ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

Accurate Coal Classification Using PAIPSO-ELM with Near-Infrared Reflectance Spectroscopy

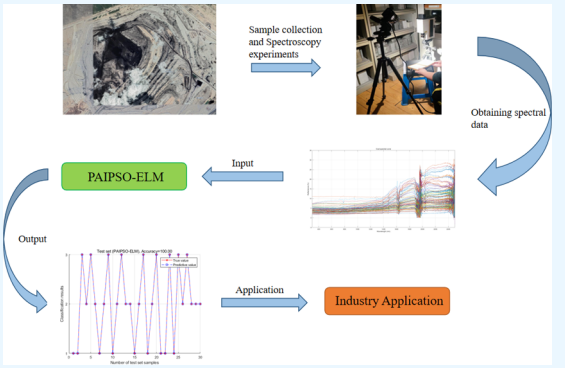
Yiyang Wang, Boyan Li, Haoyang Li, and Dong Xiao*

Cite This: *ACS Omega* 2024, 9, 47756−47764

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: China has vast proven coal reserves, encompassing a wide variety of types. However, traditional coal classification methods have limitations, often leading to inaccurate classification and inefficient utilization of coal resources. To address this issue, this paper introduces the Extreme Learning Machine (ELM) as a novel coal classification method, based on the near-infrared reflectance spectroscopy (NIRS) of coal. Initially, we collected NIRS data from coal samples using the SVC-HR-1024 spectrometer. Given the high dimensionality and strong linear correlations in NIRS data, we conducted preprocessing to enhance the usefulness of the data. In experiments, the ELM model demonstrated good classification performance. However, due to the random generation of input layer weights and hidden layer biases in the ELM model, its performance can be unstable, preventing the model from fully realizing its potential. To overcome this shortcoming, we employed the Particle Swarm Optimization (PSO) algorithm to optimize the parameters of the ELM model. Simulation results showed that the PSO-ELM model achieved a 9.68% improvement in classification accuracy compared to the original ELM model. Furthermore, we optimized the PSO algorithm by introducing exponentially decaying inertia factors and position-variant particles to further reduce the risk of the algorithm falling into local optima. The improved Position-Adaptive Inertia PSO-ELM (PAIPSO-ELM) model achieved an additional 2% increase in classification accuracy over the PSO-ELM model, without a significant increase in training time. In summary, this paper proposes a coal spectral classification method based on the PAIPSO-ELM model, effectively overcoming the limitations of traditional classification methods while meeting industrial demands for classification accuracy and speed.



1. INTRODUCTION

Coal, a common black or brown combustible sedimentary rock, is essentially the fossilized remains of plants subjected to prolonged compression and carbonization. The type of coal determines its use, making a thorough understanding of coal classification crucial for optimizing resource utilization and promoting sustainable, efficient use. Different coal types have distinct chemical properties, leading to varied industrial applications. For example, anthracite has a high ignition temperature and produces minimal smoke and odor when burned, making it essential for high-quality coke production. Bituminous coal, with its high volatile content, generates significant heat and is the primary fuel for power stations, also serving as a source for byproducts like tar, benzene, and methanol. Although lignite has a lower heat value, it can be gasified into syngas and further processed into liquid fuels such as diesel and gasoline.[1] Additionally, its high organic content makes it suitable for soil conditioners, enhancing agricultural soil quality. Therefore, accurate classification of different coal types is of great significance.

Traditional coal classification methods primarily include manual classification and chemical analysis. Manual classification is less accurate and slower, and is usually difficult to meet the demands of mass production. Although chemical analysis can theoretically provide more accurate results, it requires complex chemical testing of coal samples, which is time-consuming and costly, making it unsuitable for large-scale applications. Thus, it is clear that there is a pressing demand for more effective methods of coal classification.[2]

Spectral analysis is a unique technique for material analysis that disperses polychromatic light through a dispersive system and arranges it by wavelength to form an image, thereby revealing the optical properties of the material. In the coal industry, spectral analysis plays a crucial role by providing in-depth insights into the absorption, reflection, and transmission characteristics of coal across different wavelengths, offering

Received: August 31, 2024
Revised: October 20, 2024
Accepted: October 29, 2024
Published: November 18, 2024

valuable information about its elemental composition, mineral content, and organic structure.[3−8] Spectroscopy includes several methods with unique features. Reflectance spectroscopy, like visible spectroscopy and near-infrared reflectance spectroscopy, is used for surface analysis. NIRS is a nondestructive technique that operates in the 780−2500 nm range, commonly used to analyze surface properties of materials, particularly molecular bonds like O−H, C−H, and N−H. NIRS provides valuable insights into coal's composition, though it is often combined with other techniques to achieve more detailed analysis. Terahertz spectroscopy can penetrate nonconductive materials, while Laser-induced breakdown spectroscopy (LIBS) utilizes lasers to vaporize material for rapid elemental analysis. The weighing method or artificial selection process for ore sorting heavily relies on the experience of workers in mineral processing. Additionally, the sorting mechanism of these methods is not well-defined, making it difficult to achieve high-precision sorting for ores from different regions. In contrast, using spectroscopy for ore sorting offers higher accuracy. Compared to chemical analysis methods, the use of spectroscopy for coal analysis offers several advantages: Spectroscopic analysis provides real-time results without the need to wait for laboratory testing; it does not require chemical reagents, thereby avoiding the generation of pollutants, and it minimizes coal loss during testing, with near nondestructive capabilities; chemical analysis methods are costly, requiring both reagents and specialized equipment, whereas spectroscopic analysis requires only a spectrometer to obtain the coal's spectrum. Xiao and Yan demonstrated the advantages of intelligent sorting methods using spectroscopy over other traditional methods: spectroscopy-based intelligent sorting methods perform very well in terms of accuracy, spending, and time consumption.[13]

This technique is vital for understanding the chemical structure and reactivity of coal. By studying the interaction between coal and different wavelengths of light, we can infer the physical and chemical properties of coal, allowing for more accurate classification and regression analysis of coal composition. However, there are challenges associated with modeling coal classification using spectral data. The process of acquiring spectral data is susceptible to external factors, such as uncontrolled natural lighting conditions and errors introduced by manual operations. Additionally, the high dimensionality of spectral data within a fixed wavelength range can lead to data redundancy, complicating the modeling process. Consequently, traditional methods for modeling spectral data often fall short of practical requirements. With the advancement of technology, machine learning algorithms have gradually emerged as effective tools for handling high-dimensional spectral data. These algorithms can automatically learn and extract essential features from the data, enabling more accurate coal classification in complex environments. Especially when handling large amounts of redundant data and noise, the robust and adaptable nature of machine learning algorithms can greatly improve the precision and consistency of model predictions. As a result, an increasing number of researchers are combining machine learning techniques with spectral data for mineral analysis. Mao et al. utilized NIRS combined with a stacked autoencoder and extreme learning machine to classify various grades of magnesite, achieving favorable classification outcomes.[9] Hu et al. developed a model for detecting coal composition using NIRS and least-squares support vector machine (SVM), effectively predicting mineral content.[10] Yao et al. applied laser-induced breakdown and near-infrared reflectance spectroscopy to enhance coal characterization, creating an analytical model that accurately predicts volatiles, ash, and moisture levels.[11] Lei et al. combined NIRS with the Broad Learning (BL) algorithm to determine the source of coal, achieving favorable results.[12] Xiao et al. employed NIRS in conjunction with deep learning techniques to achieve rapid and accurate coal identification, demonstrating high effectiveness across various stages of coal utilization.[13]

Extreme Learning Machine has seen extensive use in fields like regression, classification, and feature learning. As a single-hidden layer feed-forward neural network (SLFN), ELM is known for its fast learning speed, resistance to local optima, and strong generalization capabilities.[12] Due to its strong performance in classification and regression tasks, along with robust generalization, ELM has been widely adopted and further enhanced by numerous researchers.[14−17] Huang et al. introduced ELM-LRF, which reduced the error rate from 6.5 to 2.7% compared to traditional deep learning methods, while also significantly increasing learning speed.[18] Kutlu et al. proposed the LU-ELM classification model to obtain higher classification accuracy compared to the traditional ELM classification model by using the lower-up triangularization method instead of singular value decomposition.[19] Luo et al. proposed an enhanced stacked ELM (S-ELM) that incorporates principal component analysis.[20] By integrating the sparse autoencoder ELM method into the S-ELM structure, they achieved further enhancements in learning accuracy. Bai et al. introduced the sparse ELM theory, which minimizes storage requirements and reduces testing time in large-scale applications, offering faster learning speed and improved generalization ability.[21] In mineral analysis research, ELM has demonstrated strong performance, meeting both accuracy and speed requirements, and showing great potential.

While ELM is effective for coal classification, the random configuration of parameters in ELM can lead to feature loss and reduced accuracy. In this study, Particle Swarm Optimization is introduced to improve ELM, with further optimization achieved by incorporating an exponentially decreasing inertia factor and a mutation factor. This approach allows for better allocation of weights and biases. The preprocessed coal spectral data is then used as input for the model, resulting in high-precision coal classification.

2. THEORY AND METHOD

2.1. Data Collection. The spectral data in experiment were obtained from coal types including anthracite, bituminous coal, and lignite collected from different mining areas, with the labels for anthracite, bituminous coal, and lignite set as 1, 2, and 3. The exact quantities of coal samples are listed in Table 1.

First, spectral analysis was performed on the collected coal samples using the SVC- HR-1024 spectrometer, a high-resolution portable device capable of capturing spectral data from the visible to near-infrared regions.

Table 1. Quantity of the Coal Samples

types	label	quantity
anthracite	1	37
bituminous coal	2	37
lignite	3	26

https://doi.org/10.1021/acsomega.4c08020
ACS Omega 2024, 9, 47756−47764

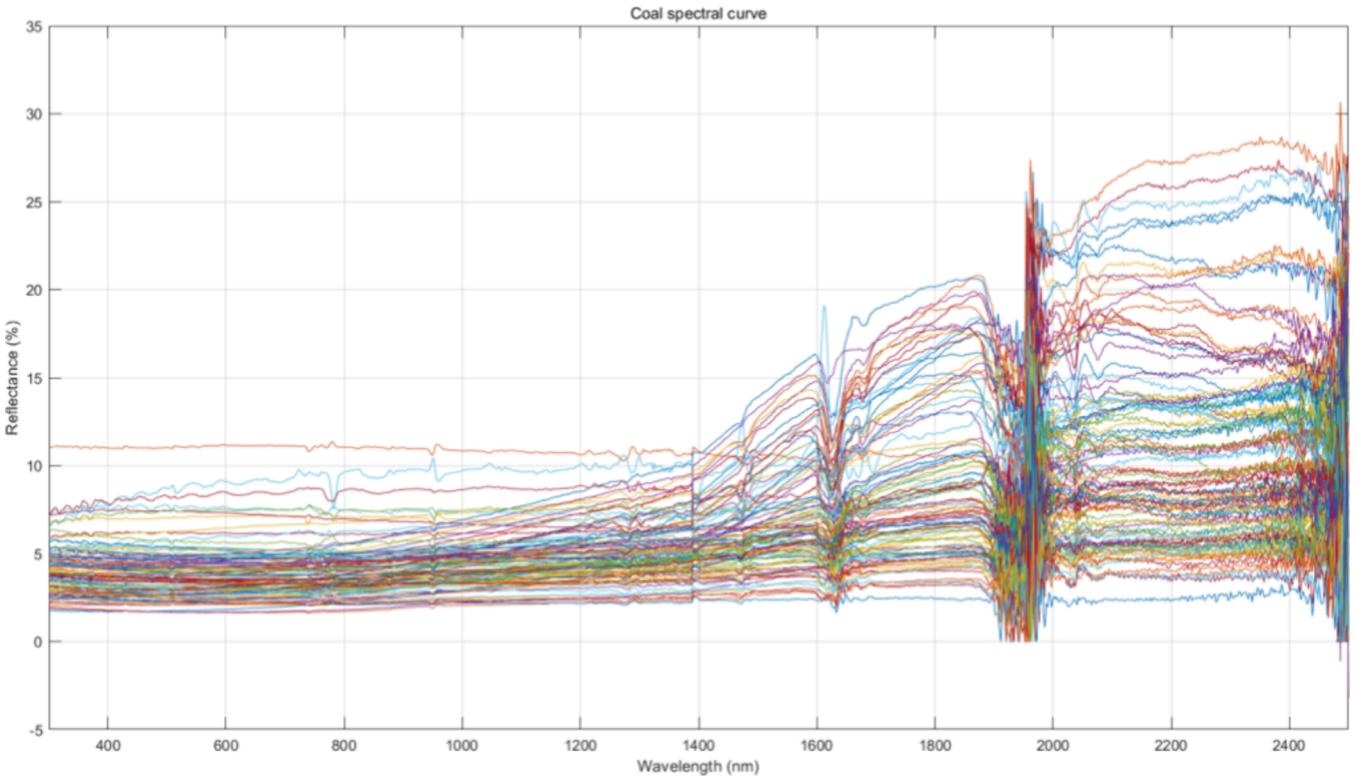


Figure 1. Coal spectral curve.

Before conducting spectral measurements on the coal samples, the samples were cleaned, cut, and ground into powder. Under clear outdoor conditions, the SVC HR-1024 portable ground spectrometer was employed to collect spectral data of the coal samples in the 350−2500 nm wavelength range. The scanning time was 1 s, and the probe was positioned 480 mm from the sample. Under the same environmental conditions, while keeping environmental factors as constant as possible over time, first calibrate the spectrometer using a standard white reference board. After adjusting the measurement angle and distance, place the prepared sample under the spectrometer to obtain the near-infrared reflectance spectrum of the sample. To minimize the influence of environmental and human interference, the same sample is measured five times, and the average is taken to generate the near-infrared reflectance spectrum data set for the sample. The resulting coal spectral curves are shown in Figure 1.

2.2. Data Preprocessing. In this experiment, spectral data with 973 dimensions were obtained using the SVC-HR-1024 spectrometer. Using high-dimensional data as model input increases complexity, potentially causing redundancy and lowering accuracy. Additionally, the spectral data exhibit high linear correlation and contain some noise, making data preprocessing crucial for reducing model complexity and improving accuracy. Principal Component Analysis (PCA) is a dimensionality reduction technique that converts high-dimensional data into a lower-dimensional form through linear projection. This approach helps retain important features, enhances model performance, and mitigates overfitting.[22]

Suppose there are m coal samples $\{x_1, x_2, \cdots, x_m\}$, each with n-dimensional features $X^i = (x^i_1, x^i_2, \ldots x^i_n)$, each with its own eigenvalue.

Standardizing the raw data begins by subtracting the mean from each feature as the first step, resulting in centered data with a mean of 0 and a variance of 1. This step puts all features on a uniform scale, preventing those with larger variances from dominating the dimensionality reduction process. The mean formula is shown in eq 1.

$$\overline{x} = \frac{1}{m}\sum_{i=1}^{m} x^i \tag{1}$$

The second step is to compute the covariance matrix:

$$\mathrm{cov}(x_1, x_2) = \frac{\sum_{i=1}^{m}(x_1^i - \overline{x}_1)(x_2^i - \overline{x}_2)}{m-1} \tag{2}$$

Using the covariance formula, the covariance matrix for this sample, based on the dimensional feature, is derived and presented in eq 3.

$$C_{n\times n} = \begin{bmatrix} \mathrm{cov}(x_1, x_1)\cdots\cdots\mathrm{cov}(x_1, x_n) \\ \vdots\ddots\vdots \\ \vdots\ddots\vdots \\ \vdots\ddots\vdots \\ \mathrm{cov}(x_n, x_1)\cdots\cdots\mathrm{cov}(x_n, x_n) \end{bmatrix} \tag{3}$$

The next step involves computing the characteristic values and vectors of the covariance matrix. Eigenvalues indicate the variance along each eigenvector direction, while the eigenvectors indicate the primary directions of data concentration.

The fourth step is to select the new data dimension k based on the cumulative contribution rate. The formula for

https://doi.org/10.1021/acsomega.4c08020
ACS Omega 2024, 9, 47756−47764

cumulative contribution rate of the k-th principal component is shown in eq 4.

$$\text{contri} = \frac{\sum_{j=1}^{K} \lambda_j}{\sum_{i=1}^{n} \lambda_i} \times 100\% \tag{4}$$

The fifth step is to map origin spectrum into a different feature space. By using first k eigenvectors as basis, an $n \times k$ matrix U is constructed. The original data set is then mapped into a new feature subspace using matrix U:

$$\begin{bmatrix} y_1^i \\ y_2^i \\ \vdots \\ \vdots \\ y_k^i \end{bmatrix} = \begin{bmatrix} u_1^T \cdot (x_1^i, x_2^i, \cdots, x_n^i)^T \\ u_2^T \cdot (x_1^i, x_2^i, \cdots, x_n^i)^T \\ \vdots \\ \vdots \\ u_k^T \cdot (x_1^i, x_2^i, \cdots, x_n^i)^T \end{bmatrix} \tag{5}$$

Where $(y^i{}_1, y^i{}_2, \cdots, y^i{}_k)^T$ represents the new feature after projection, and u_i is the column vector of the projection matrix U. The new feature space is usually smaller than the original feature space, but still retains most of the information in the data set.

2.3. Extreme Learning Machine. ELM is a SLFN introduced by Guang-Bin Huang. The core idea of ELM is to randomly assign weights and biases at the start, keeping them fixed during training. The hidden layer outputs are then linearly mapped to the output layer, with output weights typically decided by the least-squares method. ELM is known for its fast training speed, high accuracy, and strong generalization.

The input coal sample quantity is denoted as q, and the count of principal components is n. Consequently, the input layer contains n neurons. The input matrix X is shown in eq 6, where each column of X represents one feature of the entire sample set.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1q} \\ x_{21} & x_{22} & \cdots & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nq} \end{bmatrix}_{n \times q} \tag{6}$$

The hidden layer contains l neurons. Eq 7 presents the weight matrix connecting the input layer to the hidden layer,

$$w = \begin{bmatrix} w_{11} w_{12} \cdots w_{1n} \\ w_{21} w_{22} \cdots w_{2n} \\ \cdots \\ w_{l1} w_{l2} \cdots w_{ln} \end{bmatrix}_{l \times n} \tag{7}$$

where w_{ij} indicates the weight linking the i-th hidden neuron to the j-th input neuron.

The bias matrix b is shown in eq 8.

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_l \end{bmatrix}_{l \times 1} \tag{8}$$

An activation function $g(x)$ is applied in the hidden layer neurons. The model's input passes through the hidden layer, where it is processed by the weights, biases, and activation function, resulting in the hidden layer output H:

$$H(w_1, w_2, \cdots, w_l, b_1, b_2, \cdots, b_l, x_1, x_2, \cdots, x_Q) = \tag{9}$$

The output weight β is shown in eq 10.

$$\beta = \begin{bmatrix} \beta_{11} \beta_{12} \cdots \beta_{1m} \\ \beta_{21} \beta_{22} \cdots \beta_{2m} \\ \cdots \\ \beta_{l1} \beta_{l2} \cdots \beta_{lm} \end{bmatrix}_{l \times m} \tag{10}$$

Building on the results from the neural network's forward pass, the output matrix is shown in eq 11.

$$t_j = \begin{bmatrix} t_{1j} \\ t_{2j} \\ \cdots \\ t_{mj} \end{bmatrix}_{m \times 1} = \begin{bmatrix} \sum_{i=1}^{l} \beta_{i1} g\left(w_i x_j + b_i\right) \\ \sum_{i=1}^{l} \beta_{i2} g\left(w_i x_j + b_i\right) \\ \cdots\cdots \\ \sum_{i=1}^{l} \beta_{im} g\left(w_i x_j + b_i\right) \end{bmatrix}_{m \times 1} \quad (j = 1, 2, \cdots, Q) \tag{11}$$

Given that the data set labels are $Y_{1q} = [y_{11} \quad y_{12} \cdots \quad y_{1q}]^T$ the loss function is

$$\left\| H\widehat{\beta} - Y^T \right\| = \left\| HH^+ Y^T - Y^T \right\| = \min_{\beta} \left\| H\beta - Y^T \right\| \tag{12}$$

And β are solved as follows:

$$\beta = H^+ Y^T \tag{13}$$

where H^+ is the pseudoinverse of H. Figure 2 illustrates the structure of ELM.

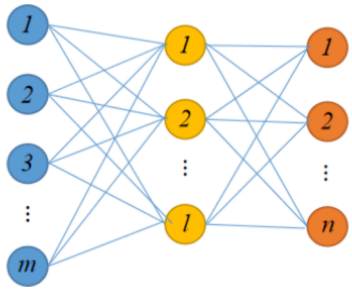


Figure 2. ELM's structure.

2.4. Particle Swarm Optimization Algorithm. PSO algorithm was introduced by Kennedy and Eberhart in 1995 as a swarm intelligence algorithm, designed for solving continuous optimization problems.[23] Inspired by the collaborative behavior of birds flocking, as illustrated in Figure 3, PSO utilizes information sharing among individuals in the swarm, allowing the group to move from disorganized exploration to a

https://doi.org/10.1021/acsomega.4c08020
ACS Omega 2024, 9, 47756−47764

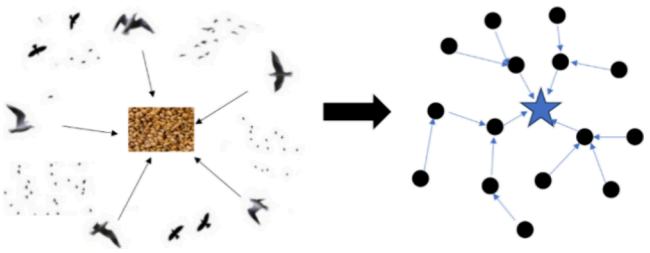


Figure 3. PSO algorithm ideas.

structured search within the solution space, ultimately converging on the optimal solution.

The parameters used in the PSO modeling process are presented in Table 2.

Table 2. Parameter Description Table

parameter	description
N	number of particles
D	dimensionality of space
X	particle position
V	particle velocity
V_{max}	maximum particle velocity
w	inertia factor
c_1	cognitive learning coefficient
c_2	social learning coefficient
r_1, r_2	random value in [0,1]
itmax	maximum number of iterations
Pbest	particle's personal best position
Gbest	global best position
indim	number of input layer neurons
hiddennum	number of hidden layer neurons
outdim	number of output layer neurons

In the space, there are N particles searching in a D-dimensional target space. The dimension D represents the solution that includes all the weights and biases. There are *indim* × *hiddennum*weights between each input and hidden layer neuron, and *hiddennum* biases. Additionally, there are *hiddennum* × *outdim* weights between the hidden and output layer neurons. Dimension D of the particles is shown in eq 14.

$$D = (indim + 1) \times hiddennum \tag{14}$$

This particle dimensionality configuration allows the PSO algorithm to optimize more adjustable parameters in ELM model, with each particle position precisely mapping to a specific weight or bias in the network. The position X_i and the velocity V_i of the i-th particle are both D-dimensional vectors, as shown in eqs 15 and 16.

$$X_i = (x_{i1}, x_{i2}, \cdots, x_{iD}),\ i = 1, 2, \cdots, D \tag{15}$$

$$V_i = (v_{i1}, v_{i2}, \cdots, v_{iD}),\ i = 1, 2, \cdots, D \tag{16}$$

Each particle determines its fitness value through the objective function, recording *Pbest* and *Gbest* within the swarm. It adjusts its velocity and direction based on these values to seek the optimal solution. The velocity update formula for particle iii is shown in eq 17.

$$v_{ij}(t+1) = wv_{ij}(t) + c_1r_1(t)[p_{ij}(t) - x_{ij}(t)] + c_2r_2(t)[p_{gj}(t) - x_{ij}(t)] \tag{17}$$

$c_1r_1(t)[p_{ij}(t) - x_{ij}(t)]$ is referred to as the cognitive component, which is a vector pointing from the current position to the particle's personal best position, representing the part of the particle's movement derived from its own experience. $c_2r_2(t)[p_{gj}(t) - x_{ij}(t)]$ is referred to as the social component, which is a vector pointing from the current position to the global best position of the swarm, reflecting the collaboration and information sharing among particles.

If a particle's velocity becomes too large after iteration, it may cause the particle to exceed the search space boundaries, reducing its ability to find the global optimum. Therefore, a velocity range is set to control the particle's movement. The velocity formula is shown in eq 18.

$$v_{ij}(t+1) = \begin{cases} v_{ij}(t+1), & v_{ij}(t+1) < v_{max} \\ v_{max}, & v_{ij}(t+1) \geq v_{max} \end{cases} \tag{18}$$

The position update formula for the particles is shown in eq 19.

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \tag{19}$$

2.5. PAIPSO-ELM. The inertia factor w in PSO is a key parameter controlling particle velocity and influencing search behavior. If w is too large, particles may overshoot the optimal solution, leading to convergence difficulties or local optima. If w is too small, particles may fail to effectively explore the search space, missing the optimal solution. A fixed w cannot flexibly switch between global and local search, nor adjust the strategy based on dynamic information during the search process. This may result in early stagnation in local optima or an inability to finely adjust toward better solutions in the later stages, reducing the algorithm's adaptability and flexibility.

To address the above issues, we replace the fixed w with an exponentially decreasing w. The formula for the exponentially decreasing inertia factor is shown in eq 20. This approach allows the inertia factor to adapt to different problems by adjusting the constant that controls the decay rate,

$$w = w_{max} \times e^{-k \times iter} \tag{20}$$

where k is the hyperparameter, *iter* is the number of iterations, and w_{max} is the initial inertia weight.

Setting a larger initial w facilitates global search, gradually decreasing w allows for fine-tuned local optimization. By introducing an adaptively decreasing w, PSO balances global and local search, improving performance and convergence speed.

Additionally, to reduce the likelihood of particles getting trapped in local optima, this paper introduces a mutation factor, giving each particle the opportunity for independent mutation. This is intended to enhance diversity and global exploration during the search process. The parameters for the mutation factor are shown in Table 3.

Table 3. Parameter Description Table for Mutation Factor

parameter	description
a	lower bound of particle space
b	upper bound of particle space
r	random number in the range [0,1]
mutRate	mutation probability
mutationFactor	magnitude of mutation factor

https://doi.org/10.1021/acsomega.4c08020
ACS Omega 2024, 9, 47756−47764

Each particle has a certain probability of undergoing mutation, with the mutation formula given in eq 21.

$$x_{ij} = x_{ij} + \text{mutationFactor} \times r \times (b - a) \tag{21}$$

Here, *mutationFactor* controls the extent of the mutation, which helps the PSO algorithm avoid premature convergence and potentially discover new promising regions in the search space, leading to the optimal solution.

Algorithm: PAIPSO-ELM

Input: Training data set $\{(x_j, y_j)\}_{j=1}^{q}$, the nodes' count l, activation function $g(\cdot)$.

Output: The predicting results $T = [t_1 \quad t_2 \quad \cdots \quad t_q]$.

1: $D = 30$; $N = 100$; $itmax = 30$; // Initialize PSO parameters

2: $k = 0.05$; $w_{\max} = 0.8$; $w = w_{max} \times e^{-k \times iter}$; // Set the exponentially decreasing inertia factor

3: $mutRate = 0.1$; $mutationFactor = 0.2$; $a = 1$; $b = -1$; // Set the Mutation Factor

4: **for** (i=1; i≤N; i++) **do**

5: Initialize particle position X_i, particle velocity V_i, $pbest$ and $gbest$;

6: **end for**

7: **for** (t=1; t≤*itmax*; t++) **do**

8: **for** (i=1; i≤N; i++) **do**

9: Update the particle velocity in each dimension according to Eq. 17;

10: Determine if the particle velocity is reasonable based on Eq. 18;

11: **if** X_i has mutated **then**

12: Update the particle positions according to Eq. 23;

13: **end if**

14: **if** X_i has not mutated **then**

15: Update the particle positions according to Eq. 19;

16: **end if**

17: Calculate the fitness function $f(X_i)$; %The fitness function is the test accuracy of the ELM

18: Update $pbest$ and $gbest$;

19: **end for**

20: **end for**

21: Obtain the optimized weight matrix W and bias matrix B from $pbest$;

22: Calculate β according to eq. 13;

23: Predict the type of the coal;

3. RESULTS AND DISCUSSION

The hardware configuration used in this experiment for coal spectral classification based on machine learning is as follows: The CPU model is a Intel(R) Core(TM) i7-12700H; the GPU is an NVIDIA GeForce RTX 3060 with 6GB of VRAM; the main storage is a DDR5 4800 MHz drive with a capacity of 512GB; and the motherboard model is LNVNB161216. The software used is MATLAB 2016b, running on a Windows 10.

3.1. Data Preprocessing. The collected coal sample spectral data consists of 973 dimensions, with some features potentially having low correlation with coal categories, or even introducing noise, which could significantly impact the performance of subsequent machine learning models. Therefore, PCA is applied to reduce the dimensionality of the coal spectral data, using cumulative contribution rate as the criterion for selecting the target dimensionality. Figure 4 illustrates the cumulative contribution rate after dimensionality reduction.

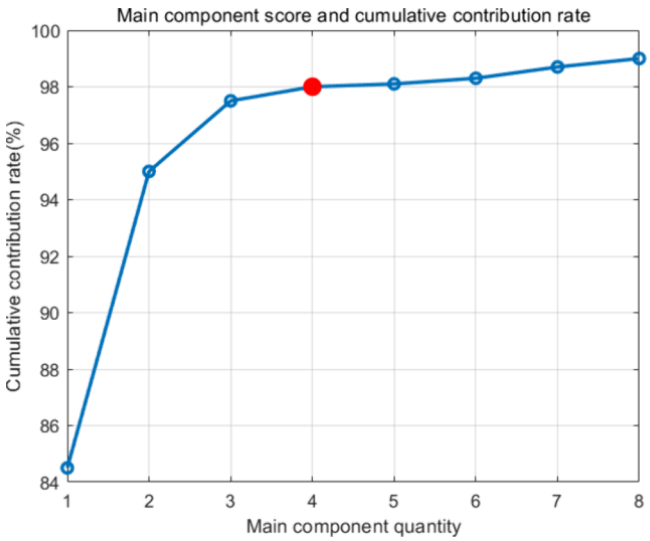


Figure 4. Cumulative contribution rate.

As shown in the Figure 4, the cumulative contribution rate of PCA reaches 98% with four principal components, indicating that reducing the coal spectral data to four dimensions still retains most of the original information. Therefore, the target dimensionality for PCA in this coal spectral classification experiment is set to 4. The reduced data is then normalized to adjust the scale, enabling the subsequent machine learning models to better learn and adapt to data patterns. Finally, the normalized data is randomly split into training and test sets with a 7:3 ratio using the holdout method, completing the data set preparation for the coal spectral classification model.

3.2. Experiment Result. *3.2.1. Simulation Validation of ELM-Based Spectral Classification Model for Coal.* The coal spectral data set was used to perform performance simulations for the Backpropagation Neural Network (BP), Support Vector Machine (SVM), and Extreme Learning Machine models. The parameters for the three machine learning algorithms are set as follows:

ELM: The number of input layer neurons is 4, the number of hidden layer neurons is 6, and the activation function for the hidden layer is the sigmoid function.

The classification accuracy of different models is as follows, with Figures 5, 6, and 7 representing the classification accuracy of the BP neural network, SVM, and ELM models on the test set coal samples, respectively.

Ten tests were carried out to assess and compare the classification accuracy and training time of the three models in predicting coal types. The results are presented in Table 4.

Experimental results indicate that ELM achieved an average classification accuracy of 85% and a training time of 0.48s; for

https://doi.org/10.1021/acsomega.4c08020
ACS Omega 2024, 9, 47756−47764

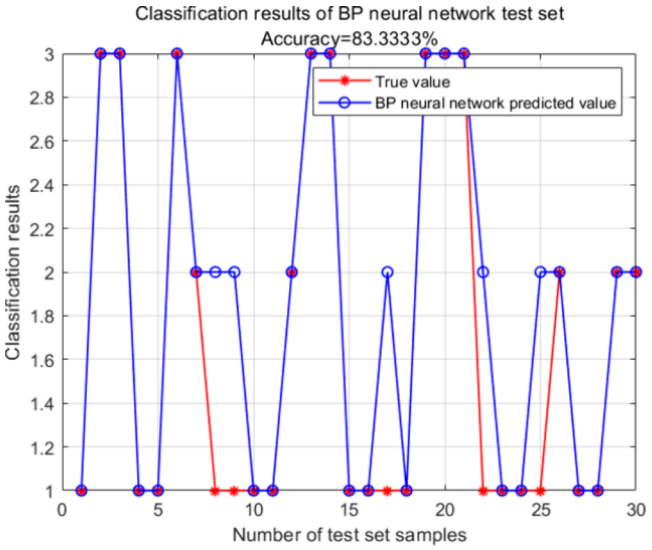


Figure 5. Classification accuracy of BP neural network.

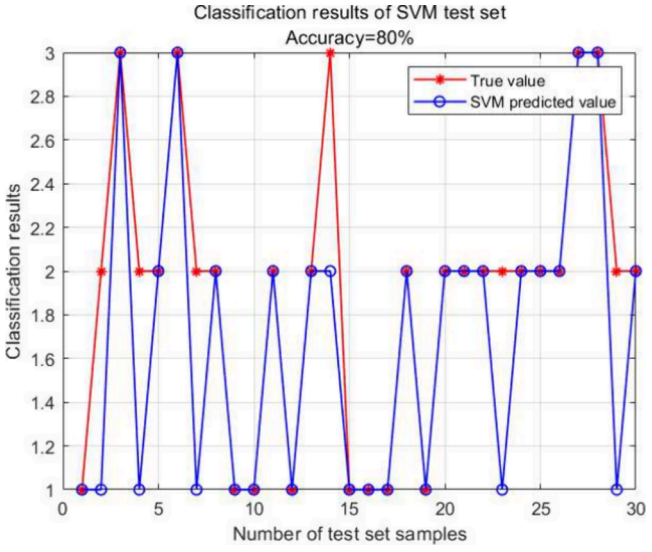


Figure 6. Classification accuracy of SVM.

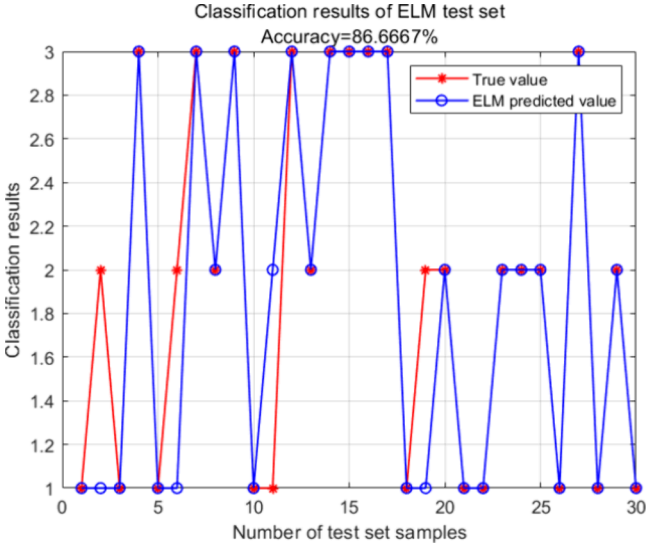


Figure 7. Classification accuracy of ELM.

Table 4. Comparison of Classification Accuracy and Training Time

	BP	SVM	ELM
average accuracy (%)	84.99	83.67	85
average time (s)	6.33	0.94	0.48

the BP neural network, they are 84.99% and 6.33 s; and for SVM, they are 83.67% and 0.94 s. When comparing the average classification accuracy and training time, SVM's performance is less satisfactory, likely due to the multi-classification nature of the problem affecting SVM's precision. While BP and ELM have similar classification accuracy, BP requires significantly more training time, likely because BP involves iterative weight adjustments between neurons, whereas ELM's weights between the input and hidden layers are randomly assigned and fixed.

These results confirm that the ELM model not only has higher classification accuracy but also trains faster compared to BP and SVM. The choice of activation function of ELM model can impact classification accuracy, requiring a comparison of various activation functions to identify the most suitable one for the model. Additionally, testing ELM with various activation functions on the same data set and comparing their classification accuracy can provide insights into the model's generalization performance. If ELM maintains high accuracy across different activation functions, it suggests strong generalization capabilities.

Ten tests were conducted to compare the average accuracy of ELM models using different activation functions, with the results presented in Table 5.

Table 5. Average Accuracy of ELM Classification Models with Different Activation Functions

	ReLU	Sigmoid	Tanh
average accuracy (%)	83.33	85	86.32

In summary, the performance of the ELM classification model has been validated, leading to the selection of the Tanh function as the activation function for this ELM model. The next step involves introducing the PSO optimization algorithm to further improve the performance of the ELM model using the Tanh function.

3.2.2. Simulation Validation of PAIPSO-ELM-Based Spectral Classification Model for Coal. In this experiment, we introduce the PSO algorithm to optimize the parameters of the ELM: the weight matrix and the bias. The PSO-ELM model results are illustrated in Figure 8.

The results show that ELM achieved an average accuracy of 86.32% with a training time of 0.48 s, while PSO-ELM improved accuracy to 96% with a training time of 1.322 s. PSO-ELM optimizes weights and biases, enhancing accuracy and generalization. Though training time increased slightly, it remains efficient, confirming PSO-ELM's superior performance. Next, the structure of PSO-ELM will be further refined.

This experiment improved PSO-ELM by introducing a decreasing inertia factor and mutation particles from genetic algorithms. In the early stage of the algorithm the global search is enhanced by the inertia factor, which has a wider search in the search space and reduces the risk of falling into a local optimum, and in the later stage a more accurate local search is used, which avoids the problem of skipping the optimal

https://doi.org/10.1021/acsomega.4c08020
ACS Omega 2024, 9, 47756−47764

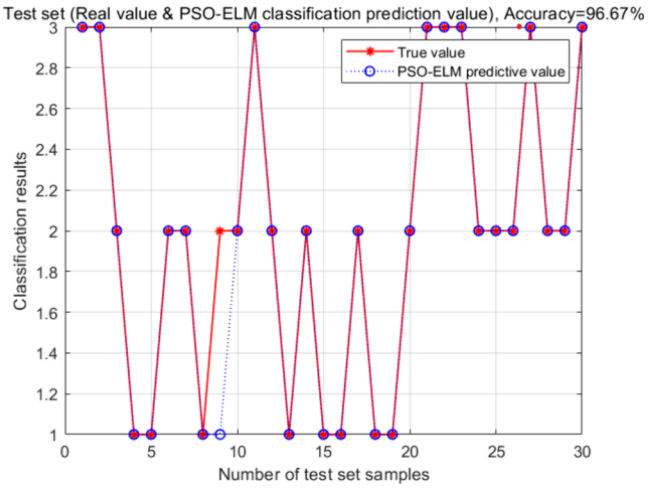


Figure 8. PSO-ELM classification model results.

solution, and the risk of falling into a local optimum is further reduced by the inclusion of a mutation factor. The results of the PAIPSO-ELM model are shown in Figure 9.

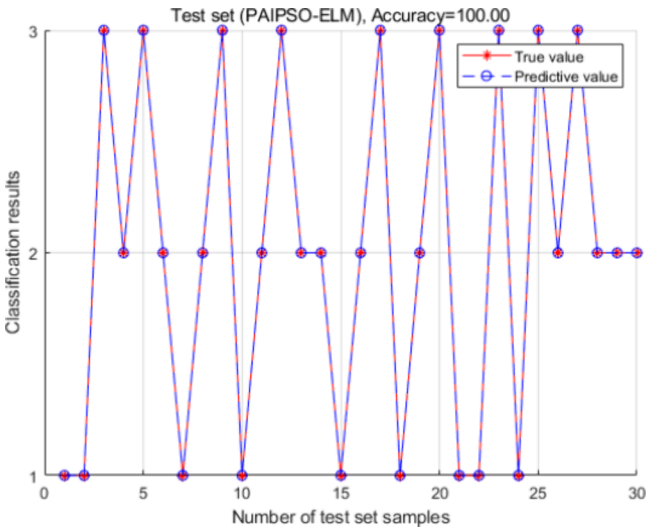


Figure 9. PAIPSO-ELM classification model results.

PAIPSO-ELM was compared with other machine learning algorithms such as BP, SVM, Random Forest (RF), and ELM. Each experiment was conducted ten times, and the average classification accuracy of PAIPSO-ELM and the other models was compared. The results are shown in Table 6.

From the above experimental results, it can be seen that PAIPSO-ELM is easier to find the optimal weights and bias than PSO-ELM, so the accuracy of classification is higher, which verifies that the performance of PAIPSO-ELM classification model is more excellent.

Table 6. Comparison of Classification Results of Different Algorithms

algorithm	average accuracy (%)
BP	84.99
SVM	83.67
RF	85.33
ELM	86.32
PSO-ELM	96
PAIPSO-ELM	98

4. CONCLUSIONS

Coal is a crucial energy resource, making its classification and utilization highly significant. This paper proposes an efficient coal classification method using reflectance spectra and machine learning. First, 100 coal samples were collected using the SVC HR-1024 spectrometer, with anthracite, bituminous coal, and lignite labeled as 1, 2, and 3. PCA was then applied to reduce data dimensionality, followed by normalization. The data set was proportionally divided into training and test sets using the holdout method.

Subsequently, ELM parameters were optimized using PSO, and further refined by introducing a decreasing inertia factor and position mutation particles to reduce the risk of local optima, resulting in the PAIPSO-ELM model. Experimental results demonstrate that the proposed PAIPSO-ELM method offers improved accuracy and generalization in coal spectral classification, with faster processing speeds, rendering it appropriate for real-time or large-scale applications. Given the strong generalization capability of SLFN, this method can also be extended to classify other types of coal and minerals. Proposed method utilizes NIRS, which primarily reflects the surface characteristics of the material. Using NIRS as the input for the model may result in a limited representation of the physicochemical properties of the ore. For more detailed sorting of commonly used ores in the future, a fusion of NIRS and LIBS spectra can be employed to represent the physicochemical properties of the ore from different perspectives.

AUTHOR INFORMATION

Corresponding Author

Dong Xiao − *School of Information Science and Engineering and Liaoning Key Laboratory of Intelligent Diagnosis and Safety for Metallurgical Industry, Northeastern University, 110819 Shenyang, China;* orcid.org/0000-0002-0401-6654; Email: xiaodong@ise.neu.edu.cn

Authors

Yiyang Wang − *School of Electrical and Automation Engineering, Liaoning Institute of Science and Technology, 117004 Benxi, China*

Boyan Li − *School of Information Science and Engineering and Liaoning Key Laboratory of Intelligent Diagnosis and Safety for Metallurgical Industry, Northeastern University, 110819 Shenyang, China*

Haoyang Li − *School of Information Science and Engineering and Liaoning Key Laboratory of Intelligent Diagnosis and Safety for Metallurgical Industry, Northeastern University, 110819 Shenyang, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.4c08020

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported in part by the National Key R&D Program of China under Grant 2022YFB2703304; in part by

https://doi.org/10.1021/acsomega.4c08020
ACS Omega 2024, 9, 47756−47764

the National Natural Science Foundation of China under Grant 52074064; in part by the Natural Science Foundation of Science and Technology Department of Liaoning Province under Grant 2021-BS-054; in part by the Fundamental Research Funds for the Central Universities under Grant N2404013, N2404015.

REFERENCES

(1) Chen, P. Study on Integrated Classification System for Chinese Coal. *Fuel Process. Technol.* **2000**, *62*, 77−87.
(2) Zou, L.; Yu, X.; Li, M.; Lei, M.; Yu, H. Nondestructive Identification of Coal and Gangue via Near-Infrared Spectroscopy Based on Improved Broad Learning. *IEEE Trans. Instrum. Meas.* **2020**, *69* (10), 1−8052.
(3) Huang, G.; Yuan, L.; Shi, W.; Chen, X.; Chen, X. Using One-Class Autoencoder for Adulteration Detection of Milk Powder by Infrared Spectrum. *Food Chem.* **2022**, *372*, No. 131219.
(4) Zhang, D.; Liang, P.; Chen, W.; Tang, Z.; Li, C.; Xiao, K.; Jin, S.; Ni, D.; Yu, Z. Rapid field trace detection of pesticide residue in food based on surface-enhanced Raman spectroscopy. *Microchim. Acta* **2021**, *188* (11), 370.
(5) Jiang, S.; He, H.; Ma, H.; Chen, F.; Xu, B.; Liu, H.; Zhu, M.; Kang, Z.; Zhao, S. Quick Assessment of Chicken Spoilage Based on Hyperspectral NIR Spectra Combined with Partial Least Squares Regression. *Int. J. Agric. Biol. Eng.* **2021**, *14* (1), 243−250.
(6) Xiao, D.; Huang, J.; Li, J.; Fu, Y.; Mao, Y.; Li, Z.; Bao, N. Inversion study of soil organic matter content based on reflectance spectroscopy and the improved hybrid extreme learning machine. *Infrared Phys. Technol.* **2023**, *128*, No. 104488.
(7) Xiao, D.; Yan, Z.; Li, J.; Fu, Y.; Li, Z. Rapid proximate analysis of coal based on reflectance spectroscopy and deep learning. *Spectrochi. Acta A, Mol. Biomol. Spectrosc.* **2022**, *287* (Pt 2), No. 122042.
(8) Xiao, D.; Xie, H.; Fu, Y.; Li, F. Mine Reclamation Based on Remote Sensing Information and Error Compensation Extreme Learning Machine. *Spectrosc. Lett.* **2021**, *54* (2), 151−164.
(9) Mao, Y.; Xiao, D.; Cheng, J.; et al. Multigrades classification model of magnesite ore based on SAE and ELM. *Journal of Sensors* **2017**, *2017*, 1−9.
(10) Hu, R.; Wang, Y.; Yang, M.; et al. Improved analysis of inorganic coal properties based on near-infrared reflectance spectroscopy. *Analytical Methods* **2015**, *7* (12), 5282−5288.
(11) Yao, S.; Qin, H.; Wang, Q.; Lu, Z.; Yao, X.; Yu, Z.; Chen, X.; Zhang, L.; Lu, J. Optimizing Analysis of Coal Property Using Laser Induced Breakdown and near-Infrared Reflectance Spectroscopies. *Spectrochim. Acta, Part A* **2020**, *239*, No. 118492.
(12) Lei, M.; Rao, Z.; Li, M.; et al. Identification of Coal Geographical Origin Using Near Infrared Sensor Based on Broad Learning. *Applied Sciences* **2019**, *9* (6), 1111.
(13) Xiao, D.; Yan, Z.; et al. Coal Identification Based on Reflection Spectroscopy and Deep Learning: Paving the Way for Efficient Coal Combustion and Pyrolysis. *ACS omega.* **2022**, *7* (27), 23919−23928.
(14) Huang, G.; Zhu, Q.; Siew, C. Extreme Learning Machine: Theory and Applications. *Neurocomputing* **2006**, *70* (1−3), 489−501.
(15) Cao, J.; Lin, Z.; Huang, G. B. Self-adaptive evolutionary extreme learning machine. *Neural processing letters* **2012**, *36* (3), 285−305.
(16) Deng, W.; Zheng, Q.; Chen, L. Research on extreme learning of neural networks. *Chinese J. Comput.* **2010**, *33* (2), 279−287.
(17) Gu, Y.; Liu, J.; Chen, Y. TOSELM: timeliness online sequential extreme learning machine. *Neurocomputing* **2014**, *128* (27), 119−127.
(18) Huang, G.; Bai, Z.; Kasun, L.; et al. Local receptive fields based extreme learning machine. *IEEE Computational intelligence magazine* **2015**, *10* (2), 18−29.
(19) Kutlu, Y.; Yayik, A.; Yıldırım, E.; Yıldırım, S. LU triangularization extreme learning machine in EEG cognitive task classification. *Neural Computation and Applications.* **2019**, *31* (4), 1117−1126.
(20) Luo, X.; Xu, Y.; Wang, W.; et al. Towards enhancing stacked extreme learning machine with sparse autoencoder by correntropy. *Journal of the franklin institute* **2018**, *355* (4), 1945−1966.
(21) Bai, Z.; Huang, G.; Wang, D.; et al. Sparse extreme learning machine for classification. *IEEE transactions on cybernetics* **2014**, *44* (10), 1858−1870.
(22) Pearson, K. On lines and planes of closest fit to systems of points in space. *London, Edinburgh, and Dublin philosophical magazine and journal of science* **1901**, *2* (11), 559−572.
(23) Kennedy, J.; Eberhart, R. Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks*; IEEE, 1995; Vol. 4, pp. 1942−1948.

https://doi.org/10.1021/acsomega.4c08020
ACS Omega 2024, 9, 47756−47764