IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 22, 2018, accepted August 23, 2018, date of publication August 29, 2018, date of current version September 21, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2867728

A New Effective Machine Learning Framework for Sepsis Diagnosis

XIANCHUAN WANG[1], ZHIYI WANG[2], JIE WENG[2], CONGCONG WEN[3], HUILING CHEN[4], AND XIANQIN WANG[3]

[1] Information Technology Center, Wenzhou Medical University, Wenzhou 325035, China
[2] Department of Emergency Medicine, The Second Affiliated Hospital and Yuying Children's Hospital, Wenzhou Medical University, Wenzhou 325000, China
[3] School of Pharmaceutical Sciences, Wenzhou Medical University, Wenzhou 325035, China
[4] Department of Computer Science, Wenzhou University, Wenzhou 325035, China

Corresponding author: Huiling Chen (chenhuiling.jlu@gmail.com) and Xianqin Wang (lankywang@163.com)

This work was supported in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LY17F020012, in part by the Science and Technology Plan Project of Wenzhou, China, under Grant ZG2017019, in part by the Wenzhou Science and Technology Bureau Project under Grant Y20170810, and in part by the Wenzhou Medical University Scientific Research and Development Fund Project.

ABSTRACT There is a lack of early specific diagnosis and effective evaluation of sepsis, and the clinical treatment is not timely. As a result, the mortality is high, which seriously threatens the health of the people. Data were collected from the human blood samples of the hospital by gas chromatography mass spectrometry. Thirty-five healthy controls and 42 sepsis patients were enrolled. Machine-learning techniques were used to diagnose the sepsis. Using the metabolic data from the sepsis patients, the proposed method has got 81.6% recognition rate, 89.57% sensitivity, and 65.77% specificity. A new learning strategy was proposed to boost the performance of the kernel extreme learning machine, known as, chaotic fruit fly optimization, and two new mechanisms were introduced into the original a fruit fly optimization, including the chaotic population initialization and the chaotic local search strategy. To further enhance the diagnosis accuracy and identify the most important biomarkers, we performed the feature selection using the random forest before the construction of the classification model. The final established model, random forest-improved fruit fly optimization algorithm-kernel extreme learning machine, was used to effectively diagnose the sepsis. Experimental results demonstrate that the proposed method obtains better results than other methods across four performance metrics. We screened out five biomarkers and performed statistical analysis on these five substances. The level of acetic acid increased ($p < 0.05$) in the sepsis group, while the level of linoleic acid and cholesterol decreased ($p < 0.05$). The promising results suggest that the developed methodology can be a useful diagnostic tool for clinical decision support.

INDEX TERMS Fruit fly optimization, sepsis diagnosis, random forest, chaos theory, kernel extreme learning machine.

I. INTRODUCTION

Sepsis is a systemic inflammatory response syndrome caused by infection [1]–[3]. Clinical manifestations are fever, tachycardia, shortness of breath, and an increase in peripheral white blood cells. Serious sepsis can cause shock and progressive multiple organ failure, and it is common in clinical critically ill patients [4], [5]. An epidemiological survey showed that the occurrence of sepsis has increased yearly (currently approximately 78 in every 100,000 people) [6]. There is a lack of early specific diagnosis and effective evaluation of sepsis [7], [8], the clinical treatment is not timely, and the mortality rate is high, which seriously threatens the health of the people [4].

After suffering from sepsis, the patient's metabolic mechanism will change [9]–[13]. The metabolic characteristics also change based on a certain rule that is difficult to recover in a specific amount of time. A comprehensive analytical study of the changes in these metabolites may identify the chemical traces of sepsis, and a number of studies have shown that metabolomics techniques can quickly obtain this information [14]–[17]. This is due to the large amount of data obtained from metabolomics and the need for robust data processing techniques.

Research has reported metabolomics in sepsis by GC-MS [18]–[20], liquid chromatography-tandem mass spectrometry analysis (LC-MS) [13], [21]–[23], and nuclear

magnetic resonance spectroscopy (NMR) [10]–[12], [24]–[27]. However, few researches have reported using machine-learning techniques based on metabolomics data. Xu *et al*. proposed to use radial basis function neural network for prognostic evaluation of sepsis using HPLC/MS-based metabolites in rat serum, and a high accuracy over 94% was obtained [23]. Lin *et al*. constructed a model for prognostic evaluation of sepsis using RBFNN by NMR-based metabolites in rat serum, with an accuracy of approximately 87% [27].

In this study, we employed the kernel extreme learning machine (KELM) [28] method, to diagnose sepsis. To the best of the authors' knowledge, it is the first to explore the potential use of KELM in sepsis prediction. Using the metabolic data from sepsis patients, the developed method achieved good results with promising accuracy of 81.6%, 89.57% sensitivity and 65.77% specificity, respectively. To improve the performance of KELM, a fruit fly optimization algorithm (FOA) [29] based on the chaos theory was utilized to tune the hyperparameters of KELM. FOA has many good characteristics, such as simple implementation and lightweight computation. Because of its excellent characteristic, FOA has been recognized as a practical tool for solving many problems [29]–[38]. However, like other nature-inspired algorithms, traditional FOA is easy to become trapped in the local minimum and converge slowly. To ameliorate the convergence speed and increase the chance to jump out of the local optimum, we introduce two chaotic mechanisms into the original FOA. First, the chaotic initialization operation was used to improve the global search capability of FOA. On the other hand, the chaotic local search was used to strengthen the local search capability of FOA. The resultant improved FOA, named CFOA, and dynamically identifies the two parameters in KELM. To further improve the diagnosis accuracy and identify the most important biomarkers, the feature selection was completed using the random forest (RF) [39] before the construction of the predictive model. The final established model, RF-CFOA-KELM, was used to effectively diagnose sepsis.

For comparison purposes, other nature-inspired algorithms-based KELM models (including particle swarm optimization (PSO)-based KELM, genetic algorithms (GA)-based KELM, FOA-based KELM, and other popular learning algorithms including artificial neural networks (ANN) and support vector machines (SVM)) were also used in this study for the diagnosis of sepsis. The contributions of this paper are three-fold: (1) the proposal of an excellent method for sepsis diagnosis based on the metabolic information from patients with sepsis; (2) the proposal of an effective approach based on a chaos enhanced FOA evolutionary KELM for diagnosing sepsis; and (3) identification of the most important biomarkers with the assistance of the random forest.

The remainder of this paper is structed as follows. Section 2 gives a description on the clinical data involved in this study. The detailed flowchart of the proposed method is presented in sections 3 and 4. Section 5 provides the detailed experimental settings. The experimental results and related discussion are presented in section 6 and Section 7, respectively. Finally, we conclude the paper in section 8.

II. BACKGROUND

A. KERNEL EXTREME LEARNING MACHINE (KELM) FOR CLASSIFICATION

The extreme learning machine (ELM) [40] is different from the traditional learning and training process of the neural network. ELMs do not need to adjust the hidden layer threshold and the connection weight between the input and hidden layers. It only needs to adjust the number of nodes in the hidden layer. The optimal solution can be obtained after the number of hidden nodes of the hidden layer is determined. The ELM is a three-level ANN, which includes an input, output, and hidden layer. The hidden layer maps the input samples from the low-dimensional space to the high-dimensional space, and it transforms the linear-inseparable problem into a linear-separable one. However, there is a ''dimensional disaster'' in the operation of the high-dimensional feature space. The kernel function can replace this mapping to transform the linear inequalities into linear-separable problems and solve the ''dimension disaster'' problem. Compared with the ELM, the KELM [28] improves the robustness and nonlinear approximation ability of the whole system. KELM includes parameter sensitivity problems, such as RBF kernel-based KELM, that re mainly affected by kernel width and a penalty coefficient. This study focuses on these two parameters and proposes a new FOA strategy for tuning the two key parameters of KELM.

B. FRUIT FLY OPTIMIZATION ALGORITHM (FOA) FOR OPTIMIZATION

Fruit fly optimization algorithm (FOA) was developed by Pan [33]. Compared with other species, the fruit fly has more advantages in terms of sensory perception, especially smell and vision. The olfactory organs of fruit flies collect all kinds of odors in the air, including food sources that are 40 kilometers away. After flying to the food position, they use sharp vision to find the location of it and the position where their companions are gathering together, and they fly to that direction. The steps of the fruit flies searching for food are outlined as follows:

Step 1: Initialize the fruit fly population randomly.

Step 2: Initialize the random distance and direction for individual flies to search for food using their olfactories.

Step 3: Since it is impossible to know the location of food, the distance from the origin is first estimated. The determination value of the flavor concentration is then calculated, which is the inverse of the value of the distance.

Step 4: Substitute the value of the flavor concentration into the flavor concentration function to determine the flavor concentration of the individual's position.

Step 5: Determine the fruit fly with the highest flavor concentration in the population.

Step 6: Keep the best flavor concentration values in the coordinates of X and Y, at which point the fruit fly population flies to the position using vision.

Step 7: Enter iterative optimization by repeating steps 2-5, and estimate whether the current flavor concentration is superior to the previous iterative one. If it is, then go to step 6.

C. RANDOM FOREST (RF) FOR FEATURE SELECTION

Random forests (RF) [39] is a classifier that contains multiple decision trees. It is a combined classification model that is composed of many decision tree models. The basic idea is to first use bootstrap sampling to extract k samples from the original training set, where the sample capacity should be consistent with the original training capacity. Second, a decision tree model is established for each sample to obtain k classification results. The final classification is performed by voting according to the k classification results. An important characteristic of the RF is the ability to compute the importance of a single feature.

The importance of a feature X in the RF is calculated as follows:

1: For each decision tree in the RF, use the corresponding out of bag (OOB) data to calculate its OOB error, denoted as $errOOB1$.

2: Randomly add noise interference to the feature X of all samples of the OOB (the value of samples at the feature X can be changed randomly), and calculate its OOB error again, which is recorded as $errOOB2$.

3: Suppose there are $Ntree$trees in the RF. $Ntree$ specifies the number of decision trees contained in random forests, with a default value of 500. Then the importance of feature X is $\sum(\boldsymbol{errOOB2} - \boldsymbol{errOOB1})/\boldsymbol{Ntree}$, and this expression can be used to measure the importance of the corresponding feature. If a certain feature randomly adds noise, then the accuracy of the OOB data is greatly reduced, which indicates that this feature has significant influence on the classification result of the sample, that is, its importance is relatively high.

III. PROPOSED CFOA STRATEGY

This study proposed the CFOA learning strategy, in which two new mechanisms were introduced into the original FOA: the chaotic population initialization and chaotic local search strategy. The flowchart of the CFOA is displayed in Figure 1.

A. CHAOTIC LOCAL SEARCH

The applications of chaotic sequences to various metaheuristic optimization algorithms have attracted more attention in recent years. In [41], a novel GA with chaotic mutation was presented by replacing the Gaussian mutation operator in real-coded with a chaotic mapping. In [42], the chaotic initialization and chaotic sequences were introduced into simulated annealing (SA) instead of the Gaussian distribution. Other metaheuristic optimization algorithms also use chaos theory [43]–[45] to enhance the search performance.

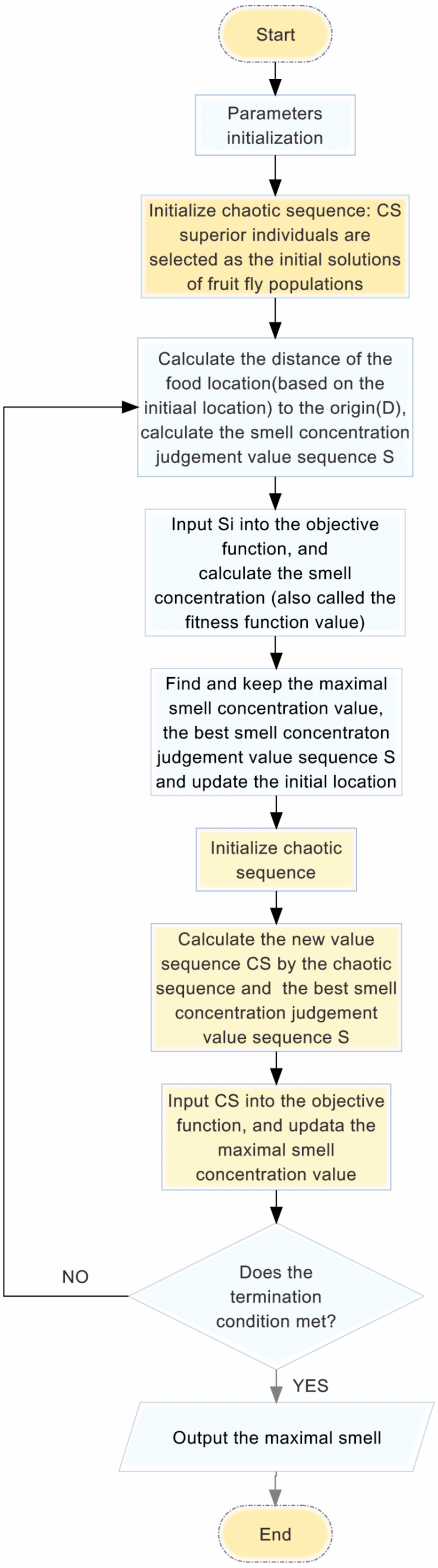


FIGURE 1. Flowchart of CFOA.

The chaotic sequence using logistic mapping can be produced as follows:

$$\boldsymbol{ch_{i+1}} = \mu \boldsymbol{ch_i} * (1 - \boldsymbol{ch_i}) \quad \boldsymbol{i} = 1, \ldots, \boldsymbol{S} - 1 \qquad (1)$$

where μ is the control parameter and $\mu = 4$. We set $0 < \boldsymbol{ch}_1 < 1$ and $\boldsymbol{ch}_1 \neq 0.25, 0.5, 0.75, 1$. It is not difficult to prove that when $\mu = 4$, the system is completely in chaos. S is the number of fruit fly populations.

It should be noted that the chaotic search usually reduces its performance when exploring a large search space. To overcome this, a strategy for the search space of the chaotic local search is introduced. Due to the randomness characteristic of the chaotic local search, the search process can eliminate premature convergence and local optima issues [46].

The expression of chaotic local search is presented as follows:

$$X_k^{\prime(c)} = (1 - \lambda)X^* + \lambda(LB + \beta_k(UB - LB)) \tag{2}$$

where $X_k^{\prime(c)}$ is the kth new position vector produced by chaotic local search. X^* is the position vector of the best solution obtained so far. β_k is kth chaotic value in chaotic sequence. LB and UB are the upper and lower bounds of the search space respectively. λ is the shrinking scale expressed as follows:

$$\lambda = \frac{Maxiter - l + 1}{Maxiter} \tag{3}$$

where $Maxiter$ represents the maximum number of iteration and l means the current number of iteration.

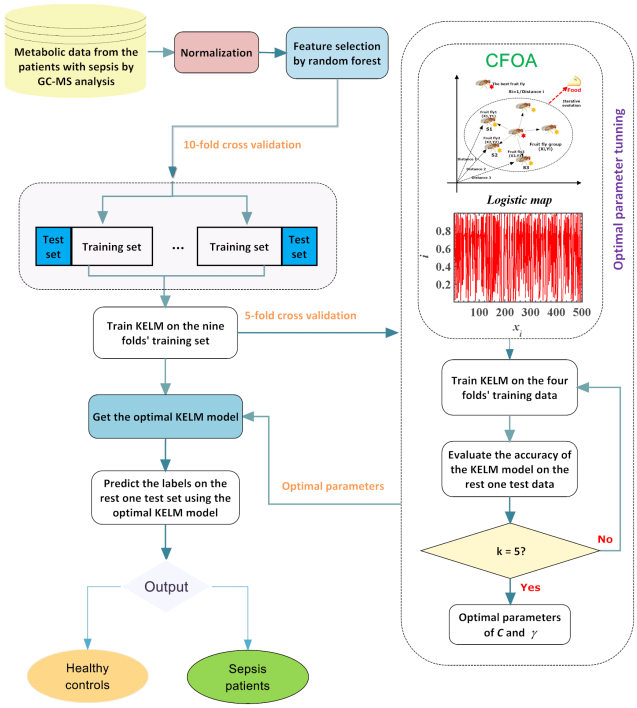


FIGURE 2. Flowchart of RF-CFOA-KELM.

IV. PROPOSED RF-CFOA-KELM DIAGNOSTIC METHOD

The flow chart of the established RF-CFOA-KELM is presented in Figure 2. The proposed method is composed of two stages. The first stage contains the feature selection and parameter optimization tasks. The feature selection was first executed by the RF method, and the two parameters of KELM were dynamically optimized by the proposed CFOA through five-fold cross validation (CV) analysis. The purpose of the first stage is to obtain the optimal feature subset and parameter pair. The main objective of the second part is to evaluate the classification performance of the proposed method. The final optimal parameters and features in the first stage were input into the KELM model in the outer loop through ten-fold CV. It should be noted that this inner fivefold CV and outer ten-fold CV scheme was adopted in many works [47], [48].

V. EXPERIMENTAL DESIGNS

A. PATIENTS AND BLOOD SAMPLES

A total of 42 sepsis patients with a mean age of 42.5 (±20.2) years were enrolled into this study (52.2% of whom were male) from the emergency department or intensive care unit (ICU) of The Second Affiliated Hospital of Wenzhou Medical University between January 2015 and January 2016.

Patients were less than 18 years old and those with venous nutrition or a metabolic disease were excluded. The samples for measurement were fasting blood that was obtained within 24 hours after sepsis diagnosis. The 35 healthy controls were randomly selected in the medical examination center over the same period.

Blood samples were collected from the healthy controls and sepsis group. The sample preparation process was previously published in our precious work [49]–[52].

B. GC-MS ANALYSIS

Analysis was conducted using Agilent Technologies 6890N-5975B GC-MS. The HP-5 MS column was at the ambient temperatures of 80°C for 5 min. The column oven temperature was increased to 260°C at a constant speed (10°C/min) and the temperature remained stable for 10 min [9]. The data gathered are exported to Microsoft Excel for analysis [53], [54].

C. STATISTICAL ANALYSIS

Version 18.0 SPSS software was applied to detect significant differences between two groups.

D. PARAMETER SETTING

MATLAB 2014a was used for simulation on a Windows 7 system. We chose the back propagation neural network (BPNN) with the Levenberg-Marquardt training algorithm. The KELM classifier was adopted from Huang at http://www3.ntu.edu.sg/home/egbhuang. The LIBSVM was developed by Chang and Lin [55]. We implemented the FOA, PSO, and GA from scratch. RF code was from http://code.google.com/p/randomforest-matlab.

Prior to classification, data normalization was used to convert all data to numbers between [−1,1]. k-fold CV was used to assess the accuracy of the classification to obtain unbiased estimates of the generalization accuracy [56].

Commonly used assessment criteria including Matthews correlation coefficients (MCC), classification accuracy (ACC), sensitivity and specificity were used to evaluate the quality of the experiments. The ACC is the most common evaluation index. It is easy to understand that the number of samples divided into the correct is divided by the number of all samples. In general, the higher the ACC was, the better the classifier obtained. The MCC characterizes the relevance of the data to a feature classification. Sensitivity represents the proportion of all positive cases that are correctly classified, which measures the ability of the classifier to recognize positive cases. Specificity represents the proportion of all negative cases that are correctly classified, and it measures the classifier's ability to identify negative cases.

VI. RESULTS

A. BENCHMARK FUNCTION VALIDATION

In this section, 12 classical functions are presented in Table 1, which are used to validate the optimization capability of CFOA. These functions are divided into two parts. The first part includes seven unimodal functions and the second part includes five multimodal functions. The boundary of the function's search space (range) and the optimal value of the functions ($f_{\min}$) are shown in Table 1.

TABLE 1. Benchmark functions.

Group	Function	Range	$f_{\min}$
Unimodal	$f_1(x)=\sum_{i=1}^{n} x_i^2$	[-100, 100]	0
	$f_2(x)=\sum_{i=1}^{n}\lvert x_i\rvert+\prod_{i=1}^{n}\lvert x_i\rvert$	[-10, 10]	0
	$f_3(x)=\sum_{i=1}^{n}(\sum_{j=1}^{i} x_j)^2$	[-100, 100]	0
	$f_4(x)=\max_i\{\lvert x_i\rvert, 1\le i\le n\}$	[-100, 100]	0
	$f_5(x)=\sum_{i=1}^{n-1}[100(x_{i+1}-x_i^2)^2+(x_i-1)^2]$	[-30, 30]	0
	$f_6(x)=\sum_{i=1}^{n}([x_i+0.5])^2$	[-100, 100]	0
	$f_7(x)=\sum_{i=1}^{n} ix_i^4+random[0,1)$	[-1.28, 1.28]	0
Multimodal	$f_8(x)=\sum_{i=1}^{n}-x_i\sin(\sqrt{\lvert x_i\rvert})$	[-500,500]	-418.9829×5
	$f_9(x)=\sum_{i=1}^{n}[x_i^2-10\cos(2\pi x_i)+10]$	[-5.12,5.12]	0
	$f_{10}(x)=-20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}\right)-\exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x\right.$	[-32,32]	0
	$f_{11}(x)=\frac{1}{4000}\sum_{i=1}^{n}x_i^2-\prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right)+1$	[-600,600]	0
	$f_{12}(x)=\frac{\pi}{n}\{10\sin(\pi y_1)+\sum_{i=1}^{n-1}(y_i-1)^2[1+10\sin^2(\pi y_{i+1})$,	[-50,50]	0
	$y_i=1+\frac{x_i+1}{4}$		
	$u(x_i,a,k,m)\begin{cases}k(x_i-a)^m & x_i>a\\ 0 & -a<x_i<a\\ k(-x_i-a)^m & x_i<-a\end{cases}$		

To explore the benefits of the CFOA, its performance has compared to the conventional FOA and nine successful metaheuristics: moth flame algorithm (MFO) [57], bat algorithm (BA) [58], dragonfly algorithm (DA) [59], flower pollination algorithm (FPA) [60], grasshopper optimization algorithm (GOA) [61], sine cosine algorithm (SCA) [62], particle swarm optimization (PSO) [63] and multi-verse optimization algorithm (MVO) [64]. In addition, the parameters that were set in all the experiments for 10 algorithms are shown in Table 2. To form a fair judgment, all algorithms are performed in the same testing environment and each benchmark function performs 30 independent runs. In this experiment, the dimension is set to 30, the population size and maximum number of iterations is set to 20 and 500 for all of the functions.

TABLE 2. Parameters setting of different algorithms.

Method	Population size	Maximum generation	Other parameters
CFOA	20	500	*ax*=20; *bx*=10; *ay*=20; *by*=10
FOA	20	500	*ax*=20; *bx*=10; *ay*=20; *by*=10
MFO	20	500	$a\in$[-1 -2]; b=1
BA	20	500	$Q\in$[0 2]; A=0.5; r=0.5
DA	20	500	$w\in$[0.9 0.2]; s = 0.1; a = 0.1; c = 0.7; f = 1; e = 1
FPA	20	500	*switch probability* p=0.8; λ =1.5
GOA	20	500	*cMax*=1; *cMin*=0.00004
SCA	20	500	a=2
MVO	20	500	*existence probability:* [0.2 1]; *travelling distance rate:*[0.6 1]
PSO	20	500	*inertial weight*=1; c_1=2; c_2=2

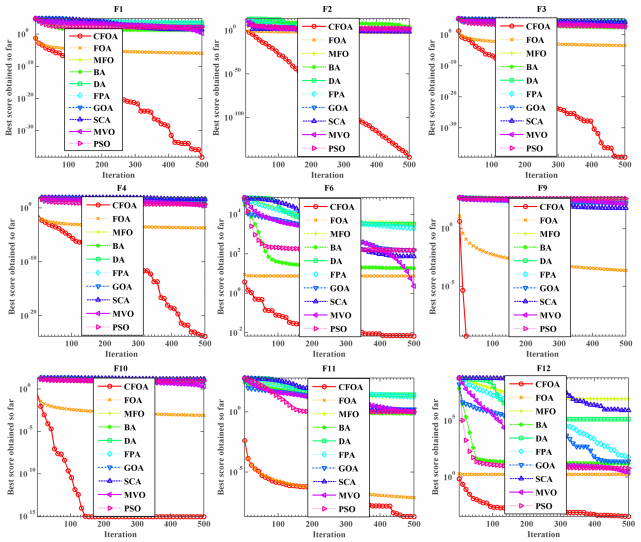


FIGURE 3. Convergence curves of some functions in F1-F12.

The results of different algorithms on the F1-F12 are shown and compared in Table 3, while Figure 3 shows the convergence tendencies in terms of the best fitness value of each algorithm. The table shows the mean (Avg) and standard deviation (Stdv) of the final values determined by each algorithm in all 30 independence runs. As shown from Table 3, the CFOA is the best algorithm and is superior to all

TABLE 3. Results of 12 benchmark functions.

	F1		F2		F3		F4	
	Avg	Stdv	Avg	Stdv	Avg	Stdv	Avg	Stdv
CFOA	5.84E-39	3.20E-38	2.14E-146	1.17E-145	3.66E-40	2.00E-39	1.59E-24	7.12E-24
FOA	1.14E-06	2.18E-08	5.85E-03	5.19E-05	3.59E-04	6.51E-06	1.94E-04	1.79E-06
MFO	2.74E+03	5.18E+03	4.43E+01	2.26E+01	2.11E+04	8.55E+03	7.37E+01	6.95E+00
BA	1.86E+01	2.51E+00	1.12E+04	5.08E+04	2.46E+02	7.66E+01	7.75E+00	4.33E+00
DA	3.41E+03	2.06E+03	2.35E+01	1.34E+01	1.99E+04	8.51E+03	3.55E+01	6.81E+00
FPA	1.76E+03	5.17E+02	3.72E+01	9.96E+00	1.75E+03	5.00E+02	2.51E+01	2.78E+00
GOA	1.95E+02	9.52E+01	3.79E+01	2.83E+01	5.13E+03	2.51E+03	1.88E+01	4.85E+00
SCA	2.71E+01	7.72E+01	5.49E-02	8.84E-02	1.50E+04	7.36E+03	4.32E+01	1.24E+01
MVO	1.86E+00	6.41E-01	8.27E+00	2.74E+01	4.47E+02	1.95E+02	2.90E+00	1.15E+00
PSO	1.52E+02	2.51E+01	2.49E+02	4.73E+02	6.94E+02	1.88E+02	5.07E+00	5.02E-01
	F5		**F6**		**F7**		**F8**	
	Avg	Stdv	Avg	Stdv	Avg	Stdv	Avg	Stdv
CFOA	1.17E+01	8.06E+00	6.77E-03	4.84E-03	9.82E-05	8.15E-05	-1.08E+04	1.99E+03
FOA	2.87E+01	9.56E-05	7.51E+00	4.81E-05	3.76E-04	1.80E-04	-8.11E+01	4.59E+01
MFO	2.70E+06	1.46E+07	3.69E+03	6.66E+03	3.08E+00	7.34E+00	-8.37E+03	8.99E+02
BA	6.53E+03	3.48E+03	1.86E+01	2.57E+00	2.38E+01	1.86E+01	-7.53E+03	5.67E+02
DA	8.22E+05	7.27E+05	3.20E+03	1.32E+03	8.87E-01	6.86E-01	-4.67E+03	5.66E+02
FPA	3.19E+05	1.76E+05	1.96E+03	6.67E+02	4.98E-01	1.86E-01	-7.33E+03	2.47E+02
GOA	4.68E+04	3.67E+04	1.54E+02	8.56E+01	1.92E+00	1.09E+00	-7.35E+03	6.26E+02
SCA	1.86E+05	6.27E+05	7.43E+01	1.55E+02	2.04E-01	3.10E-01	-3.55E+03	2.84E+02
MVO	4.29E+02	7.38E+02	2.29E+00	9.56E-01	4.83E-02	1.96E-02	-7.74E+03	5.66E+02
PSO	2.12E+05	4.73E+04	1.54E+02	1.58E+01	1.23E+02	2.97E+01	-6.51E+03	1.09E+03
	F9		**F10**		**F11**		**F12**	
	Avg	Stdv	Avg	Stdv	Avg	Stdv	Avg	Stdv
CFOA	0.00E+00	0.00E+00	8.88E-16	0.00E+00	2.07E-09	8.01E-09	3.14E-04	1.98E-04
FOA	2.26E-04	3.65E-06	7.79E-04	9.49E-06	7.57E-08	1.23E-09	1.67E+00	8.26E-06
MFO	1.76E+02	3.53E+01	1.85E+01	2.45E+00	1.62E+01	3.41E+01	8.75E+06	4.67E+07
BA	2.79E+02	2.49E+01	6.64E+00	4.84E+00	7.24E-01	5.85E-02	1.55E+01	6.59E+00
DA	1.99E+02	3.99E+01	1.23E+01	1.85E+00	2.61E+01	1.10E+01	1.34E+05	4.66E+05
FPA	1.50E+02	2.75E+01	1.20E+01	1.56E+00	1.77E+01	5.38E+00	5.84E+01	1.54E+02
GOA	1.46E+02	4.73E+01	7.76E+00	1.39E+00	1.60E+00	2.49E-01	2.36E+01	1.83E+01
SCA	5.52E+01	4.18E+01	1.53E+01	8.42E+00	1.21E+00	7.80E-01	8.32E+05	2.81E+06
MVO	1.37E+02	2.53E+01	2.08E+00	6.37E-01	9.62E-01	5.23E-02	2.58E+00	1.45E+00
PSO	3.77E+02	2.15E+01	8.88E+00	3.74E-01	1.04E+00	8.36E-03	5.94E+00	8.48E-01

other algorithms in addressing all cases (F1-F12). According to all metrics, the performance of the CFOA outperforms the conventional FOA method. Inspecting the rank results of the algorithms in Table 3, the overall ranks (in order) are the CFOA, FOA, MFO, BA, DA, FPA, GOA, SCA, MVO, and PSO algorithms, respectively.

As shown in Figure 3, the convergence tendency of the CFOA optimizer is better than other methods in all unimodal cases (F1-F7). For F2, the function value of the CFOA is smaller than the basic FOA and other popular algorithms all of the time. For F1, F3, and F4, the FOA has the effect of convergence, while the fastest convergence speed belongs to the proposed FOA-based optimizer. However, competitors such as MFO, BA, DA, FPA, GOA, SCA, MVO, and PSO algorithms fall into local optimum when solving F3 and F4 problems over the course of iterations. For F6, all of the methods are competitive, but the CFOA provides the best solution. In addition, the curve shows that F6 has fallen into a local minimum and eventually entered a stagnant state. The statistical results of the CFOA for F5 and F7 are shown in Table 3. It can be determined that the FOA and MVO give competitive results, but the best value is obtained by the CFOA.

According to the ranks of algorithms as shown in Table 4, the proposed CFOA is the best algorithm for addressing F8-F12 problems. The performance of the CFOA is superior

TABLE 4. Ranks of algorithms.

F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	All Rank	Average rank	Rank
1	1	1	1	1	1	1	1	1	1	1	1	12	1.00	1
2	2	2	2	2	3	2	10	2	2	2	2	33	2.75	2
9	8	10	10	10	10	8	2	7	10	8	10	102	8.50	10
4	10	3	5	4	4	9	4	9	4	3	5	64	5.33	4
10	5	9	8	9	9	6	8	8	8	10	8	98	8.17	9
8	6	6	7	8	8	5	6	6	7	9	7	83	6.92	8
7	7	7	6	5	7	7	5	5	5	7	6	74	6.17	5
5	3	8	9	6	5	4	9	3	9	6	9	76	6.33	6
3	4	4	3	3	2	3	3	4	3	4	3	39	3.25	3
6	9	5	4	7	6	10	7	10	6	5	4	79	6.58	7

to all algorithms in dealing with F8. As shown in Figure 3, the CFOA attained the exact optimal solutions for the 30-dimension problems F9 in all 30 runs. It is demonstrated that the chaotic strategy in the CFOA method improved the quality of solutions effectively. In F10, the function value of the CFOA is smaller than the basic FOA and other popular algorithms all of the time. F1 has a competitive performance compared to conventional FOA, but it converges to a better result after 200 iterations. The CFOA has attained the best solutions, while the basic FOA is stagnating during the whole iteration in addressing F12. In general cases, the ranks are approximately calculated as follows: CFOA>FPA>MVO>BA>GOA>SCA>PSO>FPA>DA>MFO. Comprehensively, the CFOA can realize enriched results, compared with other methods in this experiment.

TABLE 5. P-values of the Wilcoxon's signed-rank test of the CFOA results versus other algorithms.

Problem	FOA	MFO	BA	DA	FPA
F1	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F2	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F3	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F4	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F5	5.22E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F6	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F7	2.60E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F8	1.73E-06	**5.45E-02**	1.25E-04	3.18E-06	1.36E-04
F9	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F10	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F11	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
F12	1.73E-06	1.73E-06	1.73E-06	1.73E-06	1.73E-06
	GOA	**SCA**	**MVO**	**PSO**	
F1	1.73E-06	1.73E-06	1.73E-06	1.73E-06	
F2	1.73E-06	1.73E-06	1.73E-06	1.73E-06	
F3	1.73E-06	1.73E-06	1.73E-06	1.73E-06	
F4	1.73E-06	1.73E-06	1.73E-06	1.73E-06	
F5	1.73E-06	1.73E-06	1.73E-06	1.73E-06	
F6	1.73E-06	1.73E-06	1.73E-06	1.73E-06	
F7	1.73E-06	1.73E-06	1.73E-06	1.73E-06	
F8	9.71E-05	1.73E-06	1.29E-03	3.72E-05	
F9	1.73E-06	1.73E-06	1.73E-06	1.73E-06	
F10	1.73E-06	1.73E-06	1.73E-06	1.73E-06	
F11	1.73E-06	1.73E-06	1.73E-06	1.73E-06	
F12	1.73E-06	1.73E-06	1.73E-06	1.73E-06	

The Wilcoxon signed-rank test [65] is used to judge if the meaningful improvement of the proposed method of FOA is statistically significant. When the p-value is less than 0.05 (non-bold face in Table 5), it can be determined

that the results of the comparison of the proposed algorithm are statistically significant. If not, the result is not statistically significant. This comparison of the proposed method is done with respect to the competitors (as shown in Table 5), such that if the CFOA is the proposed method, comparison is performed between CFOA-FOA, CFOA-MFO, CFOA-BA, CFOA-DA, and so on. From Table 5, it can be seen that the CFOA outperforms other algorithms for the majority of the benchmark problems. For F8-F12, all p-values are less than 0.05. It can be detected that the results of the CFOA are statistically significant. Considering all p-values in Table 5, the comparison results of the CFOA and other methods, including the FOA, is statistically significant for all 12 problems, except F8. This proves that the conventional FOA has been meaningfully improved by the strategy utilized in this work.

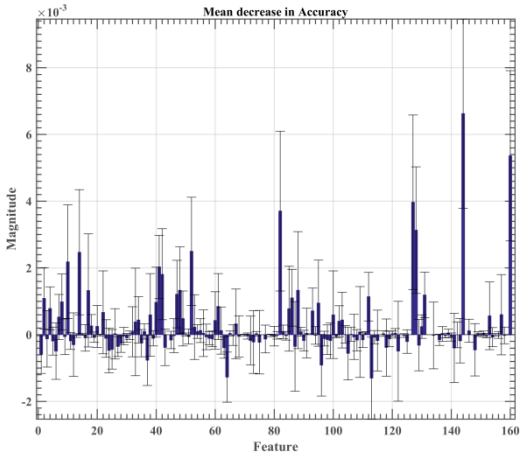


FIGURE 4. The importance of the features evaluated by the RF.

B. CLASSIFICATION RESULTS ON THE DIAGNOSIS OF SEPSIS

First of all, the RF feature selection was employed to assess the importance of the 160 features in the sepsis dataset. The result is shown in Figure 4. As shown, only a few of the 160 features have an effect on the diagnosis, and most features are redundant. Therefore, it is necessary to perform feature selection before training the model for higher diagnostic accuracy. After using the incremental selection method to test the feature combination of the best features, we found that the test result of the feature subset consisting of the five most important features (D-xylose, acetic acid, linoleic acid, D-glucopyranosiduronic acid, and cholesterol) is the best. Therefore, the best feature subset was used to train the learning model.

In this stage, we evaluated the effectiveness of the RF-CFOA-KELM model on the optimal feature subset. The detailed results are shown in Table 6. On average, RF-CFOA-KELM obtained results of 81.60% ACC, 0.7766 MCC, 89.57% sensitivity and 65.77% specificity, respectively.

TABLE 6. Classification performance of RF-CFOA-KELM in terms of ACC, MCC, sensitivity and specificity.

Fold	ACC	MCC	Sensitivity	Specificity
#1	0.7500	0.6000	0.6000	1.0000
#2	0.8750	0.7746	1.0000	0.7500
#3	0.8750	0.7454	0.8333	1.0000
#4	0.6250	0.4472	0.5000	1.0000
#5	0.7143	0.5477	1.0000	0.6000
#6	1.0000	1.0000	1.0000	1.0000
#7	0.8571	0.6455	0.5000	1.0000
#8	0.8750	0.6547	1.0000	0.8571
#9	0.7143	0.4167	0.6667	0.7500
#10	0.8750	0.7454	0.6667	1.0000
Mean	**0.8160**	**0.7766**	**0.8957**	**0.6577**
Std.	0.1110	0.2137	0.1478	0.1713

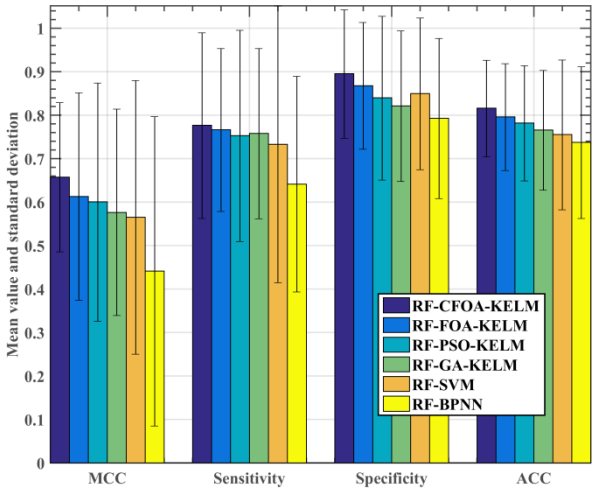


FIGURE 5. Classification performance achieved by the involved methods.

To validate the proposed approach, we conducted a comparative study of five other efficient models, including RF-FOA-KELM, RF-PSO-KELM, RF-GA-KELM, RF-SVM, and RF-BPNN. The detailed comparison of the six methods is shown in Figure 5. It is revealed that the RF-CFOA-KELM model is better than the original RF-FOA-KELM model in four evaluation indexes. For the MCC metric, RF-CFOA-KELM achieved the best result and smallest standard deviation. RF-FOA-KELM achieved second place, followed by RF-PSO-KELM, RF-GA-KELM, and RF-SVM. The result obtained by RF-BPNN is the worst. For the sensitivity metric, RF-CFOA-KELM achieved the best result. RF-FOA-KELM was the second best, followed by RF-GA-KELM, RF-PSO-KELM, and RF-SVM. The result obtained by RF-BPNN is the worst. For the specificity metric, RF-CFOA-KELM still achieved the best result and RF-FOA-KELM achieved the second best, followed by RF-SVM, RF-PSO-KELM, and RF-GA-KELM. The result obtained by RF-BPNN is still the worst. For the ACC metric, RF-CFOA-KELM obtained the best result. RF-FOA-KELM achieved the second-best result, followed by RF-PSO-KELM, RF-GA-KELM and RF-SVM. The result obtained by RF-BPNN is the worst. In summary, we observe that the

proposed RF-CFOA-KELM can obtain better results or very competitive results than other involved counterparts across four performance metrics.

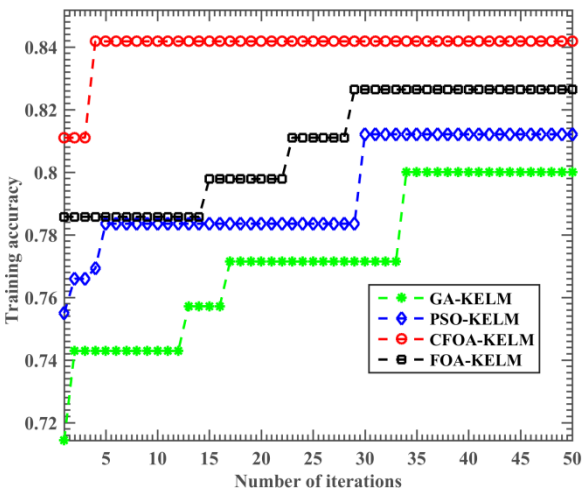


FIGURE 6. Convergence trends of the involved methods.

To verify the good convergence effect of the proposed RF-CFOA-KELM method, we recorded the convergence trend of various methods. In Figure 6, we found that RF-CFOA-KELM has achieved the best accuracy with a fast convergence speed, which means that CFOA has strong search ability and can force it to escape from local optimum. The main reason is that chaotic local search technology provides a strong search capability for FOA. Inspecting the curves in the figure, the RF-FOA-KELM model requires more iteration to converge, and the solution obtained is smaller than the RF-CFOA-KELM model. The GA algorithm has a weak global search capability, takes a long time to find the best solution and the result of RF-GA-KELM is not good.

VII. DISCUSSION

In this study, an accurate prediction model has been successfully established for discriminating between the healthy controls and the sepsis patients based on the metabolic data. Sepsis is a common clinical disease. It is very common in critically ill patients and is one of the common causes of death. The occurrence of sepsis will lead to excessive inflammation, endothelial damage, increased vascular permeability, oxygen utilization disorders, immune dysfunction, and high metabolic state, resulting in high morbidity and treatment costs. Therefore, sepsis has always been one of the focuses of medical research in critical illness. As a new subject after genomics, transcriptomics, and proteomics, metabolomics has been widely used in the study of various clinical diseases and has wide application prospects in clinical medicine.

To compensate for the lack of energy supply, the body during sepsis is forced to mobilize sugar, protein, and fat to provide more energy. Therefore, metabolites associated with the metabolism of sugar, protein, and fat can change significantly. We used the random forest algorithm to screen out five biomarkers (D-xylose, acetic acid, linoleic acid, D-glucopyranosiduronic acid, and cholesterol) and performed statistical analysis on each. The results of this study showed that the level of acetic acid increased ($p < 0.05$) in the sepsis group, while the level of linoleic acid and cholesterol decreased ($p < 0.05$), D-xylose and D-glucopyranosiduronic acid did not have statistically significance change ($p > 0.05$).

After sepsis, glucose metabolism increased significantly, resulting in large amounts of pyruvic acid [27]. Pyruvate is converted to acetyl-CoA by oxidative decarboxylation, which in turn produces acetyl phosphate, which ultimately produces acetic acid. Therefore, elevated levels of blood acetic acid may reflect hyperactivity of sugar and protein after sepsis.

After sepsis, the body is forced to carry out fat mobilization, resulting in a substantial increase in the production of free fatty acids and glycerol in the body [66], [67]. However, the results of this study indicate that, after sepsis, the levels of linoleic acid that is mainly involved in oxidative energy supply are significantly reduced. This may be related to a significant increase in free fatty acid (FFA) clearance in peripheral blood after sepsis [68]. The level of FFA in peripheral blood mainly depends on the rate of FFA production and clearance [69]. Although the FFA after sepsis is increased, the overall level of FFA is reduced in blood, due to the fact that endotoxin and inflammatory factors up-regulate the expression of fatty acid transporters and FAT/CD36 genes, allowing FFA in blood to be rapidly transported to muscle or other tissues [70]. In addition, lower FFA levels in the peripheral blood also indicate the body's energy supply and demand status. The lower the level of FFA, the more serious the imbalance is of the body's energy supply and demand. Because the imbalance of energy supply and demand is closely related to the prognosis of sepsis, the level of blood linoleic acid could be used to evaluate the condition of sepsis [23].

Cholesterol is an important structural component of cell membranes; it is also an important source of many steroid hormones and certain vitamins. During stress, the body releases a large number of steroid hormones and maintains the stability of the cell membrane, all of which need to consume cholesterol, while new cell synthesis uses total cholesterol to increase [71], [72]. Therefore, sepsis may cause a decrease in the total cholesterol level in the blood [73], [74]. The mortality of patients with sepsis is significantly associated with cholesterol levels. The lower the cholesterol concentration, the mortality rate will be the higher [73], [75]. The results of the study showed that the cholesterol level in severe sepsis patients was significantly correlated with death. The cholesterol level was progressively decreased; the risk of death was increased, indicating that cholesterol could also be used as a prognostic factor for severe sepsis.

VIII. CONCLUSIONS

We proposed to use KELM to diagnose sepsis. Using the metabolic data from the sepsis patients, the proposed method has achieved a predictive accuracy of 81.6%. To enhance

the performance of KELM, a new learning mechanism, the CFOA is proposed. We introduced two new mechanisms into the original FOA, including the chaotic population initialization and the chaotic local search strategy. To further promote the classification performance and identify the most important biomarkers, we performed the feature selection using the random forest prior to construct the classification model. The final established model, RF-CFOA-KELM, was used to diagnose the sepsis in an effective manner. We used the RF to screen out five biomarkers (D-xylose, acetic acid, linoleic acid, D-glucopyranosiduronic acid, and cholesterol), and we performed statistical analysis on these five substances. For comparison purposes, other nature-inspired algorithm-based KELM models and other popular machine-learning algorithms (including SVM and BPNN) were also used in this study for diagnosis of sepsis. In summary, we find that the proposed RF-CFOA-KELM can achieve better results or more competitive results than other involved counterparts across four performance metrics.

In the future, we will collect more relevant samples to establish more accurate sepsis severity prediction model. In addition, building an effective sepsis diagnosis decision support system based on the proposed model is also one of the future research directions.

Acknowledgment

(Huiling Chen and Xianqin Wang contributed equally to this work).

REFERENCES

[1] A. Shetty *et al.*, "Review article: Sepsis in the emergency department—Part 2: Investigations and monitoring," *Emergency Med. Australasia*, vol. 30, no. 1, pp. 4–12, 2018.

[2] H. C. Prescott and D. C. Angus, "Enhancing recovery from sepsis: A review," *Proc. JAMA*, vol. 319, no. 1, pp. 62–75, 2018.

[3] S. P. Macdonald *et al.*, "Review article: Sepsis in the emergency department—Part 1: Definitions and outcomes," *Emergency Med. Australasia*, vol. 29, no. 6, pp. 619–625, 2017.

[4] N. Alam, R. S. N. Panday, J. R. Heijnen, L. S. van Galen, M. H. H. Kramer, and P. W. B. Nanayakkara, "Long-term health related quality of life in patients with sepsis after intensive care stay: A systematic review," *Acute Med.*, vol. 16, no. 4, pp. 164–169, 2017.

[5] T. Kawasaki, "Update on pediatric sepsis: A review," *J. Intensive Care*, vol. 5, p. 47, Jul. 2017.

[6] L. Alberto, A. P. Marshall, R. Walker, and L. M. Aitken, "Screening for sepsis in general hospitalized patients: A systematic review," *J. Hospital Infection*, vol. 96, no. 4, pp. 305–315, 2017.

[7] F. F. Larsen and J. A. Petersen, "Novel biomarkers for sepsis: A narrative review," *Eur. J. Internal Med.*, vol. 45, pp. 46–50, Nov. 2017.

[8] S. S. Bansal, P. W. Pawar, A. S. Sawant, A. S. Tamhankar, S. R. Patil, and G. V. Kasat, "Predictive factors for fever and sepsis following percutaneous nephrolithotomy: A review of 580 patients," *Urol. Ann.*, vol. 9, no. 3, pp. 230–233, 2017.

[9] Z. Y. Wang *et al.*, "Metabolic study in serum from patients with sepsis and severe sepsis," *Int. J. Clin. Exp. Med.*, vol. 9, no. 3, pp. 6551–6556, 2016.

[10] V. Bakalov *et al.*, "Metabolomics with nuclear magnetic resonance spectroscopy in a *Drosophila melanogaster* model of surviving sepsis," *Metabolites*, vol. 6, no. 4, p. 47, 2016.

[11] K. A. Stringer *et al.*, "Whole blood reveals more metabolic detail of the human metabolome than serum as measured by ^{1}H-NMR spectroscopy: Implications for sepsis metabolomics," *Shock*, vol. 44, no. 3, pp. 200–208, 2015.

[12] M. Garcia-Simon *et al.*, "Prognosis biomarkers of severe sepsis and septic shock by ^{1}H NMR urine metabolomics in the intensive care unit," *PLoS ONE*, vol. 10, no. 11, p. e0140993, 2015.

[13] L. Su *et al.*, "Discrimination of sepsis stage metabolic profiles with an LC/MS-MS-based metabolomics approach," *BMJ Open Respiratory Res.*, vol. 1, no. 1, p. e000056, 2014.

[14] M. Eckerle *et al.*, "Metabolomics as a driver in advancing precision medicine in sepsis," *Pharmacotherapy*, vol. 37, no. 9, pp. 1023–1032, 2017.

[15] K. Zaitsu, Y. Hayashi, M. Kusano, H. Tsuchihashi, and A. Ishii, "Application of metabolomics to toxicology of drugs of abuse: A mini review of metabolomics approach to acute and chronic toxicity studies," *Drug Metabolism Pharmacokinetics*, vol. 31, no. 1, pp. 21–26, 2016.

[16] J. M. Shin, P. Kamarajan, J. C. Fenno, A. H. Rickard, and Y. L. Kapila, "Metabolomics of head and neck cancer: A mini-review," *Frontiers Physiol.*, vol. 7, p. 526, Nov. 2016.

[17] R. Pallares-Méndez, C. A. Aguilar-Salinas, I. Cruz-Bautista, and L. Del Bosque-Plata, "Metabolomics in diabetes, a review," *Ann. Med.*, vol. 48, nos. 1–2, pp. 89–102, 2016.

[18] V. Fanos *et al.*, "Urinary ^{1}H-NMR and GC-MS metabolomics predicts early and late onset neonatal sepsis," *Early Hum. Develop.*, vol. 90, pp. S78–S83, Mar. 2014.

[19] C. W. Seymour *et al.*, "Metabolomics in pneumonia and sepsis: An analysis of the GenIMS cohort study," *Intensive Care Med.*, vol. 39, no. 8, pp. 1423–1434, 2013.

[20] J. Hinkelbein *et al.*, "Alterations in cerebral metabolomics and proteomic expression during sepsis," *Current Neurovascular Res.*, vol. 4, no. 4, pp. 280–288, 2007.

[21] K. Sarafidis *et al.*, "Urine metabolomics in neonates with late-onset sepsis in a case-control study," *Sci Rep*, vol. 7, Apr. 2017, Art. no. 45506.

[22] Z. Lin *et al.*, "Comparison of sepsis rats induced by caecal ligation puncture or *Staphylococcus aureus* using a LC-QTOF-MS metabolomics approach," *Infection, Genet. Evol.*, vol. 43, pp. 86–93, Sep. 2016.

[23] P. B. Xu *et al.*, "A metabonomic approach to early prognostic evaluation of experimental sepsis," *J. Infection*, vol. 56, no. 6, pp. 474–481, 2008.

[24] A. Noto, M. Mussap, and V. Fanos, "Is ^{1}H NMR metabolomics becoming the promising early biomarker for neonatal sepsis and for monitoring the antibiotic toxicity?" *J. Chemotherapy*, vol. 26, no. 3, pp. 130–132, 2014.

[25] Y. Li, H. Liu, X. Wu, D. Li, and J. Huang, "An NMR metabolomics investigation of perturbations after treatment with chinese herbal medicine formula in an experimental model of sepsis," *OMICS, J. Integrative Biol.*, vol. 17, no. 5, pp. 252–258, 2013.

[26] K. A. Stringer, N. J. Serkova, A. Karnovsky, K. Guire, R. Paine, III, and T. J. Standiford, "Metabolic consequences of sepsis-induced acute lung injury revealed by plasma ^{1}H-nuclear magnetic resonance quantitative metabolomics and computational analysis," *Amer. J. Physiol.-Lung Cellular Mol. Physiol. Logo*, vol. 300, no. 1, pp. L4–L11, 2011.

[27] Z.-Y. Lin *et al.*, "A metabonomic approach to early prognostic evaluation of experimental sepsis by ^{1}H NMR and pattern recognition," *NMR Biomed.*, vol. 22, no. 6, pp. 601–608, 2009.

[28] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.

[29] L. Shen *et al.*, "Evolving support vector machines using fruit fly optimization for medical data classification," *Knowl.-Based Syst.*, vol. 96, pp. 61–75, Mar. 2016.

[30] R. Hu, S. Wen, Z. Zeng, and T. Huang, "A short-term power load forecasting model based on the generalized regression neural network with decreasing step fruit fly optimization algorithm," *Neurocomputing*, vol. 221, pp. 24–31, Jan. 2017.

[31] S. Kanarachos, A. M. Dizqah, G. Chrysakis, and M. E. Fitzpatrick, "Optimal design of a quadratic parameter varying vehicle suspension system using contrast-based fruit fly optimisation," *Appl. Soft Comput.*, vol. 62, pp. 463–477, Jan. 2018.

[32] J.-Q. Li, Q.-K. Pan, and K. Mao, "A hybrid fruit fly optimization algorithm for the realistic hybrid flowshop rescheduling problem in steelmaking systems," *IEEE Trans. Automat. Sci. Eng.*, vol. 13, no. 2, pp. 932–949, Apr. 2016.

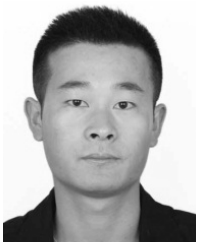
[33] W.-T. Pan, "A new Fruit Fly Optimization Algorithm: Taking the financial distress model as an example," *Knowl.-Based Syst.*, vol. 26, pp. 69–74, Feb. 2012.

[34] A. Darvish and A. Ebrahimzadeh, "Improved fruit-fly optimization algorithm and its applications in antenna arrays synthesis," *IEEE Trans. Antennas Propag.*, vol. 66, no. 4, pp. 1756–1766, Apr. 2018.

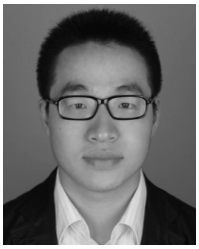
[35] T. Li, L. Gao, P. Li, and Q. Pan, "An ensemble fruit fly optimization algorithm for solving range image registration to improve quality inspection of free-form surface parts," *Inf. Sci.*, vols. 367–368, pp. 953–974, Nov. 2016.
[36] L. Wang, X. L. Zheng, and S. Y. Wang, "A novel binary fruit fly optimization algorithm for solving the multidimensional knapsack problem," *Knowl.-Based Syst.*, vol. 48, no. 2, pp. 17–23, 2013.
[37] J. Zhao and X. Yuan, "Multi-objective optimization of stand-alone hybrid PV-wind-diesel-battery system using improved fruit fly optimization algorithm," *Soft Comput.*, vol. 20, no. 7, pp. 2841–2853, 2016.
[38] X.-L. Zheng, L. Wang, and S.-Y. Wang, "A novel fruit fly optimization algorithm for the semiconductor final testing scheduling problem," *Knowl.-Based Syst.*, vol. 57, pp. 95–103, Feb. 2014.
[39] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[40] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, 2006.
[41] L. J. Yang and T. L. Chen, "Application of chaos in genetic algorithms," *Commun. Theor. Phys.*, vol. 38, no. 2, pp. 168–172, 2002.
[42] J. Mingjun and T. Huanwen, "Application of chaos in simulated annealing," *Chaos, Solitons Fractals*, vol. 21, no. 4, pp. 933–941, 2004.
[43] M. Wang *et al.*, "Toward an optimal kernel extreme learning machine using a chaotic moth-flame optimization strategy with applications in medical diagnoses," *Neurocomputing*, vol. 267, pp. 69–84, Dec. 2017.
[44] S. Saremi, S. Mirjalili, and A. Lewis, "Biogeography-based optimisation with chaos," *Neural Comput. Appl.*, vol. 25, no. 5, pp. 1077–1097, 2014.
[45] A. A. Heidari, R. Ali Abbaspour, and A. R. Jordehi, "An efficient chaotic water cycle algorithm for optimization tasks," *Neural Comput. Appl.*, vol. 28, no. 1, pp. 57–85, 2017.
[46] B. Liu, L. Wang, Y.-H. Jin, F. Tang, and D. X. Huang, "Improved particle swarm optimization combined with chaos," *Chaos, Solitons Fractals*, vol. 25, no. 5, pp. 1261–1271, 2005.
[47] H.-L. Chen *et al.*, "A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method," *Knowl.-Based Syst.*, vol. 24, no. 8, pp. 1348–1359, 2011.
[48] D. Zhao, C. Huang, Y. Wei, F. Yu, M. Wang, and H. Chen, "An effective computational model for bankruptcy prediction using kernel extreme learning machine approach," *Comput. Econ.*, vol. 49, no. 2, pp. 325–341, 2017.
[49] Q. Wu *et al.*, "Brain metabolic changes in rats after dextromethorphan," *Latin Amer. J. Pharmacy*, vol. 36, no. 9, pp. 1882–1886, 2017.
[50] C. C. Wen *et al.*, "Brain metabolomics in rats after administration of ketamine," *Biomed. Chromatogr.*, vol. 30, no. 1, pp. 81–84, 2016.
[51] M. Zhang *et al.*, "Serum metabolomics in rats models of ketamine abuse by gas chromatography–mass spectrometry," *J. Chromatogr. B*, vol. 1006, pp. 99–103, Dec. 2015.
[52] C. C. Wen, M. L. Zhang, J. S. Ma, L. F. Hu, X. Q. Wang, and G. Y. Lin, "Urine metabolomics in rats after administration of ketamine," *Drug Des. Develop. Therapy*, vol. 9, pp. 717–722, Feb. 2015.
[53] C. Wen *et al.*, "Metabolic changes in rat urine after acute paraquat poisoning and discriminated by support vector machine," *Biomed Chromatogr.*, vol. 30, no. 1, pp. 75–80, Jan. 2016.
[54] L. Yu, T. Liu, K. Liu, J. Jiang, and T. Wang, "A method for separation of overlapping absorption lines in intracavity gas detection," *Sens. Actuators B, Chem.*, vol. 228, pp. 10–15, Jun. 2016.
[55] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 389–396, 2011.
[56] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proc. 14th Int. Joint Conf. Artif. Intell.* Los Altos, CA, USA: Morgan Kaufmann, 1995, pp. 1137–1143.
[57] S. Mirjalili, "Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm," *Knowl.-Based Syst.*, vol. 89, pp. 228–249, Nov. 2015.
[58] X.-S. Yang, "A new metaheuristic bat-inspired algorithm," *Nature Inspired Cooperat. Strategies Optim.*, vol. 284, pp. 65–74, 2010.
[59] S. Mirjalili, "Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems," *Neural Comput. Appl.*, vol. 27, no. 4, pp. 1053–1073, 2016.
[60] X.-S. Yang, M. Karamanoglu, and X. He, "Flower pollination algorithm: A novel approach for multiobjective optimization," *Eng. Optim.*, vol. 46, no. 9, pp. 1222–1237, Apr. 2013.
[61] S. Saremi, S. Mirjalili, and A. Lewis, "Grasshopper optimisation algorithm: Theory and application," *Adv. Eng. Softw.*, vol. 105, pp. 30–47, Mar. 2017.
[62] S. Mirjalili, "SCA: A sine cosine algorithm for solving optimization problems," *Knowl.-Based Syst.*, vol. 96, pp. 120–133, Mar. 2016.
[63] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, vol. 4, Nov. 1995, pp. 1942–1948.
[64] S. Mirjalili, S. M. Mirjalili, and A. Hatamlou, "Multi-verse optimizer: A nature-inspired algorithm for global optimization," *Neural Comput. Appl.*, vol. 27, no. 2, pp. 495–513, 2016.
[65] S. García, A. Fernández, J. Luengo, and F. Herrera, "Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power," *Inf. Sci.*, vol. 180, no. 10, pp. 2044–2064, 2010.
[66] C. Chambrier, M. Laville, K. R. Berrada, M. Odeon, P. Boulétreau, and M. Beylot, "Insulin sensitivity of glucose and fat metabolism in severe sepsis," *Clin. Sci.*, vol. 99, no. 4, pp. 321–328, Oct. 2000.
[67] J. S. Samra, L. K. Summers, and K. N. Frayn, "Sepsis and fat metabolism," *Brit. J. Surg.*, vol. 83, no. 9, pp. 1186–1196, 1996.
[68] A. Martinez *et al.*, "Assessment of adipose tissue metabolism by means of subcutaneous microdialysis in patients with sepsis or circulatory failure," *Clin. Physiol. Funct. Imag.*, vol. 23, no. 5, pp. 286–292, 2003.
[69] R. A. Memon, K. R. Feingold, A. H. Moser, J. Fuller, and C. Grunfeld, "Regulation of fatty acid transport protein and fatty acid translocase mRNA levels by endotoxin and cytokines," *Amer. J. Physiol.-Endocrinol. Metabolism*, vol. 274, no. 2, pp. E210–E217, 1998.
[70] D. Brealey *et al.*, "Association between mitochondrial dysfunction and severity and outcome of septic shock," *Lancet*, vol. 360, no. 9328, pp. 219–223, 2002.
[71] J. Li, K. Xia, M. Xiong, X. Wang, and N. Yan, "Effects of sepsis on the metabolism of sphingomyelin and cholesterol in mice with liver dysfunction," *Exp. Therapeutic Med.*, vol. 14, no. 6, pp. 5635–5640, 2017.
[72] L. Vavrova, J. Rychlikova, M. Mrackova, O. Novakova, A. Zak, and F. Novak, "Increased inflammatory markers with altered antioxidant status persist after clinical recovery from severe sepsis: A correlation with low HDL cholesterol and albumin," *Clin. Exp. Med.*, vol. 16, no. 4, pp. 557–569, Nov. 2016.
[73] S. Yamano *et al.*, "Low total cholesterol and high total bilirubin are associated with prognosis in patients with prolonged sepsis," *J. Crit. Care*, vol. 31, no. 1, pp. 36–40, 2016.
[74] F. W. Guirgis *et al.*, "Cholesterol levels and long-term rates of community-acquired sepsis," *Crit. Care*, vol. 20, p. 408, Dec. 2016.
[75] L. Lagrost *et al.*, "Low preoperative cholesterol level is a risk factor of sepsis and poor clinical outcome in patients undergoing cardiac surgery with cardiopulmonary bypass," *Crit. Care Med.*, vol. 42, no. 5, pp. 1065–1073, 2014.

XIANCHUAN WANG received the master's degree in computer applied technology from Jiangsu University. He is currently a Senior Information Technology Engineer with the Information Technology Center, Wenzhou Medical University. He is involved in the construction of digital campus in colleges and universities. He has published two papers in the international journal. His research interest is focused on machine learning and data mining, as well as its application in medicine and other fields.

ZHIYI WANG was born in Wenzhou, China, in 1979. He received the master's degree in emergency medicine from Wenzhou Medical University in 2009. He is currently a Fellow of the Department of General Practice and the Emergency Centre, The Second Affiliated Hospital and Yuying Children's Hospital, Wenzhou Medical University. He is also a Medical Researcher with Wenzhou Medical University. In the last five years, he has published over 30 articles in international journals collected by Science Citation Index, such as *Biomedical Materials Research Part A* and *Biological and Pharmaceutical Bulletin*. His present research interests on toxicology pulmonary fibrosis disease.

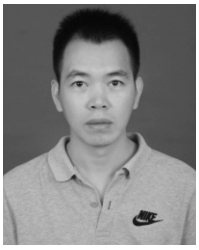


JIE WENG received the M.M. degree with the Department of Second Clinical Medicine, Wenzhou Medical University, China. He is currently a Resident Doctor with the Department of Emergency Medicine, The Second Affiliated Hospital and Yuying Children's Hospital, Wenzhou Medical University. He has published over 10 papers in international journals. His current research is focused on intoxication, toxicology, and acellularized scaffold. He is currently a Reviewer of *Human & Experimental Toxicology*.



CONGCONG WEN received the master's degree in clinical pharmacy in 2012. He was with the Laboratory Animal Center, Wenzhou Medical University, and was familiar with the establishment of experimental animal models and the drug analysis test in the body. He chaired a National Nature Fund and participated in the top three of provincial education departments and provincial projects. He can use *in vivo* drug analysis and metabonomics module GC-MS and LC-MS analysis technology, but also can be used to evaluate the changes of CYP450 enzyme activity in the body by Cocktail. Since 2012, he has been publishing over 10 papers with the first author in the *Jounar of Chromtography B*, *Xenobiotica*, *Acta Chromatographica*, *Latin American Journal of Pharmacy*, and other magazines.

HUILING CHEN received the Ph.D. degree from the Department of Computer Science and Technology, Jilin University, China. He is currently an Associate Professor with the Department of Computer Science, Wenzhou University, China. He has published over 80 papers in international journals and conference proceedings, including *Expert Systems with Applications*, *Knowledge-Based Systems*, *Neurocomputing*, *Soft Computing*, *PLOS One*, PAKDD, and among others. His present research interests include machine learning and data mining, as well as their applications such as medical diagnosis and bankruptcy prediction, among others. He is currently a Reviewer of the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, PART B.



XIANQIN WANG received the Ph.D. degree in science in agriculture from Anhui Agricultural University. He was involved in drug analysis and toxicological analysis, mastered poison detection technology, and metabolomics technology. He is currently a Senior Experimentalist with the Analysis and Testing Center, Wenzhou Medical University. He has published over 30 papers in international journals, including *Current Pharmaceutical Analysis*, the *Journal of Pharmaceutical and Biomedical Analysis*, the *Journal of Chromatography B*, *Acta Chromatographica*, and *Biomedical Chromatography*. He is currently a Reviewer of *Current Pharmaceutical Analysis*, the *Journal of Pharmaceutical and Biomedical Analysis*, the *Journal of Chromatography B*, and *Drug Development and Industrial Pharmacy*.