ACS OMEGA

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

Article

Coal Identification Based on Reflection Spectroscopy and Deep Learning: Paving the Way for Efficient Coal Combustion and Pyrolysis

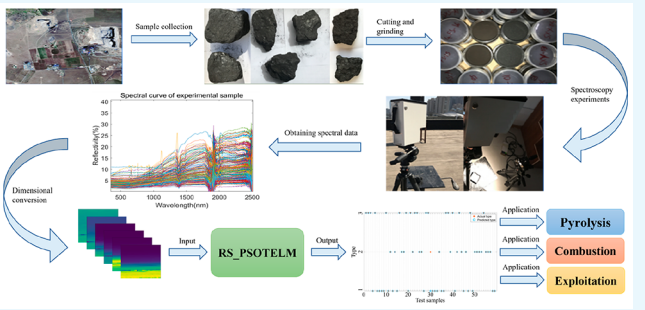
Dong Xiao,* Zelin Yan, Jian Li, Yanhua Fu, Zhenni Li, and Boyan Li

Cite This: *ACS Omega* 2022, 7, 23919−23928

Read Online

ACCESS | Metrics & More | Article Recommendations

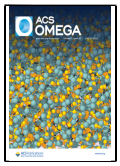
ABSTRACT: Coal plays an indispensable role in the world's energy structure. Coal converts chemical energy into energy such as electricity, heat, and internal energy through combustion. To realize the energy conversion of coal more efficiently, coal needs to be identified during the stages of mining, combustion, and pyrolysis. On this basis, different categories of coal are used according to industrial needs, or different pyrolysis processes are selected according to the category of coal. This paper proposes an approach combining deep learning with reflection spectroscopy for rapid coal identification in mining, combustion, and pyrolysis scenarios. First, spectral data of different coal samples were collected in the field and these spectral data were preprocessed. Then, an identification model combining a multiscale convolutional neural network (CNN) and an extreme learning machine (ELM), named RS_PSOTELM, is proposed. The effective features in the spectral data are extracted by the CNN, and the feature classification is realized utilizing the ELM. To enhance the identification performance of the model, we utilize a particle swarm optimization algorithm to optimize the parameters of the ELM. Experimental results show that RS_PSOTELM achieves 98.3% accuracy on the coal identification task and is able to identify coal quickly and accurately, providing a low-cost, efficient, and reliable approach for coal identification during the mining and application phases, as well as paving the way for efficient combustion and pyrolysis of coal.



1. INTRODUCTION

Coal is an essential energy source and industrial raw material and is the energy source with the largest number of varieties and the longest history of exploitation in the fossil energy family, which has made an extremely significant contribution to the development of human civilization. Coal can be classified into three main categories according to its genesis, composition, and organization: anthracite, bituminous coal, and lignite.[1] As China is the world's largest coal producer and consumer, the accurate identification of coal is of great significance to improve the efficiency of coal utilization. Chen[1] took 1100 coal samples from major coal mines in China and performed proximate, final, coal petrographic and calorific analyses, which showed that the properties of different categories of coal varied widely. Taking combustion as an example, anthracite is highly carbonized and has a high ignition point, which makes it difficult to ensure the stability of combustion, so it is not suitable for power plant boilers; bituminous coal is moderately carbonized and is mainly utilized as a boiler fuel and raw material for coking; lignite is the least carbonized and easily weathered, so it is mainly used as a boiler fuel for power plants near its origin. Coal grading conversion is one of the means for efficient, clean utilization and sustainable development of coal, while pyrolysis is the initial stage of coal conversion. In the coal pyrolysis stage, the lignite mainly adopts the method of low-temperature dry distillation, while the bituminous coal mainly adopts the method of high-temperature dry distillation, and the products mainly include gas, tar, semicoke, or coke. The products of pyrolysis are closely related to the category of coal and the pyrolysis process. Different pyrolysis processes are applicable to different categories of coal, and the ingredients of different categories of coal vary significantly, resulting in different contents and quality of pyrolysis products. As can be seen, applications such as combustion and pyrolysis of coal are closely related to the category of coal, so accurate identification of coal is particularly necessary. The current approaches for

Received: April 29, 2022
Accepted: June 17, 2022
Published: June 29, 2022



https://doi.org/10.1021/acsomega.2c02665
ACS Omega 2022, 7, 23919−23928

coal identification mainly include the artificial empirical method, weighing method, and chemical analysis method. Among them, the artificial empirical method and the weighing method have the disadvantage of unreliable accuracy, and the chemical analysis method has high accuracy but has the shortcomings of a long detection cycle and high detection cost.[2] Therefore, proposing a low-cost, high-efficiency, and high-reliability coal identification method is of great significance for developing the economy, protecting the ecological environment, and achieving sustainable development goals.

The spectral analysis technique has advantages such as high efficiency and low expense. The timely measurement information it provides can provide the basis for the identification of coal before combustion and pyrolysis, which can significantly improve efficiency and reduce the consumption of time and resources. In recent years, spectral analysis techniques have been extensively applied in the fields of soil, minerals, food, and so on.[3−8] For coal, spectroscopy is used to analyze and identify the category, composition, structure, and other characteristics of coal. Zhang et al.[9] combined the laser-induced breakdown spectra (LIBS) and independent component analysis-wavelet neural network for coal ash classification and achieved excellent performance, promoting the recovery and reuse of metallurgical waste. Lei et al.[10] proposed a coal classification model combining generalized learning and the particle swarm optimization (PSO) algorithm, which overcomes the problems of data redundancy in the original spectral data and obtained 97.05% accuracy. Zhang et al.[11] utilized the support vector machine (SVM) optimized by the genetic algorithm to categorize coal samples and utilized partial least-squares (PLS) regression to model each category of coal samples to obtain precise measurements of ash, volatile content, and calorific value. Yan et al.[12] utilized an approach that combined wavelet transform and mean impact value to abstract valuable information from LIBS, which effectively reduced computation time and improved model performance. Begum et al.[13] first improved the signal-to-noise ratio of the spectrum by preprocessing and then used the least-squares method, random forest (RF), and extreme gradient boosting to predict the coal composition and obtained the best accuracy for different compositions, respectively. Yao et al.[14] used LIBS and near-infrared spectroscopy (NIRS) to optimize coal characteristics and established an analytical model and achieved accurate predictions for volatiles, ash, and moisture content. Sun et al.[15] predicted the cutoff value of T-2 spectra representing the pore structure of coal by a back-propagation (BP) network model and obtained a better prediction, which provides a reliable approach for coal structure detection.

Because of the high dimension, strong correlation, and noise interference of spectral data, the current methods mainly use preprocessing such as principal component analysis to reduce dimensionality and denoise the spectral data and use the PLS or SVM algorithm for modeling. If the improper preprocessing method is adopted, it is not conducive to the improvement of the model performance, but also leads to the unreliable prediction accuracy of the subsequent model. Deep learning (DL) has been extensively applied in various fields.[16−18] DL is capable of building end-to-end analytical models without relying on preprocessing.[19,20] Xiao et al.[21] extracted spectral data features through a deep belief network and then constructed the coal analysis model using a derivative function with a regularization two-layer extreme learning machine (DF-RTELM) algorithm. Le et al.[22] proposed a method combining the convolutional neural network (CNN), extreme learning machine (ELM), and visible and near-infrared spectroscopy for coal identification, achieving 96.51% classification accuracy, demonstrating the effectiveness of the feature extractor and classifier in the CNN-ELM model. Azimi et al.[23] presented an approach for beta-gamma coincidence radioxenon spectra using the CNN and achieved high classification accuracy. Zhang et al.[24] integrated the one-dimensional CNN and long short-term memory networks for the detection of soil water content. Machado et al.[25] utilized a deep neural network to remove the noise in the spectrum, and the spectral resolution was effectively improved.

The ELM is a model with a structure of one-hidden layer feedforward neural network.[26] The model runs rapidly and has a simple structure, and a lot of scholars have applied and improved it.[27−29] In spectroscopy, the ELM is extensively applied in the construction of analytical and classification models. Mao et al.[30] established a multilayer ELM model for coal identification, and the experimental results show that it can achieve an identification accuracy of 92.25% while taking into account the speed. Yan et al.[31] proposed a method combining the kernel-based ELM (K-ELM) with LIBS for detecting carbon and sulfur content in coal and finally achieved an R2 of 0.994 and an RMSE of 0.3762%. Because carbon content is an important basis for coal identification, accurate detection of carbon content is also helpful for coal identification and utilization. Chen et al.[32] proposed the ensemble window ELM (EWELM) algorithm to detect the content of impurities in drugs. The experiments show that the EWELM is better than PLS and the full-spectrum ELM algorithm. Liang et al.[33] combined laser-induced breakdown spectroscopy and a particle swarm-optimized K-ELM algorithm to classify six species of Salvia miltiorrhiza in different regions. The experimental results show that the identification accuracy of this classification model is better than that of particle swarm optimization-least-squares support vector machines (PSO-LSSVM) and particle swarm optimization-random forest (PSO-RF) models, with an accuracy of 94.87%. Chen et al.[34] proposed an ensemble ELM algorithm, which can achieve multivariate calibration of NIRS, and the experimental results are better than those of the PLS algorithm.

The ELM is an effective spectral modeling method, but spectral data are generally highly interrelated. Modeling is often poor while the ELM is applied directly on the basis of raw spectral data. Compared with the ELM, DL is more capable of processing complex data. This paper combines the CNN and ELM, leverages the advantages of both algorithms, and establishes a high-performance coal identification model, which is named RS_PSOTELM. To further strengthen the predictive capability and stability of the model, we optimize the classifier parameters using the PSO algorithm. Finally, comparative experiments demonstrate the excellent performance of the RS_PSOTELM model.

2. THEORY AND METHODOLOGY

2.1. Application of DL in Coal Identification. Because different categories of coal have different products during combustion and pyrolysis, the application scenarios of different categories of coal are also different. To utilize the chemical energy in coal efficiently, it is necessary to identify the coal. In this paper, we apply a DL approach to identify coal spectral data, and the detailed application process is shown in Figure 1.

https://doi.org/10.1021/acsomega.2c02665
ACS Omega 2022, 7, 23919−23928

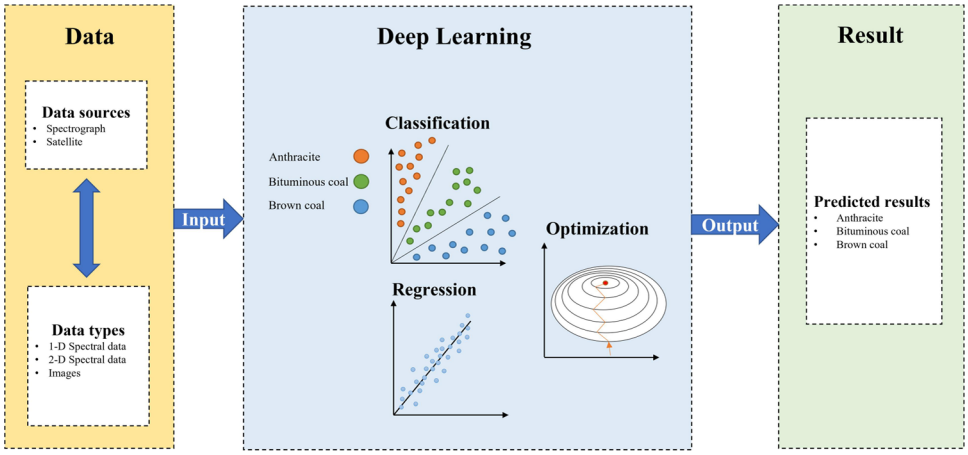


Figure 1. Application process of DL.

The first part is the data part. High-quality data are the basis for the excellent performance of the identification model. The forms of data are mainly divided into reflectance spectra, hyperspectral, and remote sensing images.

Reflectance spectra are mainly measured by spectrometers, and hyperspectral and remote sensing images are obtained by satellites. The second part is the DL part. The obtained spectral data are divided randomly in the ratio of 3:1, and 75% of the samples are regarded as the training set and the rest as the test set.

Among them, the training set is used for model training, and through multiple iterations of training, the model is made to learn the features in the spectral data. At the same time, using the preset labels to reversely correct the model parameters, continuously optimize the model parameters, and finally, build a model for coal identification. The test set is utilized to evaluate the identification model obtained from training. The data from the test set are fed into the model, and the model predicts an output based on the parameters obtained from previous training. The performance evaluation of the model can be completed by comparing the output value with the preset labels. The third part is the evaluation of the predicted results. In classification problems, accuracy is often used as an evaluation indicator, while in regression problems, RMSE and R2 indicators are more inclined to measure the performance of the model. Coal identification is a classification problem, and the final output includes anthracite, bituminous, and lignite. Therefore, this paper adopted accuracy as the evaluation index.

2.2. Convolutional Neural Network. The CNN is the most effective approach to extract features in the image processing domain currently.[35,36] Many classical networks, for instance, FCN,[37] DenseNet,[38] and U-Net,[39] utilized the CNN for image feature extraction and achieved satisfactory results. The structure of a typical CNN is composed mainly of convolutional, pooling, and fully connected (FC) layers.[40,41]

A CNN usually includes multiple convolutional layers, and feature extraction is achieved by utilizing the sliding of convolutional kernels in each convolutional layer. Assuming that the spectral data size is 4 × 4, the convolution kernel is 3 × 3, and the stride is 1, the convolution process is shown in Figure 2a. After convolution processing, the dimension of the spectral data is reduced from 4 × 4 to 2 × 2, the data are more compact, and the information in the original spectral data is included.

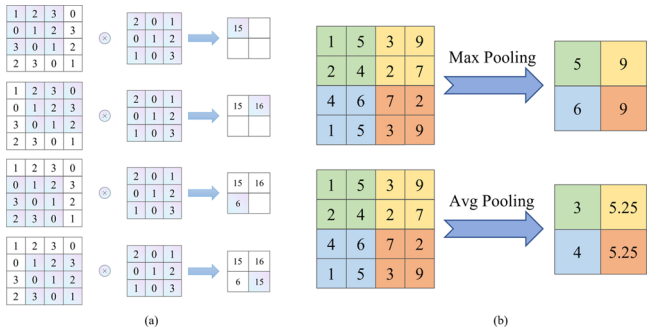


Figure 2. Process of (a) convolution and (b) pooling.

The pooling layer in the CNN mainly includes two methods: maximum pooling and average pooling. The pooling process is shown in Figure 2b. In the max pooling operation, the network selects the largest number in the filter as the output. In the average pooling operation, the network selects the mean of all numbers within the filter range as the output. After processing by the pooling layer, not only the dimensionality reduction and denoising of the data can be achieved, but also overfitting can be avoided.

The FC layer in the CNN is usually located in the last layer of the network. After processing by convolution, pooling, and activation function, the features in the original data are mapped to the hidden layer feature space. The FC layer acts as a classifier, mapping the previously learned features to the sample label space, and then obtaining the final classification result. In practical applications, the FC layer can be implemented by convolutions of different sizes according to the specific composition of the previous layer.

2.3. Residual Network. As the application scenarios of models become more and more complex, the requirements for model performance are gradually increasing. To obtain more refined features, the depth of the network is also deepening. When the network reaches a certain depth, simply adding convolutional layers does not reduce the training error. As the depth of the network increases, there may be problems such as overfitting, vanishing gradients, exploding gradients, and degradation. In addition, model deepening may also cause a decrease in the learning capability of certain shallow networks, thus limiting the learning of deeper networks. To address this problem, He et al.[42] proposed a residual network, which has few parameters, high efficiency, and enhanced feature transfer

https://doi.org/10.1021/acsomega.2c02665
ACS Omega 2022, 7, 23919−23928

between layers. A comparison of the residual network and the normal network structure is shown in Figure 3.

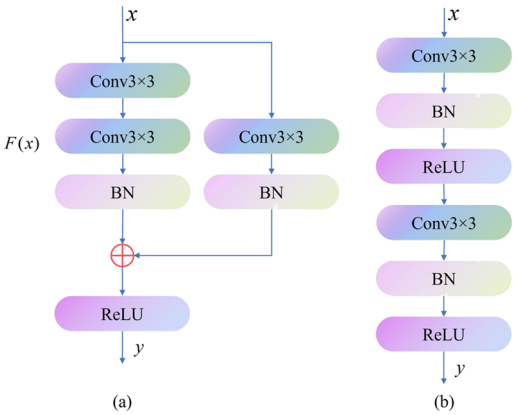


Figure 3. Structure comparison of (a) residual network and (b) ordinary network.

By adding a constant mapping branch to the original network, the residual network is able to map the input to the next layer through a function, while transferring the input directly to the next layer through a constant mapping. Finally, the two sets of features are summed element by element for integration, ensuring that the model performance does not degrade as the network depth increases. The connection between the input and output of the residual network is represented as eq 1.

$$y = f(F(x) + \omega x) \tag{1}$$

where $F(x)$ represents the mapping relationship between residual units; ω represents the linear mapping used for dimension matching; x represents the input; y represents the output; $f(x)$ represents the activation function.

2.4. Extreme Learning Machine. The ELM is a feedforward neural network, and the network structure is shown in Figure 4a. The model randomly generates weights and thresholds between the input layer and the first hidden layer and obtains the optimal output matrix through the least-squares method after obtaining the output value. Because the model does not require back-propagation to correct parameters, the computation speed is significantly improved. The calculation steps are as follows.

(1) Multiply the input matrix by the weight matrix;
(2) Add to the bias matrix;
(3) Calculate the activation function;
(4) Calculate the output value;
(5) Calculate the output matrix using the least-squares method.

The two-layer extreme learning machine (TELM) algorithm adds a second hidden layer to the ELM algorithm. Compared with the calculation steps of the ELM, the calculation steps of the TELM additionally need to combine the inverse function and Moore Penrose matrix to calculate the output weight of the second hidden layer. The network structure is shown in Figure 4b.

2.5. Particle Swarm Optimization. The PSO algorithm is proposed inspired by the foraging behavior of bird flocks. It has the advantages of rapid convergence, few parameters, and convenient realization. The effect of the algorithm is shown in Figure 5. The optimal position is at the yellow circle, and after

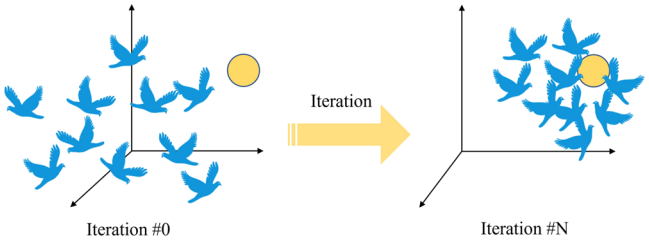


Figure 5. Effect of the PSO algorithm.

N iterations, the flock tends to the optimal position from the scattered state. In the PSO algorithm, only two properties, velocity and position, are given to all particles. Velocity represents the speed of movement, and position represents the movement orientation. The algorithm mainly includes the following four parts.

2.5.1. Initialization. First, the parameters in the algorithm are initialized. The particle group size is set to 40 and the number of maximum iterations to 50, the number of objective function arguments to 2, the speed interval to $[-1, 1]$, and the search space to $[-1, 1]$. All particles are randomly given an initial velocity and position.

2.5.2. Finding Individual Extremum and Global Optimal Value. The individual extremums are the optimal values found for each particle. According to the fitness function, a globally optimal value is found from these optimal values. Velocity and position are updated by comparison with the previous global optimal value.

2.5.3. Update Velocity and Position. The velocity of the particle is updated based on eq 2.

$$V_{id} = wV_{id} + C_1\text{random}(0, 1)(P_{id} - X_{id}) + C_2\text{random}(0, 1)(P_{gd} + X_{id}) \tag{2}$$

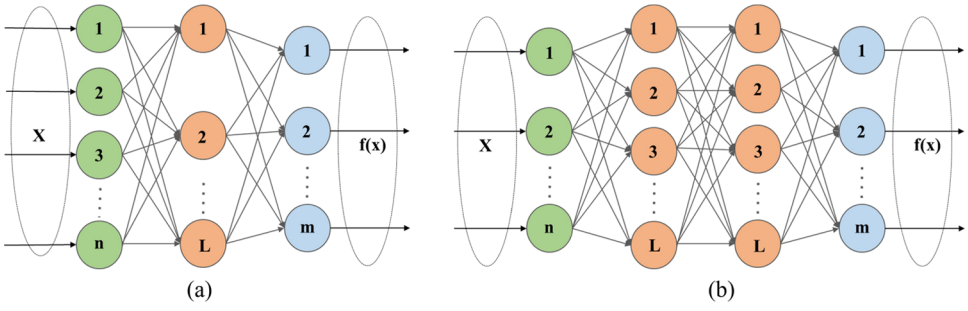


Figure 4. Network structure of (a) ELM and (b) TELM.

https://doi.org/10.1021/acsomega.2c02665
ACS Omega 2022, 7, 23919−23928

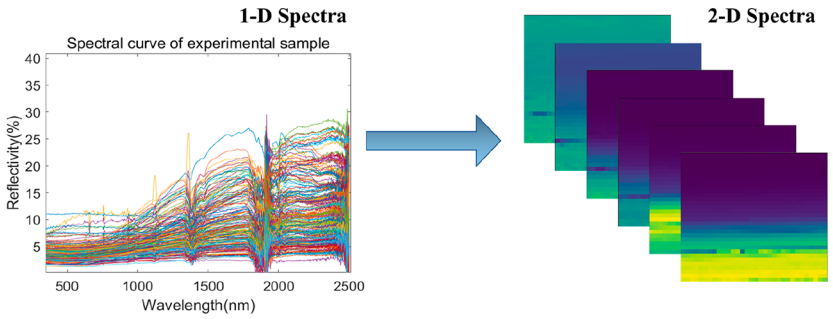


Figure 6. Conversion of spectral dimensions.

The position of the particle is updated based on eq 3.

$$X_{id} = X_{id} + V_{id} \quad (3)$$

where w is called the factor of inertia, C_1 and C_2 are called the acceleration constants, generally taken as $C_1 = C_2 \in [0,4]$, random $(0,1)$ means a random value on the range $[0,1]$, P_{id} denotes the individual extremum of the i-th variable in the d-th dimension, and P_{gd} denotes the d-th dimension of the global optimal value.

2.5.4. Termination Condition Setting. When the set amount of iterations is reached or the fitness value satisfies the difference requirement, the iteration is terminated and the optimal value is output.

2.6. Collection and Processing of Spectra. In this paper, 71 samples of anthracite, 80 samples of bituminous, and 58 samples of lignite were collected, each sample containing 973-dimensional spectral features. Spectral data are usually processed in 1-D form. 1-D spectral information mainly provides peak characteristics of different bands, which can be analyzed in a limited space. After transforming the spectrum from 1-D to 2-D, the CNN can not only obtain deeper features in the spectral data but also realize the feature fusion of adjacent bands.

To facilitate 2-D processing, some bands are randomly selected for linear combination, and the dimension is extended from 973 to 1024. Although the dimension of the spectrum is increased, no new spectral information is introduced and will not affect the identification results. After obtaining the 1024-dimensional spectral data, each sample is arranged into a 32 × 32 matrix according to the "S" shape, and then the 2-D spectral data can be obtained. As shown in Figure 6, after the spectral data are converted from 1-D to 2-D, the texture features can be better reflected, which helps the CNN extract spectral features.

2.7. Identifying Model. The TELM algorithm performs effectively on multiclassification tasks but has the following drawbacks. First, the TELM algorithm is not effective in processing complex and high-dimensional data, and the identification accuracy is hard to meet the requirements. Second, both the weight and bias of the first hidden layer in the TELM algorithm are random values, which makes the algorithm effect fluctuate greatly and not stable enough.

Aiming at the first shortcoming of the TELM algorithm, this paper proposes a new CNN model with structural parameters such as Tables 1 and 2. The model consists of eight convolutional layers, eight normalization layers, and three pooling layers combined with different categories of residual connections. It effectively realizes the multiscale fusion of spectral features, while avoiding the common problems of overfitting, gradient disappearance, gradient explosion, and degradation in deep networks.

Table 1. Structural Parameters of the CNN Branch Model

layer	type	filter size	filter number	step size	activation function
1	input				
2	convolution	1	1	16	ReLU
3	batch normalization				
4	max pooling	2			

Table 2. Structural Parameters of the CNN Backbone Model

layer	type	filter size	filter number	step size	activation function
1	input				
2	convolution	1	32	1	
3	batch normalization				
4	convolution	3	8	1	
5	batch normalization				
6	convolution	3	8	1	
7	batch normalization				
8	convolution	3	8	1	
9	batch normalization				
10	convolution	1	1	1	ReLU
11	batch normalization				
12	convolution	5	16	1	ReLU
13	batch normalization				
14	max pooling	2			
15	convolution	5	1	1	ReLU
16	batch normalization				
17	max pooling	2			

Aiming at the second disadvantage of the TELM algorithm, this paper applies the PSO algorithm to find the optimal solution for the first hidden layer weight matrix and bias vector in the TELM algorithm. A PSOTELM algorithm based on PSO and TELM is established to achieve a more stable and accurate classification effect. The PSOTELM algorithm flow is shown in Figure 7.

The FC layer usually plays the role of a classifier. The deep spectral features obtained after the convolution and pooling processing are sent to the FC layer, and the classification function is realized after processing by the FC layer. In the CNN training process, like the parameters of the convolution kernel and the pooling kernel, the parameters of the FC layer

https://doi.org/10.1021/acsomega.2c02665
ACS Omega 2022, 7, 23919−23928

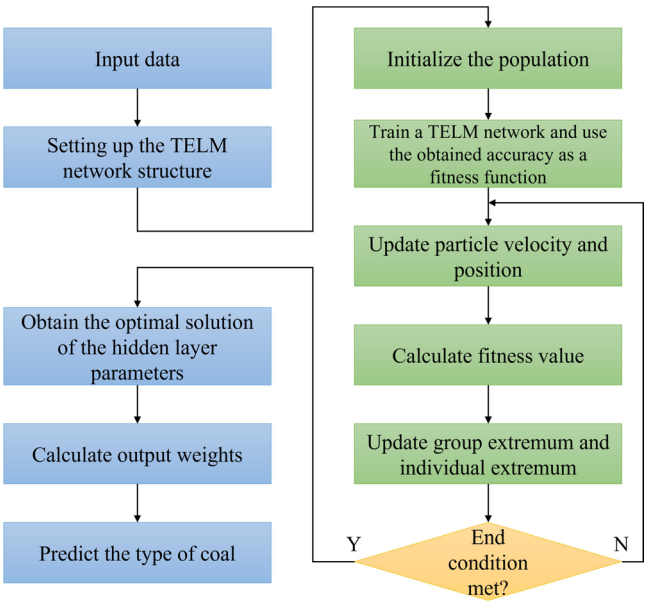


Figure 7. Flow of the PSOTELM algorithm.

also need to obtain the gradient through the back-propagation algorithm and use the gradient descent method to achieve the minimum loss. When the network model is large, the scale of parameters such as weights and biases of the FC layer also becomes larger, leading to a larger consumption of computational resources. In addition, the classification performance of PSOTELM is better than that of the FC layer, and PSOTELM can be applied as the classifier of the CNN instead of the FC layer to obtain better classification results. To address the problems of FC layers, this paper adopts the PSOTELM algorithm to replace FC layers as classifiers of CNN models, which improves the identification accuracy while saving computational resources. Figure 8 shows the overall structure of the identification model proposed in this paper.

3. RESULTS AND DISCUSSION

The model in this paper is built using the Python programming language under the Pytorch1.7 framework in Windows 11 and combined with MATLAB 2018b to visualize the training process and identification results. In this paper, all experimental networks are trained on Intel Core i7 processors with 16-GB RAM. The graphics card model used is NVIDIA RTX 3060. We set the batch size to 8, the max epoch to 128, and the learning rate to 0.001.

3.1. Experimental Result. In this paper, 209 samples in three categories are collected, and the samples are randomly divided into the training set and test set (150 samples are applied for training and 59 samples are applied for testing). Label anthracite as 1, bituminous as 2, and lignite as 3. Figure 9 shows the training process of the CNN. The line graph composed of red dots shows the accuracy change of the network, where each red dot represents the identification accuracy of an epoch, and the training error is represented by vertical line segments. It can be seen that in the first 90 iterations, there are large fluctuations in both accuracy and error. After 90 epochs, the accuracy is stable at around 99% with almost no fluctuation, and the fluctuation of the error is also significantly reduced and tends to 0. The identification results of the overall model are shown in Figure 10. Anthracite and lignite were all correctly identified, but only one sample of

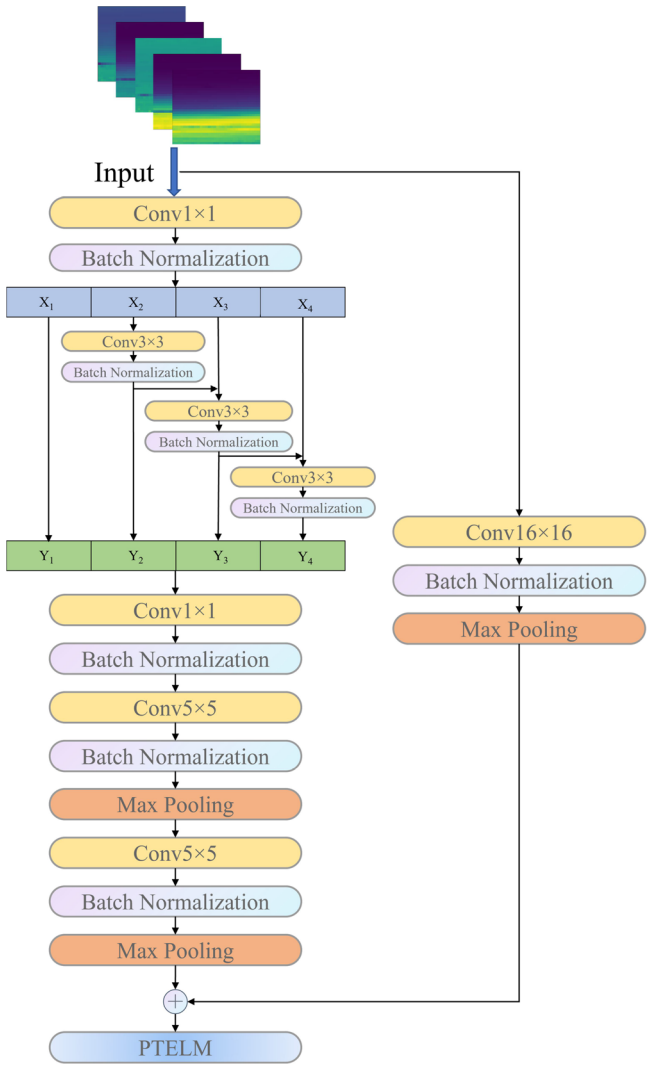


Figure 8. Structure of the RS_PSOTELM model.

bituminous was identified incorrectly, and the overall accuracy rate reached 98.3%. It can fully meet the identification requirements in the process of coal mining, combustion, and pyrolysis to ensure efficient utilization of coal resources.

3.2. Algorithm Comparison and Evaluation. To verify the effectiveness of the RS_PSOTELM model proposed in this paper, we selected several models for comparison, including ELM, TELM, PSO_TELM, BP, SVM, RF, and RS_Net. Among them, the structures of RS_Net and RS_PSOTELM are similar, and the only difference is that RS_Net adopts the FC layer as the classifier. In this paper, the feature extraction capability of CNN is demonstrated by comparing RS_PSOTELM with ELM, TELM, and PSO_TELM. By comparing with BP, SVM, and RF, we demonstrate that RS_PSOTELM outperforms the commonly applied classification methods in spectral processing; by comparing with RS_Net, we demonstrate the superior performance of PSOTELM in spectral feature identification and classification. Table 3 summarizes the specific identification accuracy of the different models. Compared to the other models, RS_PSOTELM achieves the best results in the coal identification task with an accuracy of 98.3%. The identification results of different models on anthracite, bituminous coal, and lignite respectively are depicted in Figure 11. In the anthracite identification task, all

https://doi.org/10.1021/acsomega.2c02665
ACS Omega 2022, 7, 23919−23928

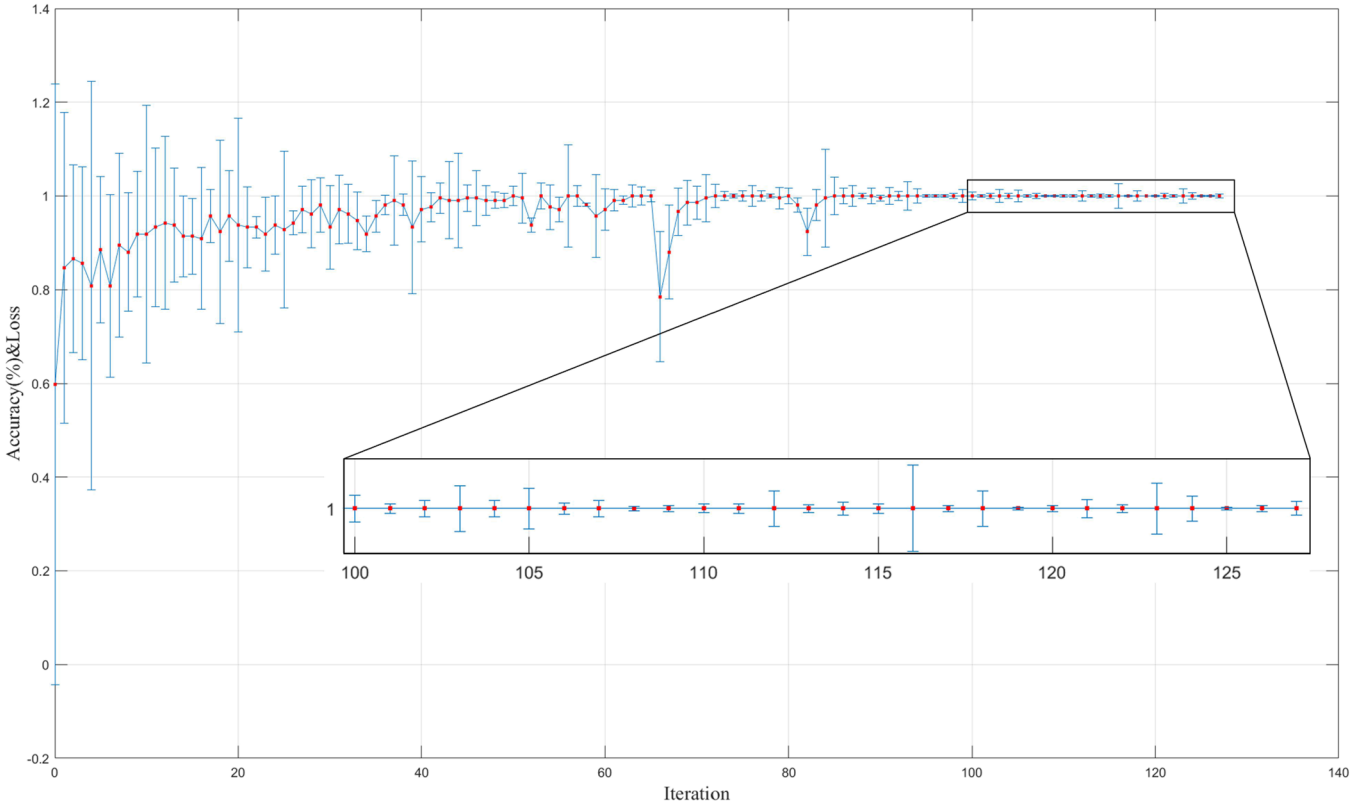


Figure 9. Training process of the model.

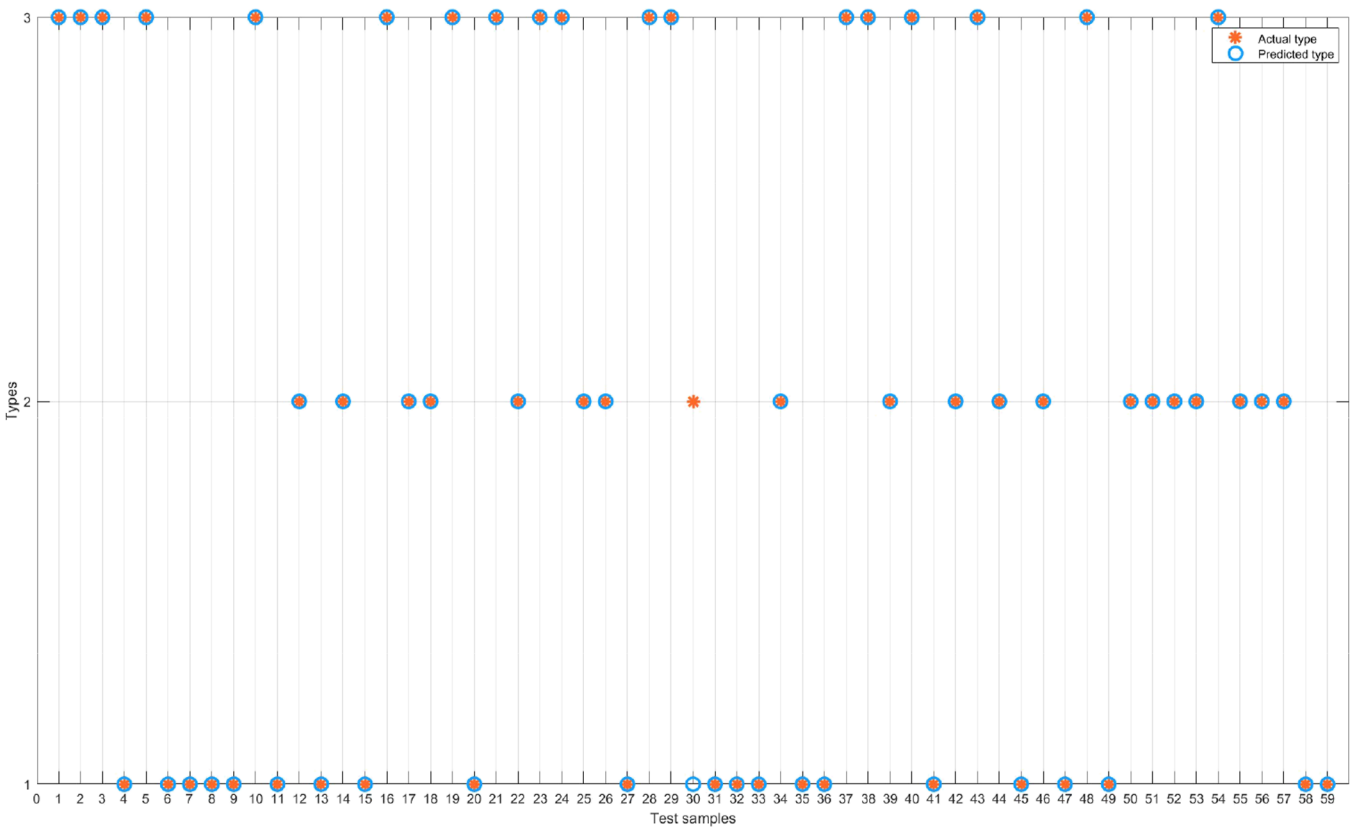


Figure 10. Identification results of the RS_PSOTELM model.

models have wrongly identified samples, and BP has the largest number of wrongly identified samples, which is 14. In the identification task of bituminous coal, PSO_TELM and BP have no wrongly identified samples, and obtained the best

https://doi.org/10.1021/acsomega.2c02665
ACS Omega 2022, 7, 23919−23928

Table 3. Comparison of Identification Results of Different Models

models	accuracy (%)
ELM	86.4
TELM	89.8
PSO_TELM	94.9
BP	67.8
SVM	83.1
RF	88.1
RS_Net	93.2
RS_PSOTELM	**98.3**

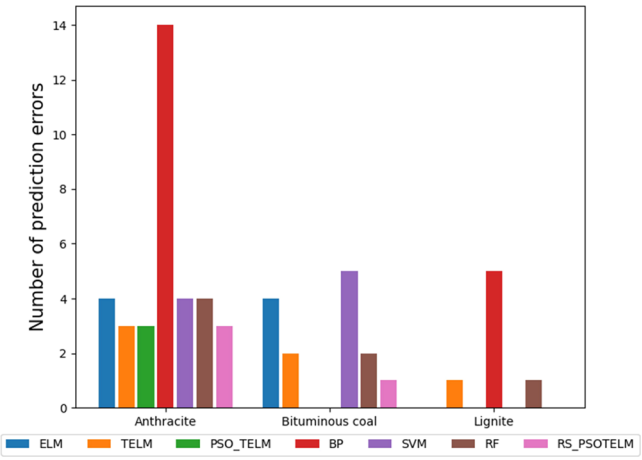


Figure 11. Number of samples misidentified by different models for anthracite, bituminous coal, and lignite.

identification results, while other models have different wrongly identified samples. In the lignite identification task, only TELM, BP, and RF have misidentified samples, of which BP has the largest number of wrongly identified samples, and other models are correct for lignite identification. Combining Table 3 and Figure 11, it can be seen that the overall performance of RS_PSOTELM is better than that of other models in the three coal identification tasks, and its excellent identification performance paves the way for the efficient utilization of coal resources.

The comparison of time cost, accuracy, and economic cost of coal identification by the chemical analysis method, manual experience method, weighing method, and RS_PSOTELM model is shown in Table 4, applied to 200 coal samples,

Table 4. Comparison of Different Coal Identification Methods

analysis methods	time (h)	accuracy (%)	costs (USD)
chemical analysis	300	100	1000
manual experience	2	60	50
weighing method	3	70	50
RS_PSOTELM	10	98.3	100

respectively. As you can see, RS_PSOTELM only costs 10 h and $100. Although the manual experience method and weighing method have lower cost, they have the disadvantage of lower accuracy. In contrast, chemical analysis methods have the highest precision but are extremely expensive and time-intensive. If the cost of chemical analysis lab equipment is included, some may cost more than $300,000. Relatively speaking, RS_PSOTELM has the advantages of high speed, high precision, and low cost.

4. CONCLUSIONS

Coal identification is a prerequisite for applications such as coal combustion and pyrolysis. For the coal identification problem in coal mining and applications, this paper proposes a coal identification method that combines DL and spectroscopy. The superior performance of the RS_PSOTELM model is demonstrated through extensive experiments, proving the superiority of the method for coal identification tasks. This provides a low-cost, efficient, and reliable identification method for coal mining, combustion, and pyrolysis processes. In industrial applications, bituminous coal is further classified into long-flame coal, gas coal, fatty coal, and so on. In the next step, we will identify different coal types within the range of bituminous coal. In addition, because both CNN and ELM have good generalization performance, the scope of future research can be not only limited to the identification of coal, but also extend the method to the identification tasks of Fe_3O_4, Fe_2O_3, and $FeCO_3$ in iron ore.

AUTHOR INFORMATION

Corresponding Author

Dong Xiao − *School of Information Science and Engineering, Northeastern University, Shenyang 110819, China;* orcid.org/0000-0002-0401-6654; Email: xiaodong@ise.neu.edu.cn

Authors

Zelin Yan − *School of Information Science and Engineering, Northeastern University, Shenyang 110819, China*

Jian Li − *Technical Service Parlor, Unit 31434 of the Chinese People's Liberation Army, Shenyang 110000, China*

Yanhua Fu − *School of JangHo Architecture, Northeastern University, Shenyang 110819, China*

Zhenni Li − *School of Information Science and Engineering, Northeastern University, Shenyang 110819, China*

Boyan Li − *School of Information Science and Engineering, Northeastern University, Shenyang 110819, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c02665

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China under Grant 52074064 and Grant 62173073, in part by the National Key Research and Development Program of China under 2020AAA0109203, in part by the Fundamental Research Funds for the Central Universities, China under Grant N2104026, Grant N2001002, and Grant N2204006, and in part by the Natural Science Foundation of Science and Technology Department of Liaoning Province under Grant 2021-BS-054.

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

https://doi.org/10.1021/acsomega.2c02665
ACS Omega 2022, 7, 23919−23928

Based on Improved Broad Learning. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 8043−8052.
(3) Schlack, T. R.; Beal, S. A.; Corriveau, E. J.; Clausen, J. L. Detection Limits of Trinitrotoluene and Ammonium Nitrate in Soil by Raman Spectroscopy. *ACS Omega* **2021**, *6*, 16316−16323.
(4) Raljević, D.; Parlov Vuković, J.; Smrečki, V.; Marinić Pajc, L.; Novak, P.; Hrenar, T.; Jednačak, T.; Konjević, L.; Pinević, B.; Gašparac, T. Machine Learning Approach for Predicting Crude Oil Stability Based on NMR Spectroscopy. *Fuel* **2021**, *305*, No. 121561.
(5) 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*, 151−164.
(6) Su, P.; Liu, S.; Min, H.; An, Y.; Yan, C.; Li, C. Accuracy Improvement on Quantitative Analysis of the Total Iron Content in Branded Iron Ores by Laser-Induced Breakdown Spectroscopy Combined with the Double Back Propagation Artificial Neural Network. *Anal. Methods* **2022**, *14*, 427−437.
(7) 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.
(8) 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*, 243−250.
(9) Zhang, T.; Yan, C.; Qi, J.; Tang, H.; Li, H. Classification and Discrimination of Coal Ash by Laser-Induced Breakdown Spectroscopy (LIBS) Coupled with Advanced Chemometric Methods. *J. Anal. At. Spectrom.* **2017**, *32*, 1960−1965.
(10) Lei, M.; Rao, Z.; Li, M.; Yu, X.; Zou, L. Identification of Coal Geographical Origin Using Near Infrared Sensor Based on Broad Learning. *Appl. Sci.* **2019**, *9*, 1111.
(11) Zhang, W.; Zhuo, Z.; Lu, P.; Tang, J.; Tang, H.; Lu, J.; Xing, T.; Wang, Y. LIBS Analysis of the Ash Content, Volatile Matter, and Calorific Value in Coal by Partial Least Squares Regression Based on Ash Classification. *J. Anal. At. Spectrom.* **2020**, *35*, 1621−1631.
(12) Yan, C.; Zhang, T.; Sun, Y.; Tang, H.; Li, H. A Hybrid Variable Selection Method Based on Wavelet Transform and Mean Impact Value for Calorific Value Determination of Coal Using Laser-Induced Breakdown Spectroscopy and Kernel Extreme Learning Machine. *Spectrochim. Acta, Part B* **2019**, *154*, 75−81.
(13) Begum, N.; Maiti, A.; Chakravarty, D.; Das, B. S. Reflectance Spectroscopy Based Rapid Determination of Coal Quality Parameters. *Fuel* **2020**, *280*, No. 118676.
(14) 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.
(15) Sun, Y.; Zhai, C.; Zhao, Y.; Xu, J.; Cong, Y.; Zheng, Y.; Tang, W. Multifractal Analysis and Neural Network Prediction of Pore Structures in Coal Reservoirs Based on NMR T_2 Spectra. *Energy Fuels* **2021**, *35*, 11306−11318.
(16) Wang, L.; Li, J.; Li, T.; Liu, H.; Wang, Y. Method Superior to Traditional Spectral Identification: FT-NIR Two-Dimensional Correlation Spectroscopy Combined with Deep Learning to Identify the Shelf Life of Fresh Phlebopus Portentosus. *ACS Omega* **2021**, *6*, 19665−19674.
(17) Urban, C. J.; Gates, K. M. Deep Learning: A Primer for Psychologists. *Psychol. Methods* **2021**, *26*, 743−773.
(18) O'Byrne, C.; Abbas, A.; Korot, E.; Keane, P. A. Automated Deep Learning in Ophthalmology: AI That Can Build AI. *Curr. Opin. Ophthalmol.* **2021**, *32*, 406−412.
(19) Acquarelli, J.; van Laarhoven, T.; Gerretzen, J.; Tran, T. N.; Buydens, L. M. C.; Marchiori, E. Convolutional Neural Networks for Vibrational Spectroscopic Data Analysis. *Anal. Chim. Acta* **2017**, *954*, 22−31.
(20) Zhang, X.; Lin, T.; Xu, J.; Luo, X.; Ying, Y. DeepSpectra: An End-to-End Deep Learning Approach for Quantitative Spectral Analysis. *Anal. Chim. Acta* **2019**, *1058*, 48−57.
(21) Xiao, D.; Le, B. T. Rapid Analysis of Coal Characteristics Based on Deep Learning and Visible-Infrared Spectroscopy. *Microchem. J.* **2020**, *157*, No. 104880.
(22) Le, B. T.; Xiao, D.; Mao, Y.; Song, L.; He, D.; Liu, S. Coal Classification Based on Visible, Near-Infrared Spectroscopy and CNN-ELM Algorithm. *Spectrosc. Spectral Anal.* **2018**, *38*, 2107−2112.
(23) Azimi, S. A.; Afarideh, H.; Chai, J.-S.; Kalinowski, M.; Gheddou, A.; Hofman, R. Classification of Radioxenon Spectra with Deep Learning Algorithm. *J. Environ. Radioact.* **2021**, *237*, No. 106718.
(24) Zhang, F.; Wu, S.; Liu, J.; Wang, C.; Guo, Z.; Xu, A.; Pan, K.; Pan, X. Predicting Soil Moisture Content over Partially Vegetation Covered Surfaces from Hyperspectral Data with Deep Learning. *Soil Sci. Soc. Am. J.* **2021**, *85*, 989−1001.
(25) Machado, L. R. P.; Silva, M. O. S.; Campos, J. L. E.; Silva, D. L.; Cançado, L. G.; Vilela Neto, O. P. Deep-learning-based Denoising Approach to Enhance Raman Spectroscopy in Mass-produced Graphene. *J. Raman Spectrosc.* **2022**, No. jrs.6317.
(26) Huang, G.-B.; Zhu, Q.-Y.; Siew, C.-K. Extreme Learning Machine: Theory and Applications. *Neurocomputing* **2006**, *70*, 489−501.
(27) Cao, J.; Zhang, K.; Yong, H.; Lai, X.; Chen, B.; Lin, Z. Extreme Learning Machine With Affine Transformation Inputs in an Activation Function. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 2093−2107.
(28) Yahia, S.; Said, S.; Zaied, M. Wavelet Extreme Learning Machine and Deep Learning for Data Classification. *Neurocomputing* **2022**, *470*, 280−289.
(29) Xiao, D.; Wan, L. Remote Sensing Inversion of Saline and Alkaline Land Based on an Improved Seagull Optimization Algorithm and the Two-Hidden-Layer Extreme Learning Machine. *Nat. Resour. Res.* **2021**, *30*, 3795−3818.
(30) Mao, Y.; Le, B. T.; Xiao, D.; He, D.; Liu, C.; Jiang, L.; Yu, Z.; Yang, F.; Liu, X. Coal Classification Method Based on Visible-Infrared Spectroscopy and an Improved Multilayer Extreme Learning Machine. *Opt. Laser Technol.* **2019**, *114*, 10−15.
(31) Yan, C.; Qi, J.; Ma, J.; Tang, H.; Zhang, T.; Li, H. Determination of Carbon and Sulfur Content in Coal by Laser Induced Breakdown Spectroscopy Combined with Kernel-Based Extreme Learning Machine. *Chemom. Intell. Lab. Syst.* **2017**, *167*, 226−231.
(32) Chen, H.; Huang, Q.; Lin, Z.; Tan, C. Detection of Adulterants in Medicinal Products by Infrared Spectroscopy and Ensemble of Window Extreme Learning Machine. *Microchem. J.* **2022**, *173*, No. 107009.
(33) Liang, J.; Yan, C.; Zhang, Y.; Zhang, T.; Zheng, X.; Li, H. Rapid Discrimination of Salvia Miltiorrhiza According to Their Geographical Regions by Laser Induced Breakdown Spectroscopy (LIBS) and Particle Swarm Optimization-Kernel Extreme Learning Machine (PSO-KELM). *Chemom. Intell. Lab. Syst.* **2020**, *197*, No. 103930.
(34) Chen, H.; Tan, C.; Lin, Z. Ensemble of Extreme Learning Machines for Multivariate Calibration of Near-Infrared Spectroscopy. *Spectrochim. Acta, Part A* **2020**, *229*, No. 117982.
(35) Wang, S.-H.; Govindaraj, V. V.; Górriz, J. M.; Zhang, X.; Zhang, Y.-D. Covid-19 Classification by FGCNet with Deep Feature Fusion from Graph Convolutional Network and Convolutional Neural Network. *Inf. Fusion* **2021**, *67*, 208−229.
(36) Yao, P.; Wu, H.; Gao, B.; Tang, J.; Zhang, Q.; Zhang, W.; Yang, J. J.; Qian, H. Fully Hardware-Implemented Memristor Convolutional Neural Network. *Nature* **2020**, *577*, 641−646.
(37) Shelhamer, E.; Long, J.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 640−651.
(38) Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K. Q. Densely Connected Convolutional Networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2017; pp 2261−2269.
(39) Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image*

https://doi.org/10.1021/acsomega.2c02665
ACS Omega 2022, 7, 23919−23928

Computing and Computer-Assisted Intervention−MICCAI 2015; 2015; Vol. *9351*, pp 234−241.
(40) Xie, H.; Yang, D.; Sun, N.; Chen, Z.; Zhang, Y. Automated Pulmonary Nodule Detection in CT Images Using Deep Convolutional Neural Networks. *Pattern Recognit.* **2019**, *85*, 109−119.
(41) Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; Chen, T. Recent Advances in Convolutional Neural Networks. *Pattern Recognit.* **2018**, *77*, 354−377.
(42) He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2016; pp 770−778.

https://doi.org/10.1021/acsomega.2c02665
ACS Omega 2022, 7, 23919−23928