Optics and Laser Technology 114 (2019) 10–15

Contents lists available at ScienceDirect

Optics and Laser Technology

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

Full length article

Coal classification method based on visible-infrared spectroscopy and an improved multilayer extreme learning machine

Yachun Mao[a], Ba Tuan Le[b,c], Dong Xiao[a,b,*], Dakuo He[b], Chongmin Liu[b], Longqiang Jiang[b], Zhichao Yu[b], Fenghua Yang[b], Xinxin Liu[b]

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

HIGHLIGHTS

- We proposed a new IAM-ELM neural network algorithm.
- We proposed a rapid coal classification method.
- Our method has the advantage of economy, speed and accuracy.
- Our method has important practical application value.

ABSTRACT

Coal classification is an indispensable task in coal mining and production. The traditional method of coal classification has the disadvantages of high cost, low speed and low accuracy. Therefore, a rapid coal classification method based on visible-infrared spectroscopy is proposed in this research. First, we collected samples of different coal types and used spectrometers to measure the spectral data of these samples. Then, we proposed an improved multilayer extreme learning machine algorithm using this algorithm to build a coal classification model. The simulation results showed that the model has good classification results. Compared with the traditional coal classification methods, this method has an unparalleled advantage in economy, speed and accuracy.

1. Introduction

Coal is the primary source of energy. In 2017, world coal production was 7.73 billion tons, and China accounted for 45.6% of the world's total output, ranking first, followed by India at 9.3%, the United States at 9.1%, Australia at 6.2% and so on. With the development of society, high-quality coal plays a decisive role in production efficiency and environmental pollution, which also increases the requirement of coal classification. There are two common kinds of traditional coal classification methods. The first is to use the artificial classification method, which is fast but has a low accuracy rate. The second is to use the chemical analysis method, which is highly accurate but has the shortcomings of being high cost and time consuming. Therefore, determining the type of coal quickly and accurately is an important problem to be solved in modern beneficiation technology. It has great significance in reducing the classification cost and improving the classification efficiency.

In recent years, advances in spectral technology have led to the advantages of fast analysis, low detection cost and high efficiency. Therefore, spectral technology has been widely used in food monitoring, ore classification, variety identification and other fields [1–4]. For coal mines, spectral technology has been successfully applied in coal analysis and classification. Andres and Bona [5,6] used diffuse reflectance infrared Fourier transform spectroscopy and partial least squares regression to estimate coal moisture, ash, volatile matter, fixed carbon, heating value and percentage of carbon, hydrogen, nitrogen and sulfur, and the results were well-predicted. Tanno et al. [7] used terahertz spectroscopy to accurately calculate moisture in coal. Kim et al. [8] used near-infrared diffuse reflectance to measure the spectroscopy analysis of coal, and positive results were obtained. Wang et al. [9,10] proposed a coal component analysis model based on a support vector machine, a partial least squares regression algorithm and near-infrared reflectance spectroscopy. The model analyzed six components of coal, including total moisture, inherent moisture, ash, volatile matter, fixed carbon, and sulfur. Zeng et al. [11,12] used the soft measurement method to provide coal moisture monitoring for utility

* Corresponding author at: Intelligent Mine Research Center, Northeastern University, Shenyang 110819, China.
E-mail address: xiaodong@ise.neu.edu.cn (D. Xiao).

https://doi.org/10.1016/j.optlastec.2019.01.005
Received 15 September 2018; Received in revised form 28 November 2018; Accepted 7 January 2019

boilers, and the results of this method were consistent with proximate analysis data. Wang et al. [13] constructed a classification model of coal based on a confidence machine, a support vector machine algorithm and near-infrared spectroscopy, and a good classification result was obtained. Gomez et al. [14] used Fourier transform infrared photoacoustic spectroscopy combined with partial least squares to predict ash content, volatile matter, fixed carbon and calorific value in coal and obtained good prediction results. Ren et al. [15] proposed a coal-like classification model based on hyperspectral data in the 500–2350 nm band combined with multilayer perceptron, which can quickly and efficiently obtain coal classification information. Le et al. [16] used visible-infrared spectroscopy and a deep neural network to analyze the moisture, ash, volatile matter, fixed carbon, sulfur and the low heating value in coal. Compared with traditional coal analysis methods, this method has unparalleled advantages in economy, speed and accuracy.

The extreme learning machine (ELM) is a single hidden layer feedforward neural network proposed by professor Huang in 2006 [17]. The algorithm has the advantages of strong learning ability, good generalization, fast training speed and high precision [18–21]. It compensates for many of the shortcomings of traditional neural networks. In this paper, the algorithm is used to construct the classifier of the coal recognition model, and the validity of this method is examined by MATLAB software simulation.

2. The Collection and processing of spectral data

This study collected a total of four sample species, including anthracite, bituminous coal, lignite and coal gangue. The number of samples is shown in Table 1. The dataset was divided into three subsets. The training set had 130 samples, the validation set had 66 samples, and the test set had 119 samples. We used the SVC HR-1024 Spectrometer of Spectra Vista Company of the United States as an experimental instrument. The spectral range of the instrument was 350–2500 nm. The spectral band was 1024. The spectral resolution was in the 1000–1850 nm range, less than 8.5 nm.

First, we cleaned the surface of the sample and then chose to make the spectral measurement experiment outside in the sun between 10:00 a.m. and 2:00p.m., during which time the sky was cloudless, or the clouds were few. The probe of the spectrometer was 480 mm from the sample surface and perpendicular to the sample surface. During the experiment, a whiteboard measurement was calibrated every 10 min. The experimenter was not allowed to move around and wore dark clothes to reduce interference with the spectral data in the surrounding environment. The spectral data acquisition experiment is shown in Fig. 1.

Fig. 1(c) is the original spectral curve, and each spectral datum has 1024 characteristics. If the data are applied directly to the classification model algorithm, the computational complexity of the model will be high. In addition, the correlation between the sample data is strong, and the redundancy information of the data will decrease the accuracy of the model. Therefore, this study used principal component analysis (PCA) [22] to reduce the dimensions of the spectral data. Finally, we replaced the original data with the first 8 principal components with a contribution rate of 99.8%. Fig. 1(d) is a sample distribution diagram of the first three principal components of the data. After the PCA dimensionality reduction, the data reflected the different characteristics of the coal types.

Table 1
Characteristics of the dataset.

Types	Label	Number of samples
Anthracite	1	71
Bituminous coal	2	80
Lignite	3	58
Coal gangue	4	106
Total	–	315

3. Classification algorithms

3.1. Extreme learning machine

For any different N samples $(\mathbf{x}_i, \mathbf{t}_i)$ in which $\mathbf{x}_i = [x_{i1}, x_{i2}, \cdots, x_{in}]^T \in \mathbf{R}^n$, $\mathbf{t}_i = [t_{i1}, t_{i2}, \cdots, t_{im}]^T \in \mathbf{R}^m$, the single hidden layer neural network of the standard hidden nerve cells at L is

$$\sum_{i=1}^{L} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{t}_j, \quad j = 1, \cdots, N;\ b_i, \beta_i \in \mathbf{R} \tag{1}$$

Here, $\mathbf{w}_i = [w_{i1}, w_{i2}, \cdots, w_{in}]^T \in \mathbf{R}^n$ represents the input weight between the nerve cell in the input layer and the nerve cell i in the hidden layer; β_i is the output weight; b_i is the bias; $g(x)$ is activation function; and $\mathbf{w}_i \cdot \mathbf{x}_j$ represents the inner product of $\mathbf{w}_i$ and $\mathbf{x}_j$.

Then, the ELM function can be expressed as

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T} \tag{2}$$

Here,

$$\mathbf{H}(\mathbf{w}_1, \cdots, \mathbf{w}_L, b_1, \cdots, b_L, \mathbf{x}_1, \cdots, \mathbf{x}_N) = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_1 + b_L) \\ \vdots & & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_N + b_L) \end{bmatrix}_{N \times L} \tag{3}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \qquad \mathbf{T} = \begin{bmatrix} \mathbf{t}_1^T \\ \vdots \\ \mathbf{t}_j^T \end{bmatrix}_{N \times m} \tag{4}$$

$\mathbf{H}$ is the matrix output by the hidden layer of the neural network. $\mathbf{T}$ is the expected output.

The method developed by Professor Huang randomly chooses the input weights and hidden layer bias. Training this network resembles the process for obtaining the least squares solution $\hat{\boldsymbol{\beta}}$ of the linear system $\mathbf{H}\boldsymbol{\beta} = \mathbf{T}$.

Finally, we calculate that the minimum of the least squares solution of the linear system is

$$\boldsymbol{\beta} = \mathbf{H}^{+}\mathbf{T} \tag{5}$$

Here, $\mathbf{H}^{+}$is the Moore-Penrose generalized inverse matrix of $\mathbf{H}$. The minimum of the least squares solution of $\mathbf{H}\boldsymbol{\beta} = \mathbf{T}$ is unique.

3.2. A multilayer extreme learning machine

In 2016, Qu et al. [23] proposed a two-hidden-layer ELM neural network (TELM). The TELM algorithm finds a way to make the output of neural network predictions infinitely close to the actual given output. At the same time, the TELM neural network also retains some advantages of the ELM neural network, such as strong generalization ability, high recognition precision and fast operation speed.

Based on the ELM and TELM algorithms, we presented a multilayer extreme learning algorithm (M-ELM) [24]. Suppose the M-ELM network has K hidden layers (K⩾ 2); then, the algorithm implementation process is as follows:

First, we calculate the output matrix $\mathbf{H}_1$ of the first hidden layer. Then, we calculate the initial output weight $\boldsymbol{\beta}_1 = \mathbf{H}_1^{+}\mathbf{T}$ of the first hidden layer and the output layer according to Eq. (5). Then, we can calculate the expected output matrix $\mathbf{H}_{2E}$ of the second layer by Eq. (6):

$$\mathbf{H}_{2E} = \mathbf{T}\boldsymbol{\beta}_1^{+} \tag{6}$$

Now, we define a matrix $\mathbf{W}_{H2} = [\mathbf{B}_2\ \mathbf{W}_2]$, where $\mathbf{B}_2$ is the input bias of the second layer, and $\mathbf{W}_2$ is the input weight of the second layer. Then, $\mathbf{W}_{H2}$ is obtained through Eq. (7).

$$\mathbf{W}_{H2} = g^{-1}(\mathbf{H}_{2E})\mathbf{M}_2^{+} \tag{7}$$

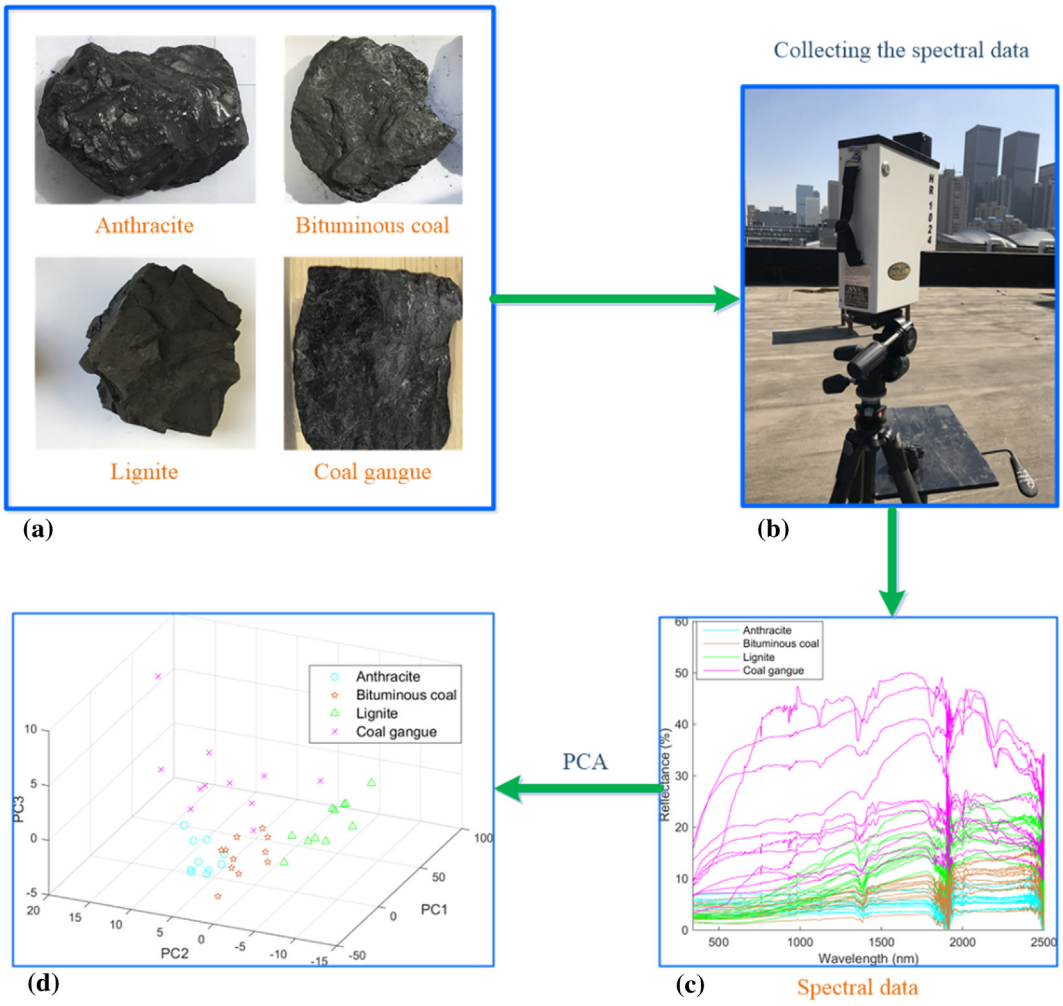

Fig. 1. Collection and processing of the spectral data.

In Eq. (7), $\mathbf{M}_2 = [\mathbf{1}\ \mathbf{H}_1]^T$, where vector $\mathbf{1}$ is a column vector consisting of N scalars 1, $\mathbf{M}_2^+$ is the generalized inverse matrix of $\mathbf{M}_2$, and $g^{-1}(x)$ is the inverse function of the activation function $g(x)$.

Next, we obtain the actual output matrix $\mathbf{H}_{2A}$ of the second layer through Eq. (8).

$$\mathbf{H}_{2A} = g(\mathbf{W}_{H2}\mathbf{M}_2) \tag{8}$$

Then, we calculate the output weight $\boldsymbol{\beta}_2$ of the second layer.

$$\boldsymbol{\beta}_2 = \mathbf{H}_{2A}^{+}\mathbf{T} \tag{9}$$

$\mathbf{H}_{2A}^{+}$ in Eq. (9) is the generalized inverse matrix of $\mathbf{H}_{2A}$.

By analogy, we obtain the expected output matrix $\mathbf{H}_{iE}$, the actual output matrix $\mathbf{H}_{iA}$, and the output weight $\boldsymbol{\beta}_i$ of the hidden layer number i ($2 \leqslant i \leqslant$ K). Finally, the output of the entire network is represented by Eq. (10).

$$f(x) = \mathbf{H}_K\boldsymbol{\beta}_K \tag{10}$$

3.3. An improved artificial bee colony algorithm for optimizing the M-ELM network

An artificial bee colony (ABC) algorithm [25] was proposed in 2007 to mimic bee foraging behavior. There are 3 kinds of bees in the algorithm: employed bees, onlooker bees, and scout bees. The location of the flower source represents a set of feasible solutions D for the staying optimization problem. The amount of nectar represents the degree of adaptability of the optimization problem. If you select the maximum number of iterations for Mk, the ABC algorithm implementation process is as follows:

Suppose the number of food sources is *SN*, and the fitness value of the i ($i = 1, 2, \ldots, SN$) source is fit_i. The employed bees search for a new food source according to Eq. (11).

$$v_{ij} = u_{ij} + \varphi(u_{ij} - u_{nj}) \tag{11}$$

In Eq. (11), v_{ij} is a new food source; u_{ij} is an existing food source; φ is a random value between $[-1, 1]$; n is a random number between $[1, SN]$, and $n \# i$, $1 \leqslant i \leqslant SN$, $1 \leqslant j \leqslant d$; d is the dimension of the feasible solution D. If the fitness value of v_{ij} is greater than that of u_{ij}, the employed bees will replace u_{ij} with v_{ij}; otherwise, it will be retained.

When all the employed bees complete the search, they return to the hive and share information about the food source with the onlooker bees. The onlooker bees select the i ($i = 1, 2, \ldots, SN$) source according to the roulette algorithm. The onlooker bees search for a new food source according to Eq. (11) within the neighborhood of the selected food source, and calculate and retain a more adaptable food source.

When the food source is in the limit cycle, the position information is not updated, and the scout bees update the food source according to Eq. (12).

$$u_{ij} = u_{\min,j} + \varphi(u_{\max,j} - u_{\min,j}) \tag{12}$$

In Eq. (12), $u_{\min,j}$ and $u_{\max,j}$ are the upper and lower limits, respectively, of the j element in feasible solution D.

In Eq. (11), we can see that the search range of the employed bees is relatively small, so ABC may not have the optimal fitness value found.

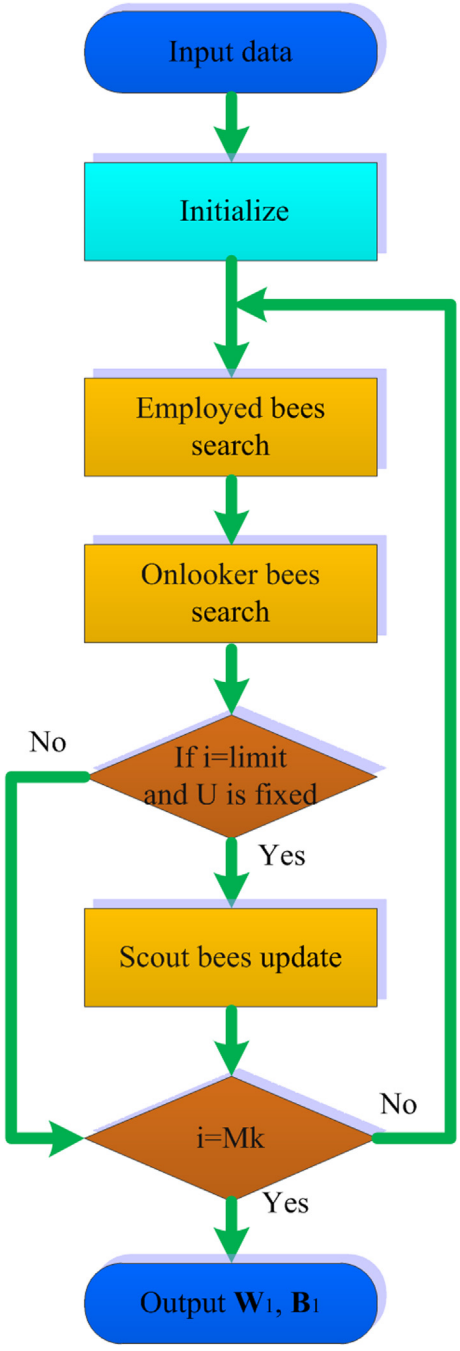

Fig. 2. Flowchart of the IAM-ELM algorithm.

Therefore, based on the idea of the particle swarm optimization algorithm [26], this study adds inertia weight ω to Eq. (11) and proposes Eq. (13):

$$v_{ij} = \omega u_{ij} + \varphi(u_{ij} - u_{nj}); \quad \omega = \omega_c - rand(0, 1)\omega_c \tag{13}$$

In Eq. (13), ω_c is a constant. We use Eq. (13) to search for new food sources and update the ABC algorithm, which we call the inertia weight ABC (IW-ABC).

Now return to the M-ELM algorithm. Under the condition that the number of hidden layer nodes and the number of hidden layers K have been determined, the main influence on the performance of the M-ELM model is the initial input weight $\mathbf{W}_1$ and initial bias $\mathbf{B}_1$. Therefore, the main concern now is determining the values of these parameters, making the classification model more stable and increasing the classification accuracy. Here, we introduce the use of IW-ABC to optimize $\mathbf{W}_1$ and $\mathbf{B}_1$ of the M-ELM algorithm (IAM-ELM algorithm). Fig. 2 is a flow chart of the algorithm. Below we introduce the implementation process of the algorithm:

Algorithm 1. IAM-ELM algorithm

1. Input data; initialize the IW-ABC algorithm, where $d = R * NE + NE$, NE is the number of hidden layer nodes of the M-ELM algorithm, the number of honey sources is $2*SN$; and the number of employed bees, onlooker bees and scout bees is SN. The D of the IW-ABC algorithm is $\mathbf{W}_H$ of the M-ELM algorithm.

for i = 1:Mk

2. The IW-ABC employed bees search for food sources and calculate the best fitness value by the M-ELM algorithm.
3. IW-ABC onlooker bees search for food sources and calculate the best fitness values by the M-ELM algorithm.
4. If i = limit and the food source position is unchanged; the scout bees update the food source location.

end

5. Retain the optimal $\mathbf{W}_1$ and $\mathbf{B}_1$ of the M-ELM algorithm.

4. Results and discussion

This experiment was conducted under a Windows 10 operating system, Intel Core™ i5-7200U CPU @ 2.7 GHz, and 8 GB of memory, NVIDIA M150 graphics, 2 GB GDDR5 memory, and MATLAB 2014b. In the experiment, the SN value of IAM-ELM is 20, the activation function is the Sigmoid function, and the number of hidden layer nodes is selected as 10. The experiment was repeated 100 times, and the average of the results was taken as the final result of the experiment.

4.1. Coal classification model based on the IAM-ELM algorithm

Based on the IAM-ELM algorithm, we established a coal classification model (IAM-ELM model). Fig. 3 shows the classification accuracy and training time of the IAM-ELM model. According to Fig. 3(a), we can see that the hidden layer of the IAM-ELM algorithm number K ≥ 3, and the accuracy of the model significantly increased from 87.1% to 92.5%. As the number of iterations Mk ranged from 20 to 40, the model was stable, and the accuracy was highest. Fig. 3(b) shows the training time of the IAM-ELM model, and the training time increased, from 26.7 s to 213.9 s, as the number of hidden layers K and iterations Mk increased.

Fig. 4 is the result of classifying coal in the IAM-ELM model. As shown in Fig. 4, we can see that the types of coal that were incorrectly identified were mainly in the first and second categories, namely anthracite and bituminous coal, because the visible and near-infrared spectra of the two coals are similar, so that a small number of coal was incorrectly identified. Based on the above analysis results, we can conclude that the IAM-ELM model has a good classification result and that the classification accuracy is over 90%, which provides a fast, efficient and high-precision classification method for coal.

4.2. Comparison of different methods

Zhang et al. [27] proposed an instance cloned extreme learning machine (IC-ELM). According to the experimental results of Zhang, IC-ELM can achieve better classification results compared with the basic ELM and other typical machine learning algorithms.

Support vector machine (SVM) is a machine learning algorithm [28–30], that can effectively solve the problems of small samples, high-dimensional data and nonlinearity, so the algorithm is widely used in many classifications. In papers [9] and [13], a coal classification model (CM-SVM) based on near-infrared spectroscopy and confidence machine-SVM was proposed. Their experimental results showed that the classification effect of the model is better than that of traditional SVM.

Random forest (RF) is an integrated machine learning algorithm [31]. It uses a sampling method to randomly extract several groups of samples, then models the samples, and finally obtains the final result by

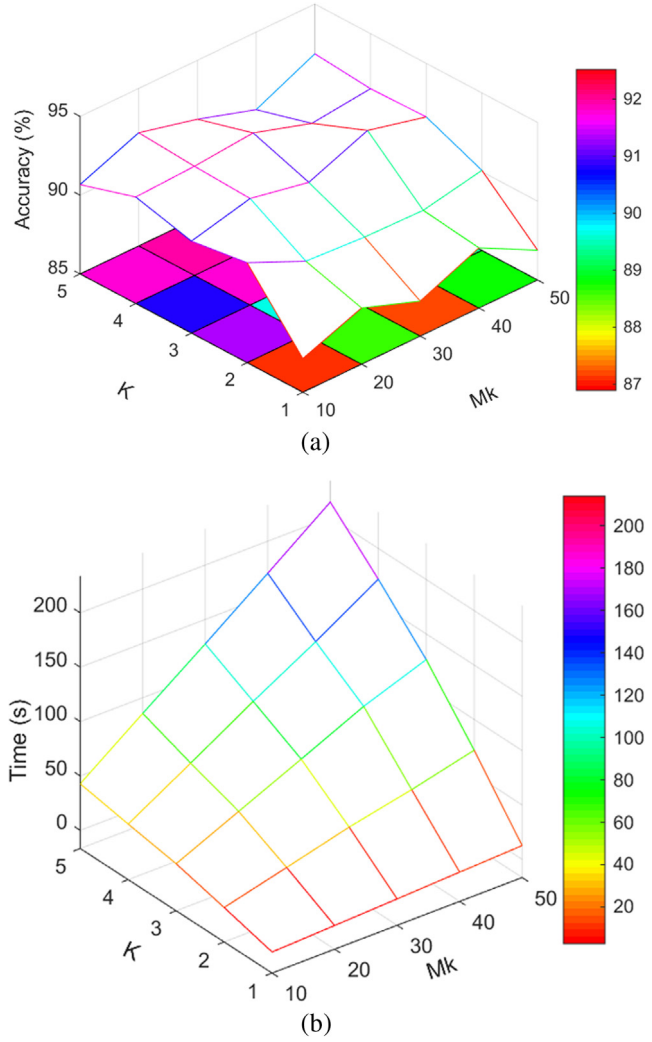

Fig. 3. Classification accuracy and training time of the IAM-ELM model. (a) classification accuracy, (b) training time.

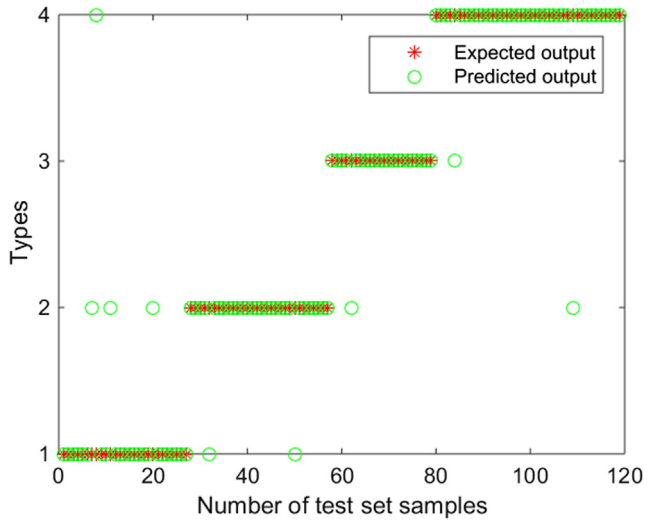

Fig. 4. Classification results of the IAM-ELM model.

the voting method. In recent years, the RF algorithm has been widely used in classification and regression problems [32–34]. As described in a previous paper [35], a SMOTE-RF algorithm was proposed in 2018, and the algorithm was used to identify coal in different countries; the results were good.

Table 2
The results of different classification methods.

Methods	Training time (s)	Accuracy rate of test set (%)
ELM	0	89.24
IC-ELM	0.85	91.26
IAM-ELM (This work)	78.81	92.25
CM-SVM	6.25	90.46
MLP	1.25	89.10
SMOTE-RF	8.30	90.18

Ren et al. [15] constructed a coal classification model by using multi-layer perceptron (MLP) algorithm. This study compares the ELM, IC-ELM, IAM-ELM, CM-SVM, MLP and SMOTE-RF algorithms. For the CM-SVM algorithm, the cross-validation number is 10, and the kernel function is the Sigmoid function. The number of trees of the BagRF algorithm is 200, the MLP algorithm uses a two-layer network.

Table 2 is the result of classifying different methods. The highest classification accuracy of the algorithm was the IAM-ELM algorithm, reaching 92.25%. The second was the IC-ELM algorithm, reaching 91.26%, and the third was the CM-SVM algorithm, reaching 90.46%. Compared with the traditional ELM algorithm, the IAM-ELM algorithm improves the classification accuracy by 3%. For training time, the fastest was the ELM algorithm, followed by the IC-ELM and MLP algorithms. The results show that the proposed IAM-ELM algorithm can improve the classification result of coal species and proves the effectiveness of the method.

5. Conclusion

The spectral technique plays an important role in the identification of coal varieties. In this research, we fused spectral and computer technology to solve the problem of mineral classification. We propose the IAM-ELM algorithm and use the algorithm to establish a rapid coal classification model. Experimental results show that the classification accuracy of the IAM-ELM algorithm is higher than that of traditional ELM, IC-ELM, CM-SVM, MLP and SMOTE-RF algorithms. Compared with the traditional method of coal classification, our method has the advantage of economy, speed and accuracy and has important practical application value.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China under Grant 61773105, Grant 61374147, Grant 41371437, Grant 61203214, Grant 61473072, and Grant 61733003, in part by the Fundamental Research Funds for the Central Universities, China, under Grant N150402001 and Grant 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] A.B.F. Camara, L.S. de Carvalho, C.L.M. de Morais, MCR-ALS and PLS coupled to NIR/MIR spectroscopies for quantification and identification of adulterant in biodiesel-diesel blends, Fuel 210 (2017) 497–506.
[2] X.G. Dong, J. Dong, Y.K. Peng, X.Y. Tang, Comparative study of albumen pH and whole egg pH for the evaluation of egg freshness, Spectr. Lett. 50 (9) (2017) 463–469.
[3] C.X. Tang, H.Y. He, E.B. Li, H.Q. Li, Multispectral imaging for predicting sugar content of 'Fuji' apples, Opt. Laser Technol. 106 (2018) 280–285.
[4] Y.C. Mao, X. Dong, J.F. Cheng, Multigrades classification model of magnesite ore based on SAE and ELM, J. Sens. 2017 (2017) 1–9.
[5] J.M. Andres, M.T. Bona, Analysis of coal by diffuse reflectance near-infrared

spectroscopy, Anal. Chim. Acta 535 (1) (2005) 123–132.
[6] M.T. Bona, J.M. Andres, Coal analysis by diffuse reflectance near-infrared spectroscopy: hierarchical cluster and linear discriminant analysis, Talanta 72 (4) (2007) 1423–1431.
[7] T. Tanno, T. Oohashi, I. Katsumata, Estimation of water content in coal using terahertz spectroscopy, Fuel 105 (2013) 769–770.
[8] D.W. Kim, O.M. Lee, J.S. Kim, Application of near infrared diffuse reflectance spectroscopy for on-line measurement of coal properties, Korean J. Chem. Eng. 26 (2) (2009) 489–495.
[9] Y.S. Wang, M. Yang, G. Wei, Improved PLS regression based on SVM classification for rapid analysis of coal properties by near-infrared reflectance spectroscopy, Sens. Actuator B-Chem. 193 (2014) 723–729.
[10] R.F. Hu, Y.S. Wang, M. Yang, Improved analysis of inorganic coal properties based on near-infrared reflectance spectroscopy, Anal. Methods 7 (12) (2015) 5282–5288.
[11] D.L. Zeng, Y. Hu, J.Z. Liu, Soft sensing of coal moisture, Measurement 60 (2015) 231–239.
[12] D.L. Zeng, Y.H. Wang, S. Gao, Modeling and control of ball mill system considering coal moisture, Measurement 90 (2016) 43–51.
[13] Y.S. Wang, M. Yang, Z.Y. Luo, Rapid coal classification based on confidence machine and near infrared spectroscopy, Spectrosc. Spectr. Anal. 36 (6) (2016) 1685–1689.
[14] Y.R. Gomez, R.C. Hernandez, J.E. Guerrero, FTIR-PAS coupled to partial least squares for prediction of ash content, volatile matter, fixed carbon and calorific value of coal, Fuel 226 (2018) 536–544.
[15] Y. Ren, X.J. Sun, X.A. Dai, Variety identification of bulk commercial coal based on full-spectrum spectroscopy analytical technique, Spectrosc. Spectr. Anal. 38 (2) (2018) 352–357.
[16] B.T. Le, D. Xiao, Y.C. Mao, D.K. He, Coal analysis based on visible-infrared spectroscopy and a deep neural network, Infrared Phys. Technol. 93 (2018) 34–40.
[17] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (1) (2006) 489–501.
[18] K. Sun, J.S. Zhang, C.X. Zhang, J.Y. Hu, Generalized extreme learning machine autoencoder and a new deep neural network, Neurocomputing 230 (2016) 374–381.
[19] B.T. Le, D. Xiao, Y.C. Mao, D.K. He, S. Zhang, X. Sun, X. Liu, Coal exploration based on a multilayer extreme learning machine and satellite images, IEEE Access 6 (2018) 44328–44339.
[20] 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 (2016) 1–9.
[21] B.T. Le, D. Xiao, D. Okello, D.K. He, J.L. Xu, T.T. Doan, Coal exploration technology based on visible-infrared spectra and remote sensing data, Spectr. Lett. 50 (8) (2017) 440–450.
[22] R. Bro, A.K. Smilde, Principal component analysis, Anal. Methods 6 (9) (2014) 2812–2831.
[23] B.Y. Qu, B.F. Lang, J.J. Liang, Two-hidden-layer extreme learning machine for regression and classification, Neurocomputing 175 (PA) (2016) 826–834.
[24] D. Xiao, B.J. Li, S.Y. Zhang, An online sequential multiple hidden layers extreme learning machine method with forgetting mechanism, Chemometrics Intell. Lab. Syst. 176 (2018) 126–133.
[25] D. Karaboga, B. Basturk, A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm, J. Glob. Optim. 39 (3) (2007) 459–471.
[26] M. Taherkhani, R. Safabakhsh, A novel stability-based adaptive inertia weight for particle swarm optimization, Appl. Soft. Comput. 38 (2016) 281–295.
[27] Y. Zhang, J. Wu, C. Zhou, Instance cloned extreme learning machine, Pattern Recognit. 68 (2017) 52–65.
[28] X.Z. Gao, L.Y. Fan, H.T. Xu, Multiple rank multi-linear kernel support vector machine for matrix data classification, Int. J. Mach. Learn. Cybern. 9 (2) (2018) 251–261.
[29] N.T. Son, C.F. Chen, C.R. Chen, V.Q. Minh, Assessment of Sentinel-1A data for rice crop classification using random forests and support vector machines, Geocarto Int. 33 (6) (2017) 587–601.
[30] Z.W. Chen, A. Gu, X. Zhang, Z.Y. Zhang, Authentication and inference of seal stamps on Chinese traditional painting by using multivariate classification and near-infrared spectroscopy, Chemometrics Intell. Lab. Syst. 171 (2017) 226–233.
[31] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[32] D. Walk, Using random forest methods to identify factors associated with diabetic neuropathy: a novel approach, Pain Med. 18 (1) (2017) 1–2.
[33] J. Heil, X. Michaelis, B. Marschner, B. Stumpe, The power of random forest for the identification and quantification of technogenic substrates in urban soils on the basis of DRIFT spectra, Environ. Pollut. 230 (2017) 574–583.
[34] A. Paul, D.P. Mukherjee, P. Das, Improved random forest for classification, IEEE Trans. Image Process. 27 (8) (2018) 4012–4024.
[35] M. Lei, X.H. Yu, M. Li, W.X. Zhu, Geographic origin identification of coal using near-infrared spectroscopy combined with improved random forest method, Infrared Phys. Technol. 92 (2018) 177–182.