

Article

Simultaneous Feature Selection and Support Vector Machine Optimization Using an Enhanced Chimp Optimization Algorithm

Di Wu ¹, Wanying Zhang ², Heming Jia ^{3,*} and Xin Leng ²¹ School of Education and Music, Sanming University, Sanming 365004, China; wudi@fjssmu.edu.cn² College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin 150040, China; zhangwanying@nefu.edu.cn (W.Z.); lengxin@nefu.edu.cn (X.L.)³ School of Information Engineering, Sanming University, Sanming 365004, China

* Correspondence: jiaheming@fjssmu.edu.cn

Abstract: Chimp Optimization Algorithm (ChOA), a novel meta-heuristic algorithm, has been proposed in recent years. It divides the population into four different levels for the purpose of hunting. However, there are still some defects that lead to the algorithm falling into the local optimum. To overcome these defects, an Enhanced Chimp Optimization Algorithm (EChOA) is developed in this paper. Highly Disruptive Polynomial Mutation (HDPM) is introduced to further explore the population space and increase the population diversity. Then, the Spearman's rank correlation coefficient between the chimps with the highest fitness and the lowest fitness is calculated. In order to avoid the local optimization, the chimps with low fitness values are introduced with Beetle Antenna Search Algorithm (BAS) to obtain visual ability. Through the introduction of the above three strategies, the ability of population exploration and exploitation is enhanced. On this basis, this paper proposes an EChOA-SVM model, which can optimize parameters while selecting the features. Thus, the maximum classification accuracy can be achieved with as few features as possible. To verify the effectiveness of the proposed method, the proposed method is compared with seven common methods, including the original algorithm. Seventeen benchmark datasets from the UCI machine learning library are used to evaluate the accuracy, number of features, and fitness of these methods. Experimental results show that the classification accuracy of the proposed method is better than the other methods on most data sets, and the number of features required by the proposed method is also less than the other algorithms.

Citation: Wu, D.; Zhang, W.; Jia, H.; Leng, X. Simultaneous Feature Selection and Support Vector Machine Optimization Using an Enhanced Chimp Optimization Algorithm. *Algorithms* **2021**, *14*, 282. <https://doi.org/10.3390/a14100282>

Academic Editor: Frank Werner

Received: 28 August 2021

Accepted: 25 September 2021

Published: 28 September 2021

Keywords: Chimp Optimization Algorithm; highly disruptive polynomial mutation; Spearman's rank correlation coefficient; Beetle Antennae Search Algorithm; support vector machine

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With society gradually becoming digitalized, the question of how to extract useful information effectively from complex and huge data has become the focus of research in recent years. Machine learning is one of the important research fields of information recognition and pattern recognition in data sets [1]. It has been widely used due to its strong data processing ability. Machine learning is evolving from computer science [2], which is a multidisciplinary and interdisciplinary major. It covers probability theory, statistics, approximate theory, and complex algorithms, and it can design efficient and accurate prediction algorithms [3]. It can be divided into supervised learning and unsupervised learning according to the learning mode [4]. SVM is one of the most widely used supervised learning algorithms. Vapnik and other researchers built the cornerstone of SVM [5]. It is widely used in pattern recognition [6,7], text recognition [8], biomedical [9–

11], imaging medicine [12], anomaly detection [13] and in other fields. SVM classifies feature spaces into two categories and defines a non-parametric method for finding decision boundaries [14]. SVM has shown excellent performance in many problems [15]. Using the support vector machines to solve problems can improve generalization performance, improve computational efficiency, reduce running time, and produce a very accurate classification model [16,17]. The Support Vector Machines can solve both linear and nonlinear classification problems. When the data set is linearly inseparable, the nonlinear problem is usually transformed into a linear problem. Then, the data set can be used to construct an optimal hyperplane in the feature space, so that the feature space can be divided into two classes more easily [18]. However, when machine learning processes high-dimensional data sets, there will be problems such as noise and redundant data [19]. In this case, feature selection [20] can be used to reduce the number of features.

Feature selection is to select the optimal feature subset from the original data set [21]. According to the form of feature selection, the feature selection methods can be divided into the following three types: filter, wrapper, and embedded [22–24]. The wrapper-based approach is used in this article. The wrapper method, which relies on the classification method [25], is more efficient than the other two, but its computational intensity is relatively greater [26]. The filter method [27] uses numerous indicators to evaluate and select high-order features according to their discriminant attributes [28]. After that it will select the subset with the richest information [29]. The main idea of the wrapper approach is to treat the selection of subsets as a search optimization problem. At first it will generate different combinations, then evaluate them, and finally compare them with other combinations. The wrapper method is often used in feature selection problems due to the superiority of its calculation results, and the wrapper method can directly interact with classifier [30]. The wrapper method can be effectively combined with the meta-heuristic optimization algorithm so to achieve better practical application effect.

In recent years, many researchers have combined optimization algorithms with SVM to solve problems. Influenced by this, this paper tries to combine ChOA with SVM for feature selection and parameter optimization. ChOA is a new meta-heuristic optimization algorithm proposed by Khishe and Mosavi in 2019 [31] which is inspired by the hierarchy mechanism of chimps in nature when they are hunting. It can solve the problem of slow convergence and local optimum when it is solving the serious dimensional problems. However, the ChOA still has some shortcomings. Firstly, the population diversity of the algorithm is insufficient in the initial stage. Secondly, there is a risk of falling into local optimal in the final search stage. Therefore, an Enhanced Chimp Optimization Algorithm is proposed in this paper. Three strategies have been introduced successively, including highly disruptive polynomial mutation [32], Spearman's correlation coefficient, and Beetle Antennae Search Algorithm [33]. In the initial stage, HDPM is used to increase the population's variety, which improved the detection ability of the optimization algorithm. Due to the updating of the position of the chimp being determined by each level of chimp, this paper will improve the global search ability of chimps by improving the position updating ability of lower level chimps. The algorithm firstly introduces the Spearman's rank correlation coefficient to calculate the distance between low grade and high grade chimp. For the lower chimps that are far away from the higher chimps, this paper will introduce the beetle antenna search algorithm. It can make the chimps with low fitness achieve the visual ability, thus they can change their movement direction according to the surrounding environment. This strategy improves the local and global search ability of the chimps with lower fitness. On this basis, this paper proposes an EChOA-SVM model. This model is used for feature selection and parameter optimization at the same time to obtain good classification accuracy and performance. The main contributions of this paper are summarized as follows:

1. An Enhanced Chimp Optimization Algorithm is proposed to solve the shortcomings of the ChOA and make it better applied to feature selection problems.
2. HDPM strategy is introduced in the initial stage to enhance the population diversity.

3. Spearman's rank correlation coefficient helps to identify the chimps that need to be improved. Then BAS is introduced to improve the positions updating ability.
4. The EChOA-SVM model is used for feature selection and SVM parameter optimization simultaneously. The model is evaluated by 17 benchmark data sets in UCI machine learning library [34]. In order to verify the effectiveness of this method, it is compared with seven optimization algorithms, such as ChOA [31], GWO [35], WOA [36], ALO [37], GOA [38], MFO [39], and SSA [40].

The structure of this article is as follows. Section two reviews relevant literatures. Section three briefly introduces the basic principle of the ChOA. Section four proposes a new EChOA, then introduces the related theory of EChOA-SVM. In Section five, compared and analyzed the experimental results.

2. Literature Review

Optimization algorithms have been widely used in different fields such as medicine, multi-objective optimization, data classification, feature selection, and Support Vector Machine optimization. Zhao and Zhang proposed a Learning-based Evolutionary Multi-objective Algorithm [41]. By comparing five algorithms, the proposed algorithm is significantly superior to the other compared algorithms in determining the convergence and the approximation of the Pareto front. Dulebenets proposed an Adaptive Polyploid Memetic Algorithm to solve the cross-docking terminal vehicle scheduling problem [42]. This algorithm can assist the correct operation planning of CDT. Liu, Wang and Huang proposed an Alternative Algorithm to deal with Multiobjective optimization problems [43] which is proven to be better than other many-objective evolutionary algorithms. Furthermore, it is easily extended to solve constrained multi-objective optimization problems. Junayed et al., established an optimization model and solution algorithm to optimize the factory-in-a-box supply chain [44]. Gianni et al., Established specific rules or formulas that can distinguish bacterial from viral meningitis by machine learning methods [45]. The method even achieved 100% accuracy in detecting bacterial meningitis. Panda and Majhi used the Salp Swarm Algorithm to train the Multilayer Perceptron for data classification [46]. Compared with other classical optimization algorithms, the results show that this method is very advantageous.

This paper mainly studies the application of optimization algorithm in feature selection and support vector machine optimization. Feature selection is to select the optimal feature subset from the original data set, which can be regarded as an optimization problem [47]. In order to obtain a better classification accuracy, it is necessary to input the optimal SVM parameters into the feature subset. The selection of the feature subset will also affect the parameters of SVM. Therefore, in order to obtain the ideal classification accuracy, both feature selection and optimization of SVM parameters are required [48]. Huang and Wang [49] proposed a GA algorithm for feature selection and support vector machine parameter optimization simultaneously. Compared with other methods, the proposed method can achieve higher classification accuracy with fewer features. Shih-Wei Lin et al. [50] proposed a feature selection and parameter optimization method based on SA, and compared the method with grid algorithm. The results showed that the proposed method could significantly improve the classification accuracy. Heming Jia and Kangjian Sun proposed an IBMO-SVM classification model [51]. This model can help SVM to find the optimal feature subset and parameters at the same time. Compared with other optimization algorithms, this model shows good performance on both low and high dimensional data sets. The EChOA-SVM model, which is proposed in this paper, is inspired by these methods.

3. Chimp Optimization Algorithm (ChOA)

ChOA is a mathematical model that is based on intelligent diversity [31]. Driving, chasing, blocking and attacking are accomplished by four different types of chimps, which

are accomplished by attackers, obstacles, chasers, and drivers. The four hunting steps are completed in two stages. The first stage is the exploration stage, and the second stage is the exploitation stage. The exploration stage includes driving, blocking, and chasing the prey. As for the exploitation stage, it has to attack the prey. In which, the driving and chasing are represented by Equations (1) and (2).

$$\mathbf{d} = |\mathbf{c} \cdot \mathbf{x}_{prey}(t) - \mathbf{m} \cdot \mathbf{x}_{chimp}(t)| \quad (1)$$

$$\mathbf{x}_{chimp}(t+1) = \mathbf{x}_{prey}(t) - \mathbf{a} \cdot \mathbf{d} \quad (2)$$

where $\mathbf{x}_{prey}$ is the vector of prey position, $\mathbf{x}_{chimp}$ is the vector of chimp position, t is the number of current iterations, $\mathbf{a}$, $\mathbf{c}$, $\mathbf{m}$ are coefficient vectors and they can be obtained through Equations (3)–(5).

$$\mathbf{a} = 2 \cdot \mathbf{f} \cdot \mathbf{r}_1 - \mathbf{f} \quad (3)$$

$$\mathbf{c} = 2 \cdot \mathbf{r}_2 \quad (4)$$

$$\mathbf{m} = \text{Chaotic_value} \quad (5)$$

where $\mathbf{f}$ non-linearly declined from 2.5 to 0, $\mathbf{r}_1$ and $\mathbf{r}_2$ is the random number between 0 and 1, and $\mathbf{m}$ is the chaotic vector. The dynamic coefficient $\mathbf{f}$ can be selected for different curves and slopes, thus chimps can use different abilities to search the prey.

Chimps can update their positions based on the other chimps, and this mathematical model can be represented by Equations (6) and (8).

$$\begin{aligned} \mathbf{d}_{Attacker} &= |\mathbf{c}_1 \mathbf{x}_{Attacker} - \mathbf{m}_1 \mathbf{x}| \\ \mathbf{d}_{Barrier} &= |\mathbf{c}_2 \mathbf{x}_{Barrier} - \mathbf{m}_2 \mathbf{x}| \\ \mathbf{d}_{Chaser} &= |\mathbf{c}_3 \mathbf{x}_{Chaser} - \mathbf{m}_3 \mathbf{x}| \\ \mathbf{d}_{Driver} &= |\mathbf{c}_4 \mathbf{x}_{Driver} - \mathbf{m}_4 \mathbf{x}| \end{aligned} \quad (6)$$

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{x}_{Attacker} - \mathbf{a}_1(\mathbf{d}_{Attacker}) \\ \mathbf{x}_2 &= \mathbf{x}_{Barrier} - \mathbf{a}_2(\mathbf{d}_{Barrier}) \\ \mathbf{x}_3 &= \mathbf{x}_{Chaser} - \mathbf{a}_3(\mathbf{d}_{Chaser}) \\ \mathbf{x}_4 &= \mathbf{x}_{Driver} - \mathbf{a}_4(\mathbf{d}_{Driver}) \end{aligned} \quad (7)$$

$$\mathbf{x}(t+1) = \frac{\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 + \mathbf{x}_4}{4} \quad (8)$$

4. Enhanced Chimp Optimization Algorithm (EChOA)

4.1. Highly Disruptive Polynomial Mutation (HDPM)

The highly disruptive polynomial mutation is the improved version of the polynomial mutation method. It can solve the shortcoming that the polynomial mutation method may fall into local optimum when the variable is close to the boundary [32]. Equations

(9)–(12) show the process of HDPM changes the x_i

$$\delta_1 = \frac{x_i - lb}{ub - lb} \quad (9)$$

$$\delta_2 = \frac{ub - x_i}{ub - lb} \quad (10)$$

$$\delta_k = \begin{cases} [(2r) + (1-2r) * (1-\delta_1)^{\eta_m+1}]^{\frac{1}{\eta_m+1}-1}, & r \leq 0.5 \\ 1 - [2(1-r) + 2(r-0.5) * (1-\delta_2)^{\eta_m+1}]^{\frac{1}{\eta_m+1}}, & \text{otherwise} \end{cases} \quad (11)$$

$$x_i = x_i + \delta_k(ub - lb) \quad (12)$$

where ub and lb represent the upper and lower boundaries of the search space. r is a random number between 0 and 1. η_m is a distribution exponential, which is a non-negative number. As can be seen from the above formula, HDPM can explore the entire search space.

4.2. Spearman's Rank Correlation Coefficient

Spearman's rank correlation coefficient is a statistical index to reflect the degree of the relationship between two groups of variables, which is figured upon a level basis. The formula for calculating as Equation (13).

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (13)$$

where d_i is the difference of grades among each pair of samples, and n is the dimension of the series. If the absolute value of the correlation coefficient is equal to 1, then the two series are monotonically correlated; otherwise, they are uncorrelated.

4.3. Beetle Antennae Search Algorithm (BAS)

The long-horned beetle has two too long antennae, which can combine the scent of prey to expand detection range and act as a protective alarm mechanism [33]. The beetle explores nearby areas by swinging its antennae on one side of its body to accept the smell. The beetle will move toward to the side where it detects a high odor concentration, as it is shown in Figure 1. The Beetle Antennae Search Algorithm is designed based on this property of beetles.



Figure 1. Schematic diagram of beetle movement [52].

The search direction of beetles is represented by Equation (14).

$$\vec{b} = \frac{rnd(k,1)}{\|rnd(k,1)\|} \quad (14)$$

where $\vec{b}$ is the direction vector of the beetle, rnd is a random function, and k is the dimension of the position.

Next, Equation (15) respectively presents the search behaviors on the left and right sides to simulate the activity tracking of the beetle.

$$\begin{aligned} x_l &= x^t + d^t \vec{b} \\ x_r &= x^t - d^t \vec{b} \end{aligned} \quad (15)$$

where x_l and x_r respectively represent the locations within the left and right search areas, x^t represents the beetle's position at t th time instant, and d^t represents the perceived length of the antenna, which will gradually decrease with the passage with time.

The position update of the beetle can be represented by Equation (16).

$$x^t = x^{t-1} + \delta^t \vec{b} \text{sign}(f(x_r) - f(x_l)) \quad (16)$$

where δ is the step size of the search, which initial scope should be equal to the search area, and the sign represents the sign function. The odor concentration at x is expressed by $f(x)$, which is also known as the fitness function.

4.4. Improvement Strategy

Although the original ChOA divides the chimps into four different levels to complete the hunting process, the algorithm still has two obvious defects when applied to higher dimensional problems. The first disadvantage is that in the initial stage of the population it lacks the population diversity. The other disadvantage is there is a risk of falling into the local optimal in the final search stage. Therefore, this article proposes an Enhanced Chimp Optimization Algorithm, which can be better applied to feature selection problems.

Three strategies including HDPM, Spearman's rank correlation coefficient, and BAS algorithm are introduced into EChOA based on the original algorithm. First, the HDPM strategy is introduced to enhance the population diversity in the initialization phase. The traditional polynomial mutation (PM) has almost no effect when the variable is close to the boundary. The HDPM strategy uses Equation (12) to generate the mutation location of the chimp, which helps further to explore the regions and boundaries of the initial space. In this way, even if the variables are on the search boundary, the search space can be fully utilized to ensure the diversity of the population. Second, the Spearman's rank correlation coefficient between the driver and attacker chimp is calculated. As it can be seen from Equation (7), position update is jointly determined by all chimps of different grades. By improving the position of the lower chimp, the population can avoid falling into local optimal effectively. The distance between the driver and the attacker can be determined by calculating the Spearman's rank correlation coefficient of the two by Equation (13). The two chimp species are negatively or uncorrelated when Spearman's rank correlation coefficient is less than or equal to 0. This will show whether the driver and the attacker are close or far away. Finally, the BAS is introduced for the chimp which is far away from the attacker. The position of the chimp is improved by using Equation (16), which makes it acquire the visual ability. Thus, they can judge the surrounding environment and decide the direction of movement. This increases the performance of the driver chimp and avoids local optimal.

4.5. EChOA for Optimizing SVM and Feature Selection

EChOA-SVM firstly uses optimization algorithm to search the optimal feature subset, and then classify the feature subset by SVM. In the feature selection, the positions of population are separated into 0 and 1 by logical function. If the feature corresponds to 1, the feature will be selected; if the feature corresponds to 0, the feature will not be selected. Since the selection of kernel function and its parameters is related to the performance of SVM, in order to obtain better classification accuracy, EChOA is also needed to optimize the parameters. The EChOA-SVM model generates the primary agents through the penalty parameter c and the nuclear parameter g , and calculates the fitness value to evaluate the chimp's positions. Finally outputs the optimal value of c , g and the best accuracy to obtain the final classification result.

The description of the EChOA-SVM is shown as follows. The flow chart is shown in Figure 2.

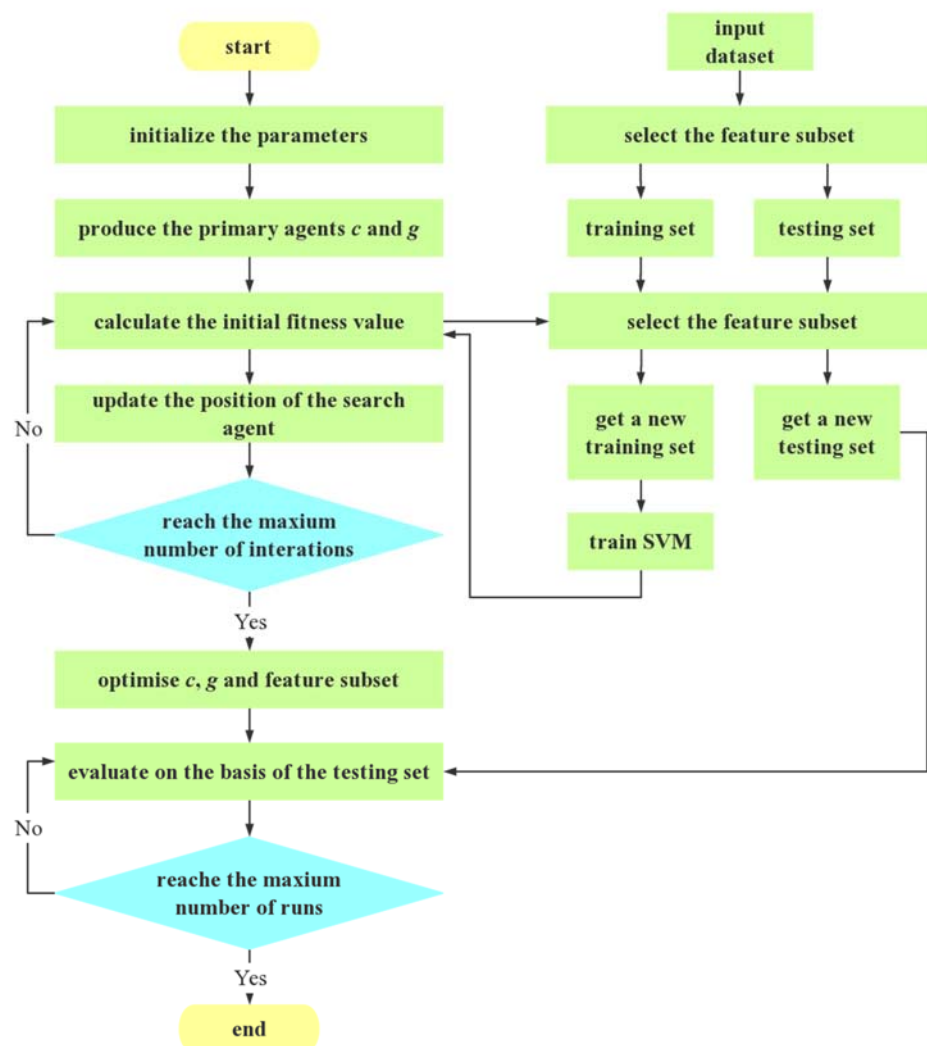


Figure 2. The flow chart of EChOA-SVM.

1. Initialize the population, position and fitness value of the chimps.
2. SVM randomly generates the values of c and g , and then calculates the initial fitness value.
3. The training model is obtained by training set in support vector machine.
4. Reorder the population to get the agent position and fitness value after sorting.

5. Experimental Results and Discussion

5.1. Datasets Details

The implementation of the proposed algorithm is done using Matlab. Seventeen data sets from the University of California at Irvine (UCI) machine learning repository [34] has been selected to evaluate the proposed EChOA-SVM approach. Table 1 describes the number of instances, the number of features, and the number of classes of the data set. Seventeen data sets of different sizes and latitudes are selected to observe the performance of the algorithm at different scales.

Table 1. List of data sets.

Data Set	Features	Instances	Classes
Balance	4	625	3
Bupa	6	345	2
Coimbra	9	116	2
Dermatology	34	358	6
Diagnostic	10	683	4
Divorce	18	170	2
Glass	10	214	6
Heart	13	270	2
Iris	4	150	3
Knowledge	5	258	4
Liver	6	345	2
Lymphography	18	148	8
Sonar	60	208	2
Teaching Assistant Evaluation (TAE)	5	151	3
Transfusion	4	748	2
Wine	13	178	3
Zoo	16	101	7

All the experimental series are carried out on Matlab R2016a, and the computer is configured as Intel(R) Core (TM) i5-1035G1 CPU @ 1.00GHz 1.19GHz, using Microsoft Windows 10 system. Each experiment is run ten times independently to reduce the random influence. In addition, the general parameters are set as follows: the population size is 30, and each run is set to 100 iterations as stopping criteria.

5.2. Parameters Setting

In order to verify the optimization performance of EChOA-SVM, this paper compares the algorithm with different meta-heuristic algorithms such as ChOA [31], GWO [35], WOA [36], ALO [37], GOA [38], MFO [39] and SSA [40].

All settings and parameters that are used in the experiments for all algorithms are presented in Table 2.

Table 2. Parameter settings of different algorithms.

Algorithm	Parameters	Value
EChOA	η_m	1
	C	2
	K	0.95
	f	[0, 2.5]
	m	Gauss chaotic
ChOA	f	[0, 2.5]
	m	Gauss chaotic

GWO	α	[0, 2]
WOA	α	[0, 2]
	a2	[-2, -1]
ALO	Number of search agents	30
	Number of iterations	100
	cMin	0.00001
GOA	cMax	1
	Number of search agents	30
	Number of iterations	100
	α	[-2, -1]
MFO	b	1
	t	[-2, -1]
	c1	$[2 \times 10^{-16}, 2]$
SSA	c2	[0, 1]
	c3	[0, 1]

5.3. Results and Discussion

In order to compare the different feature selection methods with the method proposed in this paper, three indexes are used in this paper [53].

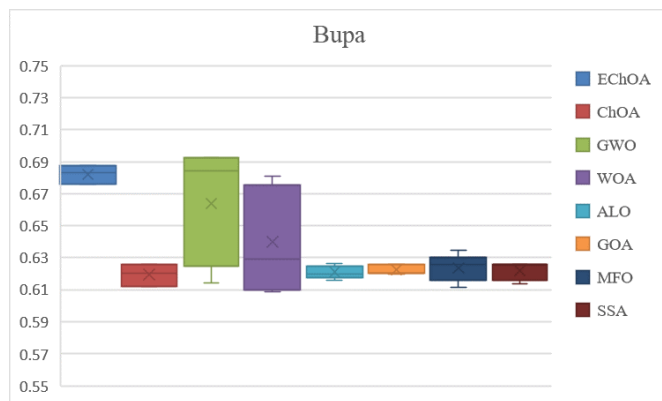
1. Accuracy: Accuracy represents the ratio between the number of correct classification and the actual number, which is reflecting the accuracy of the classifier recognition results. It is an important index to evaluate algorithm performance;
2. The number of features: The number of features reflects ability of eliminate redundancy. It shows whether the method can find the optimal feature subset;
3. Fitness value: Fitness value can reflect the advantages and disadvantages of the solution selected by the classifier. The better fitness value can get the better the solution.

Table 3 shows the mean and the standard deviation of the accuracy. Figure 3 presents the boxplot charts of the accuracy for eight used data sets. The boxplots are shown for the seventeen values of classification accuracy that are given by each algorithm at the end of the SVM training. The boxplots indicate the improved performance of EChOA for optimizing the SVM.

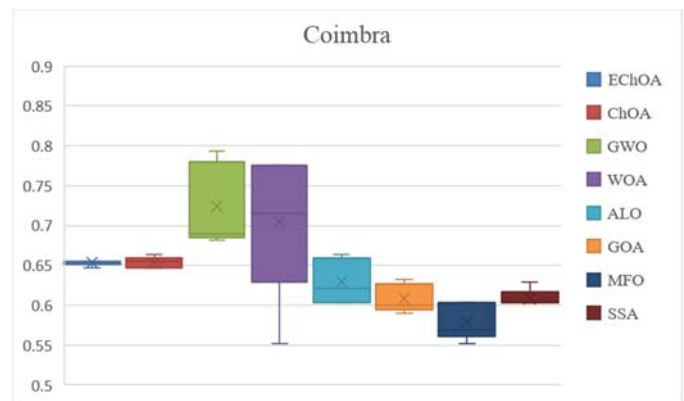
Table 3. Mean and standard deviation of classification accuracy for all data sets with applying feature selection.

		EChOA	ChOA	GWO	WOA	ALO	GOA	MFO	SSA
Balance	mean	8.9853×10^{-1}	8.9073×10^{-1}	8.8640×10^{-1}	8.9376×10^{-1}	8.9120×10^{-1}	8.9120×10^{-1}	7.7344×10^{-1}	8.9120×10^{-1}
	std	6.4825×10^{-3}	4.1199×10^{-3}	1.5513×10^{-2}	2.6349×10^{-3}	1.2413×10^{-16}	1.2413×10^{-16}	6.5830×10^{-2}	1.2413×10^{-16}
Bupa	mean	6.8217×10^{-1}	6.1943×10^{-1}	6.6380×10^{-1}	6.4002×10^{-1}	6.2106×10^{-1}	6.2260×10^{-1}	6.2378×10^{-1}	6.2204×10^{-1}
	std	5.8192×10^{-3}	7.1396×10^{-3}	3.6625×10^{-2}	3.3470×10^{-2}	3.9715×10^{-3}	3.2047×10^{-3}	8.5538×10^{-3}	5.7270×10^{-3}
Coimbra	mean	6.5348×10^{-1}	6.5348×10^{-1}	7.2414×10^{-1}	7.0518×10^{-1}	6.2930×10^{-1}	6.0818×10^{-1}	5.7930×10^{-1}	6.0858×10^{-1}
	std	3.8460×10^{-3}	7.1953×10^{-3}	5.2065×10^{-2}	9.1744×10^{-2}	2.8621×10^{-2}	1.7562×10^{-2}	2.3106×10^{-2}	1.1583×10^{-2}
Dermatology	mean	8.6590×10^{-1}	8.5010×10^{-1}	9.6087×10^{-1}	9.4638×10^{-1}	7.5922×10^{-1}	7.5752×10^{-1}	6.0560×10^{-1}	8.0838×10^{-1}
	std	1.2205×10^{-2}	6.9936×10^{-3}	2.1811×10^{-2}	3.1929×10^{-2}	2.4481×10^{-2}	4.6042×10^{-2}	8.1800×10^{-2}	2.6113×10^{-2}
Diagnostic	mean	9.6607×10^{-1}	9.5450×10^{-1}	9.6516×10^{-1}	9.6430×10^{-1}	9.5842×10^{-1}	9.5754×10^{-1}	9.3558×10^{-1}	9.5842×10^{-1}
	std	6.1598×10^{-3}	4.4034×10^{-3}	3.7997×10^{-3}	6.0828×10^{-3}	1.2969×10^{-3}	2.0683×10^{-3}	9.6469×10^{-3}	1.2969×10^{-3}
Divorce	mean	9.8044×10^{-1}	9.7768×10^{-1}	9.7650×10^{-1}	9.7298×10^{-1}	9.1764×10^{-1}	9.4236×10^{-1}	8.9176×10^{-1}	9.1412×10^{-1}
	std	6.3689×10^{-3}	2.6386×10^{-3}	4.1719×10^{-3}	1.4765×10^{-2}	2.1989×10^{-2}	1.9255×10^{-2}	6.6920×10^{-3}	3.1768×10^{-3}
Glass	mean	6.8067×10^{-1}	6.7290×10^{-1}	9.8813×10^{-1}	7.5796×10^{-1}	6.3924×10^{-1}	6.5048×10^{-1}	5.8410×10^{-1}	6.7010×10^{-1}
	std	1.1775×10^{-2}	1.2341×10^{-2}	2.3159×10^{-3}	7.1049×10^{-1}	1.5296×10^{-2}	2.5891×10^{-2}	4.9679×10^{-2}	3.7157×10^{-2}
Heart	mean	8.4230×10^{-1}	8.3207×10^{-1}	7.8594×10^{-1}	7.4666×10^{-1}	8.0370×10^{-1}	7.9738×10^{-1}	7.4740×10^{-1}	8.0888×10^{-1}
	std	4.7606×10^{-3}	8.5448×10^{-3}	2.2868×10^{-2}	8.8013×10^{-2}	6.4086×10^{-3}	9.3041×10^{-3}	3.8271×10^{-2}	1.7667×10^{-2}
Iris	mean	9.6400×10^{-1}	9.5337×10^{-1}	9.6268×10^{-1}	9.6268×10^{-1}	9.5734×10^{-1}	9.4934×10^{-1}	9.3068×10^{-1}	9.4268×10^{-1}
	std	8.9443×10^{-3}	1.1547×10^{-2}	3.6697×10^{-3}	3.6697×10^{-3}	8.9592×10^{-3}	7.5910×10^{-3}	4.0173×10^{-2}	3.6697×10^{-3}
Knowledge	mean	9.5675×10^{-1}	9.5087×10^{-1}	8.5426×10^{-1}	8.9692×10^{-1}	9.5350×10^{-1}	9.4660×10^{-1}	9.0542×10^{-1}	9.5428×10^{-1}
	std	1.8385×10^{-3}	2.3714×10^{-3}	7.9861×10^{-2}	4.6400×10^{-2}	2.7577×10^{-3}	1.2371×10^{-2}	3.9133×10^{-2}	3.2630×10^{-3}
Liver	mean	6.6937×10^{-1}	6.3383×10^{-1}	6.6844×10^{-1}	6.4350×10^{-1}	6.3248×10^{-1}	6.2490×10^{-1}	6.2142×10^{-1}	6.2610×10^{-1}

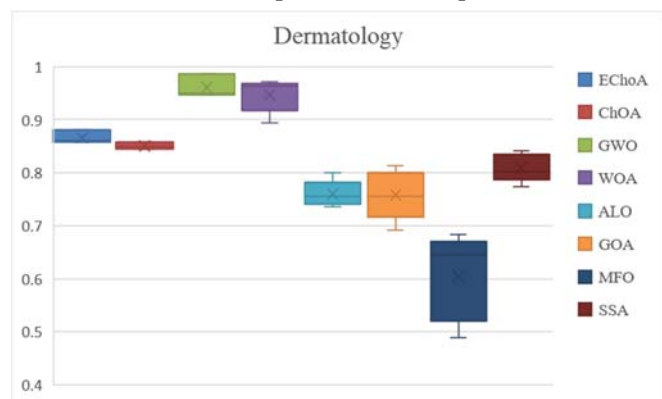
	std	3.2716×10^{-3}	1.3395×10^{-2}	1.6609×10^{-2}	6.4846×10^{-3}	1.4266×10^{-2}	4.2024×10^{-3}	5.9997×10^{-3}	0.0000
Lymphography	mean	5.1556×10^{-1}	5.0900×10^{-1}	4.3778×10^{-1}	4.3512×10^{-1}	4.7432×10^{-1}	4.7568×10^{-1}	4.2838×10^{-1}	4.7566×10^{-1}
	std	4.1004×10^{-3}	7.7942×10^{-3}	3.2275×10^{-2}	2.8920×10^{-2}	1.1088×10^{-2}	2.0061×10^{-2}	1.5544×10^{-2}	1.6980×10^{-2}
Sonar	mean	8.4137×10^{-1}	8.2373×10^{-1}	8.0578×10^{-1}	8.3368×10^{-1}	7.2978×10^{-1}	7.8076×10^{-1}	7.4040×10^{-1}	7.1252×10^{-1}
	std	9.4701×10^{-2}	4.0880×10^{-2}	4.4297×10^{-2}	3.1638×10^{-2}	5.6004×10^{-2}	1.0441×10^{-1}	1.2143×10^{-1}	1.1694×10^{-1}
TAE	mean	6.1110×10^{-1}	6.0840×10^{-1}	5.7084×10^{-1}	5.8940×10^{-1}	6.0290×10^{-1}	6.1132×10^{-1}	5.9206×10^{-1}	6.0174×10^{-1}
	std	2.3812×10^{-3}	1.1533×10^{-3}	2.9963×10^{-3}	1.6228×10^{-2}	5.4282×10^{-3}	1.7858×10^{-2}	1.7293×10^{-2}	9.0329×10^{-3}
Transfusion	mean	7.7713×10^{-1}	7.6380×10^{-1}	7.7218×10^{-1}	7.7004×10^{-1}	7.6200×10^{-1}	7.6414×10^{-1}	7.6280×10^{-1}	7.6392×10^{-1}
	std	4.1429×10^{-3}	3.1177×10^{-3}	9.3082×10^{-3}	6.1788×10^{-3}	0.0000	4.7852×10^{-3}	1.7889×10^{-3}	2.6631×10^{-3}
Wine	mean	8.5562×10^{-1}	9.5503×10^{-1}	9.3147×10^{-1}	9.0674×10^{-1}	6.9776×10^{-1}	6.4616×10^{-1}	4.5506×10^{-1}	7.7416×10^{-1}
	std	3.7552×10^{-2}	1.4873×10^{-2}	1.6064×10^{-2}	4.0183×10^{-2}	1.4329×10^{-1}	1.0406×10^{-1}	6.0622×10^{-2}	1.1504×10^{-1}
Zoo	mean	9.8119×10^{-1}	9.8010×10^{-1}	9.2377×10^{-1}	9.3565×10^{-1}	9.6139×10^{-1}	9.5743×10^{-1}	9.0892×10^{-1}	9.6337×10^{-1}
	std	3.1307×10^{-3}	9.0554×10^{-4}	3.4310×10^{-2}	1.9383×10^{-2}	8.6684×10^{-3}	1.3241×10^{-2}	2.1790×10^{-2}	1.1479×10^{-2}



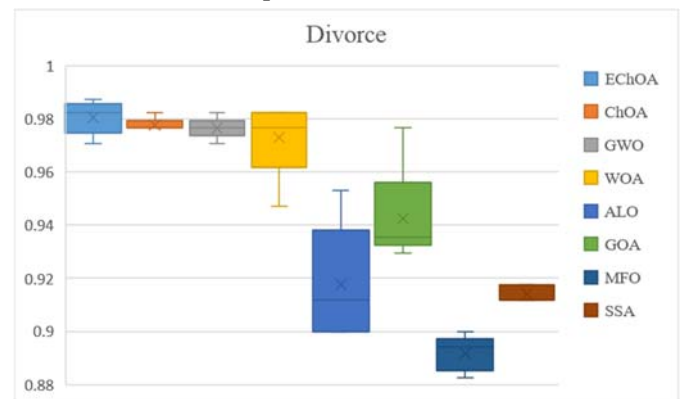
(a) Boxplot chart of Bupa



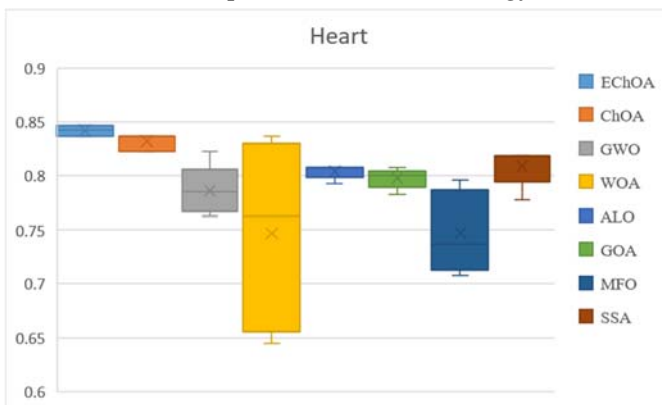
(b) Boxplot chart of Coimbra



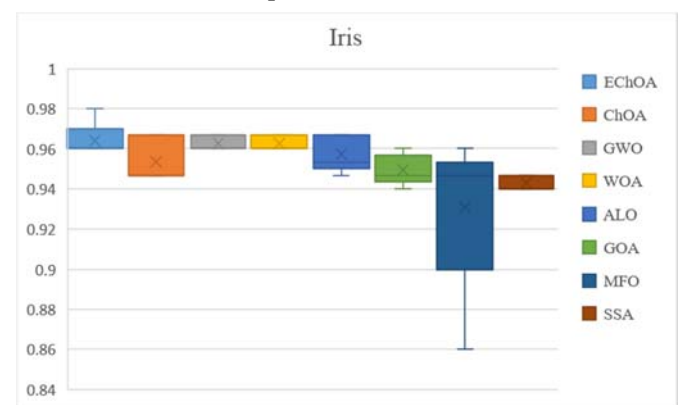
(c) Boxplot chart of Dermatology



(d) Boxplot chart of Divorce



(e) Boxplot chart of Heart



(f) Boxplot chart of Iris

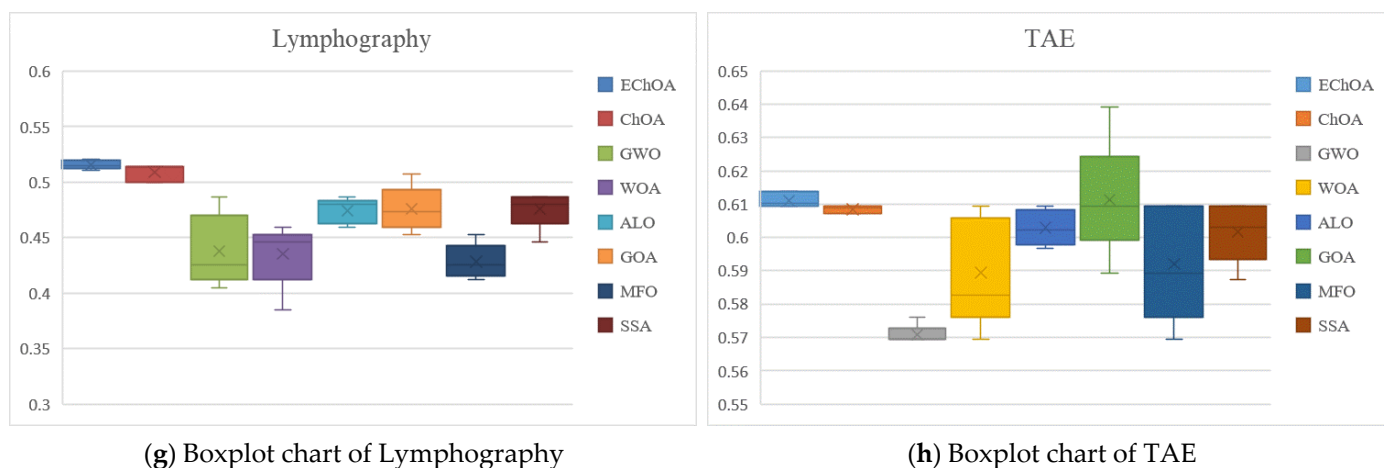


Figure 3. Boxplot charts for the accuracy of EChOA and other algorithms based on 8 data sets.

The performance of EChOA is better than other optimization algorithms in terms of average accuracy. EChOA obtains the best performance on 13 data sets out of the 17 data sets, which accounts for 76.5% of the total data sets. For the other four data sets, EChOA also has strong competitiveness. Although EChOA do not achieve the highest accuracy in Coimbra, Dermatology and Glass, its classification accuracy ranked the third among the eight optimization algorithms, just next to GWO and WOA. It is noteworthy that the accuracy of EChOA is 10 percent higher than ChOA on Bupa data sets. It is 12 percent higher than GWO on the Knowledge data set and 18 percent higher than GWO on the Lymphography data set. At the same time, it can be seen from the box chart that each optimization algorithm generates ten accuracies in the process of running each data set ten times. From the perspective of their distribution, the accuracy of EChOA is very stable. The accuracy of the EChOA optimization algorithm hardly changes significantly in the 17 data sets. Especially in the Coimbra and Heart data sets, the EChOA optimization algorithm is still very stable in the case that the accuracy of almost all optimization algorithms varies significantly each time. All these show the superiority of the EChOA algorithm.

Table 4 shows the mean and the standard deviation of the number of the features. Figure 4 presents the histogram of the number of features for all used data sets. We can intuitively see that the number of features required by EChOA is almost below the average of the number of features of all the algorithms. There are only two data sets do not get the best results (Bupa and Transfusion), which accounts for only 11.8 percent of the total. In Liver data set EChOA even requires only one feature. It shows that EChOA can use space more efficiently. In the analysis of classification accuracy, it has been found that EChOA's classification accuracy is lower than GWO and WOA in Coimbra, Dermatology and Glass. However, it is not difficult to find in Table 4 that the number of features required by EChOA in these three data sets is less than GWO and WOA. Especially in the Dermatology data set, EChOA required nearly seven fewer features than GWO and WOA. It can be seen in Figure 4, in Coimbra, Dermatology, Divorce and Liver data sets, there are two or three optimization algorithms with far higher features counts than the other optimization algorithms, but EChOA never had this situation. It should be noticed that the number of features required by ChOA is much higher than that of EChOA in Dermatology and Wine data sets, which also indicates the superiority of EChOA over ChOA and other algorithms in the quantitative aspect of features.

Table 4. Mean and standard deviation of the number of features for all data sets with applying feature selection.

		EChOA	ChOA	GWO	WOA	ALO	GOA	MFO	SSA
Balance	mean	2.3333	2.6667	3.2000	2.8000	3.8000	2.8000	3.2000	3.4000
	std	5.7735×10^{-1}	5.7735×10^{-1}	1.3038	7.1274×10^{-1}	4.4721×10^{-1}	8.3666×10^{-1}	4.4721×10^{-1}	8.9443×10^{-1}
Bupa	mean	2.0000	1.6667	3.0000	2.8000	1.2000	1.6000	2.2000	1.4000
	std	0.0000	5.7735×10^{-1}	1.8708	1.7889	4.4721×10^{-1}	5.4772×10^{-1}	4.4721×10^{-1}	5.4772×10^{-1}
Coimbra	mean	1.2000	1.4000	4.0000	4.2000	1.6000	1.4000	1.8000	1.4000
	std	4.4721×10^{-1}	5.4772×10^{-1}	1.7321	1.6432	8.9443×10^{-1}	5.4772×10^{-1}	8.3666×10^{-1}	5.4772×10^{-1}
Dermatology	mean	7.6667	8.6667	1.4000×10	1.4400×10	1.1000×10	1.0800×10	1.3000×10	9.8000
	std	5.7735×10^{-1}	5.7735×10^{-1}	6.2450	3.1305	7.0711×10^{-1}	1.0954	2.9155	1.6432
Diagnostic	mean	1.6667	2.3333	4.0000	4.4000	1.8000	2.2000	3.6000	2.0000
	std	5.7735×10^{-1}	2.3333	1.8708	2.0736	4.4721×10^{-1}	4.4721×10^{-1}	8.9443×10^{-1}	0.0000
Divorce	mean	2.0000	2.6000	7.4000	8.4000	3.8000	3.0000	6.4000	4.0000
	std	7.0711×10^{-1}	5.4772×10^{-1}	1.1402	2.4083	8.3666×10^{-1}	1.5811	1.6733	7.0711×10^{-1}
Glass	mean	3.3333	4.3333	5.3333	5.0000	4.4000	4.6000	5.2000	3.6000
	std	5.7735×10^{-1}	5.7735×10^{-1}	2.5166	4.8000	1.6733	1.3416	1.3038	1.8166
Heart	mean	2.6667	3.0000	6.2000	5.6000	4.2000	4.4000	4.0000	4.0000
	std	5.7735×10^{-1}	1.0000	1.4832	2.0736	4.4721×10^{-1}	5.4772×10^{-1}	1.5811	1.4142
Iris	mean	1.2000	2.3333	2.2000	2.0000	1.8000	1.4000	1.4000	1.6000
	std	4.4721×10^{-1}	5.7735×10^{-1}	8.3666×10^{-1}	7.0711×10^{-1}	8.3666×10^{-1}	5.4772×10^{-1}	8.9443×10^{-1}	5.4772×10^{-1}
Knowledge	mean	1.7500	2.3333	2.2000	3.8000	2.0000	2.2000	2.2000	2.0000
	std	5.0000×10^{-1}	5.7735×10^{-1}	1.7889	1.0954	0.0000	4.4721×10^{-1}	8.3666×10^{-1}	0.0000
Liver	mean	1.0000	1.3333	4.6000	3.2000	1.6000	1.2000	2.4000	1.2000
	std	1.0000	5.7735×10^{-1}	8.9443×10^{-1}	8.3666×10^{-1}	5.4772×10^{-1}	4.4721×10^{-1}	5.4772×10^{-1}	4.4721×10^{-1}
Lymphography	mean	6.4000	7.3333	7.8000	9.4000	7.6000	7.2000	8.8000	7.4000
	std	5.4772×10^{-1}	5.7735×10^{-1}	2.6833	2.6077	5.4772×10^{-1}	1.9235	1.6432	1.1402
Sonar	mean	1.4667 × 10	1.7333×10	2.6000×10	3.3600×10	2.4600×10	2.6400×10	2.8600×10	2.3400×10
	std	2.0817	2.3094	2.9155	4.0373	2.5100	4.1593	3.1305	2.3022
TAE	mean	1.6667	2.3333	2.2000	2.4000	3.0000	3.0000	3.4000	2.8000
	std	5.7735×10^{-1}	1.1547	4.4721×10^{-1}	5.4772×10^{-1}	0.0000	0.0000	8.9443×10^{-1}	4.4721×10^{-1}
Transfusion	mean	3.0000	2.3333	2.4000	2.8000	2.0000	2.2000	2.0000	2.2000
	std	0.0000	1.1547	5.4772×10^{-1}	8.3666×10^{-1}	7.0711×10^{-1}	1.0954	0.0000	8.3666×10^{-1}
Wine	mean	2.6000	6.0000	3.8000	5.8000	3.4000	4.0000	4.6000	3.0000
	std	8.4327×10^{-1}	1.0000	1.7512	1.7512	2.0736	2.0000	1.5166	1.0000
Zoo	mean	5.0000	6.2500	5.7000	7.4000	6.2000	6.2000	7.3000	6.3000
	std	6.6667×10^{-1}	9.5743×10^{-1}	3.1990	3.2042	8.6684×10^{-3}	7.8881×10^{-1}	1.4944	1.2517

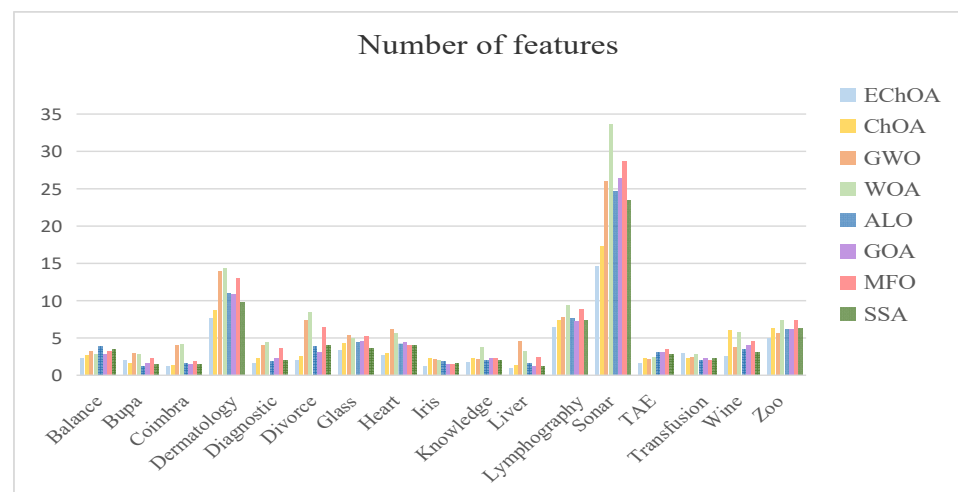
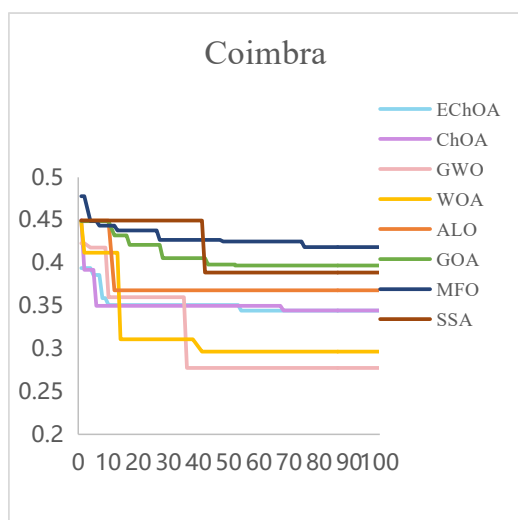
**Figure 4.** Histogram for the number of features of EChOA and other algorithms based on 17 data sets.

Table 5 shows the mean and the standard deviation of the fitness. Figure 5 presents the line chart of the fitness for six used data sets. The lines in Figure 5 represents the convergence curves of corresponding algorithm. The fitness value of EChOA is the minimum in all the 12 data sets, and the convergence time is very short. The convergence speed of

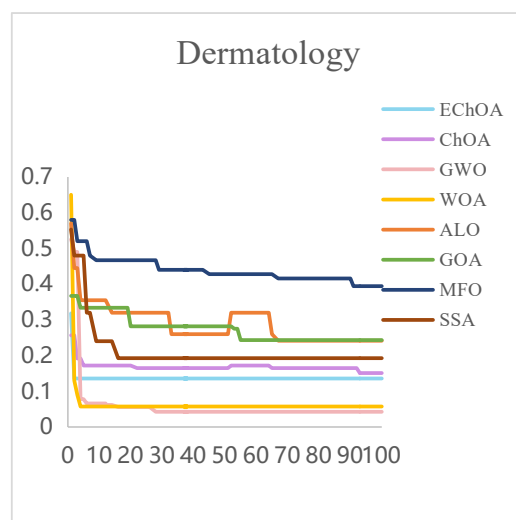
Coimbra, Dermatology, Diagnostic, Divorce and Heart data sets is all higher than ChOA, and some fitness values are also lower than ChOA. At the same time, EChOA convergence curve is relatively flat, in the performance of each data set is relatively good, which are almost below to the average fitness value, does not appear similar to GOA (e.g., Liver) in a certain data set fitness value too much higher than the other algorithms. As it can be seen from Figure 5, the convergence curve of EChOA in the Heart data set is smoother than other algorithms, which shows that the local search ability of the algorithm has been greatly improved. EChOA algorithm can achieve the maximum objective function value when the number of iterations is less than 10 on Dermatology and Divorce data sets, which is earlier than other optimization algorithms. This indicates that the convergence rate of EChOA is improved. As you can see, EChOA is better and more stable than most algorithms.

Table 5. Mean and standard deviation of fitness values for all data sets with applying feature selection.

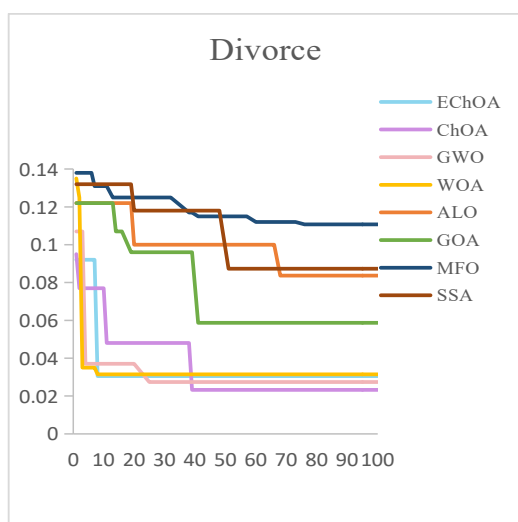
		EChOA	ChOA	GWO	WOA	ALO	GOA	MFO	SSA
Balance	mean	1.1070 × 10 ⁻¹	1.1437 × 10 ⁻¹	1.3420 × 10 ⁻¹	1.1218 × 10 ⁻¹	1.1720 × 10 ⁻¹	1.1470 × 10 ⁻¹	2.3226 × 10 ⁻¹	1.1620 × 10 ⁻¹
	std	1.9079 × 10 ⁻³	1.4434 × 10 ⁻³	3.0616 × 10 ⁻²	1.4595 × 10 ⁻³	1.1180 × 10 ⁻³	2.0917 × 10 ⁻³	6.4041 × 10 ⁻²	2.2361 × 10 ⁻³
Bupa	mean	3.3020 × 10 ⁻¹	3.7293 × 10 ⁻¹	3.3786 × 10 ⁻¹	3.6106 × 10 ⁻¹	3.7214 × 10 ⁻¹	3.7282 × 10 ⁻¹	3.7646 × 10 ⁻¹	3.7248 × 10 ⁻¹
	std	7.6531 × 10 ⁻³	9.8150 × 10 ⁻⁴	3.5019 × 10 ⁻²	3.2445 × 10 ⁻²	7.6026 × 10 ⁻⁴	9.3113 × 10 ⁻⁴	8.5483 × 10 ⁻³	9.3113 × 10 ⁻⁴
Coimbra	mean	3.4442 × 10 ⁻¹	3.4464 × 10 ⁻¹	2.7756 × 10 ⁻¹	2.9656 × 10 ⁻¹	3.6810 × 10 ⁻¹	3.9710 × 10 ⁻¹	4.1846 × 10 ⁻¹	3.8880 × 10 ⁻¹
	std	3.7090 × 10 ⁻³	7.3310 × 10 ⁻³	4.9871 × 10 ⁻²	8.9737 × 10 ⁻²	2.8289 × 10 ⁻²	3.4474 × 10 ⁻²	2.3793 × 10 ⁻²	1.0957 × 10 ⁻²
Dermatology	mean	1.3580 × 10 ⁻¹	1.5097 × 10 ⁻¹	4.2833 × 10 ⁻²	5.7340 × 10 ⁻²	2.4160 × 10 ⁻¹	2.4320 × 10 ⁻¹	3.9430 × 10 ⁻¹	1.9258 × 10 ⁻¹
	std	1.2653 × 10 ⁻²	7.1066 × 10 ⁻³	2.0294 × 10 ⁻²	3.1561 × 10 ⁻²	2.4253 × 10 ⁻²	4.5528 × 10 ⁻²	8.1597 × 10 ⁻²	2.6259 × 10 ⁻²
Diagnostic	mean	3.7460 × 10 ⁻²	4.2600 × 10 ⁻²	3.8480 × 10 ⁻²	3.9760 × 10 ⁻²	4.2980 × 10 ⁻²	4.4240 × 10 ⁻²	6.7380 × 10 ⁻²	4.3180 × 10 ⁻²
	std	5.6241 × 10 ⁻³	0.0000	2.8639 × 10 ⁻³	6.3830 × 10 ⁻³	1.4738 × 10 ⁻³	2.4006 × 10 ⁻³	1.0348 × 10 ⁻²	1.2969 × 10 ⁻³
Divorce	mean	3.0560 × 10 ⁻²	2.3240 × 10 ⁻²	2.7420 × 10 ⁻²	3.1460 × 10 ⁻²	8.3660 × 10 ⁻²	5.8720 × 10 ⁻²	1.1072 × 10 ⁻¹	8.7260 × 10 ⁻²
	std	4.3206 × 10 ⁻³	2.5938 × 10 ⁻³	4.3568 × 10 ⁻³	1.4762 × 10 ⁻²	2.2144 × 10 ⁻²	1.9400 × 10 ⁻²	7.4221 × 10 ⁻³	2.9458 × 10 ⁻³
Glass	mean	3.1980 × 10 ⁻¹	3.2820 × 10 ⁻¹	1.4633 × 10 ⁻²	2.4464 × 10 ⁻¹	3.6154 × 10 ⁻¹	3.5062 × 10 ⁻¹	4.1692 × 10 ⁻¹	3.3020 × 10 ⁻¹
	std	1.1755 × 10 ⁻²	1.2429 × 10 ⁻²	2.5166 × 10 ⁻³	2.9145 × 10 ⁻¹	1.4406 × 10 ⁻²	2.6367 × 10 ⁻²	4.8918 × 10 ⁻²	3.8471 × 10 ⁻²
Heart	mean	1.6360 × 10 ⁻¹	1.6877 × 10 ⁻¹	2.1670 × 10 ⁻¹	2.5512 × 10 ⁻¹	1.9754 × 10 ⁻¹	2.1170 × 10 ⁻¹	2.5312 × 10 ⁻¹	1.9228 × 10 ⁻¹
	std	8.6603 × 10 ⁻⁴	8.9489 × 10 ⁻³	2.2128 × 10 ⁻²	8.7301 × 10 ⁻²	6.7122 × 10 ⁻³	1.7476 × 10 ⁻²	3.8069 × 10 ⁻²	1.8465 × 10 ⁻²
Iris	mean	3.8140 × 10 ⁻²	4.8700 × 10 ⁻²	3.9460 × 10 ⁻²	4.1960 × 10 ⁻²	4.5240 × 10 ⁻²	5.3160 × 10 ⁻²	7.2140 × 10 ⁻²	5.9260 × 10 ⁻²
	std	7.7661 × 10 ⁻³	1.1432 × 10 ⁻²	2.8183 × 10 ⁻³	2.8183 × 10 ⁻³	6.8178 × 10 ⁻³	6.6842 × 10 ⁻³	4.1966 × 10 ⁻²	3.6150 × 10 ⁻³
Knowledge	mean	4.7150 × 10 ⁻²	4.9667 × 10 ⁻²	1.4868 × 10 ⁻¹	1.0966 × 10 ⁻¹	5.0020 × 10 ⁻²	5.1460 × 10 ⁻²	9.8040 × 10 ⁻²	4.9260 × 10 ⁻²
	std	1.9000 × 10 ⁻³	5.7735 × 10 ⁻⁴	7.9223 × 10 ⁻²	4.6659 × 10 ⁻²	2.7225 × 10 ⁻³	8.1076 × 10 ⁻³	3.7952 × 10 ⁻²	3.2153 × 10 ⁻³
Liver	mean	3.5083 × 10 ⁻¹	3.6473 × 10 ⁻¹	4.9214 × 10 ⁻¹	3.5930 × 10 ⁻¹	3.6650 × 10 ⁻¹	4.9214 × 10 ⁻¹	3.7932 × 10 ⁻¹	3.7214 × 10 ⁻¹
	std	1.8717 × 10 ⁻³	1.2240 × 10 ⁻²	2.6909 × 10 ⁻¹	7.1554 × 10 ⁻³	1.3778 × 10 ⁻²	2.6909 × 10 ⁻¹	6.5933 × 10 ⁻³	7.6026 × 10 ⁻⁴
Lymphography	mean	4.8520 × 10 ⁻¹	4.9017 × 10 ⁻¹	5.6088 × 10 ⁻¹	5.6442 × 10 ⁻¹	5.2464 × 10 ⁻¹	5.2308 × 10 ⁻¹	5.7058 × 10 ⁻¹	5.2320 × 10 ⁻¹
	std	2.7386 × 10 ⁻⁴	7.5692 × 10 ⁻³	3.2668 × 10 ⁻²	2.7232 × 10 ⁻²	1.0873 × 10 ⁻²	2.0454 × 10 ⁻²	1.6228 × 10 ⁻²	1.6372 × 10 ⁻²
Sonar	mean	1.6070 × 10 ⁻¹	1.7740 × 10 ⁻¹	1.9664 × 10 ⁻¹	1.7028 × 10 ⁻¹	2.7160 × 10 ⁻¹	2.2144 × 10 ⁻¹	2.6176 × 10 ⁻¹	2.8852 × 10 ⁻¹
	std	9.2529 × 10 ⁻²	4.0124 × 10 ⁻²	4.4221 × 10 ⁻²	3.1281 × 10 ⁻²	5.5205 × 10 ⁻²	1.0308 × 10 ⁻¹	1.2014 × 10 ⁻¹	1.1603 × 10 ⁻¹
TAE	mean	3.8797 × 10 ⁻¹	3.9147 × 10 ⁻¹	4.2888 × 10 ⁻¹	4.1128 × 10 ⁻¹	3.9280 × 10 ⁻¹	4.1524 × 10 ⁻¹	4.1066 × 10 ⁻¹	3.9240 × 10 ⁻¹
	std	1.8448 × 10 ⁻³	2.3094 × 10 ⁻³	2.9516 × 10 ⁻³	1.6103 × 10 ⁻²	0.0000	2.0485 × 10 ⁻²	1.8461 × 10 ⁻²	8.9443 × 10 ⁻⁴
Transfusion	mean	2.3203 × 10 ⁻¹	2.3800 × 10 ⁻¹	7.7218 × 10 ⁻¹	2.3462 × 10 ⁻¹	2.4060 × 10 ⁻¹	7.6414 × 10 ⁻¹	2.3880 × 10 ⁻¹	2.3754 × 10 ⁻¹
	std	5.0143 × 10 ⁻³	1.7321 × 10 ⁻⁴	9.3082 × 10 ⁻³	4.5329 × 10 ⁻³	1.7678 × 10 ⁻³	4.7852 × 10 ⁻³	2.4900 × 10 ⁻³	2.0780 × 10 ⁻³
Wine	mean	1.4506 × 10 ⁻¹	4.9100 × 10 ⁻²	7.1150 × 10 ⁻²	9.6780 × 10 ⁻²	3.0180 × 10 ⁻¹	2.8960 × 10 ⁻¹	5.4302 × 10 ⁻¹	2.2590 × 10 ⁻¹
	std	3.6823 × 10 ⁻²	1.4126 × 10 ⁻²	1.6038 × 10 ⁻²	3.9600 × 10 ⁻²	1.4065 × 10 ⁻¹	2.5621 × 10 ⁻²	6.0944 × 10 ⁻²	1.1317 × 10 ⁻¹
Zoo	mean	2.1740 × 10 ⁻²	2.3525 × 10 ⁻²	7.9040 × 10 ⁻²	6.8210 × 10 ⁻²	4.2120 × 10 ⁻²	4.6050 × 10 ⁻²	9.4730 × 10 ⁻²	4.0220 × 10 ⁻²
	std	3.1362 × 10 ⁻³	6.1847 × 10 ⁻⁴	3.2862 × 10 ⁻²	1.8403 × 10 ⁻²	9.0828 × 10 ⁻³	1.3499 × 10 ⁻²	2.1928 × 10 ⁻²	1.1204 × 10 ⁻²



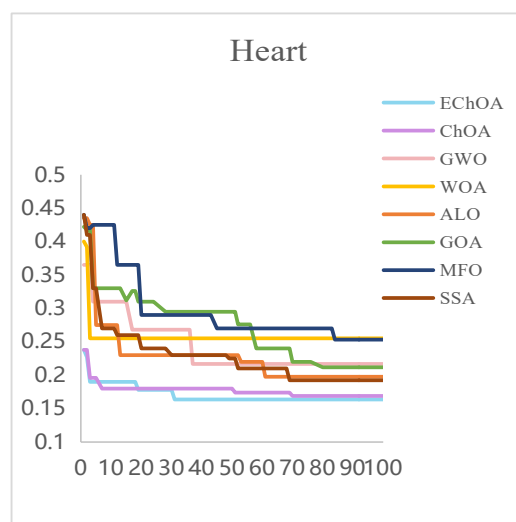
(a) Th × 10 line chart of Coimbra



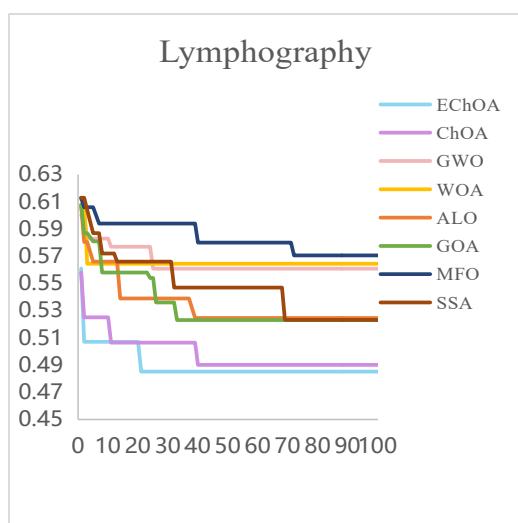
(b) The line chart of Dermatology



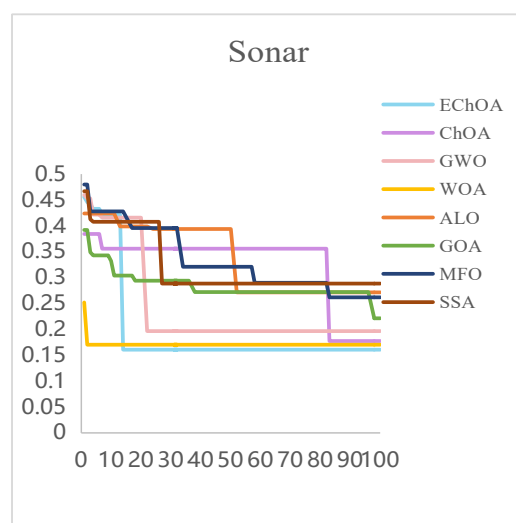
(c) The line chart of Divorce



(d) The line chart of Heart



(e) The line chart of Lymphography



(f) The line chart of Sonar

Figure 5. The line chart for the fitness of EChOA and other algorithms based on 17 data sets.

6. Conclusions

This work presents a novel hybrid method for optimizing SVM based on the EChOA. The proposed approach is able to tune the parameters of the SVM kernel and at the same time can find the best accuracy. Experimental results show that the proposed algorithm is effective in improving the classification accuracy of SVM. The experimental results show that the EChOA algorithm has certain advantages over the other seven optimization algorithms in terms of accuracy, feature number and fitness value, and its comprehensive performance is relatively stable. All these indicate that the EChOA algorithm is very competitive.

Although EChOA can effectively improve the exploitation and exploration of ChOA, EChOA introduces three strategies at the same time that increase the complexity of the algorithm. Therefore, it is necessary to consider how to reduce ineffective improvement strategies to alleviate this problem. In subsequent studies, some parallel strategies such as co-evolutionary mechanism can be introduced.

In the future research, EChOA can be applied in multi-objective problems, social manufacturing optimization, and video coding optimization. Simultaneously, the EChOA-SVM model that proposed in this paper can be studied in a larger scale. It can also be applied to other practical problems like data mining.

Author Contributions: Conceptualization, D.W. and H.J.; methodology, D.W. and H.J.; software, D.W. and W.Z.; validation, W.Z., X.L.; formal analysis, D.W. and W.Z.; investigation, D.W. and W.Z.; writing—original draft preparation, D.W. and W.Z.; writing—review and editing, H.J.; visualization, D.W. and H.J.; funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Sanming University Introduces High-level Talents to Start Scientific Research Funding Support Project (21YG01S), Bidding project for higher education research of Sanming University (SHE2101), Fujian Natural Science Foundation Project, the Guiding Science and Technology Projects in Sanming City (2021-S-8), the Educational Research Projects of Young and Middle-aged Teachers in Fujian Province (JAT200618), and the Scientific Research and Development Fund of Sanming University (B202009), Open Research Fund of Key Laboratory of Agricultural Internet of Things in Fujian Province (ZD2101), Ministry of Education Cooperative Education Project (202002064014), School level education and teaching reform project of Sanming University (J2010305), Higher education research project of Sanming University (SHE2013), Fundamental research funds for the Central Universities (2572018BF11).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Raju, B.; Bonagiri, R. A cavernous analytics using advanced machine learning for real world datasets in research implementations. *Mater. Today Proc.* **2020**, *https://doi.org/10.1016/j.matpr.2020.11.089*.
2. Jiang, J.; Yu, C.; Xu, X.; Ma, Y.; Liu, J. Achieving better connections between deposited lines in additive manufacturing via machine learning. *Math. Biosci. Eng.* **2020**, *17*, 3382–3394.
3. Kline, A.S.; Kline, T.J.B.; Lee, J. Item response theory as a feature selection and interpretation tool in the context of machine learning. *Med. Biol. Eng. Comput.* **2021**, *59*, 471–482.
4. Ebrahimi-Khusfi, Z.; Nafarzadegan, A.R.; Dargahian, F. Predicting the number of dusty days around the desert wetlands in southeastern Iran using feature selection and machine learning techniques. *Ecol. Indic.* **2021**, *125*, 1–15.
5. Tanveer, M. Robust and Sparse Linear Programming Twin Support Vector Machines. *Cogn. Comput.* **2015**, *7*, 137–149.
6. Yang, P.; Li, Z.; Yu, Y.; Shi, J.; Sun, M. Studies on fault diagnosis of dissolved oxygen sensor based on GA-SVM. *Math. Biosci. Eng.* **2021**, *18*, 386–399.
7. Aziz, W.; Hussain, L.; Khan, I.R.; Alowibdi, J.S.; Alkinani, M.H. Machine learning based classification of normal, slow and fast walking by extracting multimodal features from stride interval time series. *Math. Biosci. Eng.* **2021**, *18*, 495–517.
8. Hussain, L.; Aziz, W.; Khan, I.R.; Alkinani, M.H.; Alowibdi, J.S. Machine learning based congestive heart failure detection using feature importance ranking of multimodal features. *Math. Biosci. Eng.* **2021**, *18*, 69–91.

9. Brown, M.P.S.; Grundy, W.N.; Lin, D.; Cristianini, N.; Sugnet, C.W.; Furey, T.; Ares, M.; Haussler, D. Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proc. Natl. Acad. Sci. USA*. **2000**, *97*, 262–267.
10. Takeuchi, K.; Collier, N. Bio-medical entity extraction using support vector machines. *Artif. Intell. Med.* **2005**, *33*, 125–137.
11. Babaoğlu, I.; Findik, O.; Bayrak, M. Effects of principle component analysis on assessment of coronary artery diseases using support vector machine. *Expert Syst. Appl.* **2010**, *37*, 2182–2185.
12. Du, Z.; Yang, Q.; He, H.; Qiu, M.; Chen, Z.; Hu, Q.; Wang, Q.; Zhang, Z.; Lin, Q.; Huang, L.; et al. A comprehensive health classification model based on support vector machine for proseal laryngeal mask and tracheal catheter assessment in herniorrhaphy. *Math. Biosci. Eng.* **2020**, *17*, 1838–1854.
13. Sotiris, V.A.; Tse, P.W.; Pecht, M. Anomaly Detection through a Bayesian Support Vector Machine. *IEEE Trans. Reliab.* **2010**, *59*, 277–286.
14. Rostami, O.; Kaveh, M. Optimal feature selection for SAR image classification using biogeography-based optimization (BBO), artificial bee colony (ABC) and support vector machine (SVM): A combined approach of optimization and machine learning. *Comput. Geosci.* **2021**, *25*, 911–930.
15. Joachims, T. *Making Large-Scale Support Vector Machine Learning Practical*; MIT Press: Cambridge, MA, USA, 1999, 169–184.
16. Weston, J.; Mukherjee, S.; Chapelle, O. Feature selection for SVMs. *Adv. Neural Inf. Process Syst.* **2000**, *13*, 668–674.
17. Nguyen, M.H.; Torre, F.D.L. Optimal feature selection for support vector machines. *Pattern Recognit.* **2010**, *43*, 584–591.
18. Shahbeig, S.; Helfroush, M.S.; Rahideh, A. A fuzzy multi-objective hybrid TLBO-PSO approach to select the associated genes with breast cancer. *Signal Process.* **2017**, *131*, 58–65.
19. Wu, Q.; Ma, Z.; Fan, J.; Xu, G.; Shen, Y. A feature selection method based on hybrid improved binary quantum particle swarm optimization. *IEEE Access* **2019**, *7*, 80588–80601.
20. Souza, T.A.; Souza, M.A.; Costa, W.C.D.A.; Costa, S.C.; Correia, S.E.N.; Vieira, V.J.D. Feature Selection based on Binary Particle Swarm Optimization and Neural Networks for Pathological Voice Detection. *Int. J. Bio-Inspired Comput.* **2018**, *11*, 2.
21. Wang, H.; Niu, B.; Tan, L. Bacterial colony algorithm with adaptive attribute learning strategy for feature selection in classification of customers for personalized recommendation. *Neurocomputing* **2021**, *452*, 747–755.
22. Jha, K.; Saha, S. Incorporation of multimodal objective optimization in designing a filter based feature selection technique. *Appl. Soft Comput.* **2021**, *98*, 106823.
23. Han, M.; Liu, X. Feature selection techniques with class separability for multivariate time series. *Neurocomputing* **2013**, *110*, 29–34.
24. Nithya, B.; Ilango, V. Evaluation of machine learning based optimized feature selection approaches and classification methods for cervical cancer prediction. *SN Appl. Sci.* **2019**, *1*, 641.
25. Kohavi, R.; John, G.H. Wrappers for feature subset selection. *Artif. Intell.* **1997**, *97*, 273–324.
26. Pourpanah, F.; Lim, C.P.; Wang, X.; Tan, C.J.; Seera, M.; Shi, Y. A hybrid model of fuzzy min-max and brain storm optimization for feature selection and data classification. *Neurocomputing* **2019**, *333*, 440–451.
27. Liu, H.; Setiono, R. A probabilistic approach to feature selection—a filter solution. In Proceedings of the 9th International Conference on Industrial and Engineering Applications of AI and ES, Fukuoka, Japan, 4–7 June 1996.
28. Wang, J.; Wu, L.; Kong, J.; Li, Y.; Zhang, B. Maximum weight and minimum redundancy: A novel framework for feature subset selection. *Pattern Recognit.* **2013**, *46*, 1616–1627.
29. Sihwail, R.; Omar, K.; Ariffin, K.A.Z.; Tubishat, M. Improved Harris Hawks Optimization Using Elite Opposition-Based Learning and Novel Search Mechanism for Feature Selection. *IEEE Access* **2020**, *8*, 121127–121145.
30. Elgamal, Z.M.; Yasin, N.B.M.; Tubishat, M.; Alswaitti, M.; Mirjalili, S. An Improved Harris Hawks Optimization Algorithm with Simulated Annealing for Feature Selection in the Medical Field. *IEEE Access* **2020**, *8*, 186638–186652.
31. Khishe, M.; Mosavi, M. Chimp optimization algorithm. *Expert Syst. Appl.* **2020**, *149*, 113338.
32. Abed-Alguni, B. Island-based Cuckoo Search with Highly Disruptive Polynomial Mutation. *Int. J. Artif. Intell.* **2019**, *17*, 57–82.
33. Jiang, X.; Li, S. BAS: Beetle Antennae Search Algorithm for Optimization Problems. *arXiv* **2017**, arXiv:1710.10724.
34. Lichman, M. UCI Machine Learning Repository. Available online: <http://archive.ics.uci.edu/ml> (accessed on 15 August 2013).
35. Renita, D.B.; Christopher, C.S. Novel real time content based medical image retrieval scheme with GWO-SVM. *Multimed. Tools Appl.* **2020**, *79*, 17227–17243.
36. Yin, X.; Hou, Y.; Yin, J.; Li, C. A novel SVM parameter tuning method based on advanced whale optimization algorithm. *J. Phys. Conf. Ser.* **2019**, *1237*, 022140.
37. Zhao, S.; Gao, L.; Dongmei, Y.; Jun, T. Ant Lion Optimizer with Chaotic Investigation Mechanism for Optimizing SVM Parameters. *J. Front. Comput. Sci. Technol.* **2016**, *10*, 722–731.
38. Aljarah, I.; Al-Zoubi, A.M.; Faris, H.; Hassonah, M.A.; Mirjalili, S.; Saadeh, H. Simultaneous Feature Selection and Support Vector Machine Optimization Using the Grasshopper Optimization Algorithm. *Cogn. Comput.* **2018**, *10*, 478–495.
39. Lin, G.Q.; Li, L.L.; Tseng, M.L.; Liu, H.M.; Yuan, D.D.; Tan, R.R. An improved moth-flame optimization algorithm for support vector machine prediction of photovoltaic power generation. *J. Clean. Prod.* **2020**, *253*, 119966.
40. Sivapragasam, C.; Liong, S.-Y.; Pasha, M.F.K. Rainfall and runoff forecasting with SSA-SVM approach. *J. Hydroinformatics* **2001**, *3*, 141–152.
41. Zhao, H.; Zhang, C. An online-learning-based evolutionary many-objective algorithm. *Inf. Sci.* **2020**, *509*, 1–21.
42. Dulebenets, M.A. An Adaptive Polyploid Memetic Algorithm for scheduling trucks at a cross-docking terminal. *Inf. Sci.* **2021**, *565*, 390–421.

43. Liu, Z.-Z.; Wang, Y.; Huang, P.-Q. AnD: A many-objective evolutionary algorithm with angle-based selection and shift-based density estimation. *Inf. Sci.* **2020**, *509*, 400–419.
44. Pasha, J.; Dulebenets, M.A.; Kavooosi, M.; Abioye, O.F.; Wang, H.; Guo, W. An Optimization Model and Solution Algorithms for the Vehicle Routing Problem with a “Factory-in-a-Box”. *IEEE Access* **2020**, *8*, 134743–134763.
45. D’Angelo, G.; Pilla, R.; Tascini, C.; Rampone, S. A proposal for distinguishing between bacterial and viral meningitis using genetic programming and decision trees. *Soft Comput.* **2019**, *23*, 11775–11791.
46. Panda, N.; Majhi, S.K. How effective is the salp swarm algorithm in data classification. In *Computational Intelligence in Pattern Recognition*; Springer: Singapore, 2020, pp. 579–588.
47. Alwan, H.B.; Mahamud, K. Mixed-variable ant colony optimisation algorithm for feature subset selection and tuning support vector machine parameter. *Int. J. Bio-Inspired Comput.* **2017**, *9*, 53–63.
48. Frhlich, H.; Chapelle, O.; Schlkopf, B. Feature Selection for Support Vector Machines by Means of Genetic Algorithms. In Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2003), Sacramento, CA, USA, 5 November 2003; IEEE: New York, NY, USA, 2003.
49. Huang, C.-L.; Wang, C.-J. A GA-based feature selection and parameters optimization for support vector machines. *Expert Syst. Appl.* **2006**, *31*, 231–240.
50. Lin, S.-W.; Tseng, T.-Y.; Chen, S.-C.; Huang, J.-F. A SA-Based Feature Selection and Parameter Optimization Approach for Support Vector Machine. *Pervasive Comput. IEEE* **2006**, doi:10.1109/ICSMC.2006.384599.
51. Jia, H.; Sun, K. Improved barnacles mating optimizer algorithm for feature selection and support vector machine optimization. *Pattern Anal. Appl.* **2021**, *24*, 1249–1274.
52. Slipinski, A.; Escalona, H. *Australian Longhorn Beetles (Coleoptera: Cerambycidae)*; CSIRO Publishing: Clayton, Australia, 2013; Volume 1.
53. Mafarja, M.M.; Mirjalili, S. Hybrid Whale Optimization Algorithm with simulated annealing for feature selection. *Neurocomputing* **2017**, *260*, 302–312.