

Prediction of Postoperative Survival Level of Esophageal Cancer Patients Based on Kaplan-Meier(K-M) Survival Analysis and Gray Wolf Optimization (GWO)-BP Model

Enhao Liang

Zhengzhou University of Light Industry

Yanfeng Wang

Zhengzhou University of Light Industry

Junwei Sun (✉ junweisun@yeah.net)

Zhengzhou University of Light Industry

Lidong Wang

Zhengzhou University

Xueke Zhao

Zhengzhou University

Xin Song

Zhengzhou University

Research Article

Keywords: Esophageal squamous cell carcinoma (ESCC), Kaplan-Meier(K-M) survival analysis, BP network, K-M-GWO-BP model

Posted Date: September 29th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-762544/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Prediction of postoperative survival level of esophageal cancer patients based on Kaplan-Meier(K-M) survival analysis and Gray wolf optimization (GWO)-BP model
Enhao Liang^{1,2}, Yanfeng Wang^{1,2}, Junwei Sun^{1,2*}, Lidong Wang³, Xueke Zhao³, and Xin Song³

¹ Henan Key Lab of Information-Based Electrical Appliances, Zhengzhou University of Light Industry, Zhengzhou 450002, China

² School of Electrical and Information Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China

³ State Key Laboratory of Esophageal Cancer Prevention & Treatment and Henan Key Laboratory for Esophageal Cancer Research of The First Affiliated Hospital, Zhengzhou University, Zhengzhou, 450052, China

* Correspondence should be addressed to Junwei Sun: junweisun@yeah.net.

Abstract: Background: Esophageal squamous cell carcinoma (ESCC) is a global safety problem, especially the low 5-year survival rate of patients after surgery, and their healthy life after surgery is directly threatened. Methods: Kaplan-Meier(K-M) survival analysis is used to screen the blood indexes of patients with ESCC. The gray wolf algorithm (GWO) is introduced to optimize the weight threshold of back-propagation (BP) neural network, and a prediction model based on K-M-GWO-BP is established. Results: According to the influencing factors of postoperative survival, the postoperative survival level of patients is predicted. K-M survival analysis is used to analyze the relevant risk factors, the redundant variables are eliminated, and the whole structure of the neural network is simplified. The initial weight of BP neural network is optimized by GWO. Conclusions: BP neural network model, PSO-BP, GA-BP, SSA-BP, GWO-BP, K-M-BP, K-M-PSO-BP, K-MGA-BP, K-M-SSA-BP and K-M-GWO-BP are compared, the prediction accuracy of K-M-GWO-BP neural network model is the best.

Keywords: Esophageal squamous cell carcinoma (ESCC); Kaplan-Meier(K-M) survival analysis; BP network; K-M-GWO-BP model

1. Introduction

Esophageal cancer was one of the most common malignant tumors of digestive system in the world. Global Cancer Epidemiology Statistics (GLOBOCAN2018) shown that there are 572000 new cases of esophageal cancer worldwide, and 509000 cases were expected to die of esophageal cancer in 2018 [1]. The number of new cases of esophageal cancer in China ranked first in the world, accounting for about 50% of the global incidence of esophageal cancer [2,3]. It was one of the countries with the highest incidence of esophageal cancer in the world [4,5]. In China, squamous cell carcinoma was the main pathological type of esophageal cancer (EC), accounting for more than 90%. Surgery was the first choice for patients with resectable EC. With the progress of medical and health technology, the development of minimally invasive concept and the development of (ERAS) concept of accelerated rehabilitation surgery, the long-term prognosis of patients has been significantly improved [6,7]. However, due to the complexity of EC surgery, more postoperative complications, and a high recurrence rate after surgical resection, the 5-year survival rate was about 40% [8,9].

In fact, the survival rate of all patients with ESCC more than 5 years after operation is less than 20% [12]. According to the low accuracy of predicting the survival rate of cancer patients, recent studies have shown that a computer-aided classification method for lung cancer prediction based on evolutionary system has been proposed [13]. The work demonstrated that the proposed probabilistic genetic algorithm optimized neural network models, integrating with the t-SNE dimensionality reduction algorithm, achieved accurate prediction of patient survival [14]. The proposed GPU-based training of BP neural network was tested on a breast cancer data, which shown a significant

enhancement in training speed [16]. BP neural network model [17,18], genetic algorithm model [19,20], support vector machine model [21], decision tree method [22] and time series method [23] were commonly used prediction methods at present. However, BP neural network had some defects such as local optimization, irrelevant to physical meaning, strong dependence on training data and slow convergence speed, which hindered its application in practical engineering [17,19]. Strong macro search and global optimization capabilities were the characteristics of genetic algorithm (GA) [20]. The problem of local minimization of network could be solved to improve network performance. Therefore, GA was widely used to optimize BP neural network [19]. Due to the characteristics of multi-media and multi factors in the blood of esophageal cancer, it was difficult to determine the influencing factors which had the optimal correlation with the prediction indexes of the model. In the process of neural network modeling, it was time-consuming and difficult to optimize the neural network.

However, due to poor correlation between model input variables and output variables, and high redundancy and coupling relationship between variables, the defects of poor prediction accuracy of the model have not been well solved [24]. Some researchers adopt grey wolf optimization (GWO) to optimize the BP neural network's global search capability, which could greatly avoid trapping in local best solution [25-27]. The GWO studied in this paper was a new swarm intelligence optimization algorithm, which was proposed in 2014 [28]. The basic idea of GWO came from the mechanism of cooperative predation of gray wolf population. Compared with the previous algorithms, the setting of GWO was relatively simple, and only needs the guidance of individual gray wolf to hunt for the optimal solution. The GWO had the characteristics of simple implementation and fast convergence speed, which shown excellent results in standard test functions. At the same time, the research shown that the GWO algorithm was better than other intelligent optimization algorithms in some application fields, such as particle swarm optimization algorithm (PSO), GA [19]. The objectives of this work are summarized as follows.

- 1) A K-M-BP neural network model is proposed. The purpose of the model is to reduce the dimension of data and improve the accuracy of BP neural network prediction model. K-M analysis is used to screen the blood factors with high correlation with the survival level of patients to simplify the network structure. BP neural network is applied to predict the survival level of patients with esophageal cancer. Case study and experimental results demonstrate that K-M-BP neural network model is more effective than BP neural network model in predicting the survival level of patients.
- 2) Based on the proposed framework, a K-M-GWO-BP is proposed by adopting GWO as the optimizer for evolving the BP. GWO is used to optimize the BP neural network trained model to improve the prediction accuracy. The proposed GWO-BP is tested on a set of benchmark functions to verify its effectiveness. The prediction accuracy and applicability of BP, GWO-BP and K-M-GWO-BP prediction models are constructed to explore a new way of survival level prediction. The experimental results show that the proposed K-M-GWO-BP neural network model is superior to some of the latest BP neural network models in terms of calculation speed and prediction accuracy.

In the rest of this article, the sources of the data are described in Section II. Then, the proposed K-M and GWO and GWO-BP are given in Section III. Afterwards, the experimental results are detailed in Sections IV and V. Finally, conclusions are drawn and future work is outlined in Section VI.

2. Objects and analysis

2.1 Collect patient samples

A total of 331 patients with ESCC were treated in the affiliated Hospital of Zhengzhou University from January 2007 to December 2018, including 210 males (63.44%) and 121 females (36.56%). Patients were concentrated at age of 38 to 80 years old with average age of 60.61 years old.

2.2 Experimental data

The patient data included 17 blood indexes, such as WBC count ($10^9/L$), lymphocyte count ($10^9/L$), monocyte count ($10^9/L$), neutrophil count ($10^9/L$), eosinophil count ($10^9/L$), basophil count ($10^9/L$), red blood cell count ($10^9/L$), hemoglobin concentration (g/L), platelet count ($10^9/L$), total protein (g/L), white blood protein (g/L), globulin (g/L), PT (s), INR, APTT (s), TT (s), FIB (mg/dL).

Blood indicators were regarded as important factors in the clinical manifestations of cancer patients. The relationship between neutrophil to lymphocyte ratio (NLR), platelet to lymphocyte ratio (PLR) and lymphocyte to monocyte ratio (MLR) and the prognostic and clinicopathological significance in patients with ESCC have been reported by many studies. NLR, PLR and LMR might be served as prognostic markers in patients with ESCC [29]. Peripheral blood cell count ratio was suggested to evaluate clinical response and prognosis of patients with non-surgical ESCC [30]. Serum TT may be an important factor in prognosis of ESCC patients confirmed [31]. Preoperative serum FIB was validated to verify survival of ESCC, especially for the early pathological TNM stage (I-II) and N0 patients [32]. The nomogram combined with C-reactive protein (CRP) / ALB ratio could be used as a predictive model for the efficacy and survival outcome of thoracic ESCC treated with received chemo radiotherapy (CRT) or single radiotherapy (RT), which was found by zhang research [33].

3. Composite model for predicting survival level of patients with EC

3.1 K-M survival analyze

K-M survival analysis is a method to analyze the result and process of an event. K-M considers not only the occurrence of the event, but also the duration of the event. Therefore, survival analysis is also called time to event analysis. Survival analysis is very common for the study of survival time of cancer and other diseases in the medical field [34,35]. In order to analyze the influencing factors of survival time of EC patients, the blood indicators of patients are used as input and the survival time is used as output. The statistical software SPSS20.0 is used for K-M survival analyze. The accuracy of variable selection is determined by the size of correlation.

3.2 GWO algorithm

GWO is a new intelligent optimization algorithm proposed in 2014. The population system and predation behavior of grey wolf are imitated by GWO. In the Figure 1, the goal of optimization is achieved by simulating the hunting process of wolves. The wolf pack is composed of 5-12 wolves, which can be divided into 4 grades according to the fitness value. The wolf in the first layer of the pyramid is the leader wolf, which is expressed as α , and has the decision-making power on all major issues of the whole wolf pack. The wolf in the second layer is represented as β , which helps the leader wolf to make decisions. The wolf in the third layer is represented as δ , which is responsible for sentinel, reconnaissance and other tasks. The wolf at the bottom is denoted as ω , which is under the command of the first three levels of gray wolf. In the process of predation, α , β , δ wolves constantly change their positions to pursue prey, and the remaining gray wolf ω follows the first three, and the optimal solution is the specific location of prey. Due to the uncertainty of the location

of grey wolf, the distance between each wolf and its prey is expressed as follows :

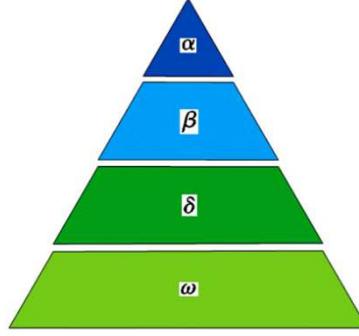


Figure 1. Hierarchy of grey wolf

$$D(t) = |C \cdot p(t) - X(t)| \quad (1)$$

The increasing number of iterations is represented by t . The prey of the t generation and the position of the wolf are expressed by $p(t)$ and $X(t)$, respectively. $C = 2r_1$, r_1 is a random number in $[0,1]$. The positions $X_\alpha(0)$, $X_\beta(0)$, $X_\delta(0)$ and $X_\omega(0)$ of each grey wolf are randomly initialized. The location of the wolf is updated in formula 2.

$$\begin{cases} X(t+1) = p(t) - A(t) \cdot D(t) \\ A(t) = 2a(t) \cdot r_2 - a(t) \\ a(t) = 2 - 2 \frac{t}{max} \end{cases} \quad (2)$$

r_2 is a random number at $[0,1]$. $a(t)$ is defined as the convergence factor. The custom maximum number of iterations is expressed as max . For the three wolves, there are the following mathematical descriptions:

$$\begin{cases} D_\alpha = |C_1 \cdot X_\alpha(t) - X(t)| \\ D_\beta = |C_2 \cdot X_\beta(t) - X(t)| \\ D_\delta = |C_3 \cdot X_\delta(t) - X(t)| \\ X_1 = X_\alpha - A_1 \cdot D_\alpha \\ X_2 = X_\beta - A_2 \cdot D_\beta \\ X_3 = X_\delta - A_3 \cdot D_\delta \end{cases} \quad (3)$$

The position of the next generation ω wolf is defined by equation (4).

$$X_\omega(t+1) = \frac{X_1 + X_2 + X_3}{3} \quad (4)$$

3.3 BP neural network algorithm

(1) Determine the input layer, hidden layer and output layer

The number of nodes in the input layer, hidden layer and output layer of the network is expressed by l , m and n , respectively. α is a random number in the range of 1-10, in Figure 2. The initial weight between input layer and hidden layer is determined by ω_{ij} , and that between hidden layer and output layer is determined by v_{jk} . The threshold of hidden layer is represented by a , $a = [a_1, a_2, \dots, a_m]$. The threshold of the output layer is expressed by b , $b = [b_1, b_2, \dots, b_n]$.

$$m = \sqrt{l + n} + \alpha \quad \alpha \in [1,10] \quad (5)$$

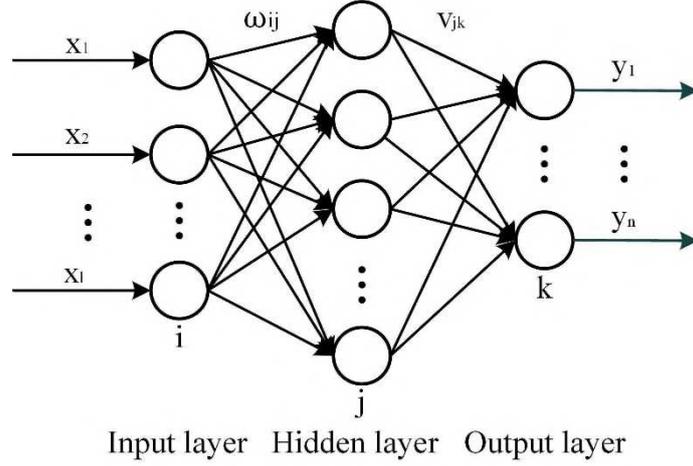


Figure.2 Structure diagram of neural network

(2) Calculate hidden layer output

$$h_j = f(\sum_{i=1}^l \omega_{ij} x_i - a_j) \quad (6)$$

In formula (6), the output of the j th neuron in the hidden layer is h_j . x_i is the input of the i th neuron in the input layer. a_j is the j th threshold of the hidden layer.

(3) Calculate output of output layer

$$O_k = \sum_{j=1}^m h_j v_{jk} - b_k \quad (k = 1, 2, \dots, n) \quad (7)$$

The k th threshold of the output layer is represented by b_k . The output of the output layer is O_k .

(4) Update connection layer weights

The following objective functions are defined:

$$J = \frac{1}{2} \sum_{s=1}^A \sum_{k=1}^n (y_k^s - o_k^s)^2 \quad (8)$$

In equation (8), A is the number of training samples and n is the number of output nodes. y_k^s is the expected output of sample s ; o_k^s is the output of the k th output node under the action of sample s .

The weight update function can be expressed as:

$$\omega_{ij}(t+1) = \omega_{ij}(t) + \mu[(1-\gamma)D(t) + \gamma D(t-1)] \quad (9)$$

$$v_{jk}(t+1) = v_{jk}(t) + \mu[(1-\gamma)E(t) + \gamma E(t-1)] \quad (10)$$

In equation (9) and equation (10), $i=1, 2, \dots, l, j=1, 2, \dots, m, k=1, 2, \dots, n$. The learning efficiency is μ , and $\mu > 0$. The inertia coefficient is γ , and $0 \leq \gamma < 1$. $D(t) = -\frac{\partial J}{\partial \omega_{ij}(t)}$; $E(t) = -\frac{\partial J}{\partial v_{jk}(t)}$.

(5) Update threshold

$$\begin{cases} a_j(t+1) = a_j(t) + \mu h_j(1-h_j) \sum_{k=1}^n v_{jk}(y_k - O_k) \\ b_k(t+1) = b_k(t) + (y_k - O_k) \end{cases} \quad (11)$$

(6) Judge

Whether the algorithm reaches the maximum number of iterations is judged. If the maximum number of iterations is not reached, return to step (2). If the maximum number of iterations is reached, the network training ends.

3.4 GWO-BP neural network algorithm

The convergence speed of BP neural network is slow and easy to fall into local minimum.

Therefore, GWO algorithm is used to enhance the global search ability. As shown in Figure 3, the gray wolf position is taken as the weight and threshold of BP neural network, and the gray wolf algorithm is iterated for many times. The location of prey is continuously judged and updated by gray wolf. The threshold and weight of BP neural network are constantly updated to calculate the global optimal result. The steps are as follows:

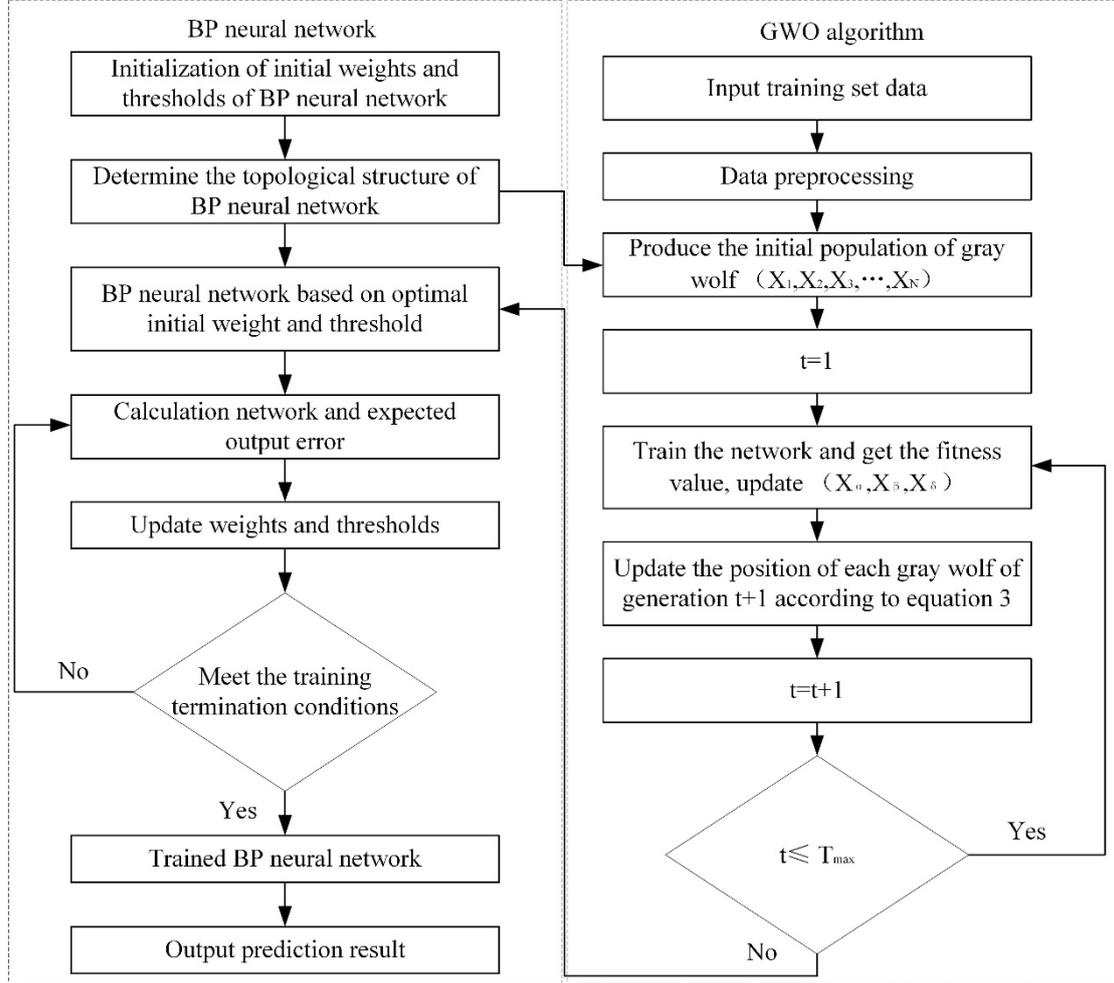


Figure. 3 Flow chart of GWO-BP neural network

Step 1: selecting appropriate training samples. The variables selected by K-M survival analysis are used as training input samples.

Step 2: the establishment of BP neural network model. The number of input layers is 1. The number of output layers is n. The number of hidden layer neuron nodes m, as shown in formula (5). a is an arbitrary constant from 1 to 10. Therefore, after many experiments, it can be concluded that when a is 5, the convergence speed and fitting accuracy of the neural network model are the most suitable in the table.

Step 3: initialization of GWO optimization algorithm. The optimal positions X_α , X_β and X_δ are initialized.

Step 4: calculating individual fitness value. The weights and thresholds of BP neural network are set as the object of GWO algorithm. The error sum of each neural node of BP neural network is used as the fitness function of GWO optimization algorithm to measure the individual position, and the position of the current optimal fitness value is obtained.

$$F = \frac{1}{A} \sum_{s=1}^A \sqrt{\sum_{k=1}^n (y_k^s - o_k^s)^2} \quad (12)$$

Step 5: updating the parameters r_1, r_2, q in GWO. According to formula (1) and equation (2), the position of each wolf was updated, and a new BP neural network is constructed and trained. According to equation (12), the fitness function value of each wolf is calculated, and the new α, β, δ are determined again.

Step 6: determining the number of iterations. When the number of iterations reaches the upper limit, GWO optimization algorithm is finished, and the optimal initial weights and thresholds of BP neural network are obtained. If the number of iterations does not reach the upper limit, return to step (5).

Step 7: output of prediction results. BP neural network is trained and evaluated according to the weights and thresholds optimized by GWO optimization algorithm, and finally the prediction results are obtained.

In the process of building the network, Matlab simulation software is used to update the individual position in GWO optimization algorithm until the number of iterations reaches the set value. As shown in the Figure 4, the optimal fitness value of GWO optimization algorithm before the number of iterations reaches 500. The optimal initial weights and thresholds of BP neural network are obtained by GWO optimization algorithm. When the number of iterations is 300, the optimal fitness value of GWO optimization algorithm before the number of iterations reaches 300. 300 iterations and 500 iterations are compared to calculate the speed and optimal value, and the optimal number of iterations 500 is obtained.

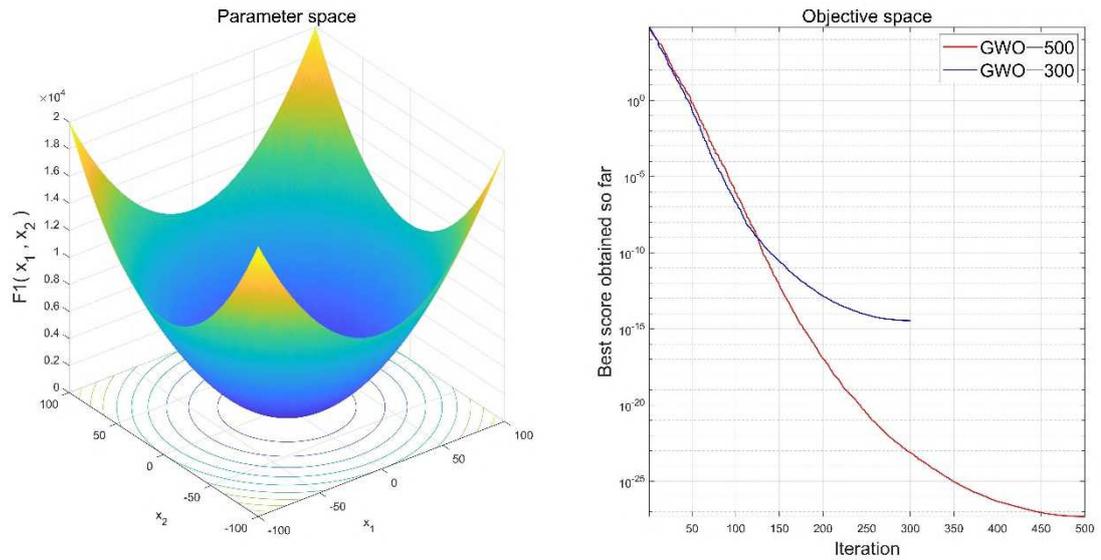


Figure 4. Optimal fitness graphs for 300 and 500 iterations

4. Case analysis

4.1 Selection of input and output parameters

Table.1 Log rank test of K-M survival analysis

Variables	Chi square	Degrees of freedom	Sig
WBCC	315.765	86	<0.001
TLC	203.852	33	<0.001
MONO	32.085	15	0.006
NEUT	411.260	81	<0.001
EOS	174.795	9	<0.001
BOS	42.030	6	<0.001
RBC	412.880	151	<0.001

HGB	163.149	71	<0.001
PLT	756.434	191	<0.001
TP	59.369	37	0.011
ALB	41.551	29	0.062
GLB	37.363	27	0.088
PT	114.245	65	<0.001
INR	110.603	65	<0.001
APTT	578.811	188	<0.001
TT	135.878	77	<0.001
FIB	262.722	115	<0.001
LMR	416.955	107	<0.001
NLR	1090.615	254	<0.001
PLR	1701.675	312	<0.001
PNI	254.672	61	<0.001

In the first step of prediction modeling, relevant data need to be obtained in Table 1. The input and output data in modeling are preprocessed to obtain accurate and applicable sample set. In view of the nonlinear complexity of the patient's blood system, K-M survival analysis is used to screen the input variables. The purpose of screening variables is related to survival level, and irrelevant variables are deleted. The significance of chi square value is less than 0.05, and the two variables are significantly correlated. The degree of freedom refers to the number of variables whose values are not limited when calculating a unified measurement. Significance refers to the risk level of rejecting zero hypothesis when zero hypothesis is true, also known as probability level, or significance level.

4.2 Performance test of GWO algorithm

In order to verify the validity and generality of GWO, 23 benchmark tests are selected to test GWO algorithm. Among them, F1 and F2 are unimodal test functions, and F9, F11 and F13 are multi peak test functions, as shown in Table 2. Salp swarm algorithm (SSA), differential evolution (DE), particle swarm optimization (PSO), ant lion optimization (ALO), dragonfly algorithm (DA) and GWO are selected for comparative study. In order to make the algorithm fairer, the parameters of the five algorithms are set as follows. The population size is set to 30 and the cutoff iterations are set to 500. In SSA, c1 is between 0 and 2, c2 and c3 are random numbers between 0 and 1. In DE algorithm, the scale factor is set to 0.5 and the crossover constant is set to 0.2. In PSO algorithm, the maximum value of inertia weight is set to 0.9, and the minimum value is set to 0.4. The learning factor of PSO algorithm is set to ca = 2.5, cb = 0.5, and the maximum limit speed is set to 1. In ALO algorithm and DA the same dimension as GWO. The convergence accuracy and convergence rate of the algorithms are evaluated.

Table. 2 Benchmark test functions

Formula	Dimension	Range	Optimal value
$F_1(x) = \sum_{i=1}^D x_i^2$	30	[-100,100]	0
$F_2(x) = \sum_{i=1}^D x_i + \prod_{i=1}^D x_i $	30	[-10,10]	0

$F_3(x) = \sum_{i=1}^D \left(\sum_{j=1}^i x_j \right)^2$	30	[-100,100]	0
$F_4(x) = \max_i \{ x_i , 1 \leq i \leq D\}$	30	[-100,100]	0
$F_5(x) = \sum_{i=1}^{D-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	30	[-30,30]	0
$F_6(x) = \sum x + 0.5 ^2$	30	[-100,100]	0
$F_7(x) = \sum_{i=1}^D ix_i^4 + \text{random}(0,1)$	30	[-1.28,1.28]	0
$F_8(x) = \sum_{i=1}^D -x_i \sin \sqrt{ x_i }$	30	[-500,500]	0
$F_9(x) = \sum_{i=1}^D [x_i^2 - 10 \cos(2\pi x_i) + 10]$	30	[-5.12, 5.12]	0
$F_{10}(x) = -20e^{(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^D x_i^2})} - e^{(\frac{1}{n}\sum_{i=1}^D \cos(2\pi x_i))} + 20 + e$	30	[-32,32]	0
$F_{11}(x) = \frac{1}{4000} \sum_{i=1}^D x_i^2 - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	[-600,600]	0
$F_{12}(x) = \frac{\pi}{n} 10 \sin(\pi y_1) + \sum_{i=1}^{D-1} (y_i - 1)^2 [1 + 10 \sin^2(\pi y_{i+1})]$ $+ (y_n - 1)^2 + \sum_{i=1}^D u(x_i, 10, 100, 4)$ $y_i = 1 + \frac{x_i + 1}{4}$	30	[-50,50]	0
$u(x_i, 10, 100, 4) = \begin{cases} k(x_i - a)^m & x_i > a \\ 0 & -a < x_i \\ k(-x_i - a)^m & x_i < -a \end{cases}$			
$F_{13}(x) = 0.1 \{ \sin^2(3\pi x_1) + \sum_{i=1}^D (x_i - 1)^2 [1 + \sin^2(3\pi x_i) + 1] + (x_n - 1)^2 [1 + \sin^2(2\pi x_n)] \} + \sum_{i=1}^D u(x_i, 5, 100, 4)$	30	[-50,50]	0
$F_{14}(x) = \left(\frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^2 (x_i - a_{ij})^6} \right)^{-1}$	2	[-65,65]	0
$F_{15}(x) = \sum_{i=1}^{11} \left[a_i - \frac{x_1(b_i^2 + b_i x_2)}{b_i^2 + b_i x_3 + x_4} \right]^2$	4	[-5,5]	0
$F_{16}(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^4 + x_1x_2 - 4x_2^2 + 4x_2^4$	2	[-5,5]	0

$F_{17}(x) = (x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6)^2 + 10(1 - \frac{1}{8\pi})\cos x_1 + 10$	2	[-5,5]	0
$F_{18}(x) = [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)] \times [30 + (2x_1 - 3x_2)^2] \times (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)$	2	[-2,2]	0
$F_{19}(x) = -\sum_{i=1}^4 c_i e^{(-\sum_{j=1}^3 a_{ij}(x_j - p_{ij})^2)}$	3	[1,3]	0
$F_{20}(x) = -\sum_{i=1}^4 c_i e^{(-\sum_{j=1}^6 a_{ij}(x_j - p_{ij})^2)}$	6	[0,1]	0
$F_{21}(x) = -\sum_{i=1}^5 [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	[0,10]	0
$F_{22}(x) = -\sum_{i=1}^7 [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	[0,10]	0
$F_{23}(x) = -\sum_{i=1}^{10} [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	[0,10]	0

Table. 3 Results of benchmark test function

Function		GWO	SSA	PSO	DE	ALO	DA
F_1	Ave	1.69e-15	4.51e-05	0.002195	0.010272	0.007754	8.208377
	Std	1.54e-16	3.43e-05	0.001589	0.022289	0.005202	5.867565
F_2	Ave	1.30e-18	0.041021	1.66e-04	0.019779	0.344763	0.485178
	Std	2.38e-19	0.114595	3.09e-04	0.040523	1.069088	0.626178
F_3	Ave	5.37e-05	9.344069	2.865525	0.012248	11.204666	17.469637
	Std	2.36e-05	6.754941	1.832225	0.027985	7.815729	16.309052
F_4	Ave	8.94e-07	5.601099	0.490286	0.306086	13.306136	16.960946
	Std	1.96e-07	2.719842	0.322861	0.169225	6.039672	8.004191
F_5	Ave	0.076699	0.013783	0.289451	0.525274	0.115306	2.742857
	Std	0.208071	0.019954	0.361101	0.357824	0.228265	1.101517
F_6	Ave	0.434285	0.499991	0.500024	0	0.500075	5.316411
	Std	0.168623	5.65e-05	8.15e-04	0	0.003769	3.671765
F_7	Ave	0.016815	0.133704	0.083586	0.230743	0.117537	0.127742
	Std	0.020713	0.071891	0.046507	0.176026	0.076103	0.1120653
F_8	Ave	2.1242e+02	3.1122e+02	1.9815e+02	0.971742	4.9405e+02	2.2939e+02
	Std	1.4854e+02	1.7010e+02	1.3821e+02	0.060516	32.556011	1.6225e+02

F_9	Ave	3.72e-09	1.392938	1.162095	0.399364	1.757771	1.551829
	Std	2.51e-09	0.809567	0.786181	0.515579	1.582437	0.613495
F_{10}	Ave	2.08e-14	0.479007	0.121888	0.559262	2.786641	1.101415
	Std	6.21e-15	0.603006	0.309137	0.258891	2.566353	0.941039
F_{11}	Ave	2.14e-08	0.074052	0.001161	0.024587	0.146765	51.978504
	Std	1.29e-08	0.095434	0.001108	0.052775	0.163191	27.669841
F_{12}	Ave	0.799699	4.689853	1.000013	0.859562	5.624831	5.823162
	Std	0.395535	3.125309	0.005251	0.457892	3.101541	3.302272
F_{13}	Ave	0.754019	1.554161	1.000018	0.886953	3.777629	5.910301
	Std	0.387561	1.239284	3.94e-04	0.146431	1.138838	2.970171
F_{14}	Ave	23.932545	23.978367	31.978334	0.029963	31.978336	31.978334
	Std	11.283171	11.302351	1.003e-06	0.003063	6.21e-06	7.23e-07
F_{15}	Ave	0.233314	2.599323	7.270406	0.394241	1.063258	3.052731
	Std	0.080406	2.096767	5.048224	0.197431	0.867394	2.361489
F_{16}	Ave	0.4012807	0.401249	0.401249	0.353553	0.401249	0.401249
	Std	0.440366	0.440396	0.440396	0.500000	0.440396	0.440396
F_{17}	Ave	2.708872	0.537405	2.708296	1	2.708296	2.708296
	Std	0.611744	0.231406	0.612773	0	0.612773	0.612773
F_{18}	Ave	0.500248	0.500001	0.500001	0	0.500001	0.500001
	Std	0.706892	0.707106	0.707106	0	0.707106	0.707106
F_{19}	Ave	0.508537	0.507603	0.507603	0.496227	0.507603	0.508225
	Std	0.369743	0.371305	0.371305	0.364716	0.371304	0.370286
F_{20}	Ave	0.345023	0.479691	0.480745	0.364069	0.460785	0.433646
	Std	0.189996	0.322890	0.352831	0.185006	0.323137	0.287861
F_{21}	Ave	8.000627	4.000085	4.000085	0.982601	4.000085	7.999633
	Std	0.006721	5.55e-05	5.54e-05	0.021352	5.57e-05	1.25e-04
F_{22}	Ave	4.001473	4.000089	4.000089	0.994969	4.000089	4.001454
	Std	0.001394	6.28e-04	6.28e-04	0.005808	6.29e-04	0.005133
F_{23}	Ave	4.007203	5.293117	4.000128	0.975137	4.007192	7.999457
	Std	2.305720	1.960713	6.31e-04	0.049725	2.306840	1.84e-05

Under the condition of the same population size and maximum iteration times, 30 tests are conducted on 23 functions by SSA, DE, PSO, ALO, DA and GWO. The average value and standard deviation are used as statistical data to observe the experimental results in Table 3. No matter in unimodal function or multimodal function, GWO is superior to other algorithms in convergence accuracy and stability. Therefore, GWO has good global convergence performance. Optimal fitness value diagram of benchmark functions are given in Figure 5, Figure 6, Figure 7. By comparing the advantages of the six algorithms, GWO has the advantages of fast computing speed and low fitness.

4.3 Prediction of survival level of patients with EC by K-M-GWO-BP

In the construction of BP network model, the input layer is determined by the number of influence variables. The number of output layers is determined by the number of prediction. The number of input layer and output layer is 5 and 1 respectively. There is no unified way to determine the number of hidden layers, but it plays an important role in the accuracy of the prediction model. The number of hidden layers is selected by comparing the training errors under different numbers of hidden layers. Select the number of hidden layers from 3 to 13 for BP network training, and get

the results as shown in Table 4 through 10 experiments. When the number of hidden layers is 11, the training error is 0.0173, and the training result is the best.

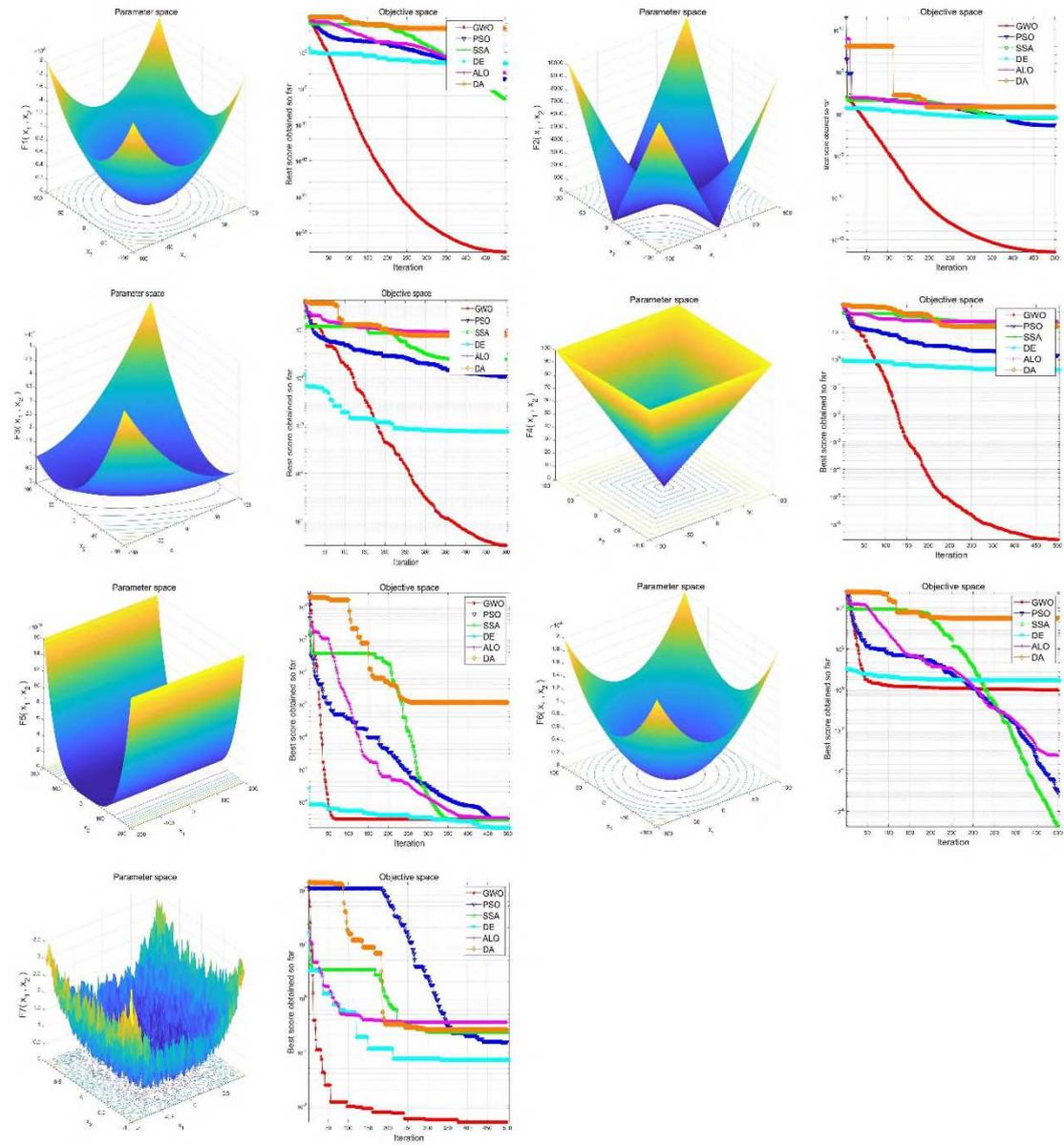


Figure. 5: Unimodal benchmark functions.

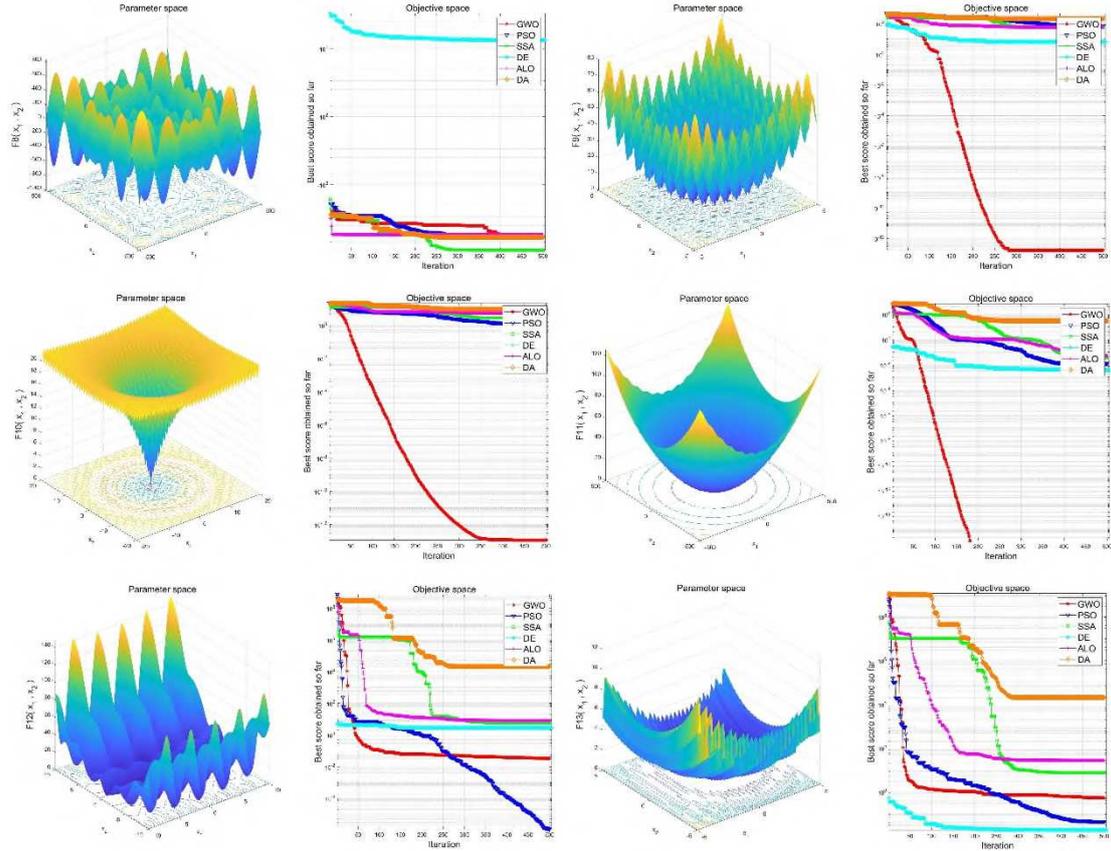


Figure. 6: Multimodal benchmark functions.

Table. 4 Network training error of node numbers in different implicit layer

Number of nodes	3	4	5	6	7	8	9	10	11	12	13
BP training error	6.6599	6.7328	6.7316	6.9328	6.6517	6.7726	6.8356	7.0237	6.9346	6.6637	6.8205
GWO-BP training error	4.8392	4.2302	5.2688	5.7011	4.6974	4.6812	4.8253	4.5574	4.8330	3.5871	5.0219

In order to comprehensively reflect the performance of the K-M-GWO-BP prediction model, the prediction results are evaluated by three indexes: the average value of the absolute error, the variance of the absolute error and the average value of the relative error. The agreement between the predicted value and the real value of test data in the prediction model is reflected by the average value of absolute error and relative error. The smaller the corresponding value is, the higher the prediction accuracy of the model is. The variance of the absolute error reflects the fluctuation of the difference, and the smaller the value is, the more stable the prediction result is. The predicted results for the normalized data are given in Table 5, with an average absolute error of 3.4156 for K-M-GWO-BP. The average relative error of 0.3277 is smaller than that of BP, PSO-BP, GA-BP, SSA-BP, GWO-BP, K-M-BP, K-M-PSO-BP, K-M-SSA-BP, K-M-GA-BP indicating that the prediction accuracy and fitting degree are higher. The average absolute error of PSO-BP is 7.3707 and the average relative error is 0.8831 higher than that of BP prediction model. The absolute error variance of BP is smaller than that of K-M-GWO-BP prediction model. The absolute error of the models without K-M analysis is given in Figure 8. The absolute error of K-M-GWO-BP model is minimum. The comprehensive results show that the K-M-GWO-BP prediction model has better training accuracy and prediction effect.

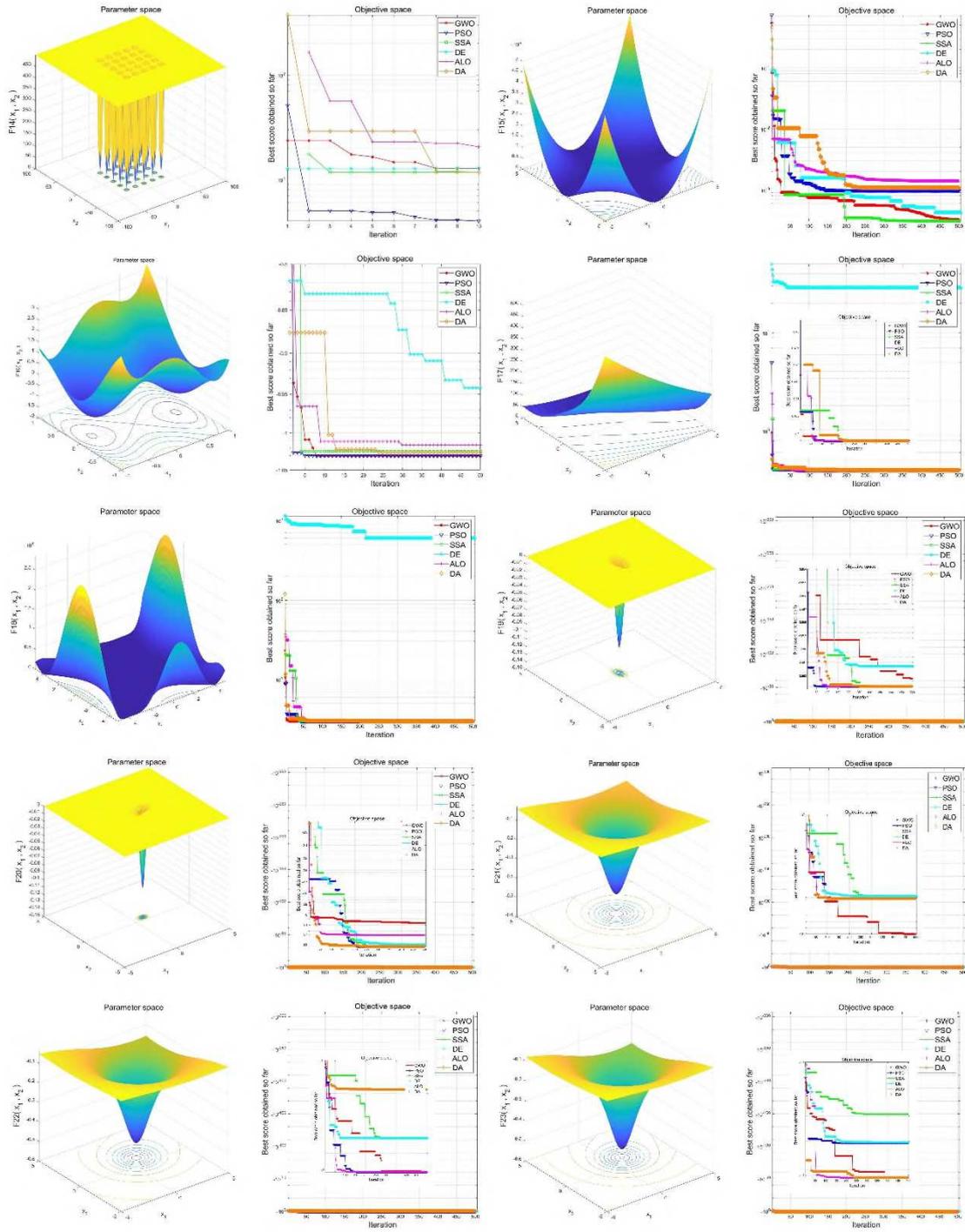


Figure. 7: Fixed-dimension multimodal benchmark functions.

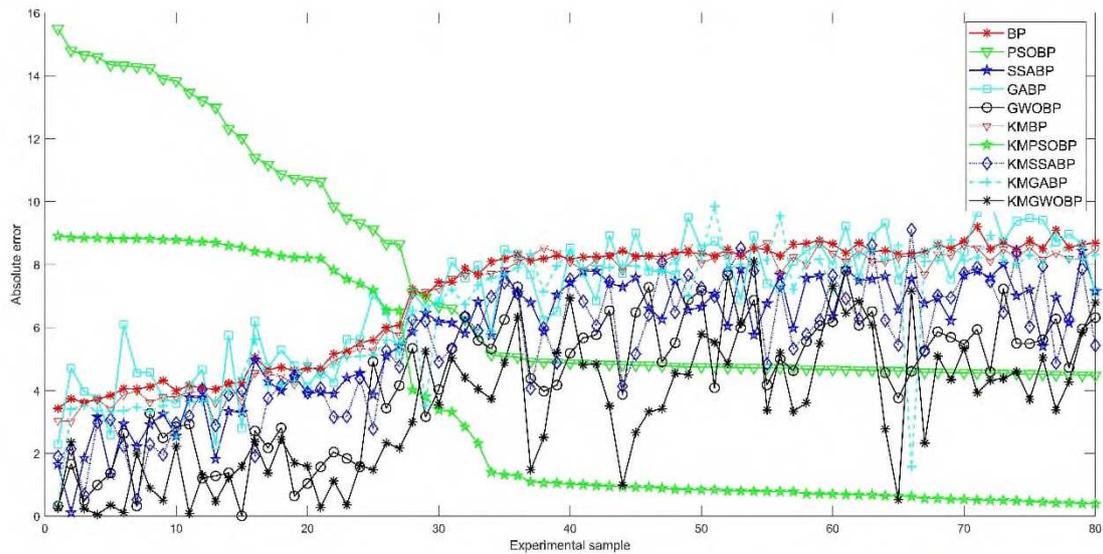


Figure 8: Absolute error of the models.

Table 5 Comparison of forecast results evaluation

Algorithm type	Average absolute error	Absolute error variance	Average relative error /%
BP	7.0483	3.6125	0.7098
PSO-BP	7.3707	13.7809	0.8831
ASSA-BP	5.7459	3.8047	0.5716
GA-BP	6.9386	3.9834	0.7007
GWO-BP	4.3486	4.2637	0.4232
K-M-BP	6.8245	3.7137	0.6852
K-M-PSO-BP	3.4979	12.5807	0.8744
K-M-ASSA-BP	5.4399	4.1366	0.5409
K-M-GA-BP	6.5191	4.6490	0.6522
K-M-GWO-BP	3.4156	4.4168	0.3277

5. Discussion

In this paper, a comprehensive model for predicting the survival level of patients with esophageal squamous cell carcinoma based on K-M survival analysis and gray wolf optimized backward propagation neural network is proposed. In view of the strong coupling and nonlinear characteristics of patient blood sample data, the sample data are analyzed by K-M survival analysis to reduce the impact of data correlation on modeling accuracy. On the basis of obtaining all kinds of sample data, the corresponding BP neural network model is distributed and constructed. The grey wolf algorithm with global optimization ability is used to optimize the parameters of error back propagation neural network, which avoids the blindness of artificial parameter selection and improves the prediction accuracy of the model.

Considering the influence factors of patients' blood, 250 groups of data are selected to train the network, and 80 groups of data are used to test the trained network. Compared with BP neural network model, PSO-BP, GA-BP, ASSA-BP, GWO-BP, K-M-BP, K-M-PSO-BP, K-M-ASSA-BP and K-M-GWO-BP, the prediction accuracy of K-M-GWO-BP is the best. The rapid and accurate prediction of patients' survival level is based on the solution method of K-M-GWO-BP neural network, which provides a new way for the healthy life of postoperative patients.

In this paper, 17 factors are found. Because there are many influence factors, the correlation is large. Reducing screened blood factors for ESCC is our next goal to improve the accuracy of predicting survival.

Declarations

Ethics approval and consent to participate:

The experimental data were approved by the First Affiliated Hospital of Zhengzhou University and used after NSFC audit. Following approval, members of the research team were authorised to access the raw data for this study. All data were anonymised prior to use. This work uses data provided by patients as part of their care and support.

Consent for publication:

Not applicable.

Availability of data and material:

The data and code for this study can be requested from the corresponding author and will be shared once approval is obtained from the data provider.

Competing interests:

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Funding:

The National Key R and D Program of China for International S and T Cooperation Projects (2017YFE0103900), in part by the Joint Funds of the National Natural Science Foundation of China (U1804262), in part by the State Key Program of National Natural Science of China under Grant 61632002, in part by the Foundation of Young Key Teachers from University of Henan Province (2018GGJS092), in part by the Youth Talent Lifting Project of Henan Province (2018HYTP016) and Henan Province University Science and Technology Innovation Talent Support Plan under Grant (20HASTIT027), and in part by the Open fund of State Key Laboratory of esophageal cancer prevention and treatment (K2020-0010, K2020-0011), and Program for Innovative Research Team (in Science and Technology) in University of Henan Province (No.20IRTSTHN017).

Author Contributions:

Conceptualization, L.D.W and X.K.Z.; methodology, E.H.L and Y.F.W.; software, E.H.L and J.W.S.; validation, E.H.L., J.W.S and X.S.; formal analysis, E.H.L., J.W.S and X.S.; data curation, L.D.W and X.K.Z.; writing—original draft preparation, E.H.L. and Y.F.W.; writing—review and editing, L.D.W and J.W.S.; funding acquisition, Y.F.W.

Acknowledgements:

This work was supported in part by the National Key R and D Program of China for International S and T Cooperation Projects (2017YFE0103900), in part by the Joint Funds of the National Natural Science Foundation of China (U1804262), in part by the State Key Program of National Natural Science of China under Grant 61632002, in part by the Foundation of Young Key Teachers from University of Henan Province (2018GGJS092), in part by the Youth Talent Lifting Project of Henan Province (2018HYTP016) and Henan Province University Science and Technology Innovation Talent Support Plan under Grant (20HASTIT027), and in part by the Open fund of State Key Laboratory of esophageal cancer prevention and treatment (K2020-0010, K2020-0011), and Program for Innovative Research Team (in Science and Technology) in University of Henan Province (No.20IRTSTHN017).

References

- [1] Bray F, Ferlay J, Soerjomataram I, et al. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries[J]. *CA Cancer J Clin*, 2018, 68(6):394-424.
- [2] Wang Y, Wang L, Yang Q, et al. Factors on prognosis in patients of stage T3N0M0 thoracic esophageal squamous cell carcinoma after two-field esophagectomy[J]. *J Cancer Res Ther*, 2015, 11 (Suppl1) :C16-23.
- [3] Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015[J]. *CA Cancer J Clin*, 2016, 66(2):115-132.
- [4] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015[J]. *CA Cancer J Clin*, 2015, 65 (1) :5-29.
- [5] Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015[J]. *CA Cancer J Clin*, 2016, 66 (2):115-32.
- [6] Yang M, Zhang H, Ma Z, et al. Log odds of positive lymph nodes is a novel prognostic indicator superior to the number based and the ratio-based N category for gastric cancer patients with

- R0resection[J]. *Cancer*, 2010, 116(11):2571-2580.
- [7] Cao J, Yuan P, Ma H, et al. Log Odds of Positive Lymph Nodes Predicts Survival in Patients After Resection for Esophageal Cancer[J]. *Ann Thorac Surg*, 2016, 102(2):424-432.
- [8] Hou HF, Meng ZX, Zhao XN, et al. Survival of Esophageal Cancer in China: A Pooled Analysis on Hospital-Based Studies From 2000 to 2018[J]. *Front Oncol*, 2019, 9:548.
- [9] Liu H, Yang J, Yuan Y, et al. Regulation of Mcl-1 by constitutive activation of NF-kappa B contributes to cell viability in human esophageal squamous cell carcinoma cells[J]. *BMC Cancer*, 2014, 14:98.
- [12] R. L. Siegel, K. D. Miller, A. Jemal, "Cancer statistics," *A Cancer Journal for Clinicians*, vol. 66, no.~1, pp. 7--30, 2017.
- [13] S. Senthil, B. Ayshwarya, "Lung cancer prediction using feed forward back propagation neural networks with optimal features," *International Journal of Applied Engineering Research*, vol. 13, no.~1, pp. 318--325, 2018.
- [14] Pan Xiaoying, Zhang Ting, Yang QingPing, et al. Survival prediction for oral tongue cancer patients via probabilistic genetic algorithm optimized neural network models[J]. *British Journal of Radiology*. 2020, 93(1112):20190825.
- [15] Wang Y, Liang E, Zhao X, et al. Prediction of Survival Time of Patients With Esophageal Squamous Cell Carcinoma Based on Univariate Analysis and ASSA-BP Neural Network[J]. *IEEE Access*, 2020, 8: 181127-181136.
- [16] Song, Wei, et al. "A GPU-Based Training of BP Neural Network for Healthcare Data Analysis." *Advanced Multimedia and Ubiquitous Engineering*[J]. Springer, Singapore, 2018, 518:193-198.
- [17] Xizhi Wu, Han Niu, Xian-Jun Li, et al. A study on the GA-BP neural network model for surface roughness of basswood-veneered medium-density fiberboard. 2020, 74(10):979-988.
- [18] Zeming Li, Yanning Li. A comparative study on the prediction of the BP artificial neural network model and the ARIMA model in the incidence of AIDS. *BMC Medical Informatics and Decision Making*. 2020, 20(1):1425-1431.
- [19] Mohamed El Amine Ben Seghier, Hermes Carvalho, Behrooz Keshtegar, et al. Novel hybridized adaptive neuro-fuzzy inference system models based particle swarm optimization and genetic algorithms for accurate prediction of stress intensity factor. *Fatigue & Fracture of Engineering Materials & Structures*. 2020, 43(11):2653-2667.
- [20] Weijun Guo, Meirong Jiang, Xueyan Li, et al. Using a genetic algorithm to improve oil spill prediction. *Marine Pollution Bulletin*. 2018, 135:386-396.
- [21] Yoko Satoh, Daiki Tamada, Yoshie Omiya, et al. Diagnostic Performance of the Support Vector Machine Model for Breast Cancer on Ring-Shaped Dedicated Breast Positron Emission Tomography Images. *Journal of Computer Assisted Tomography*. 2020, 44(44):413-418.
- [22] Jina Oh, Byungsoo Kim. Prediction Model for Demands of the Health Meteorological Information Using a Decision Tree Method. 2010, 4(3):151-162.
- [23] Yanghua Zhang, Hu Zhao. Land-Use and Land-Cover Change Detection Using Dynamic Time Warping-Based Time Series Clustering Method. *Canadian Journal of Remote Sensing*. 2020, 46(1):67-83.
- [24] Daut M.A.M., Hassan M.Y., Abdullah H. Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: a review *Renew. Renewable and Sustainable Energy Reviews*, 70 (2017), pp. 1108-1118.
- [25] Esmail H., Sobhan M., Parham S. A grey wolf optimizer-based neural network coupled with

response surface method for modeling the strength of siro-spun yarn in spinning mills. *Applied Soft Computing Journal*, 72 (1) (2018), pp. 1568-4946.

[26] Hossam Faris, Seyedali Mirjalili, Ibrahim Aljarah. Automatic selection of hidden neurons and weights in neural networks using grey wolf optimizer based on a hybrid encoding scheme. *International Journal of Machine Learning and Cybernetics*. 2019, 10(10):2901-2920.

[27] Ying Tian, Junqi Yu, Anjun Zhao. Predictive model of energy consumption for office building by using improved GWO-BP. *Energy Reports*. 2020, 6:620-627.

[28] Seyedali Mirjalili, Seyed Mohammad Mirjalili, Andrew Lewis. Grey Wolf Optimizer. *Advances in Engineering Software*. 2014, 69:46-61.

[29] Y. Sun, L. Zhang, "The clinical use of pretreatment NLR, PLR, and LMR in patients with esophageal squamous cell carcinoma: evidence from a meta-analysis," *Cancer Management and Research*, vol.10, pp.6167–6179, 2018.

[30] X. H. Zhi, K. Jiang, Y. Shen, et al, "Peripheral blood cell count ratios are predictive biomarkers of clinical response and prognosis for non-surgical esophageal squamous cell carcinoma patients treated with radiotherapy," *Journal of Clinical Laboratory Analysis*, vol.28, pp. 23468, 2020.

[31] X. H. Li, X. P. Wang, W. S. Gu, et al, "Clinical Significance of Preoperative Thrombin Time in Patients with Esophageal Squamous Cell Carcinoma following Surgical Resection," *PLoS ONE*, vol.10, no. 10, 2015.

[32] Y. Zheng, Y. Li, W. Xing, et al, "Serum Fibrinogen Is An Independent Prognostic Factor In Operable Esophageal Squamous Carcinoma: A RealWorld Study," *Cancer Management and Research*, vol. 11, pp. 8877-8883, 2019.

[33] H. Zhang, X. W. Guo, X. X. Yin, et al, "Nomogram-Integrated C-Reactive Protein/Albumin Ratio Predicts Efficacy And Prognosis In Patients With Thoracic Esophageal Squamous Cell Carcinoma Receiving Chemoradiotherapy," *Cancer Management and Research*, vol. 11, pp. 9459-9468, 2019.

[34] Xia Shujing, Tang Qi, Wang Xudong, et al. Overexpression of PSMA7 predicts poor prognosis in patients with gastric cancer. *Oncology letters*. 2019, 18(5):5341-5349.

[35] Seraj Siavash Mehdizadeh, Zadeh Mahdi Zirakchian, Werner Thomas J, et al. Pretreatment volumetric parameters of FDG-PET predict the survival after Yttrium-90 radio-embolization in metastatic liver disease. *American Journal of Nuclear Medicine and Molecular Imaging*. 2019, 9(5):248-254.

[36] Seyedali M, Amir H G, et al. Salp swarm algorithm: A bio-inspired optimizer for engineering design problems [J]. *Advances in Engineering Software*, 2017, 114: 163-191.

[37] Kennedy J, Eberhart R C. Particle swarm optimization [C]. *Proceedings of IEEE International Conference on Neural Networks*, Perth, Australia, 1995: 1942-1948.

[38] Storn R, Price K. Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces [J]. *Journal of Global Optimization*.1997, 11: 341-359.