	0.06	1.65	1.31	0.08	1.73	0.07	1.75	1.35	0.12
	Sine	1.63	0.06	1.66	1.30	0.39	1.66	0.05	1.67	1.31	0.09
	Singer	1.66	0.06	1.69	1.31	0.17	1.68	0.07	1.73	1.32	0.37
	Sinusoidal	1.65	0.06	1.68	1.29	0.15	1.68	0.06	1.71	1.30	0.14
	Tent	1.67	0.08	1.70	1.28	0.24	1.69	0.07	1.74	1.35	0.35

Table 5. Statistical Analysis for WDBC and MPED datasets (continued from previous page, see online version for color)

	Data sets →	WDBC					MPED				
	Measures →	**Mean**	**Std.**	**Best**	**Worst**	**ASS**	**Mean ⇓**	**Std.**	**Best**	**Worst**	**ASS**
WOA	⇓ Map function	1.64	0.06	1.66	1.28	0.35	1.69	0.07	22	1.34	0.54
CWOA-All	**Chebyshev**	1.54	0.12	1.63	1.35	0.41	1.59	0.13	1.69	1.36	0.34
	Circle	1.60	0.13	1.69	1.30	0.29	1.60	0.13	1.68	1.34	0.42
	Gauss/mouse	1.56	0.13	1.68	1.32	0.30	1.59	0.13	1.68	1.36	0.29
	Iterative	1.58	0.13	1.67	1.31	0.34	1.61	0.15	1.73	1.35	0.32
	Logistic	1.55	0.13	1.68	1.30	0.40	1.60	0.14	1.72	1.34	0.38
	Piecewise	1.60	0.15	1.68	1.32	0.28	1.60	0.14	1.70	1.32	0.36
	Sine	1.60	0.11	1.68	1.31	0.28	1.59	0.14	1.73	1.31	0.39
	Singer	1.58	0.13	1.68	1.29	0.30	1.60	0.14	1.69	1.34	0.38
	Sinusoidal	1.55	0.12	1.65	1.35	0.30	1.59	0.15	1.70	1.31	0.36
	Tent	1.57	0.14	1.68	1.31	0.32	1.58	0.14	1.68	1.34	0.36

of all parameters failed to improve the performance of WOA. These results are consistent with the obtained results in Table 5. Again, the circle map in most cases has the best mean, and the best and worst fitness value with a small number of features.

Table 7 shows the statistical results for hepatic and lung cancer. As can be seen for hepatic, the classification performance and average selection size for WOA and CWOA with spiral, shrinking and PI operator are almost same. Whereas, CWOA-P and CWOA-All provide the worst results. Additionally, circle and sinusoidal has the best results in terms of classification performance, stability and average selection size. For the lung cancer dataset, almost all versions of CWOA, except CWOA-All, provide better results compared with original version of WOA in terms of classification performance including worst, best and mean fitness value and average selection size, whereas CWOA-All has the highest stability quality. The results indicate the ability of chaotic to enhance the performance of WOA not only exploration but also exploitation. Again the circle map, in most cases, provides the highest classification performance.

Statistical results obtained for SPECTF heart and thoracic surgery are shown at Table 8. As can be seen, the highest stability is obtained from iterative and sinusoidal maps for SPECTF, while circle, piecewise, chebyshev and sinusoidal maps perform best for thoracic surgery. Additionally, it can be observed that CWOA outperforms the WOA in terms of stability and average selection size. Also circle and sinusoidal again obtained the highest classification performance with a high stability and a small number of

Table 6. Statistical Analysis for PDD and Cardiotocography datasets (continued on next page, see online version for color)

	Data sets →	PDD					Cardiotocography				
	Measures →	**Mean**	**Std.**	**Best**	**Worst**	**ASS**	**Mean**	**Std.**	**Best**	**Worst**	**ASS**
WOA	⇓ Map function	1.57	0.06	1.59	1.24	0.38	1.57	0.05	1.59	1.27	0.23
CWOA-SC	**Chebyshev**	1.59	0.05	1.60	1.26	0.05	1.56	0.04	1.58	1.31	0.23
	Circle	**1.59**	**0.04**	**1.64**	**1.36**	**0.10**	1.58	0.04	1.59	1.31	0.17
	Gauss/mouse	1.59	0.05	1.61	1.23	0.06	1.58	0.04	1.59	1.26	0.19
	Iterative	**1.59**	**0.04**	**1.61**	**1.30**	**0.06**	1.57	0.04	1.59	1.25	0.19
	Logistic	1.58	0.05	1.60	1.25	0.05	1.55	0.04	1.56	1.26	0.24
	Piecewise	1.58	0.05	1.60	1.20	0.08	1.59	0.04	1.60	1.30	0.14
	Sine	1.59	0.05	1.61	1.24	0.10	1.58	0.04	1.59	1.33	0.15
	Singer	1.59	0.05	1.61	1.19	0.10	1.58	0.04	1.59	1.30	0.14
	Sinusoidal	**1.59**	**0.04**	**1.61**	**1.31**	**0.07**	1.59	0.04	1.60	1.33	0.15
	Tent	1.58	0.05	1.59	1.28	0.23	1.58	0.04	1.59	1.28	0.14
CWOA-SS	**Chebyshev**	1.58	0.05	1.60	1.30	0.22	1.58	0.05	1.59	1.26	0.23
	Circle	**1.62**	**0.04**	**1.64**	**1.40**	**0.20**	**1.58**	**0.03**	**1.59**	**1.33**	**0.10**
	Gauss/mouse	1.59	0.05	1.62	1.24	0.26	1.58	0.04	1.59	1.31	0.11
	Iterative	1.58	0.05	1.60	1.23	0.06	1.58	0.04	1.59	1.26	0.21
	Logistic	1.59	0.06	1.61	1.22	0.16	1.57	0.04	1.58	1.29	0.23
	Piecewise	1.59	0.06	1.61	1.24	0.60	1.56	0.04	1.57	1.32	0.33
	Sine	1.60	0.05	1.61	1.29	0.14	1.56	0.04	1.57	1.32	0.15
	Singer	1.59	0.05	1.61	1.26	0.10	1.57	0.04	1.59	1.31	0.13
	Sinusoidal	**1.59**	**0.04**	**1.61**	**1.30**	**0.16**	1.56	0.05	1.58	1.32	0.16
	Tent	1.65	0.08	1.67	1.20	0.74	1.56	0.04	1.57	1.27	0.19
CWOA-PI	**Chebyshev**	1.61	0.05	1.62	1.24	0.14	1.58	0.04	1.59	1.29	0.15
	Circle	**1.64**	**0.04**	**1.65**	**1.28**	**0.06**	1.58	0.04	1.59	1.29	0.11
	Gauss/mouse	1.64	0.06	1.65	1.26	0.06	1.56	0.05	1.58	1.25	0.07
	Iterative	1.59	0.06	1.61	1.20	0.49	**1.58**	**0.03**	**1.59**	**1.35**	**0.10**
	Logistic	1.59	0.05	1.60	1.27	0.05	1.58	0.05	1.59	1.30	0.12
	Piecewise	1.60	0.05	1.62	1.27	0.06	1.58	0.04	1.59	1.34	0.11
	Sine	1.63	0.06	1.65	1.29	0.06	1.58	0.05	1.57	1.26	0.20
	Singer	**1.58**	**0.04**	**1.60**	**1.26**	**0.06**	1.58	0.05	1.59	1.26	0.15
	Sinusoidal	**1.59**	**0.04**	**1.60**	**1.31**	**0.05**	1.57	0.05	1.59	1.29	0.15
	Tent	1.63	0.06	1.65	1.27	0.06	1.57	0.05	1.59	1.29	0.74
CWOA-P	**Chebyshev**	1.62	0.06	1.66	1.26	0.11	1.56	0.04	1.57	1.29	0.19
	Circle	1.57	0.07	1.61	1.19	0.23	1.57	0.05	1.59	1.25	0.16
	Gauss/mouse	1.57	0.07	1.61	1.36	0.37	1.59	0.05	1.60	1.32	0.15
	Iterative	1.59	0.07	1.61	1.25	0.21	1.58	0.04	1.59	1.32	0.10
	Logistic	1.64	0.07	1.67	1.27	0.10	1.55	0.04	1.57	1.29	0.23
	Piecewise	1.55	0.08	1.61	1.25	0.21	1.56	0.04	1.57	1.27	0.15
	Sine	1.58	0.06	1.60	1.24	0.11	1.55	0.04	1.57	1.33	0.47
	Singer	1.62	0.07	1.65	1.29	0.09	**1.58**	**0.03**	**1.59**	**1.26**	**0.15**
	Sinusoidal	1.63	0.07	1.65	1.28	0.07	1.58	0.05	1.59	1.30	0.11
	Tent	1.63	0.06	1.65	1.27	0.07	1.54	0.04	1.58	1.33	0.36

Table 6. Statistical Analysis for PDD and Cardiotocography datasets (continued from previous page, see online version for color)

	Data sets →	PDD					Cardiotocography				
	Measures →	**Mean**	**Std.**	**Best**	**Worst**	**ASS**	**Mean**	**Std.**	**Best**	**Worst**	**ASS**
WOA	⇓ Map function	1.57	0.06	1.59	1.24	0.38	1.57	0.05	1.59	1.27	0.23
CWOA-All	**Chebyshev**	1.54	0.16	1.65	1.22	0.31	1.44	0.06	1.50	1.32	0.29
	Circle	1.56	0.13	1.66	1.26	0.30	1.49	0.08	1.54	1.29	0.38
	Gauss/mouse	1.55	0.15	1.65	1.30	0.28	1.42	0.07	1.48	1.26	0.31
	Iterative	1.55	0.15	1.65	1.22	0.34	1.46	0.07	1.52	1.30	0.39
	Logistic	1.55	0.16	1.65	1.20	0.28	1.45	0.07	1.50	1.33	0.30
	Piecewise	1.55	0.15	1.66	1.29	0.27	1.44	0.09	1.50	1.29	0.31
	Sine	1.55	0.16	1.65	1.22	0.32	1.49	0.06	1.53	1.35	0.35
	Singer	1.55	0.16	1.66	1.24	0.30	1.47	0.09	1.53	1.30	0.43
	Sinusoidal	1.55	0.15	1.65	1.23	0.32	1.43	0.06	1.47	1.29	0.32
	Tent	1.56	0.14	1.66	1.28	0.32	1.46	0.10	1.53	1.30	0.35

features for both datasets. Also CWO with modification of PI provides the higher results compared with original WOA and other versions of CWOA. In addition, it can be seen that obtained results that the CWOA-All algorithm is still performing the worst. The statistical results for statlog heart and Indian liver patient datasets are presented in Table 9. The minimum standard deviation values representing the highest stability are highlighted. The highest stability for the statlog dataset is provided from chebyshev, while iterative, sinusoidal and logistic provide the most stability for the Indian liver patient dataset. However, in most cases, circle map provided the best stability with high classification performance and a small number of features. Additionally, it can be observed for statlog dataset, all versions of the CWOA algorithm except CWOA-All provide superior results compared to the original WOA. For the Indian liver dataset, it can be observed that the classification performance of CWOA versions is very similar to WOA; however, the CWOA algorithms (except the CWOA-All version) enhance the stability of WOA.

5.5 Convergence Curves Analysis of the Data Sets with Different Chaotic Maps

Figures 4, 5, 6, 7, 8, 9, 10, 11, 12 and 13 evaluate the stability of the algorithms in terms of convergence rate through visualizing the best-obtained fitness value during the course of iterations. In addition, these figures are

Table 7. Statistical Analysis for Hepatitis and Lung Cancer datasets (continued on next page, see online version for color)

	Data sets →	Hepatitis					Lung Cancer				
	Measures →	**Mean**	**Std.**	**Best**	**Worst**	**ASS**	**Mean**	**Std.**	**Best**	**Worst**	**ASS**
WOA	⇓ Map function	1.78	0.09	1.80	1.23	0.06	1.80	0.17	1.85	1.00	0.05
CWOA-SC	**Chebyshev**	1.78	0.08	1.80	1.31	0.07	1.82	0.15	1.85	1.08	0.03
	Circle	1.79	0.07	1.81	1.26	0.06	1.80	0.12	1.85	1.04	0.03
	Gauss/mouse	1.77	0.08	1.80	1.27	0.08	1.83	0.12	1.85	1.06	0.03
	Iterative	1.77	0.08	1.80	1.39	0.07	1.81	0.14	1.85	1.00	0.04
	Logistic	1.78	0.07	1.80	1.26	0.06	1.81	0.15	1.85	0.98	0.04
	Piecewise	1.78	0.07	1.80	1.32	0.06	1.81	0.15	1.85	1.00	0.04
	Sine	1.76	0.08	1.80	1.37	0.10	1.82	0.13	1.85	0.96	0.03
	Singer	1.77	0.09	1.80	1.27	0.07	1.81	0.16	1.85	0.94	0.03
	Sinusoidal	**1.79**	**0.05**	**1.80**	**1.37**	**0.05**	1.82	0.13	1.85	1.02	0.08
	Tent	1.77	0.08	1.80	1.29	0.07	1.82	0.13	1.85	1.02	0.03
CWOA-SS	**Chebyshev**	1.77	0.08	1.80	1.36	0.08	1.81	0.14	1.85	0.97	0.04
	Circle	1.78	0.07	1.80	1.37	0.08	1.82	0.13	1.85	0.99	0.04
	Gauss/mouse	1.77	0.09	1.80	1.27	0.08	1.82	0.13	1.85	1.04	0.04
	Iterative	1.76	0.12	1.80	1.24	0.09	1.81	0.15	1.85	0.98	0.04
	Logistic	1.78	0.08	1.80	1.36	0.06	1.80	0.16	1.85	1.03	0.07
	Piecewise	1.78	0.07	1.80	1.35	0.06	1.81	0.15	1.85	1.05	0.04
	Sine	1.77	0.08	1.80	1.31	0.08	1.82	0.13	1.85	1.04	0.03
	Singer	1.79	0.10	1.80	1.31	0.14	1.82	0.14	1.85	0.95	0.03
	Sinusoidal	1.78	0.08	1.80	1.30	0.06	1.73	0.24	1.85	1.05	0.12
	Tent	1.78	0.08	1.80	1.25	0.07	1.81	0.16	1.85	1.04	0.04
CWOA-PI	**Chebyshev**	1.78	0.06	1.80	1.35	0.11	1.82	0.16	1.85	0.96	0.04
	Circle	1.78	0.08	1.80	1.24	0.07	1.82	0.13	1.85	0.98	0.04
	Gauss/mouse	1.77	0.09	1.80	1.27	0.07	1.80	0.19	1.85	1.01	0.05
	Iterative	1.78	0.07	1.80	1.29	0.06	1.80	0.16	1.85	1.10	0.05
	Logistic	1.78	0.07	1.80	1.29	0.06	1.81	0.17	1.85	0.96	0.05
	Piecewise	1.78	0.08	1.80	1.28	0.07	1.82	0.13	1.85	1.02	0.03
	Sine	1.78	0.07	1.80	1.29	0.06	1.82	0.13	1.85	1.08	0.03
	Singer	1.79	0.06	1.80	1.31	0.06	1.82	0.14	1.85	1.06	0.04
	Sinusoidal	1.77	0.08	1.80	1.28	0.07	1.82	0.14	1.85	0.98	0.03
	Tent	1.78	0.07	1.80	1.30	0.06	1.82	0.14	1.85	0.98	0.03
CWOA-P	**Chebyshev**	1.75	0.10	1.80	1.29	0.07	1.78	0.19	1.85	0.98	0.08
	Circle	1.77	0.10	1.80	1.27	0.08	1.82	0.14	1.85	1.02	0.04
	Gauss/mouse	1.77	0.09	1.80	1.31	0.08	1.81	0.17	1.85	0.90	0.04
	Iterative	1.72	0.15	1.80	1.30	0.23	1.81	0.15	1.85	1.00	0.04
	Logistic	1.77	0.08	1.80	1.36	0.08	1.80	0.17	1.85	1.03	0.05
	Piecewise	1.76	0.09	1.80	1.27	0.09	1.80	0.18	1.85	0.09	0.04
	Sine	1.77	0.09	1.80	1.26	0.07	1.78	0.17	1.85	0.92	0.09
	Singer	1.77	0.09	1.80	1.29	0.07	1.77	0.18	1.85	0.98	0.06
	Sinusoidal	1.76	0.10	1.80	1.35	0.11	1.82	0.13	1.85	1.05	0.04
	Tent	1.74	0.13	1.80	1.20	0.11	1.78	0.18	1.85	0.88	0.06

Table 7. Statistical Analysis for Hepatitis and Lung Cancer datasets (continued from previous page, see online version for color)

	Data sets →	**Hepatitis**					**Lung Cancer**				
	Measures →	**Mean**	**Std.**	**Best**	**Worst**	**ASS**	**Mean**	**Std.**	**Best**	**Worst**	**ASS**
WOA	⇓ Map function	1.78	0.09	1.80	1.23	0.06	1.80	0.17	1.85	1.00	0.05
CWOA-All	**Chebyshev**	1.53	0.10	1.60	1.30	0.25	1.24	0.05	1.27	1.03	0.92
	Circle	1.50	0.15	1.60	1.22	0.30	1.32	0.12	1.43	1.08	0.32
	Gauss/mouse	1.53	0.10	1.60	1.29	0.31	1.20	0.04	1.22	0.99	0.85
	Iterative	1.60	0.13	1.69	1.37	0.34	1.27	0.10	1.36	1.03	0.28
	Logistic	1.52	0.11	1.60	1.26	0.24	**1.20**	**0.03**	**1.24**	**1.00**	**0.88**
	Piecewise	1.56	0.10	1.60	1.34	0.32	1.28	0.11	1.44	0.94	0.42
	Sine	1.51	0.12	1.60	1.32	0.28	1.31	0.08	1.41	1.04	0.28
	Singer	1.51	0.12	1.60	1.28	0.27	1.15	0.04	1.19	0.96	0.90
	Sinusoidal	1.55	0.07	1.60	1.32	0.23	1.23	0.06	1.26	0.92	0.30
	Tent	1.51	0.10	1.59	1.35	0.37	1.26	0.11	1.36	1.02	0.30

used to evaluate the ability of WOA and CWOA algorithms to search extensively promising regions in the search space and the ability to converge faster toward the optimum. As can be observed from these figures, in the early stages of the optimization process, the search agents change abruptly and then gradually converge. Figure 4 shows the convergence curves of all chaos maps with different version of WOA. The results of both CWOA with modification spiral method and with PI are comparable, however, in most cases COWA-PI obtains the highest scores as in Table 5. Additionally, the circle is the most stable map and have the fastest convergence rate while singer and sine maps have the lowest convergence rate. CWOA-All has the worst convergence rate. Considering the obtained results at Table 5 and Figure 4, it is shown that the circle map can improve the performance of WOA in terms of both exploitation and exploration.

The convergence curves of chaos maps for the MPED dataset are depicted at Figure 5. This figure shows that circle, guassian/mouse and sinusoidal maps have the fastest convergence rate, while singer, sine and logistics have the slowest convergence rate. In addition, CWOA-PI converges faster than the others. While CWOA-All again is the worst one, and it can be observed that the best score increases faster and may converge around iteration number 15. These results indicate the performance of WOA can be boosted by circle CWOA-PI.

The convergence curves of all chaotic maps for the PDD dataset are depicted in Figure 6. This figure shows that circle, iterative, logistic and sinusoidal have the fastest convergence rates. However, the circle maps show

Table 8. Statistical Analysis for SPECTF Heart and Thoracic Surgery Datasets (continued on next page, see online version for color)

	Data sets →	**SPECTF Heart**					**Thoracic Surgery**				
	Measures →	**Mean**	**Std.**	**Best**	**Worst**	**ASS**	**Mean**	**Std.**	**Best**	**Worst**	**ASS**
WOA	⇓ Map function	1.77	0.13	1.80	1.14	0.05	1.63	0.09	1.71	1.39	0.53
CWOA-SC	**Chebyshev**	1.77	0.12	1.81	1.26	0.04	1.61	0.07	1.71	1.34	0.25
	Circle	1.78	0.10	1.81	1.15	0.04	1.61	0.04	1.62	1.33	0.12
	Gauss/mouse	1.78	0.12	1.81	1.13	0.04	1.57	0.04	1.61	1.30	0.10
	Iterative	1.78	0.11	1.81	1.16	0.04	1.58	0.04	1.61	1.44	0.13
	Logistic	1.78	0.11	1.81	1.13	0.04	1.58	0.05	1.61	1.33	0.08
	Piecewise	1.78	0.11	1.81	1.12	0.04	**1.59**	**0.02**	**1.61**	**1.39**	**0.08**
	Sine	1.78	0.11	1.81	1.13	0.04	1.55	0.03	1.56	1.37	0.06
	Singer	1.78	0.10	1.81	1.14	0.04	1.62	0.06	1.71	1.38	0.26
	Sinusoidal	1.79	0.10	1.81	1.15	0.03	1.66	0.08	1.71	1.37	0.19
	Tent	1.78	0.10	1.81	1.14	0.04	1.58	0.05	1.61	1.34	0.16
CWOA-SS	**Chebyshev**	1.79	0.10	1.81	1.15	0.03	1.58	0.04	1.61	1.39	0.68
	Circle	1.79	0.11	1.81	1.17	0.03	**1.66**	**0.02**	**1.71**	**1.37**	**0.11**
	Gauss/mouse	1.79	0.10	1.81	1.11	0.03	1.55	0.03	1.56	1.40	0.17
	Iterative	**1.79**	**0.09**	**1.81**	**1.13**	**0.03**	1.66	0.06	1.71	1.43	0.40
	Logistic	1.78	0.12	1.81	1.17	0.05	1.61	0.04	1.62	1.39	0.36
	Piecewise	1.79	0.11	1.81	1.23	0.03	1.61	0.03	1.62	1.43	0.94
	Sine	1.77	0.12	1.81	1.13	0.04	1.55	0.03	1.57	1.43	0.08
	Singer	1.77	0.13	1.81	1.18	0.05	1.68	0.06	1.71	1.38	0.26
	Sinusoidal	1.78	0.12	1.81	1.15	0.04	**1.55**	**0.02**	**1.56**	**1.38**	**0.09**
	Tent	1.77	0.12	1.81	1.13	0.04	1.56	0.03	1.57	1.39	0.34
CWOA-PI	**Chebyshev**	1.79	0.11	1.81	1.11	0.04	1.73	0.07	1.76	1.38	0.13
	Circle	1.79	0.12	1.81	1.12	0.04	1.71	0.07	1.74	1.33	0.07
	Gauss/mouse	1.79	0.10	1.81	1.15	0.03	1.59	0.03	1.61	1.39	0.06
	Iterative	1.77	0.14	1.81	1.12	0.05	1.71	0.10	1.76	1.33	0.08
	Logistic	1.79	0.11	1.81	1.13	0.03	1.59	0.05	1.61	1.34	0.11
	Piecewise	1.75	0.14	1.81	1.16	0.05	1.70	0.07	1.76	1.41	0.20
	Sine	1.77	0.13	1.81	1.18	0.04	1.74	0.06	1.76	1.39	0.47
	Singer	1.78	0.11	1.81	1.16	0.03	1.56	0.03	1.57	1.38	0.08
	Sinusoidal	**1.79**	**0.09**	**1.81**	**1.15**	**0.03**	1.61	0.03	1.62	1.38	0.26
	Tent	1.77	0.12	1.81	1.19	0.04	1.61	0.04	1.62	1.39	0.52
CWOA-P	**Chebyshev**	1.76	0.13	1.81	1.13	0.05	**1.55**	**0.02**	**1.56**	**1.42**	**0.07**
	Circle	1.78	0.13	1.81	1.11	0.04	1.71	0.08	1.76	1.38	0.49
	Gauss/mouse	1.74	0.15	1.81	1.13	0.10	1.63	0.06	1.67	1.37	0.90
	Iterative	1.72	0.14	1.81	1.14	0.05	1.71	0.08	1.76	1.32	0.13
	Logistic	1.70	0.12	1.78	1.15	0.09	1.54	0.05	1.57	1.39	0.21
	Piecewise	1.76	0.14	1.81	1.12	0.05	1.68	0.07	1.71	1.39	0.18
	Sine	1.67	0.23	1.81	1.10	0.19	1.56	0.09	1.62	1.36	0.45
	Singer	1.76	0.14	1.81	1.14	0.06	1.53	0.05	1.57	1.39	0.36
	Sinusoidal	1.74	0.13	1.81	1.14	0.04	1.72	0.07	1.76	1.39	0.13
	Tent	1.75	0.12	1.81	1.11	0.13	1.68	0.06	1.71	1.37	0.49

Table 8. Statistical Analysis for SPECTF Heart and Thoracic Surgery Datasets (continued from previous page, see online version for color)

	Data sets →	SPECTF Heart					Thoracic Surgery				
	Measures →	**Mean**	**Std.**	**Best**	**Worst**	**ASS**	**Mean**	**Std.**	**Best**	**Worst**	**ASS**
WOA	⇓ Map function	1.77	0.13	1.80	1.14	0.05	1.63	0.09	1.71	1.39	0.53
CWOA-All	**Chebyshev**	1.46	0.15	1.57	1.14	0.28	1.54	0.04	1.57	1.38	0.26
	Circle	1.44	0.16	1.57	1.17	0.26	1.51	0.08	1.57	1.37	0.20
	Gauss/mouse	1.48	0.16	1.61	1.08	0.31	1.52	0.07	1.57	1.36	0.30
	Iterative	1.46	0.14	1.56	1.16	0.28	1.52	0.08	1.58	1.38	0.29
	Logistic	1.47	0.15	1.57	1.13	0.30	1.52	0.08	1.57	1.36	0.31
	Piecewise	1.44	0.14	1.54	1.13	0.36	1.52	0.08	1.57	1.37	0.28
	Sine	1.43	0.16	1.53	1.17	0.28	1.52	0.08	1.57	1.39	0.27
	Singer	1.43	0.16	1.56	1.14	0.29	1.51	0.10	1.58	1.34	0.29
	Sinusoidal	1.46	0.17	1.57	1.14	0.30	1.53	0.07	1.58	1.37	0.28
	Tent	1.47	0.16	1.59	1.18	0.28	1.51	0.10	1.57	1.33	0.31

the overall fastest convergence rate where the most of CWOA versions converge around iteration 10. Also CWOA with modification PI outperforms than other CWOA versions and the original WOA. This indicates that the WOA'performance can be boosted by the circle PI operator. Figure 7 shows the convergence curves for all WOA versions with 10 chaos maps. As can be observed, there is no distinguishing convergence behavior of the algorithms; almost all CWOAs convergence behavior are very close to the obtained results from Table 6.

Figure 8 shows the convergence curves for the hepatic dataset. This figure shows there is no superiority for any algorithms except CWOA-All algorithm; the classification performance and convergence speed are very close to each other. However, CWOA-P algorithm is in second place of convergence speed after CWOA-All algorithm. The same observations are found in Figure 9, where all algorithms except CWOA-All obtains similar classification performance. These results are consistent with the obtained results from Table 7.

Figure 10 compares the convergence curves of the SPECTF heart dataset for all chaotic maps for different versions of WOA algorithms. As can be observed, the highest convergence speeds are obtained from checbyshev, piecewise, circle and sinusoidal maps, however circle map outperform other chaotic maps. In addition, it can be noticed that the classification performance of all algorithms are close to each other, which is consistent with the results of Table 8. Finally, CWOA-PI converges faster than the others in most cases and CWOA-All has the slowest convergence rate.

Table 9. Statistical Analysis for Statlog Heart and Indian Liver Patient Datasets (continued on next page, see online version for color)

	Data sets →	Statlog Heart					Indian Liver Patient				
	Measures →	**Mean**	**Std.**	**Best**	**Worst**	**ASS**	**Mean**	**Std.**	**Best**	**Worst**	**ASS**
WOA	⇓ Map function	1.38	0.05	1.39	1.08	0.20	1.38	0.06	1.40	1.05	0.18
CWOA-SC	**Chebyshev**	1.44	0.04	1.46	1.15	0.78	1.40	0.04	1.40	1.12	0.29
	Circle	1.40	0.04	1.42	1.30	0.31	1.39	0.04	1.40	1.10	0.20
	Gauss/mouse	1.41	0.05	1.45	1.15	0.40	1.40	0.03	1.41	1.22	0.10
	Iterative	1.41	0.07	1.45	1.02	0.43	**1.40**	**0.02**	**1.41**	**1.18**	**0.10**
	Logistic	1.43	0.06	1.45	1.13	0.19	1.40	0.04	1.41	1.12	0.11
	Piecewise	1.42	0.06	1.45	1.06	0.23	1.39	0.03	1.40	1.15	0.10
	Sine	1.42	0.07	1.46	1.05	0.18	1.40	0.05	1.41	1.04	0.11
	Singer	1.44	0.03	1.45	1.20	0.17	1.39	0.04	1.41	1.12	0.11
	Sinusoidal	1.41	0.06	1.45	1.18	0.62	**1.40**	**0.02**	**1.41**	**1.12**	**0.11**
	Tent	1.41	0.07	1.45	1.08	0.27	1.40	0.04	1.41	1.28	0.10
CWOA-SS	**Chebyshev**	**1.39**	**0.02**	**1.41**	**1.23**	**0.22**	1.39	0.05	1.40	1.12	0.15
	Circle	1.37	0.04	1.39	1.16	0.31	1.40	0.04	1.40	1.13	0.11
	Gauss/mouse	1.43	0.06	1.45	1.18	0.20	1.40	0.05	1.41	1.10	0.66
	Iterative	1.45	0.04	1.46	1.14	0.28	1.39	0.04	1.41	1.17	0.11
	Logistic	1.44	0.05	1.46	1.16	0.18	1.39	0.04	1.41	1.18	0.26
	Piecewise	1.43	0.04	1.16	1.28	0.24	1.40	0.03	1.40	1.14	0.10
	Sine	1.42	0.04	1.44	1.14	0.22	1.40	0.04	1.42	1.11	0.11
	Singer	1.37	0.04	1.39	1.11	0.44	1.39	0.05	1.41	1.10	0.11
	Sinusoidal	1.41	0.07	1.46	1.02	0.83	1.39	0.05	1.40	1.04	0.27
	Tent	1.44	0.06	1.46	1.02	0.17	1.39	0.05	1.40	1.10	0.12
CWOA-PI	**Chebyshev**	1.44	0.04	1.45	1.18	0.29	1.39	0.05	1.41	1.13	0.13
	Circle	1.44	0.05	1.46	1.10	0.16	1.40	0.04	1.41	1.12	0.10
	Gauss/mouse	1.44	0.05	1.45	1.11	0.16	1.44	0.04	1.45	1.16	0.10
	Iterative	1.44	0.05	1.46	1.09	0.23	1.43	0.04	1.45	1.15	0.11
	Logistic	1.42	0.07	1.46	1.08	0.63	**1.44**	**0.02**	**1.45**	**1.26**	**0.10**
	Piecewise	1.44	0.05	1.46	1.06	0.16	1.40	0.04	1.41	1.12	0.11
	Sine	1.44	0.05	1.45	1.09	0.17	1.44	0.05	1.44	1.06	0.20
	Singer	1.35	0.03	1.36	1.15	0.24	1.43	0.06	1.45	1.11	0.12
	Sinusoidal	1.45	0.03	1.46	1.20	0.38	1.43	0.05	1.44	1.14	0.11
	Tent	1.44	0.05	1.46	1.06	0.16	1.39	0.04	1.40	1.12	0.12
CWOA-P	**Chebyshev**	1.44	0.05	1.46	1.22	0.18	1.38	0.05	1.41	1.15	0.22
	Circle	1.44	0.04	1.46	1.14	0.16	1.38	0.04	1.40	1.19	0.13
	Gauss/mouse	1.44	0.06	1.46	1.01	0.17	1.36	0.06	1.41	1.18	0.25
	Iterative	1.41	0.08	1.45	1.01	0.38	1.43	0.04	1.44	1.17	0.11
	Logistic	1.45	0.04	1.46	1.16	0.22	1.37	0.06	1.40	1.04	0.17
	Piecewise	1.34	0.04	1.36	1.08	0.17	1.38	0.04	1.40	1.14	0.33
	Sine	1.44	0.06	1.46	1.12	0.19	1.39	0.05	1.41	1.09	0.12
	Singer	1.42	0.07	1.46	1.05	0.20	1.43	0.07	1.45	1.10	0.32
	Sinusoidal	1.32	0.09	1.39	1.00	0.29	1.43	0.05	1.45	1.11	0.12
	Tent	1.42	0.04	1.44	1.17	0.61	1.36	0.07	1.40	1.15	0.37

Table 9. Statistical Analysis for Statlog Heart and Indian Liver Patient Datasets (continued from previous page, see online version for color)

	Data sets →	Statlog Heart					Indian Liver Patient				
	Measures →	**Mean**	**Std.**	**Best**	**Worst**	**ASS**	**Mean**	**Std.**	**Best**	**Worst**	**ASS**
WOA	⇓ Map function	1.38	0.05	1.39	1.08	0.20	1.38	0.06	1.40	1.05	0.18
CWOA-All	**Chebyshev**	1.34	0.13	1.45	1.12	0.42	1.38	0.10	1.44	1.08	0.36
	Circle	1.40	0.07	1.44	1.07	0.41	1.39	0.10	1.44	1.12	0.34
	Gauss/mouse	1.38	0.11	1.45	1.03	0.35	1.36	0.13	1.44	1.13	0.27
	Iterative	1.30	0.04	1.34	1.20	0.33	1.37	0.12	1.44	1.11	0.29
	Logistic	1.30	0.04	1.32	1.11	0.31	1.36	0.12	1.44	1.14	0.23
	Piecewise	1.24	0.11	1.33	1.00	0.34	1.37	0.12	1.44	1.11	0.33
	Sine	1.31	0.16	1.43	1.05	0.38	1.35	0.14	1.44	1.09	0.36
	Singer	1.40	0.06	1.46	1.30	0.60	1.39	0.08	1.44	1.13	0.34
	Sinusoidal	1.37	0.10	1.46	1.16	0.43	1.36	0.12	1.44	1.16	0.34
	Tent	1.37	0.11	1.45	1.06	0.31	1.37	0.11	1.44	1.11	0.34

Figure 13 shows the convergence curves for the thoracic surgery dataset. As can be observed, the highest convergence speed is for CWOA with modification of spiral and PI operator algorithms. Additionally, circle map outperforms other chaotic maps.

Figure 11 shows the convergence curves for the statlog dataset. As can be observed the fastest convergence speed is obtained from the circle map. Again CWOA-PI converges faster than the other algorithms. The convergence curves for the Indian liver dataset are shown at Figure 12. As can be seen, the classification performance of each algorithm are very close with results from Table 9.

To sum up, the experimental results show that the iterative, circle, sinusoidal, piecewise, chebyshev, tent, singer, sine, logistic and Guas/mouse can improve the original WOA. Also the results of CWOA-SS, CWOA-SC and CWOA-PI are much better than CWOA-All, whereas CWOA-PI outperform other chaotic versions. This indicates the weakness of exploration versus exploitation for WOA. In addition, the chaotic version of WOA with modification of PI is able to alleviate this weakness. The results of the CWOA using the circle map show that the majority of this map to improve the performance of WOA in terms of highest stability quality, highest classification performance and small feature subset. The highest classification performance and a small number of features are highly preferred in medicine and biology. This is due to fewer experiments needed for a certain disease or cancer, which may be difficult sometimes on patient and can cause side effects (Zawbaa, Emary, and Grosan, 2016). Also, it can reduce the cost involved for each experiment. Generally speaking, the results of the chaotic

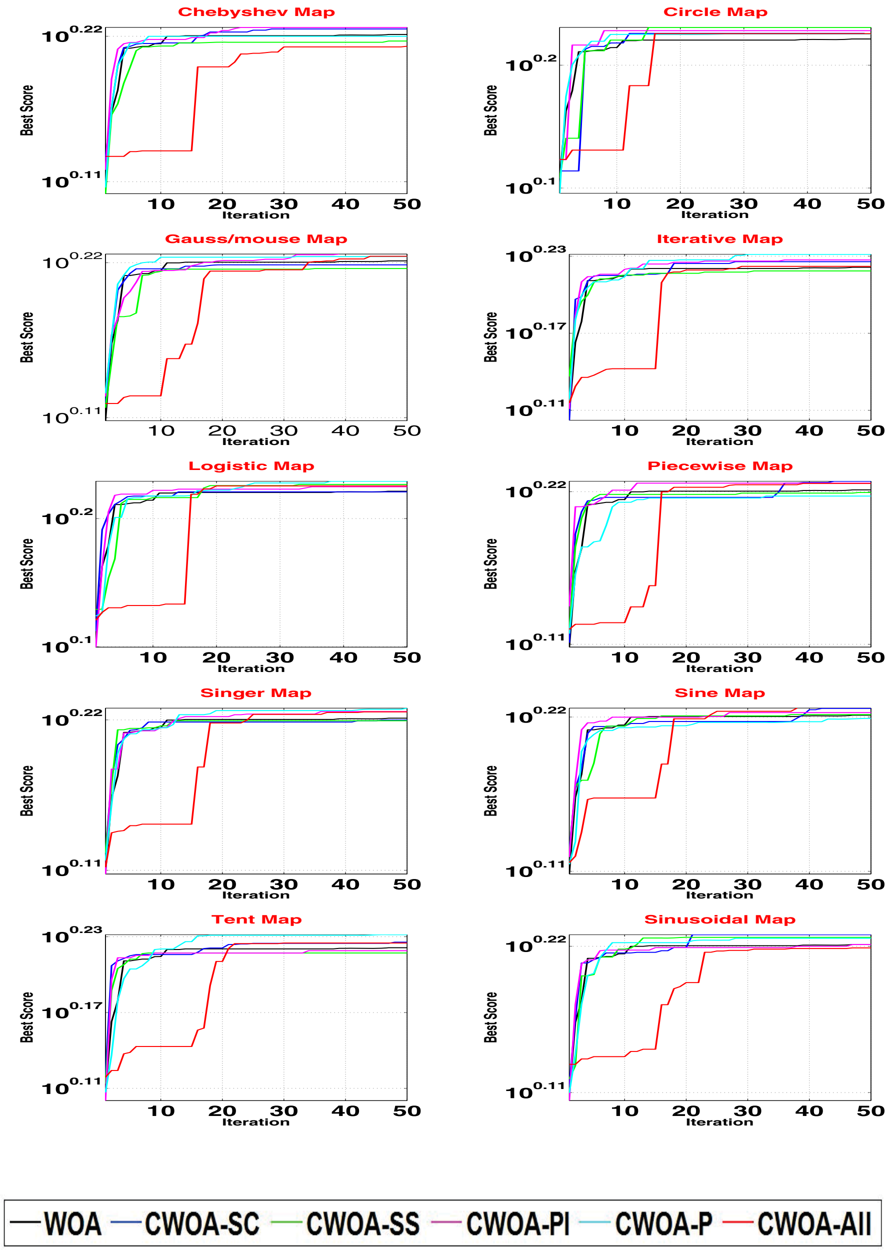

Figure 4. Convergence Curves for **WDBC Dataset** using Different Chaotic Maps (see online version for color)

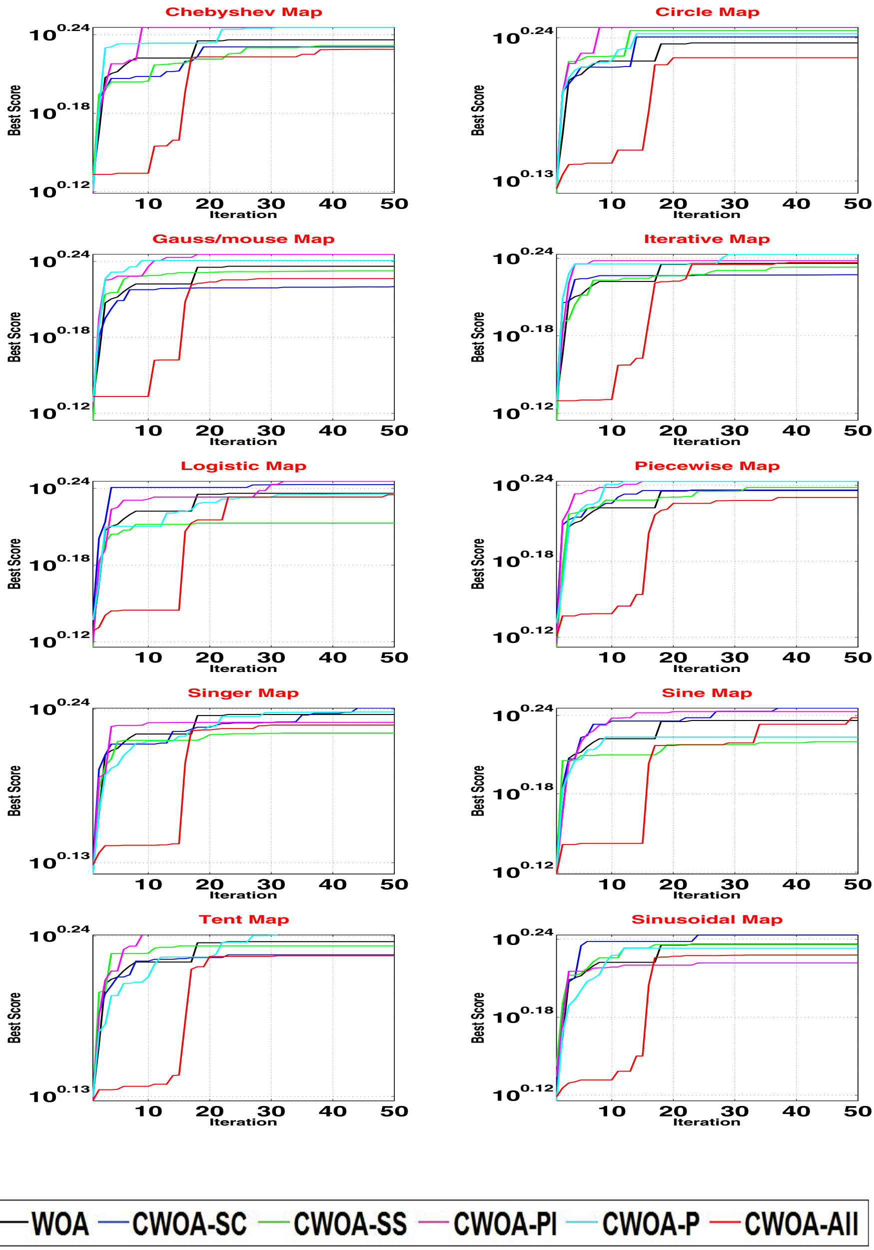

Figure 5. Convergence Curves for**MPED Dataset** using Different Chaotic Maps (see online version for color)

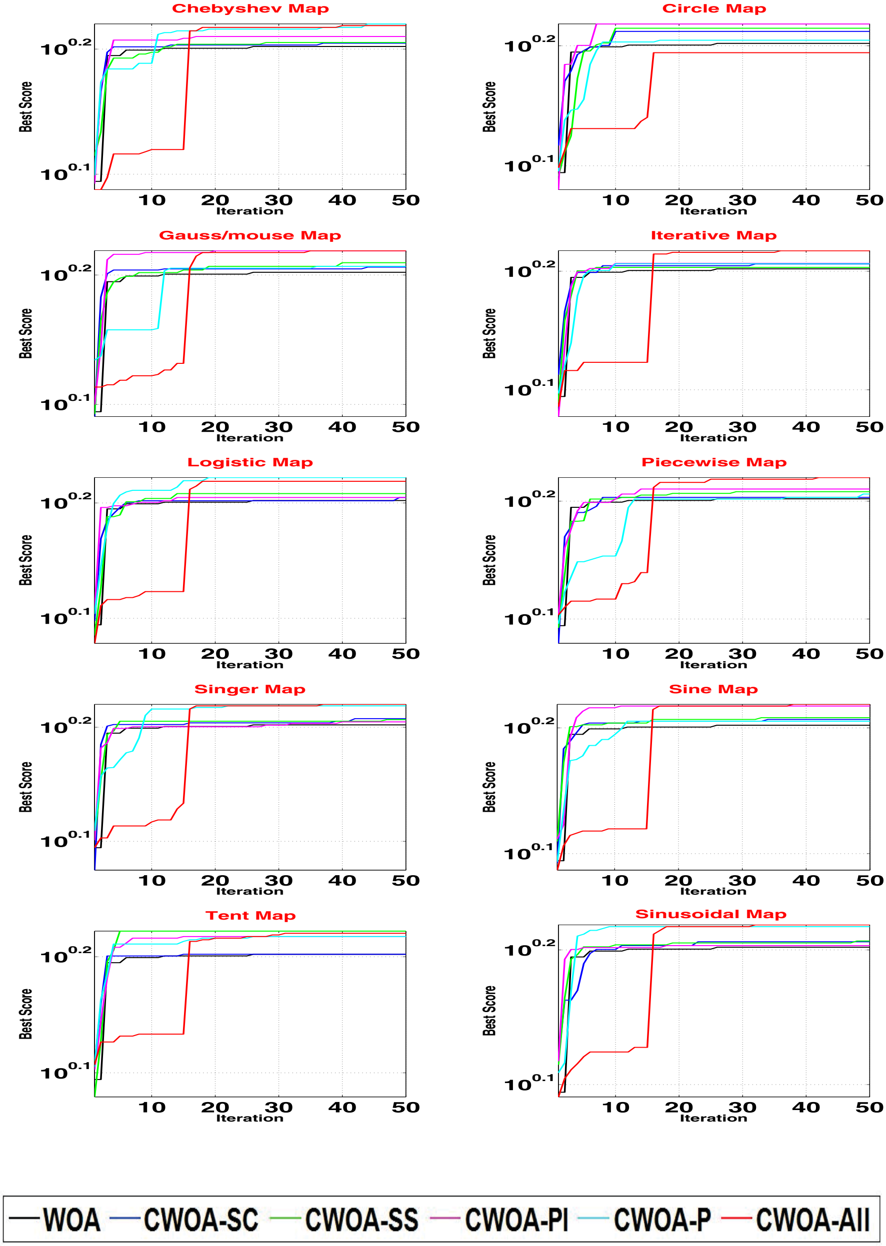

Figure 6. Convergence Curves for **PDD Dataset** using Different Chaotic Maps (see online version for color)

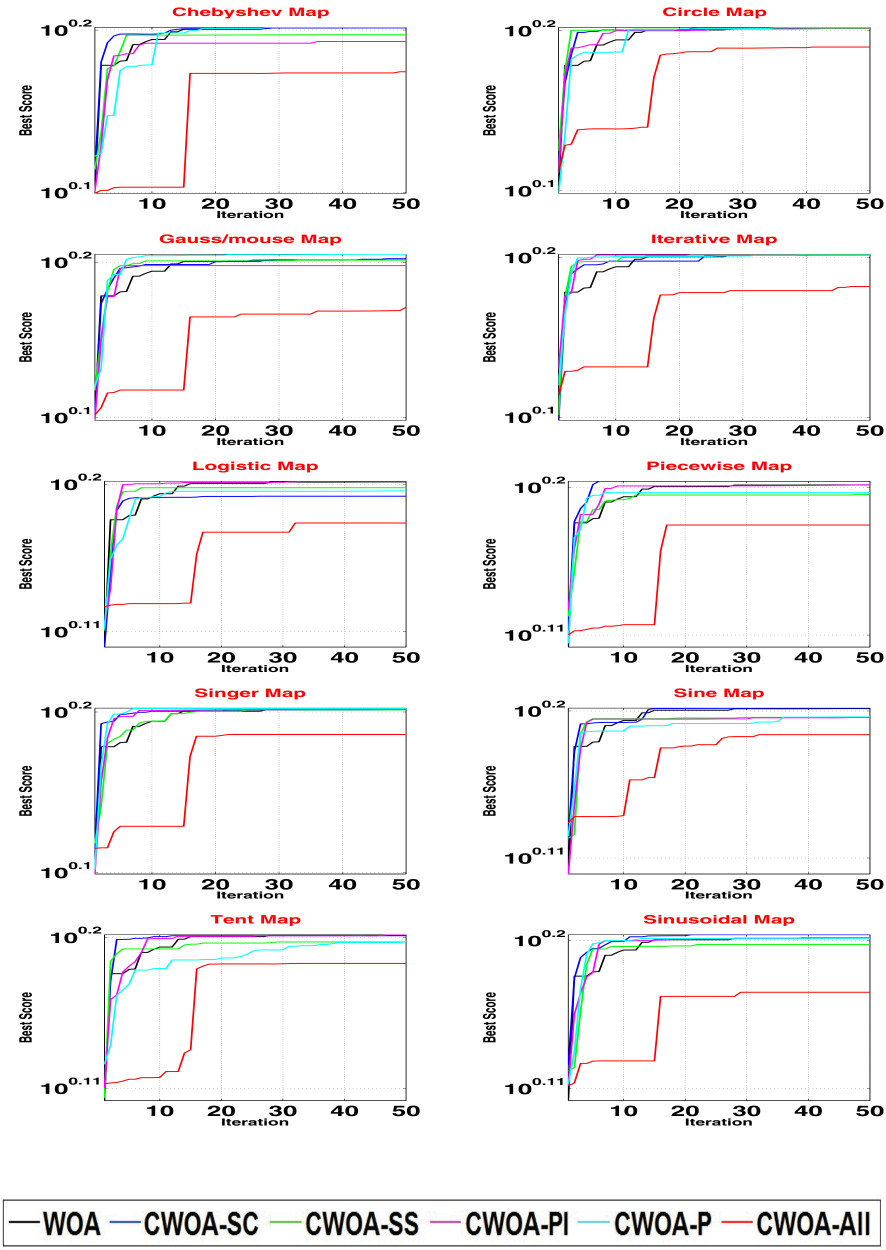

Figure 7. Convergence Curves for **Cardiotocography Dataset** using Different Chaotic Maps (see online version for color)

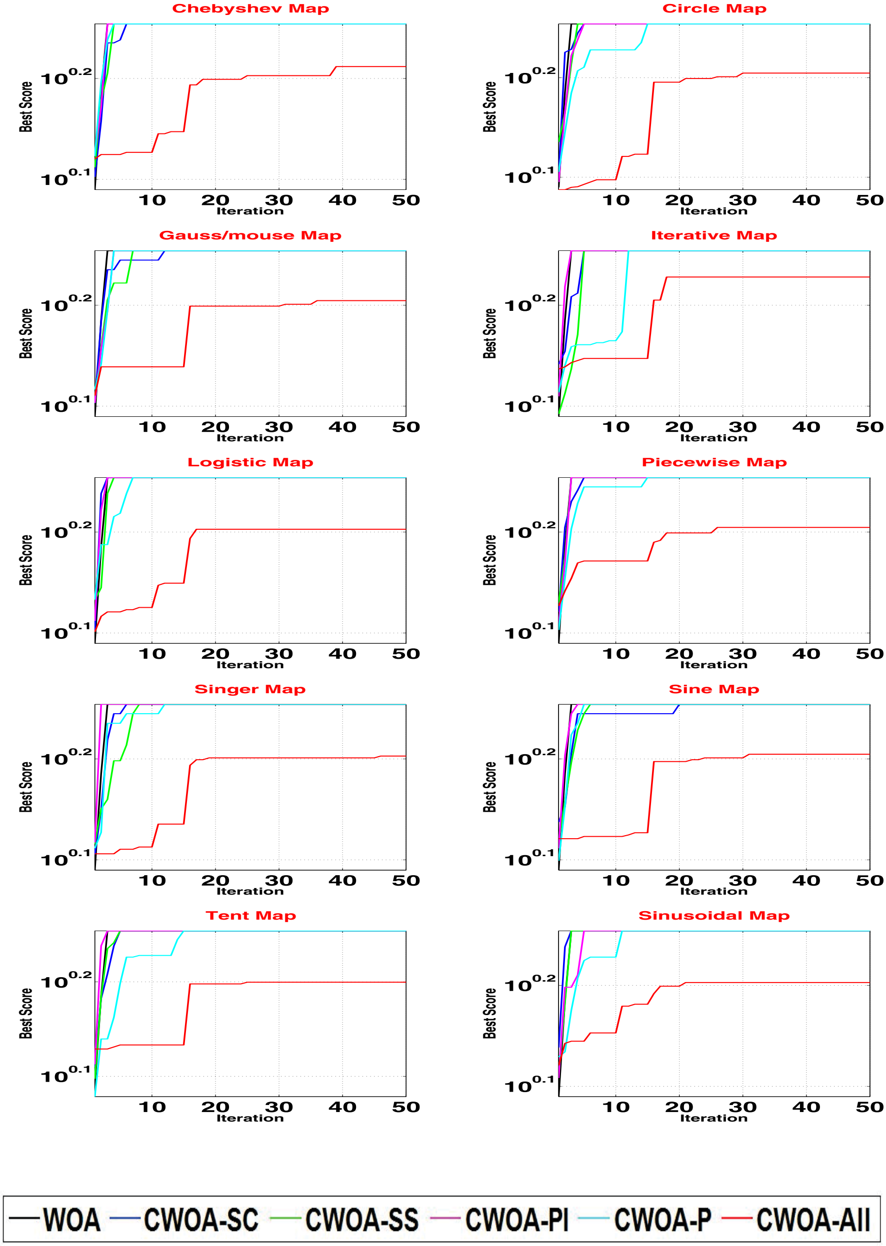

Figure 8. Convergence Curves for **Hepatitis Dataset** using Different Chaotic Maps (see online version for color)

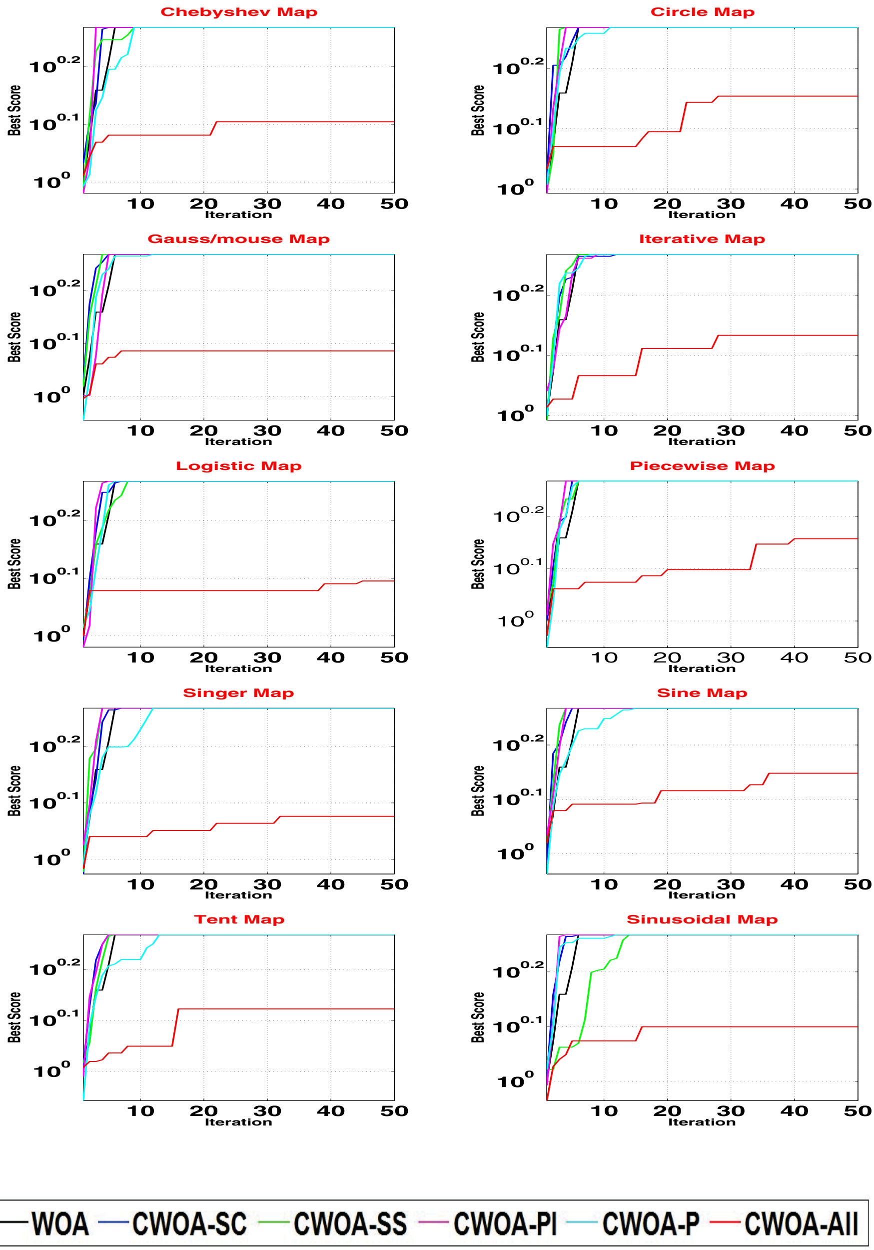

Figure 9. Convergence Curves for**Lung Cancer Dataset** using Different Chaotic Maps (see online version for color)

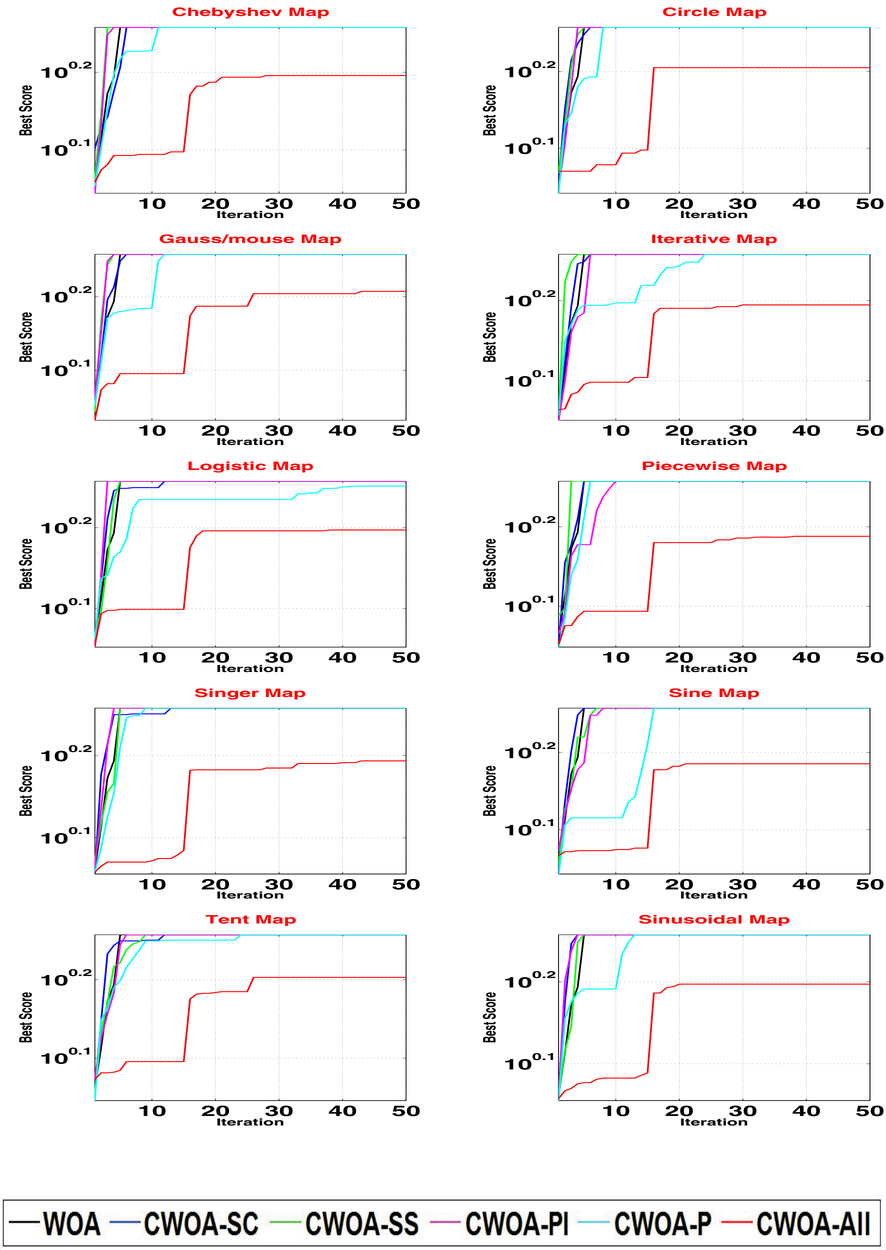

Figure 10. Convergence Curves for **SPECTF Heart Dataset** using Different Chaotic Maps (see online version for color)

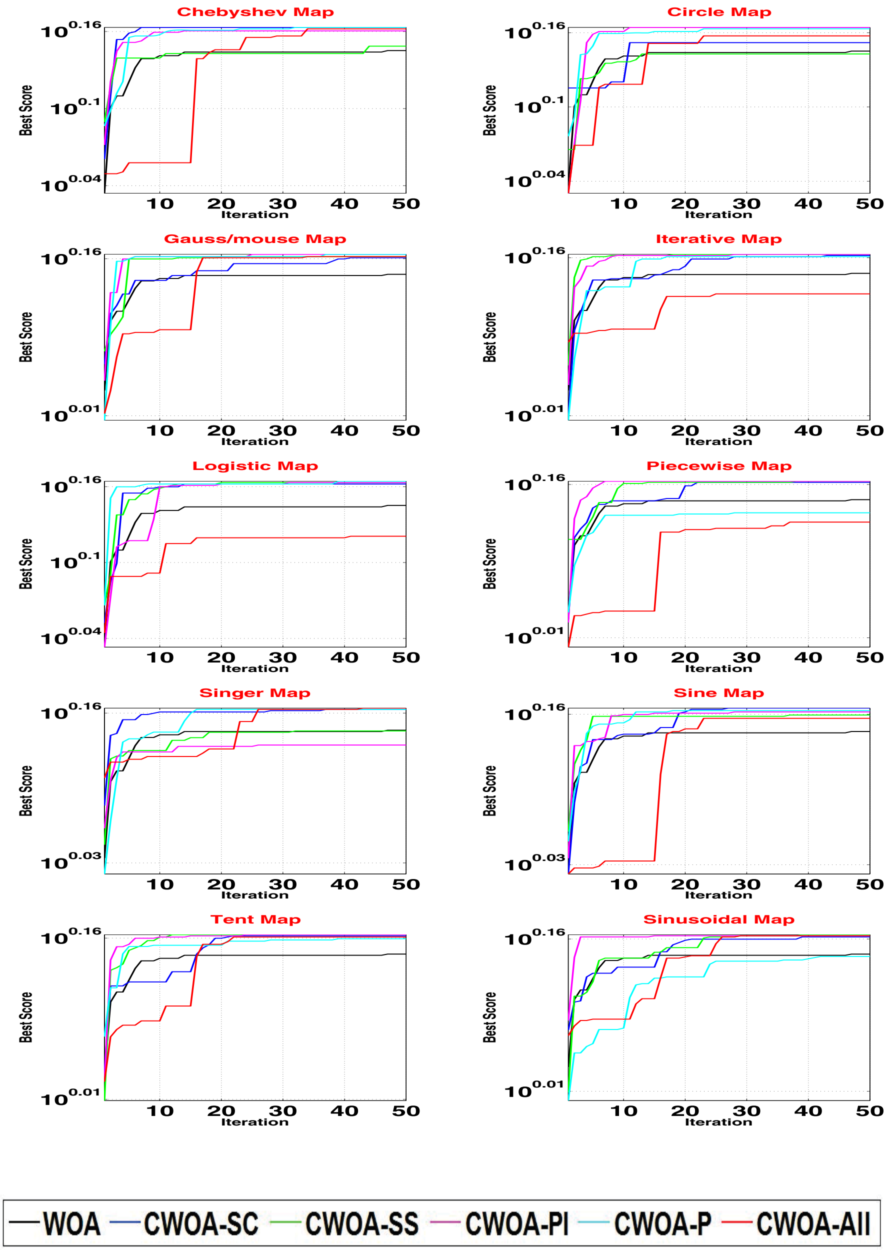

Figure 11. Convergence Curves for **Stat-log Heart Dataset** using Different Chaotic Maps (see online version for color)

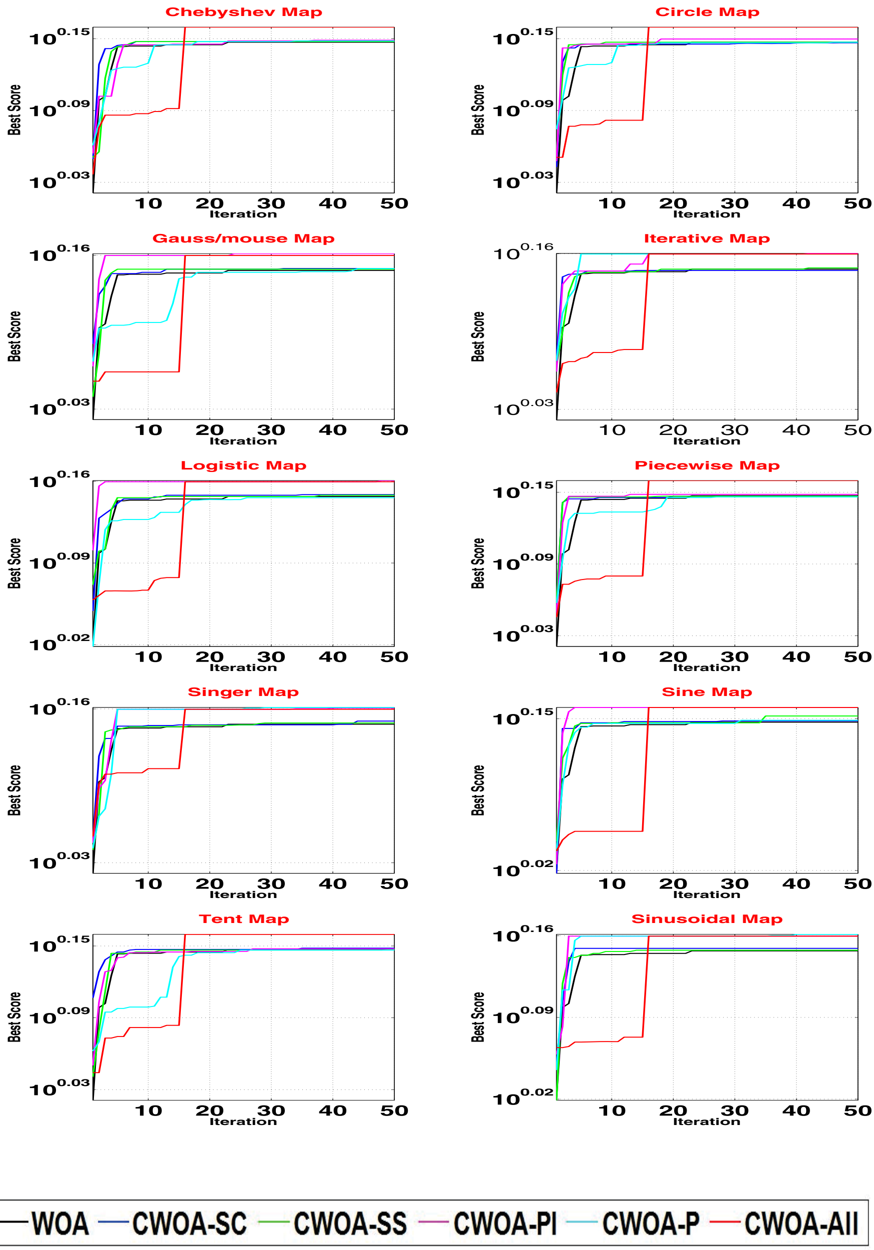

Figure 12. Convergence Curves for **Indian Liver Patient Dataset** using Different Chaotic Maps (see online version for color)

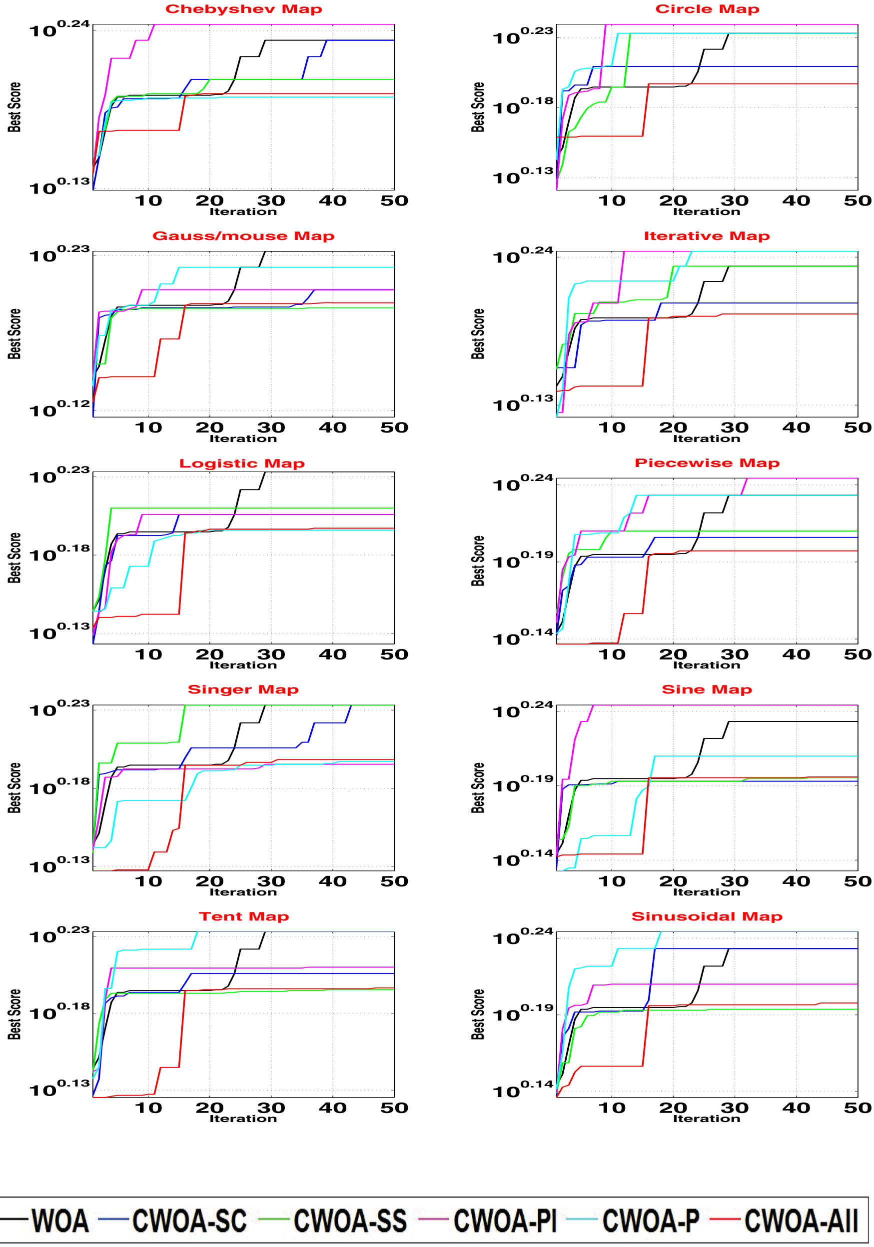

Figure 13. Convergence Curves for **Thoracic Surgery Dataset** using Different Chaotic Maps (see online version for color)

Table 10. Comparison CWOAs with Other Optimization Algorithms in Terms of **Average Features Selection Siz**, where D1-D10 are WDBC, MPED, PDD, Cardiotocography, hepatitis, Lung Cancer, SPECTF Heart, Thoracic Surgery, Statlog Heart and ILP datasetes

Data sets →	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**	**D9**	**D10**
PSO	0.76	0.37	0.46	0.35	0.41	0.55	0.47	0.48	0.30	0.32
ABC	0.48	0.58	0.49	0.45	0.47	0.62	0.52	0.43	0.45	0.29
CSO	0.34	0.33	0.34	0.34	0.41	0.36	0.43	0.42	0.28	0.40
BBO	0.51	0.19	0.11	0.12	0.09	**0.02**	0.04	0.22	0.20	0.32
EHO	0.55	0.64	0.56	0.59	0.60	0.64	0.62	0.56	0.59	0.64
KH	0.16	0.45	0.10	**0.10**	0.05	0.03	**0.03**	0.08	0.31	**0.10**
BSA	0.31	0.19	0.37	0.40	0.46	0.34	0.12	0.47	0.38	0.29
FPA	0.14	0.21	0.18	0.18	0.16	0.15	0.17	0.22	0.25	0.12
MFO	0.80	0.75	0.79	0.80	0.73	0.56	0.80	0.53	0.76	0.73
GWO	0.47	0.53	0.53	0.44	0.51	0.59	0.51	0.50	0.51	0.41
WOA	0.35	0.54	0.38	0.23	0.06	0.05	0.05	0.53	0.20	0.18
CWOA_SC	0.12	0.11	**0.06**	**0.10**	**0.05**	0.03	**0.03**	0.08	0.17	**0.10**
CWOA_SS	0.10	0.13	0.20	**0.10**	0.08	0.04	**0.03**	0.11	0.22	**0.10**
CWOA_PI	**0.04**	**0.09**	**0.06**	**0.10**	0.06	0.03	**0.03**	0.08	0.38	**0.10**
CWOA_P	0.22	**0.09**	0.07	0.15	0.08	0.04	0.13	**0.07**	**0.16**	0.11
CWOA_All	0.28	0.42	0.30	0.35	0.37	0.88	0.36	0.26	0.31	0.34

maps on all the used datasets follow the order of circle, sinusoidal, iterative, piecewise, chebyshev, singer, logistic, gauss/mouse, tent and sine, respectively. Finally, the results from convergence curves suggest the superior exploration of the circle map does not have a negative impact on the exploitation.

5.5.1 Comparative Analysis Against Related Optimization Algorithms

Tables 10, 11 and 12 record the average best score, average features selection size and standard deviation from 50 runs, respectively, where D1, D2, D3, D4, D5, D6, D7, D8, D9 and D10 are WDBC, MPED, PDD, Cardiotocography, hepatitis, Lung Cancer, SPECTF Heart, Thoracic Surgery, Statlog Heart and ILP datasets. Table 12 compares the average best score obtained during the runs. As can be seen, the performance of CWOAs and KH are comparable. CWOAs algorithms perform better than other optimization algorithms on seven (D1-2, D4-6, D8 and D10) of the ten benchmark datasets while KH performs better for the rest of the datasets (D3, D7 and

Table 11. Comparison CWOAs with Other Optimization Algorithms in Terms of **Stability Quality**, where D1-D10 are WDBC, MPED, PDD, Cardiotocography, hepatitis, Lung Cancer, SPECTF Heart, Thoracic Surgery, Statlog Heart and ILP datasetes

Data sets →	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**	**D9**	**D10**
PSO	0.04	0.05	0.05	0.05	0.09	0.05	0.03	0.03	0.03	0.03
ABC	0.03	0.02	0.04	0.04	0.12	0.03	0.03	0.12	0.03	0.06
CSO	0.06	0.05	0.07	0.04	0.06	0.14	0.07	0.03	0.04	0.03
BBO	0.08	0.10	0.04	0.04	0.04	0.10	0.08	0.04	0.05	0.02
EHO	0.07	**0.01**	0.01	0.01	0.01	0.00	**0.01**	0.02	0.02	0.02
KH	0.15	0.05	0.08	0.03	0.02	0.12	0.06	0.03	0.24	0.03
BSA	0.11	0.09	0.07	0.05	0.12	0.18	0.15	0.03	0.05	0.06
FPA	0.05	0.04	0.07	0.04	0.08	0.13	0.12	0.03	0.05	0.02
MFO	**0.02**	**0.01**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	**0.01**	**0.00**
GWO	0.04	0.02	0.04	0.04	0.08	0.15	0.04	0.03	0.03	0.04
WOA	0.06	0.07	0.05	0.05	0.09	0.17	0.13	0.09	0.05	0.06
CWOA_SC	0.05	0.05	0.04	0.03	0.05	0.12	0.10	0.02	0.03	0.02
CWOA_SS	0.04	0.04	0.04	0.03	0.07	0.13	0.09	0.02	0.02	0.03
CWOA_PI	0.04	0.04	0.04	0.03	0.06	0.13	0.09	0.03	0.03	0.02
CWOA_P	0.05	0.05	0.06	0.03	0.08	0.13	0.12	0.02	0.04	0.04
CWOA_All	0.11	0.13	0.13	0.06	0.10	0.30	0.14	0.04	0.04	0.08

D9). WOA and BBO are found to be the second most effective algorithms, while MFO and EHO, in most cases, obtained the worst results. In addition, it can be observed CWOA-PI provides the best results for most cases. Table 11 compares the stability of CWOA algorithms with different optimization algorithms. As can be observed, MFO and KH perform better than other optimization algorithms. However, COWAs and GWO are in second place and in the most cases, BSA provides the worst stability quality. Table 10 compares the average feature subset size of CWOAs and other optimization algorithms. As can be seen, CWOA, especially CWOA-PI, outperforms other optimization algorithms. KH and BBO are in second place, while BSA obtains the worst results in the most cases.

In this work, we use a similar approach for evaluation as in Steinley and Brusco (2007). A non-parametric statistical test, namely Wilcoxon's rank-sum test is adopted to evaluate and compare CWOA of modified PI parameter using circle chaotic maps (as it is found the optimal) with other algorithms. Wilcoxon's rank-sum test is adopted to judge whether the produced

Table 12. Comparison CWOAs with Other Optimization Algorithms in Terms of **Mean Best Score**, where D1-D10 are WDBC, MPED, PDD, Cardiotocography, hepatitis, Lung Cancer, SPECTF Heart, Thoracic Surgery, Statlog Heart and ILP datasetes

Data sets →	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**	**D9**	**D10**
PSO	1.43	1.49	1.40	1.42	1.47	1.15	1.21	1.42	1.27	1.25
ABC	1.37	1.33	1.34	1.37	1.42	1.10	1.22	1.46	1.09	1.28
CSO	1.46	1.50	1.40	1.44	1.40	1.36	1.27	1.44	1.24	1.19
BBO	1.55	1.62	1.63	1.52	1.79	1.81	1.80	1.58	1.35	1.39
EHO	1.38	1.33	1.33	1.31	1.36	1.12	1.21	1.39	1.19	1.14
KH	1.66	**1.75**	**1.65**	**1.58**	**1.79**	1.82	**1.82**	1.56	**1.47**	1.40
BSA	1.49	1.64	1.40	1.41	1.45	1.39	1.59	1.47	1.24	1.30
FPA	1.61	1.61	1.54	1.56	1.70	1.60	1.51	1.63	1.34	1.38
MFO	1.38	1.33	1.37	1.36	1.40	1.16	1.24	1.40	1.09	1.08
GWO	1.36	1.37	1.32	1.37	1.40	1.41	1.22	1.41	1.17	1.18
WOA	1.64	1.69	1.57	1.57	1.78	1.80	1.77	1.63	1.38	1.38
CWOA_SC	1.66	1.73	1.59	**1.58**	**1.79**	**1.83**	1.79	1.59	1.44	1.40
CWOA_SS	1.64	1.70	1.62	**1.58**	1.78	1.82	1.79	**1.66**	1.39	1.40
CWOA_PI	**1.67**	**1.75**	1.64	**1.58**	**1.79**	1.82	1.79	1.61	1.45	**1.44**
CWOA_P	1.66	1.73	1.63	**1.58**	1.77	1.82	1.75	1.55	1.44	1.43
CWOA_All	1.60	1.60	1.56	1.49	1.55	1.20	1.44	1.54	1.30	1.39

Table 13. CWOA vs. other meta-heuristic algorithms in terms of P-value of Wilcoxon's rank-sum test

Data sets →	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**	**D9**	**D10**
CWOA vs. PSO	1.61E-06	1.17E-06	2.01E-06	1.23E-07	9.35E-12	4.05E-11	4.65E-07	2.57E-04	1.89E-05	6.86E-04
CWOA vs. ABC	7.81E-07	3.19E-07	1.01E-05	1.54E-06	1.03E-09	2.14E-11	1.22E-10	1.90E-04	1.94E-05	2.70E-05
CWOA vs. CSO	1.79E-06	1.17E-06	4.36E-04	1.48E-07	2.12E-09	1.25E-09	2.23E-07	2.35E-03	1.80E-05	1.51E-05
CWOA vs. BBO	7.38E-03	1.13E-06	2.01E-03	2.23E-03	1.23E-05	8.00E-01	2.46E-01	1.23E+04	5.41E-04	1.24E-03
CWOA vs. EHO	7.72E-07	1.56E-07	1.60E-01	1.23E-06	2.56E-11	1.25E-07	1.24E-07	2.13E-03	2.92E-06	4.64E-05
CWOA vs. KH	1.75E-03	2.46E-03	2.01E-03	1.23E-02	5.41E-01	8.00E-01	1.23E-01	2.23E-03	4.24E-06	3.39E-01
CWOA vs. BSA	9.54E-07	1.24E-04	1.15E-05	2.26E-05	1.25E-06	1.23E-03	1.25E-06	1.46E-04	1.24E-06	2.35E-05
CWOA vs. FPA	2.05E-05	2.15E-06	2.01E-05	1.24E-05	1.27E-08	1.24E-07	2.34E-06	5.33E-03	2.33E-05	5.32E-04
CWOA vs. MFO	7.35E-07	1.22E-07	1.26E-06	1.88E-06	1.25E-09	1.29E-05	5.32E-08	1.25E-02	1.80E-06	1.25E-04
CWOA vs. GWO	7.79E-07	2.23E-06	1.23E-05	2.35E-06	1.24E-06	2.22E-05	5.35E-07	2.47E-03	5.41E-07	5.27E-05
CWOA vs. WOA	1.17E-04	2.35E-02	1.23E-03	1.25E-05	1.20E-04	2.35E-05	5.33E-03	1.24E-01	1.23E-05	3.15E-01

results of the algorithms differ from each other in a statistically significant way. Table 13 shows P-values of CWOA verse other ten meta-heuristic algorithms. As it can be observed from this table, the P values acquired prove that the superiority of CWOA is statistically significant. Moreover, it can be observed that the performance of KH and BBO are close to CWOA. In addition, it can be seen that CWOA is statistically significant compared to the original WOA. These results are consistent with the obtained results from Tables 10 and 12.

From previous results, CWOAs algorithms showed for ten benchmark datasets that they are capable of improving the performance of WOA in terms of local optima avoidance and convergence speed. Thus demonstrating the capability of CWOAs to find optimal feature combinations, which provides highest classification performance with a small number of features. Additionally, it can be seen, CWOAs can solve feature selection problems better than comparable algorithms.

Finally, it should be noted that, CWOA has another advantage, as it needs fewer parameters to adjust compared with other population-based optimization algorithms. As the algorithm's performance relies mainly on those parameters where it is more difficult to select the best parameter settings for any optimization algorithm (Gandomi et al., 2013). Therefore, this paper presents a new algorithm to deal with this problem by using chaotic maps instead of these parameters. In CWOA, chaos theory is adopted to adjust the key parameters of WOA.

6. Conclusion and Future Work

In this paper, a novel optimization algorithm based on chaotic and whale optimization algorithms (CWOA) is proposed using ten chaotic maps for global optimization of feature selection. The properties of the chaotic systems are used, such as regularity and semi-stochastic, to improve the performance of the WOA algorithm. Ten chaotic maps are adopted and compared to improve the performance of the WOA algorithm. Ten benchmark datasets collected from UCI repository are employed to compare the performance of CWOA on enhancing exploration and exploitation in terms of classification performance, stability quality, number of selected features and convergence speed. In addition, the performance of CWOA is compared and evaluated with original WOA and ten other optimization algorithms. The experimental results show that tuned WOA with chaotic maps CWOA is able to significantly improve the performance of WOA and enhance the quality of the solutions. Moreover, the results of the comparative study of different version of CWOA with different chaotic maps suggest that the circle map is the best map. Also, chaotic WOA with modification of PI parameter obtains the highest results. The results showed that chaotic version of WOA with modification of all parameters is not able to improve significantly the performance of the WOA. Finally, CWOA outperforms other comparative optimization algorithms. In the future, it will be needed to evaluate CWOA in solving several complex sciences and engineering optimization problems.

References

ABDULLAH, A., ENAYATIFA, R., and LEE, M. (2012), "A Hybrid Genetic Algorithm and Chaotic Function Model for Image Encryption", *Journal of Electronics and Communication, 66*, 806–816.

ALATAS, B., AKIN, E., and OZER, A. (2009), "Chaos Embedded Particle Swarm Optimization Algorithms", *Journal of Chaos Soliton and Fractals, 4*, 1715–1734.

BACHE, K., and LICHMAN, M. (2016), "UCI Machine Learning Repository", http://archive.ics.uci.edu/ml.

CHAOSHUN, L., XUELI, A., and RUHAI, L. (2015), "A Chaos Embedded Gsa-Svm Hybrid System for Classification", *Neural Computing and Applications, 26*, 713–721.

CHUANG, L., YANG, C., and LI, J. (2011), "Chaotic Maps Based on Binary Particle Swarm Optimization for Feature Selection", *Applied Soft Computing, 11*, 239–248.

DERRAC, J., GARCÍA, S., MOLINA, D., and HERRERA, F. (2011), "A Practical Tutorial on the Use of Nonparametric Statistical Tests as a Methodology for Comparing Evolutionary and Swarm Intelligence Algorithms", *Swarm Evoloutionary Computation, 1*, 3–18.

EBRAHIMI, A., and KHAMEHCHI, E. (2016), "Sperm Whale Algorithm: An Effective Metaheuristic Algorithm for Production Optimization Problems", *Journal of Natural Gas Science and Engineering, 29*, 211–222.

EMARY, E., ZAWBAA, H., and HASSANIEN, A. (2016), "Binary Grey Wolf Optimization Approaches for Feature Selection", *Neurocomputing, 172*, 371–381.

GADAT, S., and YOUNES, L. (2007), "A Stochastic Algorithm for Feature Selection in Pattern Recognition", *Journal of Machine Learning Research, 8*, 509–547.

GAI-GE, W., SUASH, D., LEANDRO, D., and COELHO, S. (2015), "Elephant Herding Optimization", in *3rd International Symposium on Computational and Business Intelligence (ISCBI)*, Bali, pps 1–5.

GANDOMI, A., and ALAVI, A. (2012), "Krill Herd: A New Bio-Inspired Optimization Algorithm", *Communications in Nonlinear Science and Numerical Simulation, 17(12)*, 4831–4845.

GANDOMI, A., GUN, G., YANG, X., and TALATAHARI, S. (2013), "Chaos-Enhanced Accelerated Particle Swarm Optimization", *Communication Nonlinear Science and Numerical Simulation, 18(2)*, 327–340.

GOLDBOGEN, J., FRIEDLAENDER, A., CALAMBOKIDIS, J., MCKENNA, M., SIMON, M., and NOWACEK, D. (2013), "Integrative Approaches to the Study of Baleen Whale Diving Behavior, Feeding Performance, and Foraging Ecology", *BioScience, 69*, 90–100.

HOF, P., and VAN, E. (2007), "Structure of the Cerebral Cortex of the Humpback Whale", *Megaptera Novaeangliae (Cetacea, Mysticeti, Balaenopteridae), 290*, 1–31.

HOLLAND, J.H. (1992), *Adaptation in Natural and Artificial Systems*, Cambridge MA: MIT Press.

HOSSEINPOURFARD, R., and JAVIDI, M. (2015), " Chaotic Pso Using the Lorenz System: An Efficient Approach for Optimizing Nonlinear Problems", *Çankaya University Journal of Science and Engineering, 12(1)*, 40–59.

HUANG, C., and WANG, C. (2006), "A Ga-Based Feature Selection and Parameters Optimization for Support Vector Machines", *Expert Systems with Applications, 31*, 231–240.

KARABOGA, D. (2005), "An Idea Based on Honey Bee Swarm for Numerical Optimization", Technical Report TR06, Erciyes University, Engineering Faculty, Computer Engineering Department.

KENNEDY, J., and EBERHART, R. (1995), "Particle Swarm Optimization", in *IEEE International Conference on Neural Networks*, pps. 1942–1948.

LI, C., ZHOU, J., and KOU, P. (2012a), " A Novel Chaotic Particle Swarm Optimization Based Fuzzy Clustering Algorithm", *Neurocomputing, 83*, 98–109.

LI, C., ZHOU, J., XIAO, J., and XIAO, H. (2012b), "Parameters Identification of Chaotic System by Chaotic Gravitational Search Algorithm", *Chaos Solitons Fractals, 45(4)*, 559–547.

LIU, B., WANG, L., and JIN, Y. (2005), "Improved Particle Swarm Optimization Combined with Chaos", *Chaos Solitons Fractals, 25*, 1261–1271.

LIU, F., and ZHOU, Z. (2015), "New Data Classification Method Based on Chaotic Particle Swarm Optimization and Least Square-Support Vector Machine", *Chemometrics and Intelligent Laboratory Systems, 147*, 147–156.

MENG, X., LIU, Y., GAO, X., and ZHANG, H. (2014), "A New Bio-Inspired Algorithm: Chicken Swarm Optimization", in *Advances in Swarm Intelligence: 5th International Conference, ICSI*, pps. 86–94.

MENG, X., GAO, X.Z., LU, L., LIU, Y., and ZHANG, H. (2016), "A New Bio-Inspired Optimisation Algorithm: Bird Swarm Algorithm", *Journal of Experimental and Theoretical Artificial Intelligence, 28(4)*, 673–687.

MIRJALILI, S. (2015), "Moth-Flame Optimization Algorithm: A Novel Nature-Inspired Heuristic Paradigm", *Knowledge-Based Systems, 89*, 228 – 249.

MIRJALILI, S., and LEWIS, A. (2016), " The Whale Optimization Algorithm", *Advances in Engineering Software, 95*, 51–67.

MIRJALILI, S., SEYED, M., and LEWIS, A. (2014), "Grey Wolf Optimizer", *Advanced Engineering Software, 69*, 46–61.

ÖZKAYNAK, F. (2015), "A Novel Method to Imrove the Performance of Chaos Based Evolutionary Algorithms", *Optik, 126*, 5434–5438.

RATTENBORG, N., AMLANER, C., and LIMA, S. (2000), "Behavioral, Neurophysiological and Evolutionary Perspectives on Unihemispheric Sleep", *Neuroscience and Biobehavioral Reviews, 24*, 817–842.

SANTOS, D., LUVIZOTTO, L., MARIANI, V., and COELHO, L. (2012), "Least Squares Support Vector Machines with Tuning Based on Chaotic Differential Evolution Approach Applied to the Identification of a Thermal Process", *Expert Systems with Applications, 39*, 4805–4812.

SAREMI, S., MIRJALILI, S., and LEWIS, A. (2014a), "Biogeography-Based Optimization with Chaos", *Neural Computing and Applications, 25*, 1077–1097.

SAREMI, S., MIRJALILI, S., and LEWIS, A. (2014b), "Chaotic Krill Herd Optimization Algorithm", *Procedia Technology, 12*, 180–185.

SHEIKHPOUR, R., SARRAMA, M., and SHEIKHPOUR, R. (2016), "Particle Swarm Optimization for Bandwidth Determination and Feature Selection of Kernel Density Estimation Based Classifiers in Diagnosis of Breast Cancer", *Applied Soft Computing, 40*, 113–131.

SHOUBAO, S., YU, S., and MINGJUAN, X. (2014), "Comparisons of Firefly Algorithm with Chaotic Maps", *Computer Modeling and New Technologies, 18(12)*, 326–332.

SIMON, D., and CLEVELAND, O. (2008), "Biogeography-Based Optimization", *IEEE Transactions on Evolutionary Computation, 12(6)*, 702–713.

SPROTT, J. (2010), *Elegant Chaos Algebraically Simple Chaotic Flows*, Singapore: World Scientific.

STEINLEY, D., and BRUSCO, M.J. (2007), "Initializing K-Means Batch Clustering: A Critical Evaluation of Several Techniques", *Journal of Classification, 24(1)*, 99–121.

STROGATZ, S. (1994), *Nonlinear Dynamics and Chaos*, Singapore: Perseus Books Publishing.

WANG, G., GUO, L., AMIR, H., HAO, G., and WANG, H. (2014), "Chaotic Krill Herd Algorithm", *Information Sciences, 274*, 17–34.

WANG, N., LIU, L., and LIU, L. (2001), "Genetic Algorithm in Chaos", *OR Transaction, 5*, 1–10.

WATKINS, W., and SCHEVILL, W. (1979), "Aerial Observation of Feeding Behavior in Four Baleen Whales: Eubalaena Glacialis , Balaenoptera Borealis , Megaptera Novaean-Gliae , and Balaenoptera Physalus", *Journal of Mammalogy, 60(1)*, 155–163.

WILCOXON, F. (1945), "Individual Comparisons by Ranking Methods", *Biometrics Bulletin, 1*, 80–83.

WU, Q. (2011), "A Self-Adaptive Embedded Chaotic Particle Swarm Optimization for Parameters Selection of Wv-Svm", *Expert Systems with Applications, 38*, 184–192.

YANG, L., and CHEN, T. (2002), "Application of Chaos in Genetic Algorithms", *Communications in Theortical Physics, 38*, 168–172.

YANG, X. (2012), "Flower Pollination Algorithm for Global Optimization", in *Proceedings of the 11th International Conference on Unconventional Computation and Natural Computation*, Berlin, Heidelberg: Springer, pp. 240–249.

ZAWBAA, H., EMARY, E., and GROSAN, C. (2016), "Feature Selection via Chaotic Antlion Optimization", *Plos One, 11(3)*.