Infrared Physics and Technology 97 (2019) 341–348

Contents lists available at ScienceDirect

Infrared Physics & Technology

journal homepage: www.elsevier.com/locate/infrared

Regular article

Detection method of TFe content of iron ore based on visible-infrared spectroscopy and IPSO-TELM neural network

Dong Xiao[a,b,*], Chongmin Liu[a,b], Ba Tuan Le[a,b,c]

[a] Information Science & Engineering School, Northeastern University, Shenyang 110004, China
[b] Intelligent Mine Research Center, Northeastern University, Shenyang 110819, China
[c] College of Control Technology, Le Quy Don Technical University, Hanoi 100000, Viet Nam

ARTICLE INFO

Keywords:
Visible-infrared spectroscopy
TFe
Iron ore
Improved Particle Swarm Optimization (IPSO)
Two hidden layer extreme learning machine (TELM)

ABSTRACT

As the main material for industrial production, the TFe content of iron ore determines the grade and quality of iron ore. The existing methods for measuring the TFe content of iron ore either have large errors or take a long time. Therefore, this paper proposes a method for detecting TFe content of iron ore based on IPSO-TELM (two hidden layer extreme learning machine optimized by the improved particle swarm optimization) algorithm using visible-infrared spectroscopy. The IPSO (improved particle swarm optimization) is used to optimize the first hidden layer parameters and the number of the hidden layer nodes in the TELM network. We first obtained iron ore samples from the Anshan mining area. Then, measured the spectral data through the spectral analysis instrument and spectral features are analyzed by PCA (principal component analysis). Finally, we applied the IPSO-TELM algorithm to establish a detection model for TFe. Experiments shows that the IPSO-TELM model has higher detection accuracy and better generalization ability than the TELM and PSO-TELM models. Compared with traditional chemical analysis methods and instrumental analysis methods, this method has great advantages in economy, speed and accuracy.

1. Introduction

The determination of TFe content of iron ore will have a direct impact on process costs and economic benefits. The existing method for determining the TFe content of iron ore has chemical analysis method and instrumental analysis method [1,2]. The chemical analysis method is also called titration analysis. The sample pretreatment process is complicated and the process is long, which is difficult to meet the requirements of modern production for rapid analysis. The instrumental analysis method has the advantages of fast detection speed, but the complex matrix of iron ore is easy to influence the measurement, how to reduce the interference of complex matrix as much as possible, and improve the accuracy of instrument detection, involving the pre-processing of the samples and the optimization of the instrumental analysis software, which are still to be further improved. Therefore, finding a method for quickly and accurately detecting the TFe content of iron ore is of great significance for the rational exploitation and utilization of iron ore.

Spectral analysis method is an analytical method for determining the structure and chemical composition of a substance by using the principles and experimental methods of spectroscopy. Due to its fast analysis speed, simple operation and high sensitivity, it has been widely used in many aspects [3–7]. Owing to the complex content of chemical elements in iron ore, oxides such as silicon, magnesium and calcium have an effect on the spectral data of iron ore. Therefore, the visible and infrared spectral data of iron ore often contain many chemical information unrelated to the detection of iron ore TFe content, which makes the iron ore spectral data dimension high, information redundancy, correlation. Direct modeling will make the input data dimension too high and the model structure complex, so it is necessary to reduce the dimensionality of the iron ore spectral data.

The main idea of Principal Component Analysis [8] (PCA) is to transform high-dimensional raw data from high-dimensional space to low-dimensional space through linear transformation, and construct a new set of variables to replace the original variables. Principal component analysis is a basic mathematical analysis method, and its practical application is very extensive [9–11]. Therefore, this paper uses principal component analysis to reduce the spectral data of iron ore.

In recent years, due to the advantages of efficient learning and general adaptation of machine learning models, the use of machine learning for material discovery and design has received more and more attention, and has made great improvements in terms of time efficiency

* Corresponding author at: Information Science & Engineering School, Northeastern University, Shenyang 110004, China.
E-mail address: xiaodong@ise.neu.edu.cn (D. Xiao).

https://doi.org/10.1016/j.infrared.2019.01.005
Received 8 November 2018; Received in revised form 3 January 2019; Accepted 4 January 2019
Available online 08 January 2019
1350-4495/

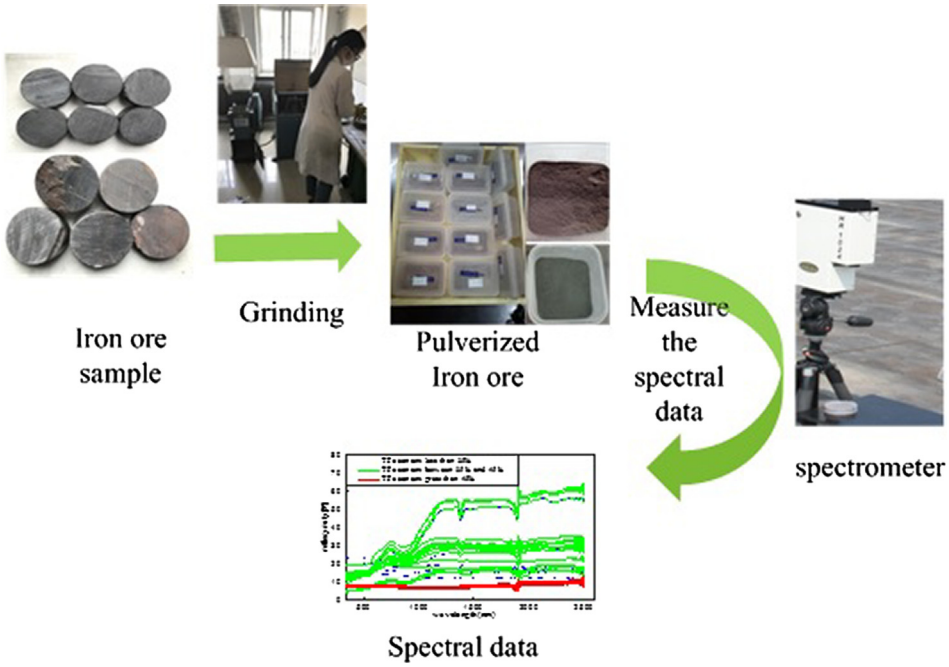

Fig. 1. The process of processing the sample.

and prediction accuracy. Literature [12] outlines typical patterns and basic procedures for applying machine learning in materials science. It also gives an introduction to the application of machine learning in material property prediction, new material discovery and other fields, and discusses machine learning related issues in materials science. Literature [13] for the research and development of lithium-ion batteries, expounding the theory and multi-scale modeling and simulation as a supplement to the experimental work, can greatly help to make up for some of the current experimental and technical gaps, and predict the path-independent characteristics. In [14], topological and regression analysis was used to predict the Tg of the As_xSe_{1-x} glass system, achieving a higher precision prediction.

The Extreme Learning Machine [15] (ELM) as an artificial neural network algorithm in machine learning has the advantages of fast computing speed and strong generalization ability, and has been widely studied and applied in recent years [16–19]. However, since the ELM randomly selects the input weight and the hidden layer bias, the accuracy of the ELM is generally low. In order to achieve the desired accuracy improvement, BY Qu proposed a two hidden layer extreme learning machine (TELM) algorithm [20], is effectively improve the accuracy of the ELM algorithm. Due to the high dimensional and nonlinear characteristics of the iron ore spectral data matrix and the advantages of the TELM network, this paper uses the TELM algorithm to establish the TFe content detection model of iron ore based on the spectral data of iron ore.

The TELM algorithm proposed in [20] randomly selects the input weight matrix and the hidden layer deviation vector of the first hidden layer, and there will be some weights and deviations of 0, which will cause TELM to require a large number of hidden layer nodes to achieve the desired results and there will be large differences in each run. The Particle Swarm Optimization (PSO) algorithm is a global optimization algorithm proposed by Dr. Eberhart and Dr. Kennedy in 1995 [21]. It has been paid attention to by many scholars, and new achievements have been made in the performance improvement and analysis of algorithms, and widely used in many fields [22–25]. In 2006, Xu and Shu proposed an extreme learning machine based on particle swarm optimization (PSO-ELM) [26]. To improve the performance of the algorithm, Fei Han improved the PSO-ELM and improved the generalization ability of PSO-ELM by considering the norm of the output weight matrix [27]. However, the above improved method does not consider the selection of the number of nodes in the hidden layer of the ELM neural network. Li and Chen [28] introduced the mutation operator to optimize the number of hidden layer nodes, enhanced the diversity of the population, and improved the convergence speed of the algorithm, but its generalization ability was poor.

The above improved methods are based on the single hidden layer extreme learning machine neural network structure, the accuracy needs to be improved. Therefore, this paper uses the improved particle swarm optimization (IPSO) algorithm to optimize the TELM network. Firstly, a linearly decreasing inertia weight is introduced for the particle swarm optimization algorithm. Secondly, in the particle update process, not only the norm of the output weight is considered, but also the idea of mutation is used to change the length of the particle. It is possible to find the optimal number of hidden layer nodes while optimizing the first hidden layer parameters of the TELM neural network. This method improves the convergence speed of the algorithm while improving the generalization ability. The spectral data of iron ore was processed by the two hidden layer extreme learning machine neural network optimized by the improved particle swarm optimization algorithm (IPSO-TELM) to obtain the model for measuring TFe content of iron ore. The experimental results show that the proposed IPSO-TELM algorithm has advantages in terms of average accuracy and generalization ability compared with the traditional TELM model and the PSO-ELM model.

The rest of this paper is organized as follows: Section 2 introduces the process of sample preparation and spectral testing. Section 3 presents a brief review of the basic concepts of ELM, TELM and PSO and two hidden layer extreme learning machine based on improved particle swarm optimization is put forward. Section 4 reports and analyzes the experimental results, and finally, Section 5 summarizes key conclusions of the present study.

2. Sample preparation and spectral testing

Liaoning Anshan iron ore district is one of the major iron ore districts in China. The iron ore is widely distributed and has large reserves. Therefore, this study selected the Anshan mining area as an experimental area, collected iron ore samples on site, and milled the collected samples for spectral testing of iron ore, the process of processing the

sample is shown in Fig. 1.

The SVC HR-1024 ground spectrometer from Spectra Vista, USA was used as an experimental instrument. The instrument has a spectral range of 350–2500 nm, built-in memory: 500 scans, weight: 3 kg, number of channels: 1024, spectral resolution: (FWHM ≦ 8.5 nm) 1000–1850 nm, minimum integration time: 1 ms.

In this experiment, a total of 123 iron ore samples were collected, including 91 haematite and 32 magnetite. Firstly, the exact TFe content of iron ore is obtained by chemical test. The TFe content of haematite is between 11.4% and 37.52%, the magnetite has a TFe content of 27.12–66.83%, and the sample size is evenly distributed.

The experimental sample was subjected to a spectral test with an instrument scan time of 1 s/time, and the probe of the spectrometer was 300 mm from the surface of the iron ore powder sample piece and perpendicular to the surface of the sample piece. In order to reduce the influence of aerosol and solar radiation propagation paths during the experiment, the spectral test time is controlled during the day from 10:00 to 14:00, and the sky is clear and cloudless. Each sample was subjected to five spectral tests using a spectrometer (SVC HR-1024), and the spectral data results were averaged as spectral data of the sample. After the spectrum test is completed, the spectral data of the test is subjected to data pre-processing such as gross error elimination and band fitting, and the spectrum of the iron ore experimental sample is obtained. Figs. 2 and 3 are spectral plots of haematite and magnetite, respectively.

Iron ore spectral data is a set of high-dimensional, nonlinear data matrices, the artificial neural network has strong adaptability when dealing with nonlinear and highly coupled data samples. As an algorithm of artificial neural network, extreme learning machine has the characteristics of fast computing speed and strong generalization ability. TELM has improved on the basis of ELM and improved the accuracy of the algorithm. Therefore, this paper uses the principle of TELM neural network combined with the principle of spectroscopy, collects data, builds models, and conducts prediction experiments.

3. Model establishment

3.1. Extreme learning machine

The Extreme Learning Machine algorithm was proposed by Professor Huang Guangbin of Nanyang Technological University. This algorithm is a supervised learning algorithm for SLFNs (a feedforward neural network with a single hidden layer). The main idea is: the weight parameter between the input layer and the hidden layer, and the offset

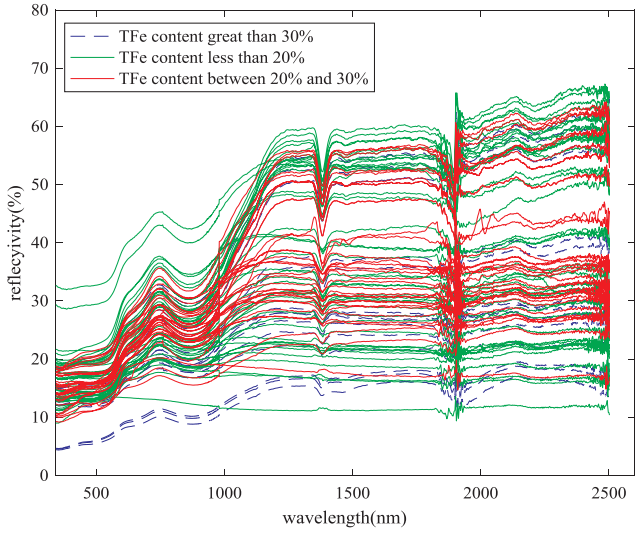

Fig. 2. Spectral curve of haematite samples.

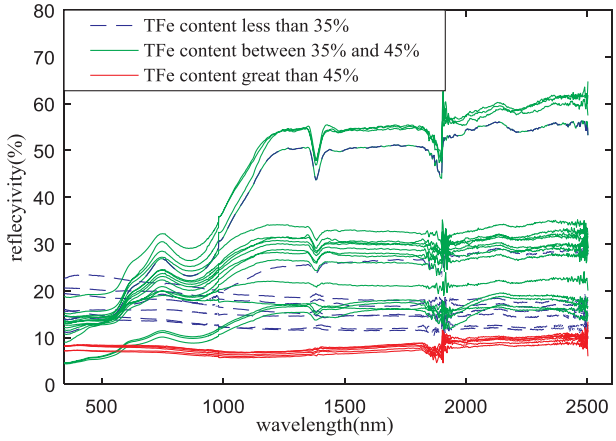

Fig. 3. Spectral curve of magnetite samples.

vector parameter on the hidden layer are determined once, and only one minimum norm least squares problem needs to be solved. Therefore, the algorithm has the advantages of less training parameters and very fast speed. Later, Professor Huang did more research on this basis, for example, extending ELM to the complex domain and proposing an online timing algorithm based on ELM.

For a single hidden layer neural network, assume that there are N arbitrary data (x_i, t_i), $x_i = [x_{i1} \quad x_{i2}\cdots x_{in}]^T \in R^n$, $t_i = [t_{i1} \quad t_{i2}\cdots t_{im}]^T \in R^m$. The output of a single hidden layer neural network with L hidden layer nodes can be expressed as:

$$\sum_{i=1}^{L} \beta_i G(\omega_i, b_i, x_j) = \delta_j \qquad j = 1, 2, \cdots, N \tag{1}$$

where $\omega_i = [\omega_{i1} \quad \omega_{i2}\cdots\omega_{in}]^T$ is the weight vector connecting the i-th hidden neuron and input neurons. β_i is the weight vector connecting the i-th hidden neuron and output neurons, and b_i is the bias of the i-th hidden neuron. $G(x)$ represents the output of hidden layer neurons, for the additive hidden nodes $G(\omega_i, b_i, x_j) = g(\omega_i \cdot x_j + b_i)$.

The goal of single hidden layer neural network learning is to minimize the error of the output, which can be expressed as:

$$\sum_{j=1}^{N} \|o_j - t_j\| = 0 \tag{2}$$

That is, there are β_i, W_i and b_i such that

$$\sum_{i=1}^{L} \beta_i G(\omega_i, b_i, x_j) = t_j \qquad j = 1, 2, \cdots, N \tag{3}$$

Can be expressed in a matrix as:

$$H\beta = T \tag{4}$$

where H is the output matrix of hidden layer, β is the output weight matrix of the hidden layer, and T is the expected output.

$$H(\omega_1, \ldots\omega_L, b_1, \ldots b_L, x_1, \ldots x_N) = \begin{bmatrix} g(\omega_1 . x_1 + b_1) & \cdots & g(\omega_L . x_L + b_L) \\ \vdots & & \vdots \\ g(\omega_1 . x_N + b_1) & \cdots & g(\omega_L . x_N + b_L) \end{bmatrix}_{N\times L} \tag{5}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L\times m} \qquad \mathrm{T} = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N\times m} \tag{6}$$

The traditional ELM algorithm randomly selects input weights and hidden biases. Training this network is equivalent to solving the least-squares solution of linear system: $H\beta = T$.

$$\underset{\omega,b,\beta}{Min:} \quad \|H\beta - T\| \tag{7}$$

Prof. Huang has proved that the minimum value of the least-squares solution of the linear system is as follows:

$$\hat{\beta} = H^{+}T \tag{8}$$

where H^{+} is the Moore-Penrose generalized inverse of H, and the minimum of the least squares solution of $H\beta = T$ is unique.

3.2. Tow hidden layer extreme learning machine

The two hidden layer extreme learning machine (TELM) is a two hidden layer network structure proposed by Qu to improve the accuracy of ELM. The algorithm adds a hidden layer to the single hidden layer ELM network. The network structure is formed by connecting the input layer, two intermediate layers (hidden layers), and the output layer. Each network layer has an indefinite number of neuron nodes, and the neurons between the layers are all connected together. And using a novel method to calculate parameters related to the second hidden layer (the connection weight between the first and second hidden layers and the bias of the second hidden layer), the method inherits the characteristics of the ELM algorithm randomly selecting the first hidden layer parameters, and effectively improves the accuracy of the ELM algorithm.

Assume that the TELM neural network has N training sample data (x_i, t_i). Where $x_i = [x_{i1} \quad x_{i2} \cdots x_{in}]^T \in R^n$, $t_i = [t_{i1} \quad t_{i2} \cdots t_{im}]^T \in R^m$. And assume that all hidden layers in the TELM model contain the same number of hidden layer neuron nodes and activation functions.

First, consider the two hidden layers of TELM as ELM neural networks with a single hidden layer, so that the output of the hidden layer can be obtained:

$$H = g(WX + B) \tag{9}$$

From Section 3.1 and Eq. (8), the weight matrix between the hidden layer and the output layer can be obtained:

$$\beta = H^{+}T \tag{10}$$

Now, the second hidden layer is added to the TELM neural network, and the neural network structure with two hidden layers is restored, and the two hidden layers are fully connected. The predicted output of the second hidden layer can be obtained:

$$H_1 = g(W_1 H + B_1) \tag{11}$$

where W_1 is the weight matrix between the first hidden layer and the second hidden layer, and B_1 is the threshold of the second hidden layer.

The expected output of the second hidden layer is:

$$H_{1*} = T\beta^{+} \tag{12}$$

where β^{+} is the generalized inverse matrix of β.

In order to satisfy the predicted output of the final hidden layer infinitely close to the expected output, let $H_1 = H_{1*}$.

Assuming the matrix $W_{HE} = [B_1 \quad W_1]$, the weight W_1 and threshold of the second hidden layer B_1 can be solved:

$$B_{HE} = g^{-1}(H_{1*})H_E^{+} \tag{13}$$

where H_E^{+} is the generalized inverse matrix of the matrix $H_E = [1 \quad H]^T$, 1 represents a vector with Q elements, and each element is 1, and $g^{-1}(x)$ is the inverse of the activation function $g(x)$.

When the weight parameter W_1 and the threshold parameter B_1 of the second hidden layer are all solved, the predicted output H_2 of the second hidden layer may be updated:

$$H_2 = g(W_1 H + B_1) = g(W_{HE} H_E) \tag{14}$$

Therefore, the output matrix β of the hidden layer can be updated to:

$$\beta_{new} = H_2^{+}T \tag{15}$$

Finally, the final neural network output $f(x)$ can be obtained:

$$f(x) = H_2\beta_{new} \tag{16}$$

3.3. Particle Swarm Optimization Algorithm

Particle Swarm Optimization (PSO) is a global optimization algorithm proposed by Kennedy and Eberhart. It is a kind of group intelligence algorithm designed by simulating the predation behavior of birds. Suppose there is only one piece of food in the area (that is, the optimal solution in the optimization problem), and the task of the flock is to find this food source. In the whole process of searching, the birds pass each other's information to let other birds know their position. Through such cooperation, they can judge whether they find the optimal solution and also pass the information of the optimal solution to entire flock of birds, eventually, the entire flock of birds can gather around the food source, finding the optimal solution.

The particle swarm optimization algorithm is implemented as follows. In a group, each bird is abstracted as a particle and is extended to the N-dimensional space. The location of particle i in the N-dimensional space is $X_i = (X_{i1}, X_{i2}, \cdots X_{iN})$. The speed of the particle's flight is $V_i = (V_{i1}, V_{i2}, \cdots, V_{iN})$, $i = 1, 2, \cdots, m$. Each particle has a fitness value determined by the objective function.

In each iteration, the particle passes through the best position $pbest_i = (pbest_{i1}, pbest_{i2}, \cdots, pbest_{iN})$ that the particle itself goes through and the best position $gbest_i = (gbest_{i1}, gbest_{i2}, \cdots, gbest_{iN})$ that the entire population passes through. The process continuously update the speed and position according to Eqs. (17) and (18).

$$V_i^{k+1} = \omega V_i^k + c_1 r_1 (pbest_i - X_i^k) + c_2 r_2 (gbest_i - X_i^k) \tag{17}$$

$$X_i^{k+1} = X_i^k + V_i^k \tag{18}$$

where k is the current number of iterations, c_1 and c_2 are learning factors, and ω is the inertia weight.

3.4. Two hidden layer extreme learning machine based on improved particle swarm optimization

The TELM algorithm inherits the characteristics of the single hidden layer extreme learning machine, and randomly gives the first layer hidden layer parameters (input weights and thresholds). This may result in some parameters being zero, and the parameters of the second hidden layer of the TELM network and the output weight matrix of the network are calculated based on the parameters of the first hidden layer. Therefore, randomly selecting the first hidden layer parameter of TELM will increase the number of hidden layer nodes, resulting in slower response speed and generalization ability of TELM for the test set. The choice of hidden layer nodes will directly affect the structure of the network and affect the performance of the TELM network. However, in the TELM algorithm, the number of hidden layer nodes of the network is selected empirically. Therefore, it is necessary to find a way to optimize the first hidden layer parameters of TELM and the number of nodes in the network hidden layer.

In this paper, the improved particle swarm optimization algorithm is used to optimize the input weight matrix and the deviation vector of the first hidden layer and the number of hidden layer nodes of the TELM network. The specific details of the algorithm are as follows:

First, initialize the particle population. The length of each particle in the population is $D = L\cdot(n + 1)$, n is the number of input layer neurons, and L is the number of hidden layer nodes. Each particle in the swarm is composed of a set of input weights and hidden biases: $P = \begin{bmatrix} \omega_{11} & \omega_{12} & \cdots & \omega_{1L} \\ \vdots & \vdots & \vdots & \vdots \\ \omega_{n1} & \omega_{n2} & \cdots & \omega_{nL} \\ b_1 & b_2 & \cdots & b_L \end{bmatrix}$. All components in the particle are randomly

initialized within the range of [−1, 1].

Second, in order to avoid overfitting the TELM, using the root mean square error RMSE of the verification set as the fitness value of each particle.

Third, in the traditional PSO algorithm's structure, the inertia weight is a fixed value. It has been proved that although PSO with the constant inertia weight ω has faster convergence speed, it tends to fall into local optimum in later periods [29]. This paper uses the linearly decreasing inertia weight to improve the PSO optimization performance.

$$\omega(k) = \omega_{start} - (\omega_{start} - \omega_{end}) \cdot k / T_{\max} \tag{19}$$

where $\omega_{\text{start}} = 0.9$, $\omega_{\text{end}} = 0.4$, k is the current number of iterations and $T_{\max}$ is the maximum number of iterations. At the beginning of the iteration, the inertia weight is larger to ensure the global search ability of the algorithm, and the inertia weight of the iteration is smaller. As a result the algorithm can effectively perform more accurate local optimization.

Fourth, The smaller the output matrix norm of a neural network, the better its generalization ability [30,31], Therefore, in order to improve the generalization capability of the TELM, in the iterative process, the individual extreme position and the outlier of the group extreme position are updated. This paper considers not only the fitness value of the particle but also the output matrix norm of the TELM network. Use the formulae (20) and (21) to update individual best position and global best position of all particles.

$$P_{ib} = \begin{cases} P_i & f(P_{ib}) > f(P_i) \text{ and } \|\omega o_{P_i}\| < \|\omega o_{P_{ib}}\| \\ P_{ib} & else \end{cases} \tag{20}$$

$$P_g = \begin{cases} P_i & f(P_g) > f(P_i) \text{ and } \|\omega o_{P_i}\| < \|\omega o_{P_g}\| \\ P_g & else \end{cases} \tag{21}$$

where $f(P_i)$, $f(P_{ib})$ and $f(P_g)$ are the corresponding fitness values for the i-th particle, the best position of the i-th particle and the global best position of all particles, respectively. ωo_{P_i}, $\omega o_{P_{ib}}$ and ωo_{P_g} are the corresponding output weights that are set the input weights and deviation vectors of the first hidden layer of TEML as the i-th particle, the best position of the i-th particle and the global best position of all particle, respectively.

Fifth, in order to find the optimal number of hidden layer nodes, this paper introduces the idea of mutation in the update process of particles, changes the length of the particles in each update process, and then changes the number of hidden layer nodes of the TELM. Change the number of hidden layer nodes according to Eq. (22).

$$L^{t+1} = L^t[1 + (rand(\,) - 1/2)] \tag{22}$$

where L^t and L^{t+1} are the number of hidden layer nodes for the t-th and $(t + 1)$-th iterations, respectively.

Sixth, update the speed and position of the particles. By changing the number n of hidden layer nodes in the ELM network, the next generation particle length of the particle group is $L^{t+1} \cdot (n + 1)$.

After the particle is mutated, if the length of the particle increases, the particle velocity and position adjustment rules are as follows:

$$v_{k,j}^{t+1}(i) = \begin{cases} v_{k,j}^t(i), & 1 \leqslant k \leqslant n+1 \quad 1 \leqslant j \leqslant L^t \\ random(-1, 1), & L^{t+1} \leqslant j \leqslant L^{t+1} \end{cases} \tag{23}$$

$$x_{k,j}^{t+1}(i) = \begin{cases} x_{k,j}^t(i), & 1 \leqslant k \leqslant n+1 \quad 1 \leqslant j \leqslant L^t \\ 0, & L^{t+1} \leqslant j \leqslant L^{t+1} \end{cases} \tag{24}$$

After the particle is mutated, if the particle length is shortened, the L^{t+1}-column elements of the previous generation particle velocity $V^t(i)$ and position $X^t(i)$ are randomly selected.

After the particle is mutated, if the particle length does not change, the particle's velocity and position are not adjusted.

Finally, the iteration is repeated until the maximum number of iterations is reached or the searched optimal position satisfies a predefined minimum fitness value.

The TELM neural network optimized by the above improved particle swarm optimization algorithm has the best input weights, hidden biases and hidden layer node number.

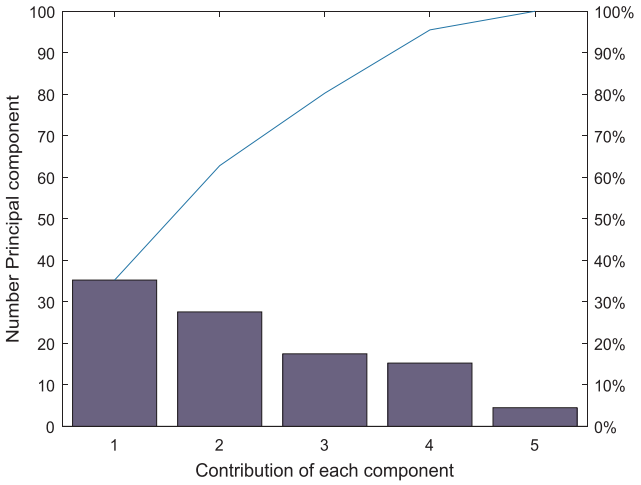

Fig. 4. Principal component contribution rate of the spectral data.

4. Experiment results and discussion

(1) Dimension reduction of iron ore spectral data.

The spectral data of iron ore sample obtained by the spectral test is 973-dimensional, the dimension is too high, and the data has strong correlation. Therefore, the obtained spectral data is reduced by principal component analysis (PCA). Take the dimension of its principal component as 5 so that the cumulative contribution rate reaches 99.9%, (Fig. 4 shows the contribution rate of each principal component) and use it as the input of the network. Therefore, the initial data is converted from the matrix structure of 973 × 123 to the matrix structure of 5 × 123, which greatly facilitates the establishment of experimental models and saves training time.

(2) Classification of iron ore samples.

Before establishing the iron content detection model, in order to facilitate the establishment of iron content detection models for each type of iron ore, it is first necessary to classify iron ore. In this paper, ELM, BP, SVM and RF are used to establish the classification model of iron ore. For each model, the 10-fold cross-validation method is used to test the performance of the model.123 iron ore samples were randomly divided into 10 parts, 9 of which were taken as a training set and 1 as a test set. The average of the accuracy of the 10 results is used as an estimate of the accuracy of the classification model. And the model is compared from the classification accuracy of the model and the time consumption of the model. The results are shown in Table 1.

It can be seen from Table 1 that the classification accuracy of ELM is 96.86%, which is the highest classification accuracy among all models. The classification accuracy rate of BP is 96.67%, which is similar to the

Table 1
Test accuracy of different classification models.

Model type	Time consuming (s)	accuracy
SVM	2.2599 s	92.31%
BP	0.1376 s	96.67%
ELM	0.0353 s	96.86%
RF	0.0536 s	96.73%

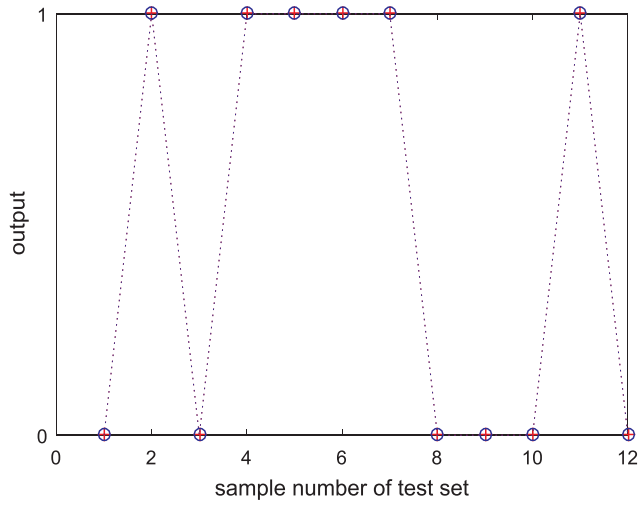

Fig. 5. Classification results of iron ore test set.

accuracy of ELM, but the training time of this model is much longer than that used by ELM. Therefore, ELM is selected as the classification model of iron ore. Fig. 5 shows the classification results for a set of test sets in the ten-fold cross-validation of the ELM classification model.

(3) TFe content detection model

Based on the above classification results, the TFe content detection models of magnetite and hematite were established respectively. Firstly, ELM, BP, TELM and RF were used to establish the TFe content detection models of different ores. Due the number of samples is small, in order to make full use of the existing data sets to test the effect of the algorithm, this paper uses a cross-validation method. 10 times cross validation and 4 times cross validation were used for haematite and magnetite, respectively. Tables 2 and 3 give the test results of different models, and compare the models from three aspects: time consumption, root mean square error (RMSE) and correlation coefficient (R).

It can be seen from Tables 2 and 3 that the root-square error of TELM is the smallest among all models, and the correlation is also the best. Therefore, TELM is used as the detection model of TFe.

In order to verify the feasibility of the IPSO-TELM proposed in this paper, this paper uses TELM, PSO-TELM and IPSO-TELM to establish the prediction model of TFe content of haematite and magnetite respectively. The number of hidden layer nodes of TELM and PSO-TELM is the optimal case for multiple trials, and the activation function of the network uses $f(x) = 1/(1 + e^{-x})$. The population size and maximum number of iterations of the particle swarm algorithm are set to 40 and 50, respectively, based on experience. Huang [32] has proved that the acceleration factor has a value range of (0.5, 3.0) and (0.5, 3.5) can more effectively avoid premature convergence and increase stability to the network. Therefore, the acceleration coefficient of particle swarm optimization algorithm $c1 = 2$, $c2 = 2$.

The performance of the established TFe content of iron ore detection model was evaluated by comparing the mean square error of the test set, the number of hidden layer nodes, the prediction time, and the output matrix norm.

Table 2
Test results of TFe content detection model of magnetite.

Model type	Time consuming (s)	RMSE	R
TELM	0.0042 s	2.2235	0.9580
BP	0.8510 s	2.8591	0.9478
ELM	0.0033 s	2.6385	0.9566
RF	0.0824 s	5.7139	0.9224

Table 3
Test results of TFe content detection model of haematite.

Model type	Time consuming (s)	RMSE	R
TELM	0.0191 s	1.9857	0.9178
BP	0.3291 s	2.2016	0.8616
ELM	0.0037 s	2.2798	0.8833
RF	0.0890 s	2.5772	0.7644

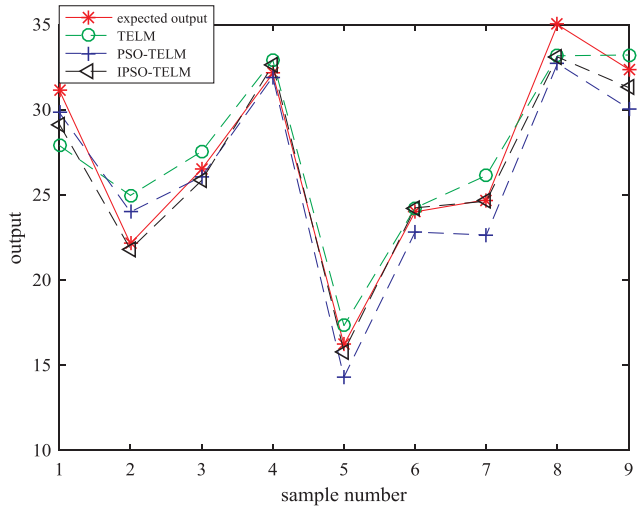

Fig. 6. Output of haematite test set.

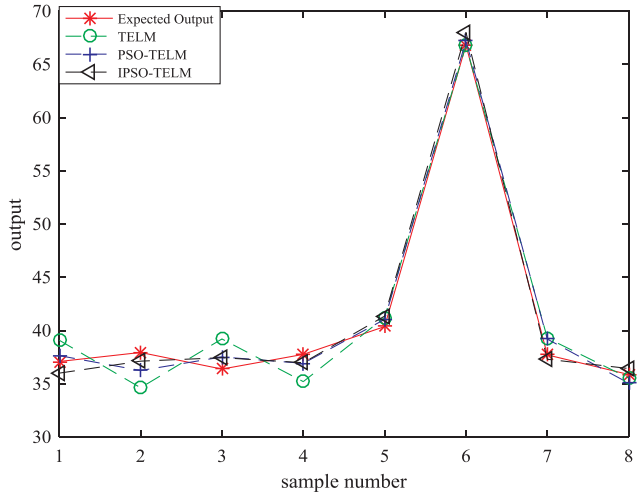

Fig. 7. Output of magnetite test set.

The simulations of this experiment were carried out in the environment of MATLAB R2016a. Figs. 6 and 7 show the results of one cross-validation of the TFe content models of haematite and magnetite TELM, PSO-TELM and IPSO-TELM. Tables 4 and 5 list the performance of TELM, PSO-TELM and IPSO-TELM from four aspects: root mean square error (RMSE) of the test set, time consumption of the model, number of hidden layer nodes of the network, correlation coefficient (R) and norm of output weight. The specific analysis results are as follows:

First, for the TFe content of haematite detection model, the number of hidden layer nodes required for PSO-TELM and IPSO-TELM is 18, while the number of hidden layer nodes required for TELM is 28. For the TFe content of magnetite detection model, the number of hidden layer nodes required for PSO-TELM and IPSO-TELM is 8, and the

Table 4
Comparison of performance TFe content of haematite detection model.

Model type	RMSE	Time consuming (s)	Number of hidden layer nodes	R	Output matrix norm
TELM	1.9857	0.019066 s	28	0.9178	5.6752e+14
PSO-TELM	1.4580	5.920568	18	0.9520	1.7769e+13
IPSO-TELM	1.3233	4.720048 s	18	0.9815	3.5227e+9

Table 5
Comparison of performance TFe content of magnetite detection model.

Model type	RMSE	Time consuming (s)	Number of hidden layer nodes	R	Output matrix norm
TELM	2.2235	0.004080	12	0.9580	1.8555e+13
PSO-TELM	1.3486	4.769674	8	0.9893	6.2317e+11
IPSO-TELM	1.1073	3.765815 s	8	0.9940	3.6747e+10

Table 6
Comparison of TFe content of iron ore detection methods.

Method	Accuracy	Time consuming (h)	Cost (yuan)
Instrument test	70%	4	About 5000
Chemical test	100%	72	About 20,000
IPSO-ELM	95%	3	About 200

number of hidden layer nodes required for TELM is 18. Compared with TELM, PSO-TELM and IPSO-TELM can get smaller prediction errors with fewer hidden layer nodes, which shows that PSO-TELM and IPSO-TELM can achieve better prediction results with a simpler network structure.

Second, the training time of PSO-TELM and IPSO-TELM is significantly larger than the training time of TELM, and the time spent is mainly used to find the optimal input weight and deviation vector. Since IPSO-TELM optimizes the hidden layer nodes, the running speed of the algorithm is improved compared to PSO-TELM.

Third, from the norm of the output weight matrix, the output matrix norm of IPSO-TELM is significantly smaller than the output matrix norm of PSO-TELM and TELM, indicating that IPSO-TELM has better generalization performance.

Finally, from the root mean square error of the test set, compared with TELM and PSO-TELM, the root mean square error of the IPSO-TELM test set is smaller, which can more accurately detect the TFe content of iron ore. And from the correlation coefficient, compared with TELM and PSO-TELM, the correlation coefficient of the IPSO-TELM model is closer to 1, which can better fit the true value.

Table 6 gives the chemical analysis method, instrumental analysis method and the proposed method based on iron ore spectral data and IPSO-TELM network, which compares the detection accuracy, detection time and detection cost. It can be seen that the instrument analysis method has low detection precision, and the chemical test method has high precision but high cost and long time. The method based on spectral data and IPSO-TELM algorithm for detecting TFe content of iron ore in this paper is shorter in time, lower in cost and higher in prediction accuracy, which can meet the needs of industrial production.

In summary, the proposed model for detecting TFe content of iron ore based on spectral data is a modified two-implicit layer extreme learning machine with improved particle swarm optimization algorithm with optimal input weight matrix, bias vector and hidden layer nodes. The generalization ability is better than that of TELM and PSO-TELM, which runs faster than PSO-TELM, and the prediction accuracy is higher than that of TELM and PSO-TELM. Compared with the traditional instrumental analysis method, the method has simple operation and high precision, and the detection speed is fast and the cost is low compared with the chemical analysis method.

5. Conclusion

In this paper, a detection model of TFe content in iron ore based on spectral data and IPSO-TELM network is proposed, and the iron ore samples collected on site for model verification. First, 123 iron ore samples were collected, and perform spectral tests on them to obtain corresponding spectral data and principal component analysis was used for dimensionality reduction. Then the classification of iron ore by ELM network shows that the classification accuracy of the method can reach 100%, which can replace the traditional manual classification method. The TFe content detection model was established by using TELM and PSO-TELM respectively for the classified haematite and magnetite. In order to further improve the detection accuracy and the generalization ability of the model, the IPSO-TELM model proposed in this paper. The experimental results show that the IPSO-TELM model can accurately detect the TFe content of iron ore. Compared with the performance of the TELM and PSO-TELM models, it is concluded that the IPSO-TELM network can better detect the TFe content of iron ore and has better generalization ability. Compared with the traditional TFe content detection method for iron ore, the model based on spectral data and IPSO-TELM network has the advantages of high precision, fast detection speed and low cost, which provides a new method for the detection of TFe content in iron ore.

Conflict of interest

The authors declared that there is no conflict of interest.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grants 61773105, 61374147, 41371437, 61203214, 61473072, and 61733003, in part by the Fundamental Research Funds for the Central Universities, China, under Grants N150402001 and N160404008, in part by the National Key Research and Development Plan, China, under Grant 2016YFC0801602, and in part by the National Twelfth Five-Year Plan for Science and Technology Support, China, under Grant 2015BAB15B01.

References

[1] Hong Yan, Liping Ren, Yuqiong Qin, Research progress on analysis of total iron content in iron ore, Metall. Anal. 34 (2014) 21–26.
[2] Lei Ma, Study on detection technology of total iron content in iron ore, Chem. Manage. 8 (2013) 196.
[3] Y.C. Mao, D. Xiao, J.P. Cheng, J.H. Jiang, B.T. Le, S.J. Liu, Research in magnesite grade classification based on near infrared spectroscopy and ELM algorithm, Spectrosc. Spect. Anal. 37 (2017) 89–94.
[4] A.P. Li, Z.Y. Li, J.P. Jia, Chemical comparison of coat and kernel of mung bean by nuclear magnetic resonance-based metabolic fingerprinting approach, Spectrosc. Lett. 49 (2016) 276.

[5] Yi Yang, Shujuan Zhang, Yong He, Dynamic classification detection of fresh jujube based on ELM and visible/near infrared spectroscopy, Spectrosc. Spectral Anal. 35 (2015) 1870–1874.
[6] D. Yang, D.D. He, A.X. Lu, D. Ren, J.H. Wang, Combination of spectral and textural information of hyperspectral imaging for the prediction of the moisture content and storage time of cooked beef, Infrared Phys. Technol. 83 (2017) 206–216.
[7] H. Kaya-Celiker, P.K. Mallikarjunan, A. Kaaya, Mid-infrared spectroscopy for discrimination and classification of Aspergillus, spp. contamination in peanuts, Food Control 52 (2015) 103–111.
[8] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemomet. Intell. Lab. Syst. 2 (1987) 37–52.
[9] R.S.D. Souza, U. Maio, V. Biffi, Robust PCA and MIC statistics of baryons in early minihaloes, MNRAS 440 (2018) 240–248.
[10] W. Huizinga, D.H.J. Poot, J.M. Guyader, PCA-based groupwise image registration for quantitative MRI, Med. Image Anal. 29 (2016) 65–78.
[11] Q. Liu, C. Zhang, Q. Guo, Adaptive sparse coding on PCA dictionary for image denoising, Visual Computer. 32 (2016) 1–15.
[12] Yue Liu, Tian Lu Zhao, Wangwei Ju, Siqi Shi, Materials discovery and design using machine learning, J. Materiomics 3 (2017) 159–177.
[13] Siqi Shi, Jian Gao, Yue Liu, Yan Zhao, Qu Wu, Wangwei Ju, Chuying Ouyang, Ruijuan Xiao, Multi-scale computation methods: their applications in lithium-ion battery research and development, Chin. Phys. B 25 (2016) 018212.
[14] Yue Liu, Tian Lu Zhao, Guang Yang, Wangwei Ju, Siqi Shi, The onset temperature (Tg) of As_xSe_{1-x} glasses transition prediction: a comparison of topological and regression analysis methods, Comput. Mater. Sci. 140 (2017) 315–321.
[15] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: a new learning scheme of feedforward neural networks, International Joint Conference on Neural Networks, 2 2014, pp. 985–990.
[16] D. Xiao, B.T. Le, Y.C. Mao, J.H. Jiang, L. Song, S.J. Liu, Research on coal exploration technology based on satellite remote sensing, J. Sens. (2016) 1–9.
[17] X. Sun, J. Xu, C. Jiang, Extreme learning machine for multi-label classification, Entropy 18 (2016) 1–12.
[18] W. Li, C. Chen, H. Su, Local binary patterns and extreme learning machine for hyperspectral imagery classification, IEEE Trans. Geosci. Remote Sens. 53 (2015) 3681–3693.
[19] K. Cao, G. Wang, D. Han, J. Ning, X. Zhang, Classification of uncertain data streams based on extreme learning machine, Cogn. Comput. 7 (2015) 150–160.
[20] B.Y. Qu, B.F. Lang, J.J. Liang, Two-hidden-layer extreme learning machine for regression and classification, Neurocomputing 175 (2016) 826–834.
[21] J. Kennedy, Particle swarm optimization, Proc IEEE International Conference on Neural Networks, 4 1995, pp. 1942–1948.
[22] Yong Zhang, Dun-Wei Gong, Wan-Qiu Zhang, A simplex method based improved particle swarm optimization and analysis on its global convergence, Acta Autom. Sin. 35 (2009) 289–298.
[23] Pan Feng, Zhou Qian, Li Wei-Xing, Gao Qi, Analysis of standard particle swarm optimization algorithm based on Markov chain, Acta Autom. Sin. 39 (2013) 381–389.
[24] S. Fong, R. Wong, A.V. Vasilakos, Accelerated PSO swarm search feature selection for data stream mining big data, IEEE Trans. Serv. Comput. 9 (2016).
[25] M. Masdari, F. Salehi, M. Jalali, A survey of PSO-based scheduling algorithms in cloud computing, J. Netw. Syst. Manage. 25 (2017) 1–37.
[26] Y. Xu, Y. Shu, Evolutionary extreme learning machine – based on particle swarm, International Conference on Advances in Neural Networks, Springer-Verlag, 2006, pp. 644–652.
[27] F. Han, H.F. Yao, Q.H. Ling, An improved evolutionary extreme learning machine based on particle swarm optimization, Neurocomputing. (2013) 87–93.
[28] Yuhua Li, Yuzhong Chen, Kun Guo, Parallel extreme learning machine based on improved particle swarm optimization, Pattern Recogn. Artific. Intell. (2016) 840–849.
[29] Y. Shi, R. Eberhart, Modified particle swarm optimizer, Proc. of IEEE ICEC Conference, Anchorage, 1998, pp. 69–73.
[30] Q.Y. Zhu, A.K. Qin, P.N. Suganthan, Evolutionary extreme learning machine, Pattern Recogn. 38 (2005) 1759–1763.
[31] P.L. Bartlett, The sample complexity of pattern classification with neural networks: the size of the weights is more important than the size of the network, IEEE Trans. Inf. Theory 44 (1998) 525–536.
[32] Shaorong Huang, A new particle swarm optimization with random parameters, J. Chongqing Normal Univ. 30 (2013) 123.