

Improved marine predators algorithm for feature selection and SVM optimization

Heming Jia^{1,2*}, Kangjian Sun², Yao Li², and Ning Cao¹

¹ College of Information Engineering, Sanming University
Sanming, 365004, China

[e-mail: jiaheminglucky99@126.com, ning.cao2008@hotmail.com]

² College of Mechanical and Electrical Engineering, Northeast Forestry University
Harbin, 150040, China

[e-mail: sunkangjian@nefu.edu.cn, liyao@nefu.edu.cn]

*Corresponding author: Heming Jia

*Received January 26, 2021; revised June 29, 2021; accepted April 14, 2022;
published April 30, 2022*

Abstract

Owing to the rapid development of information science, data analysis based on machine learning has become an interdisciplinary and strategic area. Marine predators algorithm (MPA) is a novel metaheuristic algorithm inspired by the foraging strategies of marine organisms. Considering the randomness of these strategies, an improved algorithm called co-evolutionary cultural mechanism-based marine predators algorithm (CECMPA) is proposed. Through this mechanism, search agents in different spaces can share knowledge and experience to improve the performance of the native algorithm. More specifically, CECMPA has a higher probability of avoiding local optimum and can search the global optimum quickly. In this paper, it is the first to use CECMPA to perform feature subset selection and optimize hyperparameters in support vector machine (SVM) simultaneously. For performance evaluation the proposed method, it is tested on twelve datasets from the university of California Irvine (UCI) repository. Moreover, the coronavirus disease 2019 (COVID-19) can be a real-world application and is spreading in many countries. CECMPA is also applied to a COVID-19 dataset. The experimental results and statistical analysis demonstrate that CECMPA is superior to other compared methods in the literature in terms of several evaluation metrics. The proposed method has strong competitive abilities and promising prospects.

Keywords: Marine predators algorithm, co-evolutionary cultural mechanism, feature selection, support vector machine, hyperparameters optimization

This work was supported by the Sanming University introduces high-level talents to start scientific research funding support project under Grant No.20YG14, No.19YG02, Sanming University National Natural Science Foundation Breeding Project No.PYT2105, Fujian Natural Science Foundation Project No.2021J011128, Key project of school level education and teaching reform of Sanming University No.J2010305

1. Introduction

Machine learning (ML) is an inevitable outcome of the development of artificial intelligence research. In the past two decades, the ability of humans to collect, store, transmit, and process data has been rapidly improved. ML attracts much attention since it can effectively analyze and utilize massive data [1]. According to whether the training data has labels, learning algorithms are divided into unsupervised learning and supervised learning [2]. Clustering [3] is representative of the former, while regression and classification are representative of the latter. The most common supervised learning algorithms include decision tree [4], naïve Bayes [5], k-Nearest Neighbor [6], neural networks [7], and support vector machine [8]. Among them, SVM was formally published in 1995 [9]. SVM is based on the statistical learning theorem and structural risk minimization principle to design the learning machine, which can not only tackle linear problems, but also tackle nonlinear problems using kernel methods. So far, SVM has provided powerful technical support for many areas due to excellent performance, including face recognition [10], intrusion detection [11], sound recognition [12], text classification [13], etc. Although SVM has many advantages, it also has some limitations. For instance, it is sensitive to the initial values of hyperparameters. These hyperparameters can affect the complexity and generalization ability of the model. Besides, the serious obstacle that all ML algorithms face is the curse of dimensionality [14]. If a dataset contains n features, the number of available solutions increases exponentially with the number of features, resulting in 2^n solutions being generated and evaluated. It increases the computational cost and affects the performance of the classifier.

Feature selection (FS) is a significant data preprocessing process. FS can reduce the difficulty of learning tasks by eliminating irrelevant or redundant features. If the subsequent learning process only constructs the model on these selected features, the curse of dimensionality can be significantly alleviated. Traditional subset search algorithms such as complete search and sequential search cannot achieve better output for the high computational cost. Many researchers have demonstrated the effectiveness of using metaheuristic algorithms (MAs) to solve complex optimization problems [15]. In 2020, Faramarzi et al. proposed the marine predators algorithm, which simulated the predation behavior of marine organisms based on Brownian motion and Lévy flight [16]. MPA has the features of few parameters, simple implementation, and flexibility in improving the version. Al-qaness et al. [17] presented an improved version of the adaptive neuro-fuzzy inference system (ANFIS) model to forecast the number of people infected by COVID-19 in four countries. In this work, MPA was used to optimize the ANFIS weights that lie between Layers 4 and 5. Over all the experiments, MPA-ANFIS outperformed all compared models on several measures. Zhong et al. [18] proposed a teaching-learning-based marine predator algorithm (TLMPA). This algorithm made full use of the good global search capability of teaching-learning-based optimization (TLBO) and the fast convergence capability of MPA. In addition, the cross mutation strategy of the differential evolution algorithm was also added to increase the diversity of the population. The comprehensive evaluation on IEEE CEC-2017 functions and engineering problems verified the outstanding performance of TLMPA. Elaziz et al. [19] hybridized MPA and moth flame optimization (MFO) based on their features. The proposed algorithm was called MPAMFO. This algorithm is first applied to segment medical images, such as COVID-19 CT images. Evaluation outcomes showed that the MPAMFO outperformed other algorithms in terms of structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), and fitness value.

Due to the more flexible and complex application areas, researchers have higher requirements for the native algorithm. No Free Lunch also turns out that the world is dominated by real-world problems without a known provably efficient algorithm [20]. Hence, researchers either continue to study new algorithms or improve the native ones [21]. In this paper, a co-evolutionary cultural mechanism-based marine predators algorithm is proposed. This mechanism divides the population into different sub-populations. The sub-populations evolve in parallel in their respective spaces based on MPA. The designed operators can coordinate the knowledge and experience among sub-populations. This improved version CECMPA increases the diversity of the population and the probability of avoiding trapping into the local optimum. For performance evaluation, CECMPA is used as a subset search algorithm while optimizing the hyperparameters of SVM. The simultaneous implementation of the two phases not only reduces the computational cost, but also improves the performance of the classifier. Specifically, the evaluation on the COVID-19 dataset is also of real great value.

The remainder of this paper is organized as follows: Section 2 presents some related works. Some theoretical background is introduced in Section 3. Section 4 highlights the proposed method. Experiments are implemented and results are analyzed in Section 5. Section 6 applies the proposed method to a COVID-19 dataset. Finally, conclusions and future researches are given in Section 7.

2. Related Work

Throughout the years, MAs have outstanding performance in various real-world or academic problems. A research extends the application of MAs to data classification and feature selection.

GA is one of the most classical evolutionary algorithms that simulates Darwinian natural selection. Gauthama Raman et al. [22] introduced the hypergraph into GA and proposed HG-GA for FS and parameter optimization in SVM. The hyper clique property was used to generate the initial population to prevent GA from trapping into the local optimum. The performance on intrusion datasets showed that HG-GA SVM performed better than comparators. Huang et al. [23] proposed a method called PSO-SVM to achieve the FS and kernel parameter setting of SVM. The proposed method was implemented with a distributed parallel system to reduce the training time. Baliarsingh et al. [24] proposed a method known as memetic algorithm-based SVM (M-SVM), which was inspired by embedding social engineering optimizer (SEO) in emperor penguin optimizer (EPO). SEO was considered a local search strategy, and EPO was used as a global optimization framework. The experiment was analyzed from two aspects, including binary-class classification and multi-class classification. The numerical results showed that M-SVM outperformed other comparison methods regarding gene selection and accuracy for microarray datasets. In 2018, a hybrid method based on the grasshopper optimization algorithm (GOA) was presented by Aljarah et al. [25] to find the optimal features and appropriate hyperparameters in SVM. The proposed method outperformed traditional grid search and seven well-regarded methods. The designed experiments used only one dataset with 100 features. For large-scale classification problems, the proposed method has not been further tested in this paper.

In [26], simulated annealing (SA) was hybridized into whale optimization algorithm (WOA) to enhance the global search ability of WOA. The two versions are called WOASAT-1 and WOASAT-2, respectively. Through validation, WOASAT-2 achieved better results regarding classification accuracy and feature reducts. In [27], the hybrid of seagull optimization

algorithm (SOA) and thermal exchange optimization (TEO) was applied to function optimization and FS. The authors proposed three different levels of hybrid versions. Through the comparison on UCI datasets, the third version SOA-TEO3 could better train the classifier to distinguish classes. MAs applied in this area also include salp swarm algorithm (SSA) [28], TLBO [29], harris hawks optimization (HHO) [30], etc.

3. Theoretical Background

3.1 Support vector machine

Given the dataset $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, $y_i \in \{-1, +1\}$, the basic idea of classification learning is to find a hyperplane and divide the samples into different classes. The optimal hyperplane can be described by the following linear equation:

$$h(x) = \omega^T x + b = 0 \quad (1)$$

where ω determines the direction of the hyperplane, and b determines the distance between the hyperplane and the origin. If $y_i = +1$, then $h(x) \geq +1$; if $y_i = -1$, then $h(x) \leq -1$. The sample point where $h(x) = \pm 1$ is true is called the support vector. The sum of the distances of two different classes of support vectors to the hyperplane is called the margin. Fig. 1 shows an example of a two-class SVM classification model using an optimal hyperplane.

Further, maximizing margin is equal to minimizing ω . Introducing the slack variable ξ , $\xi > 0$ represents that there are a small number of outliers. The penalty factor c is a hyperparameter that represents the tolerance to outliers. The basic model of SVM is as following:

$$\begin{cases} \min_{\omega, b, \xi} \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^m \xi_i \\ \text{s.t. } y_i (\omega^T x_i + b) \geq 1 - \xi_i, i = 1 \text{ to } m \end{cases} \quad (2)$$

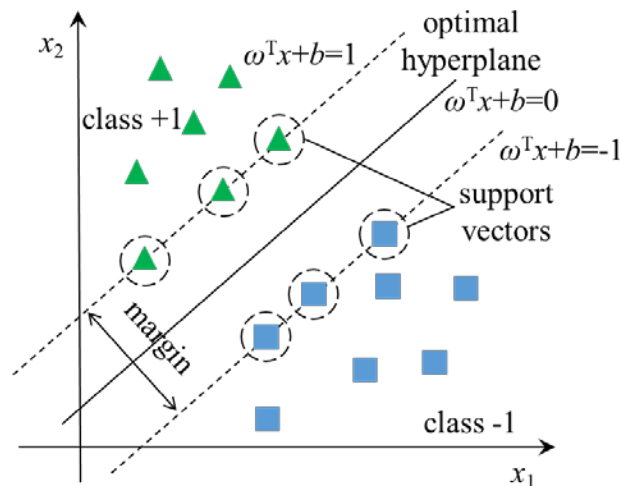


Fig. 1. Linear SVM.

For the nonlinear case, sample x can be mapped from the original space to the high-dimensional feature space. The mapped feature vector is $\phi(x)$. Because it is hard to calculate

the inner product $\phi(x_i)^T \phi(x_j)$, the kernel function $k(x_i, x_j)$ is defined. Specifically, the inner product of x_i and x_j in the feature space is equal to what they calculated in the original space by the kernel function. Then, the model of SVM can be transformed into the following dual problem:

$$\begin{cases} \min_{\alpha} & \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j k(x_i, x_j) - \sum_{i=1}^m \alpha_i \\ \text{s.t.} & \sum_{i=1}^m \alpha_i y_i = 0, 0 \leq \alpha_i \leq c, i = 1 \text{ to } m \end{cases} \quad (3)$$

where α_i represents the Lagrange multiplier.

In this paper, we adopt a widely useful radial basis function (RBF) kernel. Its expression is:

$$k(x_i, x_j) = e^{(-\gamma \|x_i - x_j\|^2)} \quad (4)$$

where γ is a kernel parameter which represents the width of the kernel function.

The penalty factor c and kernel parameter γ can affect the complexity and generalization ability and of SVM. This paper will use the proposed improved algorithm to optimize the two hyperparameters.

3.2 Marine predators algorithm

MPA is a novel swarm-inspired MA, which simulates the phenomenon of marine organisms looking for food. While the predator in the ocean is searching for prey, the prey is also searching for its food. According to the velocity ratio of predator and prey, the search process is defined as three phases. Here, the predator with the most robust foraging ability is called the elite. Like other swarm intelligence optimization algorithms, the final aim of MPA is to obtain the optimal solution (elite) according to the designed position update methods. The specific implementation is as follows:

Phase 1: When the velocity of the prey is higher than that of the predator, the predator is in a waiting state and observes the movement of the prey. This process occurs in the first third of iterations and is in the exploration phase of MPA.

$$\begin{aligned} \text{stepsize}_i &= R_B \otimes (\text{Elite}_i - R_B \otimes \text{Prey}_i) \quad i = 1, \dots, n \\ \text{Prey}_i &= \text{Prey}_i + P.R \otimes \text{stepsize}_i \end{aligned} \quad (5)$$

where $P=0.5$, Elite represents the position of the elite, Prey represents the position of the prey, n represents the size of the population, R_B represents a vector based on Brownian motion, and R represents a random number in $[0,1]$.

Phase 2: When the velocity of the prey matches that of the predator, the prey follows Brownian motion, and the elite follows Lévy flight. This process occurs in the middle of iterations and is in the exploration and exploitation phases of MPA. Half of the population is responsible for exploration using (6). The other half is responsible for exploitation using (7).

$$\begin{aligned} \text{stepsize}_i &= R_L \otimes (\text{Elite}_i - R_L \otimes \text{Prey}_i) \quad i = 1, \dots, n \\ \text{Prey}_i &= \text{Prey}_i + P.R \otimes \text{stepsize}_i \end{aligned} \quad (6)$$

$$\begin{aligned} \text{stepsize}_i &= R_B \otimes (R_B \otimes \text{Elite}_i - \text{Prey}_i) \quad i = 1, \dots, n \\ \text{Prey}_i &= \text{Elite}_i + P.CF \otimes \text{stepsize}_i \end{aligned} \quad (7)$$

where R_L represents a vector based on Lévy flight, and $CF=(1-t/t_{\max})^{(2-t/t_{\max})}$.

Phase 3: In the last third of iterations, the velocity of the predator is higher than that of the prey. The exploitation phase of the algorithm is further extended.

$$\begin{aligned} \text{stepsize}_i &= R_L \otimes (R_L \otimes \text{Elite}_i - \text{Prey}_i) \quad i = 1, \dots, n \\ \text{Prey}_i &= \text{Elite}_i + P.CF \otimes \text{stepsize}_i \end{aligned} \quad (8)$$

In addition to the above three phases, MPA also considers the Fish Aggregating Devices (FADs) effects. Equation (9) is used to simulate the prey jumping from one environment to another with a certain probability. It can effectively prevent the algorithm from trapping at the local optimum.

$$\text{Prey}_i = \begin{cases} \text{Prey}_i + CF[x_{\min} + R \otimes (x_{\max} - x_{\min})] \otimes U \\ \text{Prey}_i + [FADs(1-R) + R](\text{Prey}_{r_1} - \text{Prey}_{r_2}) \end{cases} \quad (9)$$

where $FADs=0.2$, x_{\min} and x_{\max} represent the lower and upper boundaries, r_1 and r_2 represent two random indexes of the prey. If the random number in $[0,1]$ is less than $FADs$, then $U=0$; otherwise, $U=1$.

3.3 Co-evolutionary cultural mechanism

The co-evolutionary cultural mechanism, as shown in Fig. 2, divides the population into different spaces. Each subpopulation evolves respectively in its own space through knowledge sources. The spaces also share accumulated experience with the lapse of iteration. The mechanism can be described as two basic operators: cultural operator and co-evolutionary operator. Firstly, the population of size n is divided into four subpopulations using (10). There are n_{11} agents in the belief space B1, n_{21} agents in the belief space B2, n_{12} agents in the population space P1, and n_{22} agents in the population space P2. The accept behaviors are implemented from the population space to the belief space according to the accept probability Acp . If $rand \leq Acp$, the best agent in P1 replaces the worst agent in B1; the best agent in P2 replaces the worst agent in B2. The affect behaviors are also implemented from the belief space to the population space according to the affect probability Afp . If $rand \leq Afp$, the Afn excellent agents in P1 replaces the Afn bad agents in B1; the Afn excellent agents in P2 replaces the Afn bad agents in B2 [31].

$$\begin{cases} n = (n_{11} + n_{12}) + (n_{21} + n_{22}) \\ n_{11} + n_{12} = n_{21} + n_{22} \end{cases} \quad (10)$$

Secondly, the shared global belief (SGB) space is established based on (11). Then, the average fitness of agents in SGB space is calculated. The agents whose fitness is larger than the average fitness are reserved, and the agents whose fitness is less than the average fitness are reinitialized. The affect behaviors are implemented from the SGB space to the population space according to the affect probability $SAfp$. If $rand \leq SAfp$, the $SAfn_1$ excellent agents in SGB replaces the $SAfn_1$ bad agents in P1, the $SAfn_2$ reinitialized agents in SGB replaces the $SAfn_2$ bad agents in P1; the $SAfn_1$ excellent agents in SGB replaces the $SAfn_1$ bad agents in P2, the $SAfn_2$ reinitialized agents in SGB replaces the $SAfn_2$ bad agents in P2. After the above affect behaviors, the experience exchange behaviors are implemented between P1 and P2 according to the exchange probability EEp . If $rand \leq EEp$, the EE_n excellent agents in P1 replaces the EE_n bad agents in P2; the EE_n excellent agents in P2 replaces the EE_n bad agents in P1 [32].

$$n_0 = n_{11} + n_{21} \quad (11)$$

In [33], the parameter guideline is given through a series of orthogonal experiments. (0.7, 0.2, 4, 0.7, 10, 0.5, 10) is an appropriate combination of (Acp , Afp , Afn , $SAfp$, $SAfn_1$, $SAfn_2$, EEp , EE_n).

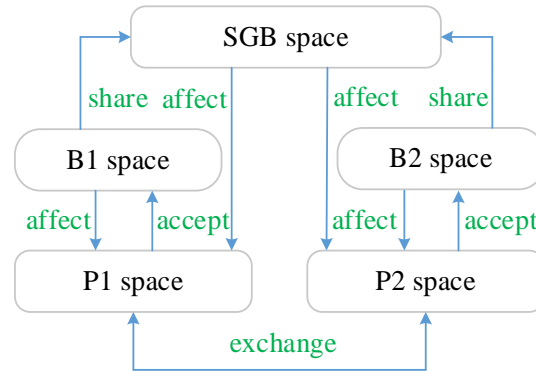


Fig. 2. Flowchart of the co-evolutionary cultural mechanism.

4. Proposed Method

4.1 Improved marine predators algorithm

MPA has such features of few parameters, simple implementation, and flexible improvement. However, it can be observed from (5)-(8) that the movement of prey is mainly based on Brownian motion and Lévy flight. These strategies have a certain degree of blindness. The solution space in real-world problems is usually more complex, which increases the probability of the native algorithm trapping at the local optimum. The co-evolutionary cultural mechanism can enhance the diversity of the population through "accept operators" and "affect operators", and provide the exchange of experience and knowledge for the exploitation of the native algorithm. Under this framework, the improved algorithm CECMPA can robustly avoid local optimum and improve the accuracy of the final solution. Besides, the solution of each space evolves respectively based on MPA. CECMPA does not damage the exploration of the native MPA. **Algorithm 1** provides the pseudo-code of CECMPA.

Algorithm 1 Pseudo-code of CECMPA.

- 1: Initialize the prey population n and parameters
- 2: Establish B1, P1, B2 and P2 spaces using (10) and SGB space using (11)
- 3: Evaluate the population
- 4: **While** $t < t_{\max}$
- 5: **For** each prey in B1, P1, B2 and P2
- 6: **If** $t < 1/3 * t_{\max}$
- 7: Update each prey using (5)
- 8: **Else if** $1/3 * t_{\max} < t < 2/3 * t_{\max}$
- 9: Update each prey in the first half using (6)
- 10: Update each prey in the latter half using (7)
- 11: **Else if** $t > 2/3 * t_{\max}$
- 12: Update each prey using (8)
- 13: **End if**
- 14: Apply *FADs* effects to each prey using (9)
- 15: Calculate the fitness and update the elite
- 16: **End for**

- 17: Implement accept behaviors from population spaces to belief spaces
 - 18: Implement affect behaviors from belief spaces to population spaces
 - 19: Update SGB space
 - 20: Implement affect behaviors from SGB space to population spaces
 - 21: Implement experience exchange between population spaces
 - 22: $t=t+1$
 - 23: **End while**
-

4.2 CECMPA for simultaneous FS and SVM optimization

No matter what kind of swarm-inspired MAs, the phenomenon simulated by the algorithm is to solve how to generate a new population with the lapse of iteration. The updating method of the population has been presented in the above section. Then, the two core issues of applying the optimization algorithm to the real-world area are the representation of the solution and the definition of the objective function. Once the two issues are solved, the application can be transformed into an optimization algorithm to solve the objective function.

In this paper, each solution consists of the penalty factor c , the kernel parameter γ , and features in the dataset as shown in (12). When designing a SVM model based on the RBF kernel, the penalty factor c controls the tolerance for errors of the model, and the kernel parameter γ determines the number of support vectors. Therefore, the hyperparameters combination (c, γ) can significantly affect the classification performance and generalization ability of the model. In FS tasks, all features (f_1, f_2, \dots, f_D) are encoded $[0,1]$. If the value of the element is within $[0.5,1]$, the feature is reserved; if not, the feature is excluded. Most traditional methods only perform FS or tune hyperparameters. But in fact, both of them affect each other in improving the performance of the classifier. Simultaneous FS and hyperparameters optimization can not only reduce the experimental cost, but also fully explore the potential of the classifier.

The objective function can evaluate the quality of the solution. FS aims to minimize the number of selected features and maximize the classification accuracy. The objective function in (13) that combines these two factors is used in this paper.

$$x_i = [c, \gamma, f_1, f_2, \dots, f_D], i = 1 \text{ to } n \quad (12)$$

$$F_{object} = \alpha E_R + (1 - \alpha) \frac{|R|}{|D|} \quad (13)$$

where f represents the feature in the dataset, E_R is the error rate of SVM, $|R|$ is the number of the selected features, $|D|$ is the total number of the features of the dataset, and $\alpha = 0.99$ [34].

The proposed method first preprocesses the dataset. After that, the dataset is divided into the training set and the testing set based on 10-fold cross validation. CECMPA executes the random vectors generated by (14). Then, SVM performs the training process on the training set. During this phase, the inner cross validation is carried out to produce unbiased classification results. CECMPA receives the fitness value at the end of the training process. All steps are repeated until the stop criterion is met. Finally, the proposed method reports the optimal solution. The final optimized hyperparameters and selected features are applied to the testing phase. The flowchart is shown in [Fig. 3](#).

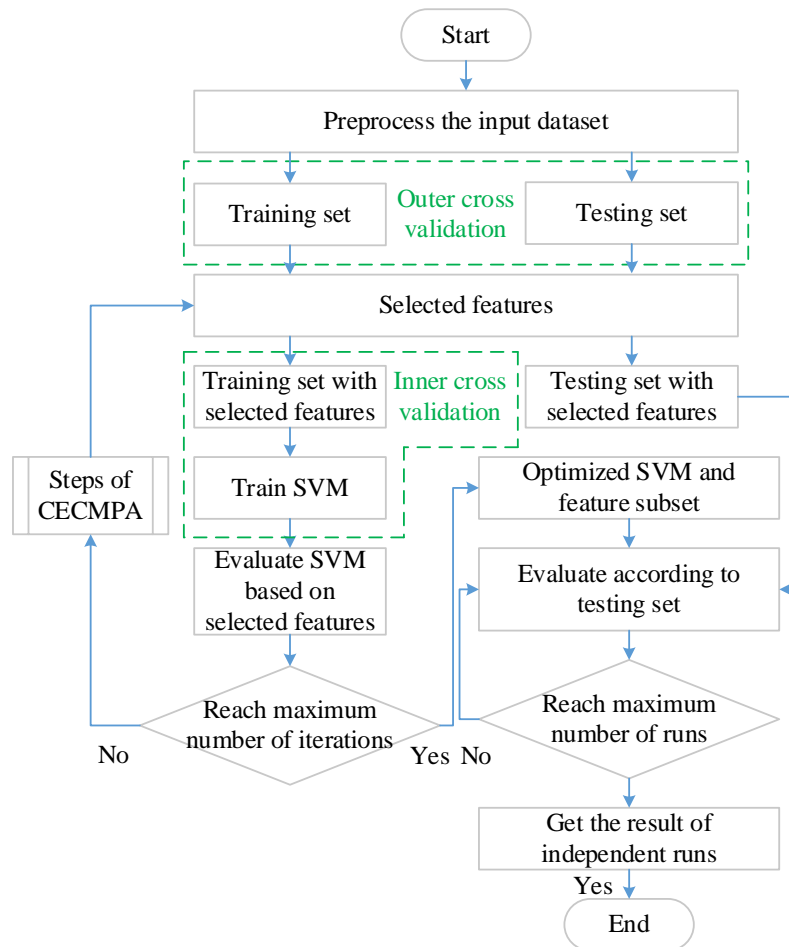


Fig. 3. Flowchart of the proposed method.

5. Experimental results and analysis

5.1 Dataset and parameters setting

Twelve UCI datasets [35] are used to evaluate the performance of the proposed method. Table 1 shows the number of features, instances, and classes in each dataset. Before executing the program, the dataset is preprocessed through two phases. Firstly, the logical features are converted into numeric form. And then, the min-max normalization scales the numeric features to [0,1]. In this way, the effect of numerical magnitude on feature weights can be avoided. LIBSVM is a well-known SVM toolbox [36]. 10-fold cross validation stratified resampling is performed to alleviate overfitting and underfitting in ML.

The proposed method is compared with MPA and five other FS methods including PSO [23], GOA [25], SSA [28], WOASAT-2 [26], and SOA-TEO3 [27]. Their parameters are shown in Table 2. The agent size of P1 and P2 are all 30, and the agent size of B1 and B2 are all 15. For the eight parameters in CECMPA, we follow the guideline in [35], as given in subsection 3.3. The maximum number of iterations is 100. Each experiment is run 10 times and carried out on MATLAB R2016b.

Table 1. Details of each UCI dataset.

Dataset	No. of features	No. of instances	No. of classes
Iris	4	150	3
ILPD	10	583	2
CongressEW	16	435	2
Zoo	16	101	7
Lymphography	18	148	4
Parkinsons	22	195	2
Flags	30	194	8
Dermatology	34	366	6
Ionosphere	34	351	2
Divorce Predictors	54	170	2
Lung cancer	56	32	3
Sonar	60	208	2

Table 2. Parameter settings of each compared method.

Method	Parameter	Value
PSO	Learning factors c_1 and c_2	2
	Maximum weight w_{\max}	0.95
	Minimum weight w_{\min}	0.05
	Maximum velocity v_{\max}	+200
	Minimum velocity v_{\min}	-200
GOA	Minimum coefficient c_{\min}	0.00001
	Maximum coefficient c_{\max}	1
SSA	Control Parameter c_1	[2,e-16]
	Random parameters c_2, c_3	(0,1)
WOASAT-2	Constant number b	1
	Initial temperature t_0	0.1
	Temperature reduction rate a	0.93
SOA-TEO3	Control Parameter A	[2, 0]
	Random number k	[0,2 π]
	Constant number u	1
	Constant number v	1

5.2 Evaluation metrics

1. Classification accuracy: this metric evaluates the accurate of SVM in predicting the right class using selected features. And the mathematical equation is as follows:

$$Accuracy = \frac{1}{m} \sum_{i=1}^m \Pi(h(x_i)=y_i) \quad (14)$$

where $\Pi(\cdot)$ is the indicator function. When the input is true, the output is 1, and when the input is false, the output is 0.

2. Feature size: this metric evaluates the size of the optimal feature subset obtained by the search algorithm. And the mathematical equation is as follows:

$$Size = \sum_{i=1}^D \Pi(f_i=1) \quad (15)$$

3. Fitness value: this metric combines the above two factors as the objective function. Please refer to (13) for the mathematical expression.
4. Computational time: this metric reflects the memory cost and execution speed of the method.

5. P-value: this metric provides information about whether a statistical hypothesis test is significant or not, and it also indicates something about how significant the result is.

5.3 Comparison with other methods

Table 3 presents the average classification accuracy of each method on each dataset. The proposed method achieves the highest results on nine out of twelve cases. It is quite impressive that CECMPA obtains 100% classification accuracy on "Divorce Predictors" dataset. Taking "Sonar" dataset as an example, the classification accuracy of the improved algorithm is at least 6.57% higher than that of the native MPA. It demonstrates that the SVM model optimized by CECMPA can more accurately predict the classes of samples. In the case of the "Lung cancer" dataset, WOASAT-2 is better than CECMPA with a slightly different result. **Table 4** presents the average feature size of each method on each dataset. The proposed method provides the smallest feature reduction on 66.67% of the datasets used. This shows that the proposed subset search algorithm CECMPA can more effectively eliminate irrelevant or redundant features. On the "Zoo" dataset, CECMPA provides 99.59% classification accuracy by using 4.8 features only, whereas SSA provides 94.05% classification accuracy and 5.6 features. **Table 5** presents the average fitness value of each method on each dataset. The objective function in (13) is a minimization problem. Observing the results obtained, the proposed method contributes to the lowest fitness value on most datasets. It proves that CECMPA can deeply explore areas in the feature space to obtain the optimal solution. The introduction of the co-evolutionary cultural mechanism reduces the risk of the native MPA trapping into the local optimum. The stability of the solutions across the different runs is a very important issue when MAs are used. **Table 6** presents the standard deviation of the fitness value. It can be seen that the optimal values (8 out of 12 cases) can be found by CECMPA. Observing the results of this comprehensive metric, the standard deviation of the solutions achieve a consistent ranking, indicating the stability of CECMPA. The average computational time of each method on each dataset is given in **Table 7**. In terms of the "CongressEW" dataset, the average computational time is sorted as follows: PSO>WOASAT-2>MPA>CECMPA>SSA>SOA-TEO3>GOA. From **Table 7**, the computational time of CECMPA is ranked in the middle of the seven methods on nearly all datasets. Besides, the computational time of CECMPA is slightly lower than that of MPA. Due to parallel evolution in different spaces, the proposed method can search the global optimum quickly while balancing exploration and exploitation. The computational time is especially important for real-world problems. Therefore, the proposed method guarantees high performance in an acceptable time.

Finally, some nonparametric statistical tests are carried out. In the area of inferential statistics, the null hypothesis H_0 represents the distribution of no difference. The level of significance α decides the probability of H_0 being rejected. In this section, pairwise comparison such as Wilcoxon rank-sum test detects significant differences between the two methods. α is stipulated as 0.05 [37]. When $p\text{-value} < 0.05$, the results are statistically different. The smaller p-values are stronger against H_0 . The p-values are reported in **Table 8**. P-values obtained are less than 0.05 on most cases, seen from **Table 8**. It demonstrates that the superiority of our improvement is statistically significant. So the null hypothesis H_0 can be rejected. Considering the comprehensive results of this section, the evaluation of the proposed method in this paper on the UCI dataset is positive.

Table 3. Comparison the average classification accuracy.

Dataset	PSO	GOA	SSA	WOAS AT-2	SOA- TEO3	MPA	CECM PA
Iris	0.9400	0.9467	0.9600	0.9800	0.9666	0.9866	0.9874
ILPD	0.7204	0.7390	0.7339	0.7341	0.7392	0.7375	0.7521
CongressEW	0.9698	0.9735	0.9827	0.9741	0.9740	0.9738	0.9750
Zoo	0.9504	0.9801	0.9405	0.9900	0.9603	0.9900	0.9959
Lymphography	0.8100	0.7837	0.8108	0.8783	0.8445	0.8648	0.8835
Parkinsons	0.9014	0.9179	0.8923	0.9025	0.8974	0.9175	0.9216
Flags	0.4278	0.4690	0.4587	0.6185	0.4948	0.6237	0.6377
Dermatology	0.9692	0.7653	0.8016	0.9832	0.7569	0.9850	0.9888
Ionosphere	0.9230	0.6780	0.6837	0.9715	0.7692	0.9544	0.9700
Divorce Predictors	0.9823	0.8000	0.8529	0.9882	0.8117	0.9941	1.0000
Lung cancer	0.8062	0.8926	0.8687	0.9062	0.8661	0.8500	0.8986
Sonar	0.8028	0.8701	0.7403	0.9519	0.9230	0.9086	0.9683

Table 4. Comparison the average feature size.

Dataset	PSO	GOA	SSA	WOAS AT-2	SOA- TEO3	MPA	CECM PA
Iris	1.9	1.5	3.0	2.1	2.8	3.0	2.0
ILPD	5.4	2.7	2.3	7.1	2.0	3.1	2.0
CongressEW	12.8	6.2	5.3	7.5	5.2	4.4	4.3
Zoo	6.1	6.6	5.6	10.3	5.1	7.7	4.8
Lymphography	10.9	6.7	6.8	11.0	7.0	13.2	6.5
Parkinsons	10.6	10.5	7.9	11.7	10.8	11.2	6.9
Flags	14.2	11.1	11.5	15.9	8.7	13.4	7.6
Dermatology	14.1	12.5	8.9	22.6	11.4	18.1	10.0
Ionosphere	19.0	11.6	10.4	19.6	7.2	19.3	12.0
Divorce Predictors	23.4	17.3	24.7	28.6	16.2	27.7	15.8
Lung cancer	26.8	27.5	26.1	19.8	30.4	26.9	22.4
Sonar	33.3	21.2	24.6	29.7	24.5	30.1	20.7

Table 5. Comparison the average fitness value.

Dataset	PSO	GOA	SSA	WOAS AT-2	SOA- TEO3	MPA	CECM PA
Iris	0.0619	0.0552	0.0471	0.0248	0.0380	0.0207	0.0199
ILPD	0.2817	0.2601	0.2652	0.2702	0.2600	0.2628	0.2474
CongressEW	0.0373	0.0293	0.0201	0.0299	0.0287	0.0281	0.0269
Zoo	0.0527	0.0233	0.0619	0.0160	0.0423	0.0141	0.0135
Lymphography	0.1928	0.2173	0.1906	0.1265	0.1577	0.1410	0.1130
Parkinsons	0.1010	0.0857	0.1097	0.1014	0.1060	0.0862	0.0859
Flags	0.5714	0.5291	0.5397	0.3829	0.5029	0.3771	0.3620
Dermatology	0.0345	0.2358	0.1986	0.0230	0.2438	0.0163	0.0156
Ionosphere	0.0817	0.3219	0.3160	0.0337	0.2305	0.0507	0.0392
Divorce Predictors	0.0217	0.2011	0.1500	0.0168	0.1893	0.0108	0.0100
Lung cancer	0.1918	0.0995	0.1055	0.0962	0.1031	0.1521	0.0987
Sonar	0.2006	0.1320	0.2610	0.0524	0.0801	0.0954	0.0443

Table 6. Comparison the standard deviation of fitness value.

Dataset	PSO	GOA	SSA	WOAS AT-2	SOA-TEO3	MPA	CECMA PA
Iris	0.0146	0.0117	0.0067	0.0017	0.0091	0.0016	0.0014
ILPD	0.0028	0.0024	0.0020	0.0026	0.0024	0.0025	0.0018
CongressEW	0.0056	0.0045	0.0023	0.0049	0.0034	0.0029	0.0027
Zoo	0.0214	0.0107	0.0228	0.0020	0.0118	0.0017	0.0015
Lymphography	0.0310	0.0358	0.0240	0.0103	0.0176	0.0150	0.0084
Parkinsons	0.0153	0.0060	0.0168	0.0154	0.0166	0.0095	0.0078
Flags	0.0360	0.0282	0.0327	0.0120	0.0241	0.0116	0.0114
Dermatology	0.0245	0.1132	0.1120	0.0131	0.1145	0.0115	0.0111
Ionosphere	0.0094	0.0860	0.0807	0.0042	0.0509	0.0060	0.0053
Divorce Predictors	0.0106	0.1201	0.1047	0.0090	0.1076	0.0089	0.0088
Lung cancer	0.0313	0.0108	0.0132	0.0090	0.0125	0.0229	0.0106
Sonar	0.1068	0.0936	0.1659	0.0109	0.0505	0.0669	0.0096

Table 7. Comparison the average computational time (s).

Dataset	PSO	GOA	SSA	WOAS AT-2	SOA-TEO3	MPA	CECMA PA
Iris	30.7	26.0	30.6	18.9	25.5	22.8	20.1
ILPD	243.4	210.1	205.8	300.0	206.0	199.9	188.9
CongressEW	43.0	22.6	23.9	28.4	22.8	26.4	24.5
Zoo	34.1	25.3	26.4	32.8	26.0	30.6	27.0
Lymphography	40.6	36.9	34.4	36.5	41.8	35.9	35.2
Parkinsons	34.2	28.5	28.8	39.9	28.0	32.9	31.6
Flags	110.3	90.1	93.7	107.7	92.4	101.3	93.3
Dermatology	252.8	198.1	206.6	142.1	195.0	201.8	190.4
Ionosphere	148.5	128.0	127.0	107.2	134.3	126.2	120.8
Divorce Predictors	44.0	31.8	28.4	22.8	29.7	28.3	28.0
Lung cancer	6.3	6.9	6.1	8.0	8.4	7.9	6.7
Sonar	83.3	69.2	67.2	82.6	64.8	77.0	71.2

Table 8. P-value obtained based on Wilcoxon rank sum test.

Dataset	CECMA vs PSO	CECMA vs GOA	CECMA vs SSA	CECMA vs WOASAT-2	CECMA vs SOA-TEO3	CECMA vs MPA
Iris	5.7E-40	4.5E-37	4.5E-38	2.1E-01	5.4E-34	9.6E-12
ILPD	9.5E-41	3.2E-29	7.6E-31	9.5E-41	4.2E-32	1.0E-30
CongressEW	4.8E-40	3.8E-01	9.6E-13	1.4E-07	1.6E-09	2.1E-10
Zoo	2.6E-18	2.3E-01	3.2E-07	2.3E-13	1.3E-02	4.9E-03
Lymphography	3.4E-39	1.8E-36	2.6E-36	6.0E-01	1.2E-35	1.3E-31
Parkinsons	2.5E-22	3.0E-09	1.1E-07	3.9E-01	2.8E-01	3.7E-10
Flags	3.0E-36	2.1E-35	6.4E-39	3.6E-02	8.9E-35	4.8E-23
Dermatology	8.7E-40	1.3E-37	8.9E-38	2.9E-24	7.4E-37	2.3E-16
Ionosphere	1.9E-38	2.2E-36	2.1E-35	1.1E-19	7.5E-35	9.1E-15
Divorce Predictors	7.2E-39	2.2E-38	3.6E-38	3.1E-37	1.1E-36	2.6E-10
Lung cancer	3.3E-34	4.6E-32	4.6E-32	5.0E-14	4.6E-32	4.9E-27
Sonar	4.4E-39	3.9E-35	4.6E-35	3.8E-05	2.8E-14	2.7E-34

6. Real-world application

The COVID-19 pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged in late 2019. Even over one year has passed since the COVID-19 onset, the confirmed cases and deaths still continue to increase significantly, and more than 90,000,000 people have been affected all over the world as of Jan. 2021. Besides, public health systems in most countries have been affected unprecedentedly. The clinical symptoms for individuals with confirmed COVID-19 include fatigue, cough, fever, and difficulty breathing within an incubation period of 2-14 days after exposure to the virus. At present, ML is the most common treatment for predicting large-scale pandemics. Through the possible prediction outcomes, the infected patients can be treated appropriately [38].

In this section, the proposed method uses the COVID-19 dataset to predict the severity of patients and is further compared and analyzed with other methods. The dataset used is public and available at [39]. To meet this study, the dataset has been pre-processed. It finally includes thirteen features and two targets as shown in Table 9.

Fig. 4 shows the classification accuracy and feature size for all the methods stated above. It reveals that CECMPA has the highest classification accuracy of 92.58%. We also discover that the proposed method achieves the prediction of the severity based on roughly four features. Besides, the results show that CECMPA does not select the symptom 4, symptom 5, and symptom 6. Judging from the result, these features are unprofitable for patient health prediction. The purpose of this case study is to provide radiologists and other clinicians with a diagnostic aid. Under no circumstances is this pilot study a substitute for medical advice.

Table 9. Details of the COVID-19 dataset.

No.	Feature	Description
1	location	the location of the patient
2	country	the native country of the patient
3	gender	the gender of the patient
4	age	the age of the patient
5	sym_on—hosp_vis	the number of days from symptoms being noticed to the patient visiting the hospital
6	vis_wuhan	whether the patient visited Wuhan, China
7	from_wuhan	whether the patient from Wuhan, China
8	symptom 1	symptom of patients
9	symptom 2	symptom of patients
10	symptom 3	symptom of patients
11	symptom 4	symptom of patients
12	symptom 5	symptom of patients
13	symptom 6	symptom of patients
Target: recovery/death		



Fig. 4. Accuracy and feature size on a COVID-19 dataset.

7. Conclusion

In this paper, a co-evolutionary cultural mechanism-based marine predators algorithm is proposed for simultaneous FS and hyperparameters optimization in SVM. Considering the features of MPA, the co-evolutionary cultural mechanism is used to improve the performance of the native algorithm in tackling ML problems. Twelve UCI datasets are used to evaluate the performance of CECMPA. Our findings show that the proposed method achieves similar or even higher classification accuracy in 75% of datasets, and searches the smaller optimal feature subset in 66.67% of datasets. Wilcoxon rank-sum tests are conducted to validate the significant improvement. Besides, the proposed method is also used to predict the COVID-19 pandemic. According to the obtained results, roughly four factors mainly affect the death and recovery conditions of patients for CECMPA. Therefore, we achieve the real-world goals of proposing a novel ML method in this paper.

For further research, the framework can try other belief spaces based on any evolutionary or swarm intelligence algorithm to improve the computational performance and solutions quality of MPA. In the COVID-19 era, CECMPA can also be applied to medical image segmentation to extract regions containing clinical features.

References

- [1] M. Shafiq, Z. Tian, A. K. Bashir, A. Jolfaei, and X. Yu, "Data mining and machine learning methods for sustainable smart cities traffic classification: A survey," *Sustainable Cities and Society*, vol. 60, pp. 102177, Sep. 2020. [Article \(CrossRef Link\)](#)

- [2] A. U. Haq, J. Li, A. Saboor, J. Khan, S. Wali, S. Ahmad, A. Ali, G. A. Khan, and W. Zhou, "Detection of Breast Cancer Through Clinical Data Using Supervised and Unsupervised Feature Selection Techniques," *IEEE Access*, vol. 9, pp. 22090-22105, 2021. [Article \(CrossRef Link\)](#)
- [3] G. A. Khan, J. Hu, T. Li, B. Diallo, and Q. Huang, "Weighted Multi-View Data Clustering via Joint Non-Negative Matrix Factorization," in *Proc. of 2019 IEEE 14th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*, pp. 1159-1165, 2019. [Article \(CrossRef Link\)](#)
- [4] Y. Zheng, "Decision tree algorithm for precision marketing via network channel," *Computer Systems Science and Engineering*, vol. 35, no. 4, pp. 293-298, 2020. [Article \(CrossRef Link\)](#)
- [5] K. Zhu, N. Zhang, S. Ying, and X. Wang, "Within-project and cross-project software defect prediction based on improved transfer naive bayes algorithm," *Computers, Materials & Continua*, vol. 63, no. 2, pp. 891-910, 2020. [Article \(CrossRef Link\)](#)
- [6] C. Zhang, J. Yao, G. Hu, and T. Schött, "Applying feature-weighted gradient decent k-nearest neighbor to select promising projects for scientific funding," *Computers, Materials & Continua*, vol. 64, no. 3, pp. 1741-1753, 2020. [Article \(CrossRef Link\)](#)
- [7] P. H. Thike, Z. Zhao, P. Liu, F. Bao, and Y. Jin, "An early stopping-based artificial neural network model for atmospheric corrosion prediction of carbon steel," *Computers, Materials & Continua*, vol. 65, no. 3, pp. 2091-2109, 2020. [Article \(CrossRef Link\)](#)
- [8] A. N. Khan, H. Nazarian, N. A. Golilarz, A. Addeh, J. P. Li, and G. A. Khan, "Brain Tumor Classification Using Efficient Deep Features of MRI Scans and Support Vector Machine," in *Proc. of 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, pp. 314-318, 2020. [Article \(CrossRef Link\)](#)
- [9] C. Cortes, V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, pp. 273-297, 1995. [Article \(CrossRef Link\)](#)
- [10] G. P. Nam, B. J. Kang, and K. R. Park, "Robustness of Face Recognition to Variations of Illumination on Mobile Devices Based on SVM," *KSII Transactions on Internet and Information Systems*, vol. 4, no. 1, pp. 25-44, 2010. [Article \(CrossRef Link\)](#)
- [11] Y. Liu, and D. Pi, "A Novel Kernel SVM Algorithm with Game Theory for Network Intrusion Detection," *KSII Transactions on Internet and Information Systems*, vol. 11, no. 8, pp. 4043-4060, 2017. [Article \(CrossRef Link\)](#)
- [12] Y. He, I. Ahmad, L. Shi, and K. Chang, "SVM-based Drone Sound Recognition using the Combination of HLA and WPT Techniques in Practical Noisy Environment," *KSII Transactions on Internet and Information Systems*, vol. 13, no. 10, pp. 5078-5094, 2019. [Article \(CrossRef Link\)](#)
- [13] R. Chen, C. Hsieh, "Web page classification based on a support vector machine using a weighted vote schema," *Expert Systems with Applications*, vol. 31, no. 2, pp. 427-435, 2006. [Article \(CrossRef Link\)](#)
- [14] Y. Xue, T. Tang, W. Pang, and A. X. Liu, "Self-adaptive parameter and strategy based particle swarm optimization for large-scale feature selection problems with multiple classifiers," *Applied Soft Computing*, vol. 88, pp. 106031, 2020. [Article \(CrossRef Link\)](#)
- [15] H. Jia, K. Sun, W. Song, X. Peng, C. Lang, and Y. Li, "Multi-Strategy Emperor Penguin Optimizer for RGB Histogram-Based Color Satellite Image Segmentation Using Masi Entropy," *IEEE Access*, vol. 7, pp. 134448-134474, 2019. [Article \(CrossRef Link\)](#)
- [16] A. Faramarzi, M. Heidarinejad, S. Mirjalili, and A. H. Gandomi, "Marine Predators Algorithm: A nature-inspired metaheuristic," *Expert Systems with Applications*, vol. 152, pp. 113377, Aug. 2020. [Article \(CrossRef Link\)](#)
- [17] M.A.A. Al-qaness, A.A. Ewees, H. Fan, L. Abualigah, and M. Abd Elaziz, "Marine Predators Algorithm for Forecasting Confirmed Cases of COVID-19 in Italy, USA, Iran and Korea," *International Journal of Environmental Research and Public Health*, vol. 17, no. 10, pp. 3520, 2020. [Article \(CrossRef Link\)](#)
- [18] K. Zhong, Q. Luo, Y. Zhou, and M. Jiang, "TLMPA: Teaching-learning-based Marine Predators algorithm," *AIMS Mathematics*, vol. 6, no. 2, pp. 1395-1442, 2021. [Article \(CrossRef Link\)](#)

- [19] M. A. Elaziz, A. A. Ewees, D. Yousri, H. S. N. Alwerfali, Q. A. Awad, S. Lu, and M. A. A. Al-Qaness, "An Improved Marine Predators Algorithm With Fuzzy Entropy for Multi-level Thresholding: Real World Example of COVID-19 CT Image Segmentation," *IEEE Access*, vol. 8, pp. 125306-125330, 2020. [Article \(CrossRef Link\)](#)
- [20] D. Wolpert, "No free lunch theorems for optimization," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 67-82, 1997. [Article \(CrossRef Link\)](#)
- [21] K. Sun, H. Jia, Y. Li, and Z. Jiang, "Hybrid Improved Slime Mould Algorithm with Adaptive β Hill Climbing for Numerical Optimization," *Journal of Intelligent & Fuzzy Systems*, vol. 40, no. 1, pp. 1667-1679, Jan. 2021. [Article \(CrossRef Link\)](#)
- [22] M.R. Gauthama Raman, N. Somu, K. Kirthivasan, R. Liscano, and V.S. Shankar Sriram, "An efficient intrusion detection system based on hypergraph - Genetic algorithm for parameter optimization and feature selection in support vector machine," *Knowledge-Based Systems*, vol. 134, pp. 1-12, Oct. 2017. [Article \(CrossRef Link\)](#)
- [23] C. Huang, and J. Dun, "A distributed PSO-SVM hybrid system with feature selection and parameter optimization," *Applied Soft Computing*, vol. 8, no. 4, pp. 1381-1391, Sep. 2008. [Article \(CrossRef Link\)](#)
- [24] S. Baliarsingh, W. Ding, S. Vipsita, and S. Bakshi, "A memetic algorithm using emperor penguin and social engineering optimization for medical data classification," *Applied Soft Computing*, vol. 85, pp. 105773, 2019. [Article \(CrossRef Link\)](#)
- [25] I. Aljarah, A.M. Al-Zoubi, H. Faris, M.A. Hassonah, S. Mirjalili, and H. Saadeh, "Simultaneous Feature Selection and Support Vector Machine Optimization Using the Grasshopper Optimization Algorithm," *Cognitive Computation*, vol. 10, pp. 478-495, 2018. [Article \(CrossRef Link\)](#)
- [26] M.M. Mafarja, and S. Mirjalili, "Hybrid Whale Optimization Algorithm with simulated annealing for feature selection," *Neurocomputing*, vol. 260, pp. 302-312, Oct. 2017. [Article \(CrossRef Link\)](#)
- [27] H. Jia, Z. Xing, and W. Song, "A New Hybrid Seagull Optimization Algorithm for Feature Selection," *IEEE Access*, vol. 7, pp. 49614-49631, 2019. [Article \(CrossRef Link\)](#)
- [28] A.M. Al-Zoubi, A.A. Heidari, M. Habib, H. Faris, I. Aljarah, and M.A. Hassonah, "Salp Chain-Based Optimization of Support Vector Machines and Feature Weighting for Medical Diagnostic Information Systems," *Evolutionary Machine Learning Techniques*, pp. 11-34, 2020. [Article \(CrossRef Link\)](#)
- [29] S.P. Das, and S. Padhy, "A novel hybrid model using teaching-learning-based optimization and a support vector machine for commodity futures index forecasting," *International Journal of Machine Learning and Cybernetics*, vol. 9, pp. 97-111, 2018. [Article \(CrossRef Link\)](#)
- [30] E. H. Houssein, M. E. Hosney, D. Oliva, W. M. Mohamed, and M. Hassaballah, "A novel hybrid Harris hawks optimization and support vector machines for drug design and discovery," *Computers & Chemical Engineering*, vol. 133, pp. 106656, 2020. [Article \(CrossRef Link\)](#)
- [31] R.G. Reynolds, "An Introduction to Cultural Algorithms," in *Proc. of the Third Annual Conference on Evolutionary Programming*, pp. 131-139, Feb. 1994. [Article \(CrossRef Link\)](#)
- [32] B. Liu, H. Ma, and X. Zhang, "A Co-evolutionary Differential Evolution Algorithm for Constrained Optimization," in *Proc. of Third International Conference on Natural Computation (ICNC 2007)*, pp. 51-57, 2007. [Article \(CrossRef Link\)](#)
- [33] Y. Sun, L. Zhang, and X. Gu, "A hybrid co-evolutionary cultural algorithm based on particle swarm optimization for solving global optimization problems," *Neurocomputing*, vol. 98, pp. 76-89, Dec. 2012. [Article \(CrossRef Link\)](#)
- [34] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary ant lion approaches for feature selection," *Neurocomputing*, vol. 213, pp. 54-65, Nov. 2016. [Article \(CrossRef Link\)](#)
- [35] UCI Machine Learning Repository. Accessed: Jan. 4, 2021. [Online]. Available: <http://archive.ics.uci.edu/ml/index.php>.
- [36] C. Chang, and C. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1-27, Apr. 2011. [Article \(CrossRef Link\)](#)
- [37] J. Derrac, S. García, D. Molina, F. Herrera, "A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms," *Swarm and Evolutionary Computation*, vol. 1, no. 1, pp. 3-18, 2011. [Article \(CrossRef Link\)](#)

- [38] C. Iwendi, A.K. Bashir, A. Peshkar, R. Sujatha, J.M. Chatterjee, S. Pasupuleti, R. Mishra, S. Pillai, and O. Jo, "COVID-19 patient health prediction using boosted random forest algorithm," *Frontiers in Public Health*, vol. 8, Jul. 2020. [Article \(CrossRef Link\)](#)
- [39] The COVID-19 dataset. Accessed: Jan. 4, 2021. [Online]. Available: <https://github.com/Atharva-Peshkar/Covid-19-Patient-Health-Analytics>.



Heming Jia received the Ph.D. degree in system engineering from Harbin Engineering University, China, in 2012. He is currently a professor in Sanming University. His research interests include nonlinear control theory and application, swarm optimization algorithm, image segmentation, and feature selection.



Kangjian Sun received the B. E. degree from Northeast Forestry University, China, in 2019. He is currently pursuing the M. E. degree in control engineering from Northeast Forestry University. His research interests include swarm intelligence optimization, image segmentation, and feature selection.



Yao Li was born in Yichun, China, in 1997. She is currently pursuing the M.S. degree in control theory and control engineering from Northeast Forestry University, China. Her research interests include swarm intelligence optimization and feature selection.



Ning Cao received his Ph.D. in computer science from University College Dublin in 2015. Up to date he has published over 100 publications in computer science. His research interests include IoT, security, and big data.