applied sciences

MDPI

Article

Online Coal Identification Based on One-Dimensional Convolution and Its Industrial Applications

Shaochen Ma [1], Kaixun He [1,*] and Xin Peng [2]

[1] College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China; chinamardenma@163.com
[2] Key Laboratory of Smart Manufacturing in Energy Chemical Process, East China University of Science and Technology, Shanghai 200237, China; xinpeng@ecust.edu.cn
* Correspondence: kaixunhe@sdust.edu.cn

Abstract: In order to improve the utilization rate of coal generation and reduce carbon emissions from coal-fired boilers, the operation parameters of power plant boilers should be matched with the actual burning coal. Due to the complex and high-risk blending process of multiple coal types, the actual application of coal types inconsistent with expectations may lead to low combustion efficiency of boilers, cause disturbances to the normal operation of thermal power units, increased energy waste and carbon emissions, and even lead to serious explosion accidents. Therefore, the online identification of coal types for thermal power units is of great significance. To obtain the precise type of coal online, in the present work, a data-driven coal identification method is proposed based on convolutional networks that do not necessitate additional hardware detection equipment and are easy to implement. Experimental results indicate that the proposed method exhibits superior performance in comparison to traditional methods, thus ultimately improving the performance of thermal power plant.

Keywords: coal identification; convolutional networks; thermal power plant

check for updates

Citation: Ma, S.; He, K.; Peng, X. Online Coal Identification Based on One-Dimensional Convolution and Its Industrial Applications. *Appl. Sci.* **2023**, *13*, 9867. https://doi.org/10.3390/app13179867

Academic Editors: Jerry Chun Wei Lin, Chunwei Tian, Rafał Cupek and Marcin Fojcik

Received: 8 July 2023
Revised: 25 August 2023
Accepted: 28 August 2023
Published: 31 August 2023

1. Introduction

Recently, thermal power generation plays an important role in electric energy supply. In order to reduce costs and improve boiler efficiency, coal blending technology is widely used in thermal power plants (TPP), using different types of coal in specific proportions for combustion [1,2]. Coal blending technology refers to the process of burning coal in a boiler for power generation by blending several types of coal with different characteristics in a predetermined proportion. Considering that original coal structure has a significant effect on burnout, and abnormal changes in coal species can affect the low load stability and combustion capacity of boilers, improper blending or changes in coal types may result in a decrease in efficiency and, in severe cases, safety accidents [3–5]. Hence, online identification of coal is important to improve the efficiency of power plants.

In recent years, online coal identification has received extensive attention from both industry and academia. Chemical analysis-based, image-based, and machine-learning-based methods are widely used for coal identification. The chemical analysis-based method is reliable and relatively accurate. However, this method requires stringent analysis techniques for elemental content, and it is easily affected by foreign chemical substances. The image-based method primarily focuses on coal's shape, texture, and color, and it has gained considerable attention in recent years. For example, in 2019, Chengzhao Liu et al. [6]. proposed a coal quality recognition method based on electric coal images, which detects coal quality indexes more accurately using color and texture features. Huiling Meng [7] proposed an image analysis method for analyzing power coal quality, which accelerated the model recognition speed. Moreover, Yuanyuan Pu et al. [8] proposed a method in 2021, which combines image processing and convolutional neural network (CNN) to perform binary classification of coal and gangue. However, such methods are primarily used for the identification of pulverized coal before combustion and are easily influenced by light and imaging, which makes the identification of coal types challenging.

The online coal identification is a classification recognition problem. Machine learning methods, such as support vector machine (SVM), Naive Bayes, decision tree, and CNN, have been adopted to address this issue [9,10]. Andrew Ng and Michael Jordan [11] compared plain Bayesian and logistic regression methods, reflecting the advantages of plain Bayesian in small data volumes. However, as the era of big data emerged, the classification effectiveness of the Bayesian method began to diminish, leading to the development of improved methods, such as the weighted Bayesian classifier method proposed by Hanchuan Peng et al. [12]. Later, other improved methods such as semi-parsimonious Bayes classifier and Gaussian parsimonious Bayes classifier [13] were developed to achieve better classification results. The decision tree approach was first proposed by Quinlan J R [14], and later, in 1993, the shortcomings of the previous method were addressed in the literature [15], which addressed the shortcomings of previous methods, such as the ability to handle missing values. Over time, the accuracy of decision trees has gradually improved [16,17]. Alex Krizhevsky et al. [18]. were the first to apply CNN for image classification and classify 1.2 million images into 1000 categories with a low error rate. Xiangyu Zhang et al. [19]. improved a deeper level CNN visual geometry group (VGG) in 2015, which has been widely used in the field of target detection. Kaiming He et al. [20]. introduced residual computation in neural networks in 2015, which have addressed, to a certain extent, the gradient disappearance phenomenon that can easily occur in deep neural networks. Google subsequently proposed InceptionNet [21] and MobileNet [22], which further improved the classification accuracy of CNNs.

With the improvement of computational precision, the time cost and memory usage have also increased. Yu Xue et al. [23] addressed the issue of excessive memory consumption during training by proposing a partial channel connection based on channel attention for differentiable neural architecture search (ADARTS) in 2023, which enhanced the efficiency of the search process and optimized memory usage. In the same year, Yu Xue et al. [24] proposed a multi-objective evolutionary algorithm with a probability stack (MOEA-PS), which effectively reduces the time cost. Additionally, Zicheng Cai et al. [25] introduced a novel and efficient channel attention mechanism termed EPC-DARTS, which allocates weights based on channel importance and selectively employs channels with higher weights. This approach has contributed to mitigating the aforementioned issues to a certain extent.

CNN, as a deep learning network, provides significant advantages compared with traditional neural networks. Given its multiple convolutional layers, it can automatically extract data features, avoiding the time-consuming and inefficient manual feature extraction process. In traditional neural networks, overfitting can easily occur because of the excessive number of parameters, which negatively affects network performance. By contrast, CNNs can mitigate overfitting to a certain extent by leveraging the weight-sharing feature. In addition, the robustness of CNNs is effective in combating noise interference. Data-driven classification recognition has been primarily applied for image processing and text-to-speech recognition; however, it has not been applied to coal species recognition. In real power plant scenarios, massive amounts of high-dimensional data are often generated, accompanied by noise interference, which makes it challenging to complete the coal identification task using traditional or manual methods. Dongjun Li et al. [26] proposed a deep learning framework for coal and gangue detection based on image recognition. Qiang Liu et al. [27] presented an enhanced YOLOv4 algorithm for coal and gangue recognition using deep learning, showing better performance than YOLOv3. Ziqi Lv et al. [28] proposed a cascade network with a detector and a discriminator, enhancing coal and gangue detection under complex conditions by designing a multi-channel feature fusion layer and optimizing the CNN in the discriminator. The aforementioned existing methodologies are designed to address coal quality recognition by constructing classification models based on coal powder images. However, to the best of our knowledge, there is currently a dearth of research focused on utilizing combustion process data, specifically derived from sensor measurements, to perform online coal species identification.

To deal with the aforementioned problems, this paper proposes a data-driven algorithm for coal identification based on one-dimensional convolution, leveraging the advantages of the CNN algorithm in the field of classification and recognition. The proposed method collects process data from coal-fired boilers and feeds the pre-process data into a CNN model to construct a coal classification model. During online application, the process data are preprocessed first and then input into the CNN model, and the coal type is identified on the basis of model output. Industrial data-based verification demonstrates that the proposed method achieves high recognition accuracy, which can be easily implemented in the field.

The remainder of this paper is organized as follows. Section 2 provides a description of the target boiler used in this work. Preliminaries regarding Maximum information coefficient (MIC) and CNN are provided in Section 3. Section 4 presents the proposed modeling method. Section 5 discusses experiments and results obtained using real world data. Finally, conclusions are provided in Section 6.

2. The Profile of the Thermal Power Plant

This study is focused on a 1030 MW ultra-supercritical coal-fired power generation unit, which is a Spiral Wound Universal Pressure boiler (SWUP) featuring balanced ventilation, ultra-supercritical parameters, one reheating, and a spiral furnace. The unit is equipped with a solid slag discharge method and an open layout. A medium-speed coal mill, a positive pressure direct blowing cold primary air pulverizing system, front and rear wall opposing combustion mode, and low NOx dual-adjustable swirl burners and low NOx nozzles are employed in the system. The furnace chamber has a cross-sectional width of 33,128.7 mm, a depth of 16,308.7 mm, and a height of 64,500 mm. Boiler pulverization and the boiler system are crucial parts of a power plant, directly affecting power generation efficiency and energy utilization (Figure 1).

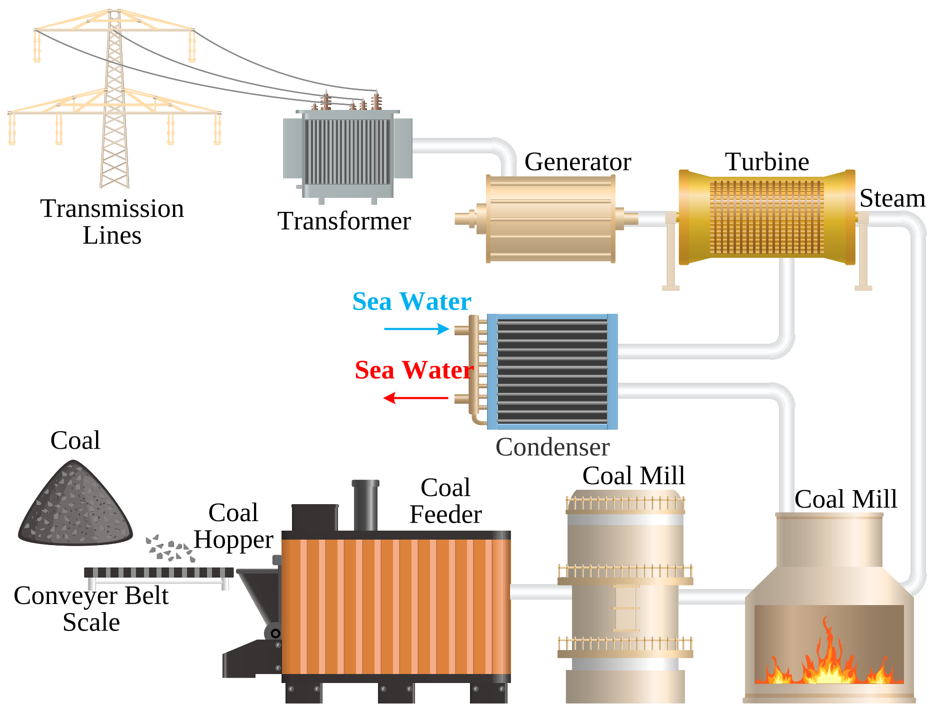

Figure 1. Process flow diagram of thermal power plant.

The thermal power unit comprises various components, including a coal feeder, coal mill, coal-fired boiler, air supply, steam piping, turbine, generator, and auxiliary equipment.

Raw coal is delivered to a raw coal hopper via a coal conveying belt, weighed by using a weighing belt, and then passed through the coal feeder to the coal mill, where it is ground into suitable pulverized coal. Given the huge amount of pulverized coal and the difficulty in keeping the grinding degree consistent, unqualified pulverized coal is returned to the mill for further grinding to ensure moderate diameter of the pulverized coal particles, while magazines that cannot be ground are collected in the stone coal hopper. In preventing coal powder leakage, the pulverizing system is equipped with a sealed air system, consisting of centrifugal fans that supply air to the rotary separator and other equipment after pressurization. Then, the treated pulverized coal is introduced into the boiler for combustion via the powder pipe. The main equipment of the boiler system includes a burner, boiler, turbine, generator, condenser, and auxiliary equipment. The burner converts the chemical energy of the pulverized coal into heat energy through combustion, generating a large amount of steam in the boiler. The steam pushes the turbine, converting heat energy into mechanical energy, which drives the generator to produce electricity. Thus, based on the processed data of the coal-fired boiler, the online identification of coal type, timely detection of abnormal changes in coal quality, and timely adjustment of boiler optimization control parameters are important to improve the efficiency of coal-fired power generation.

3. Preliminaries

3.1. Maximum Information Coefficient (MIC)

The MIC was proposed by Reshef et al. [29] based on information entropy to analyze the degree of linear and nonlinear dependencies among different features, which can be used for redundant feature rejection. MIC has been widely used for the extraction of linear and nonlinear feature variables.

The advantages of MIC are as follows:

(1) Universality: applicable to all types of data, regardless of data distribution issues.
(2) Autonomy: automatically mines the relationship among different features.
(3) Robustness: strong anti-interference ability, not affected by outliers and missing values.
(4) Interpretability: MIC results are in the range of [0, 1], allowing the strength of the correlation to be visualized.

The MIC primarily aims to discretize two features in space and divide them into several intervals to calculate the joint probability and determine the correlation.

Given a dataset: Consider two discrete random variables $X = \{x_1, x_2, ..., x_n\}$ and $Y = \{y_1, y_2, ..., y_n\}$, where n is the number of samples. The mutual information of X and Y can be calculated by using the following equations:

The formula for calculating mutual information is as follows :

$$I(x;y) = \int p(x,y) \log_2 \frac{p(x;y)}{p(x)p(y)} dxdy \tag{1}$$

The calculation formula for MIC is as follows:

$$mic(x;y) = \max_{a*b<B(n)} \frac{I(x;y)}{\log_2 \min(a,b)} \tag{2}$$

where a and b are the number of intervals partitioned in space; *B(n)* is the interval partition coefficient with the value of $B(n) = n^{0.6}$, and n is the amount of data. The MIC calculation process is shown in Algorithm 1.

Algorithm 1 Feature Selection through Maximum Information Coefficient.

Input: A data matrix **D**
Output: Selected k features $\hat{\mathbf{D}}_\mathbf{i} = (\hat{d}_{i1}, \hat{d}_{i2}...\hat{d}_{in})$, i = 1,2,. . . , k;
// One of the random feature in **D**
1:$x = \mathbf{D_k} = (d_{k1}, d_{k2}...d_{kn})$, k = 1,2,. . . , n;
// Another random feature in **D**
2:$y = \mathbf{D_l} = (d_{l1}, d_{l2}...d_{ln})$, l = 1,2,. . . , n;
Step 1: Compute $I(x; y)$ according to Equation (3)
Step 2: Compute $mic(x; y)$ according to Equation (4)
Step 3: Sort the result $mic(x; y)$ in descending order
Step 4: Select the best k features as the output result.
$\hat{\mathbf{D}}_\mathbf{i} = (\hat{d}_{i1}, \hat{d}_{i2}...\hat{d}_{in})$, i = 1,2,. . . , k;

3.2. One-Dimensional Convolution Network

Convolutional neural networks are important algorithms for deep learning [30,31], which are widely used in areas such as image processing and natural language processing. They can extract features from input data and classify them into different categories. The input data are preprocessed and then fed into multiple convolutional layers in which features are extracted. The data are then compressed through pooling layers to reduce computational complexity during down sampling. Finally, the features are combined nonlinearly in the fully connected layer, and the output layer produces the category to which they belong. Structure diagram of the one-dimensional convolution network is shown in Figure 2.

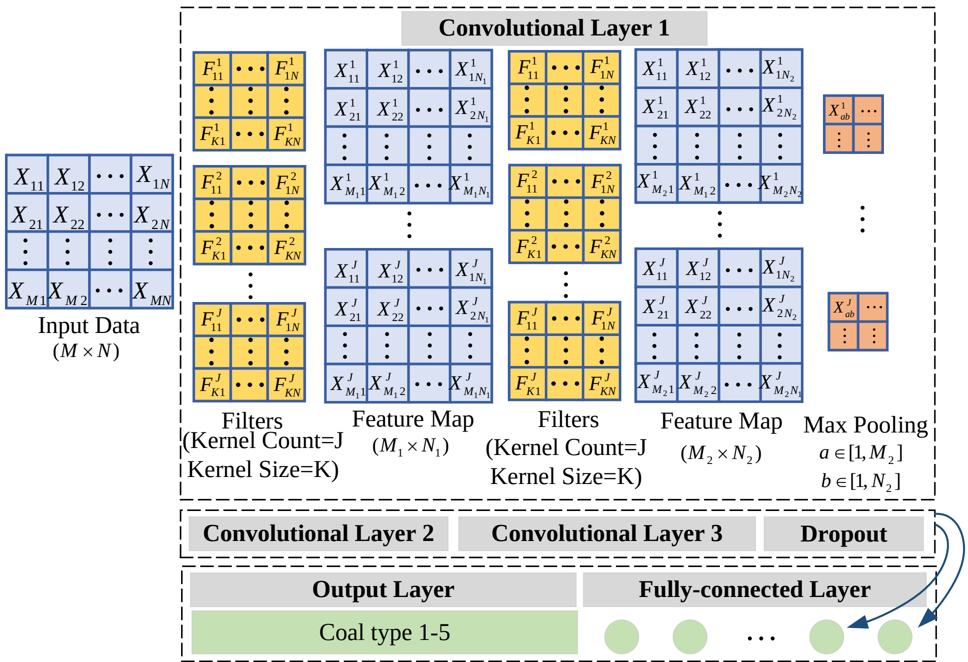

Figure 2. Structure diagram of the one-dimensional convolution network.

In each convolutional layer, the output is related to the input, convolutional kernel size, feature map, and bias of the previous layer. The output of the convolutional layer l is formulated as follows:

$$x_j^l = f(\sum_{i \in M_j} x_i^{l-1} k_{ij}^l + b_j^l) \tag{3}$$

where x_j^{l-1} is the input of the layer l; M_j is the feature map; k_{ij}^l is the convolutional kernel size, and b_j^l is the bias term of the layer.

In the pooling layer, a large number of data features are down sampled after convolution to reduce the amount of data in the next layer while preserving the features and preventing model overfitting. This process can be formulated as follows:

$$Output^l = f(b_o + w_o f_v) \tag{4}$$

where b_o and w_o are the deviation vectors and weight matrix, respectively.

4. Methodology

4.1. Variable Selection Strategy

Processed data contains numerous variables, and their direct use in building CNN models not only prolongs the network training time but also increases noise because of many redundant variables, thereby decreasing model accuracy. Therefore, preprocessing the original process variables to extract important features for classification recognition is necessary before building CNN classification recognition models. In this paper, variable screening aims to measure the strength of correlation among different variables by calculating the MIC between each selected variable and other variables and eliminating redundant variables based on the MIC metric. The specific steps of the proposed MIC-based variable screening are as follows:

(1) Calculation of MIC values: MIC values are computed to measure the correlation among different variables, which belong to a non-parametric method that can adjust data size and dimensionality adaptively. This method is highly adaptable to different data sets, and each pair of variables can be calculated to determine the correlation strength relationship.
(2) Sorting of MIC values: The MIC values calculated for each variable are sorted from largest to smallest, with the ranking indicating the strength of the correlation between the characteristic variables and the response variables. A value closer to 1 indicates a stronger correlation, whereas a value closer to 0 indicates a weaker correlation.
(3) Variable selection: The variables with a strong correlation with the remaining variables are selected from the ranked MIC values for elimination, and the remaining variables are retained as feature variables for subsequent modeling.

4.2. Construction of the Model for Coal Identification

After eliminating redundant features, the feature extraction capability of CNN is utilized to achieve the coal identification task of coal type data from coal-fired boilers in TPPs. The redundant features of CNN learning, which may affect model accuracy, are effectively addressed. The model workflow is illustrated in Figure 3. The pseudo code of the algorithm is shown in Algorithm 2. The details are as follows:

(1) Data acquisition: The experimental data used in this study are related to the actual application of coal in a coal-fired boiler in a TPP, consisting of a total of 288,000 data.
(2) Data pre-processing: Directly collected data from TPPs often contain missing and duplicated data caused by the industrial environment and process flow, making them unsuitable for direct use. Therefore, data cleaning is performed to remove or correct missing data, normalize data to reduce the computational load, and improve data labeling to meet the supervised learning conditions of 1D-CNN.
(3) Data set partitioning: The processed data are divided into training and test sets.
(4) MIC-CNN model construction: The input data are fed into the model for training and testing, and model accuracy is validated using the validation set. In addition, the model is compared with back propagation (BP) neural network and 1D-CNN methods.

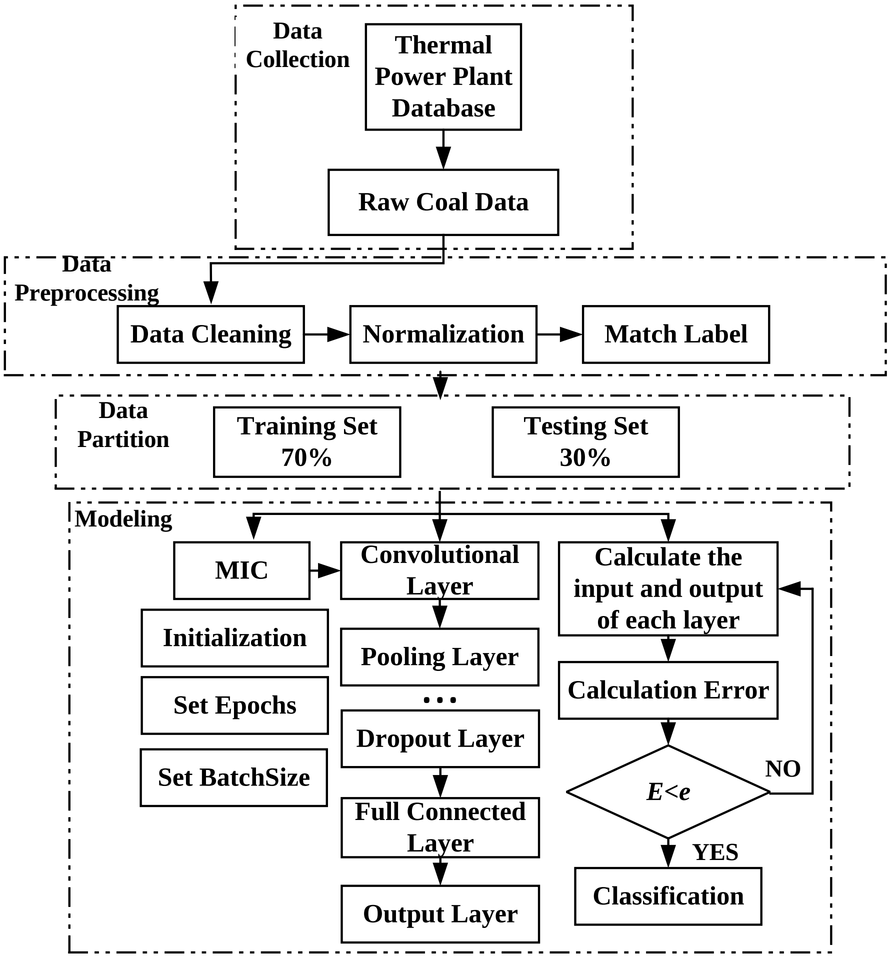

Figure 3. Flowchart of the proposed method.

Algorithm 2 Coal identification algorithm based on MIC-CNN.

Input: A data matrix $\mathbf{D} = (D_1, D_2 ... D_n)$
Output: An array R; // Classification Result
// Complete data preprocessing
Step 1:Def function (data preprocessing):
    Return Data cleaning, Normalization, Match Label;
// Complete data partition
Step 2: Def function (data partition):
    Training Set: 70%, Testing Set: 30%
    Return Training Set, Testing Set;
// Complete modeling
Step 3: Def function (MIC): // feature selection
    Select best k features $\hat{\mathbf{D}}_{\mathbf{i}} = (\hat{d}_{i1}, \hat{d}_{i2} ... \hat{d}_{in})$, $i = 1,2, \ldots, k$;
    Return $\mathbf{D}_{\mathbf{new}} = (D'_1, D'_2 ... D'_n)$; // New data
Step 4: Def function (CNN):
    Pooling Layer, Dropout Layer, Full Connected Layer, Output Layer
Initialization, Set Epochs, Set BatchSize
    Calculation Error E
    If $E < e$:
    Return R;
    Else:
    Calculate again;

4.3. Performance Measurement Criteria

The loss function is used to assess the discrepancy between the predicted and actual values during model training. In classification tasks, various loss functions, including entropy, cross-entropy, K-L divergence, dice loss [32], and focal loss [33], are typically utilized. In this experiment, cross-entropy is utilized to evaluate the prediction accuracy of the model. The computation of cross-entropy is performed as follows:

$$L = -\frac{1}{N}\sum_{i}\sum_{j=1}^{M} y_{ij}\log(p_{ij}) \tag{5}$$

where M represents the total number of categories, which takes the value of either 0 or 1, based on whether it corresponds to the true category, and p_{ij} represents the probability that the current sample i belongs to category j.

5. Case Study

5.1. Data Description

The data utilized in this experiment are derived from an industrial collection at a coal-fired TPP, particularly from a coal-fired boiler, yielding a total of 288,000 data points. These data were collected in 2020, which can be used to address coal identification challenges that have arisen from 2020 until present. The raw data comprise 47 variables and originates from sensor collections at the TPP. The original data were sampled once per second, but given its large volume, downsampling was used to reduce the sampling rate to once every 50 s. All process variables are shown in Table A1.

In this study, the data sets were generated with a 7:3 ratio between the training set and the test set. The training set comprises 201,600 data points, which are used to build and train the model. Meanwhile, the test set, consisting of 86,400 data points, is used to evaluate the accuracy and generalization ability of the trained model. All the approaches mentioned in this work are conducted in PyCharm (Community Edition 2022.2.4).

5.2. Data Pre-Processing

Coal combustion and the technology used in TPPs are intricate, leading to the presence of "dirty data" in actual coal data collection, such as missing or redundant information. Eliminating these invalid data can enhance the accuracy of the coal identification model and prevent data interference.

The coal data contains numerous feature vectors, each with a distinct evaluation index and singular samples with significant differences from other feature vectors. Such issues can be resolved by restricting data normalization within the range of [0, 1]. Data normalization is calculated as follows:

$$X = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{6}$$

The coal type data utilized in this study were collected through 47 sensors situated at various locations. This approach recorded numerous individual features of the coal type signal, providing a significant advantage for using a CNN model.

Considering data characteristics, supervised learning is necessary when using 1D-CNN. Hence, labeling the coal type signals is essential. The corresponding labels for the coal type dataset are presented in Table 1.

Table 1. Coal type dataset.

Number of Samples Training Set/Test Set	Coal Type	Label
40,320/17,280	Coal type1	0
40,320/17,280	Coal type2	1
40,320/17,280	Coal type3	2
40,320/17,280	Coal type4	3
40,320/17,280	Coal type5	4

5.3. Model Performance Analysis

A total of 288,000 data points were collected from TPPs and classified into five categories. The resulting 47-dimensional matrix was obtained by processing the data based on the number of features (47) and sampling frequency (50 s/time). Given the presence of redundant variables, which can interfere with model accuracy and increase model operation time, feature selection was performed using MIC. Each of the 47 features was individually computed to derive the corresponding MIC value against the remaining features ($mic, mic \in [0, 1]$), where a value close to 1 indicates a stronger correlation among the features, indicating the need for feature elimination. The top k features with the strongest correlation were removed by sorting the MIC values in descending order, and 30 features were retained with a k value of 17, which was determined through trial and error to yield the best results.

In order to establish robustness and optimal performance of the proposed approach, a hyperparameter tuning process was conducted using a trial and error method. The hyperparameters, including batch size, dropout rate, and learning rate, were systematically adjusted to identify the configuration that yields the best results. Batch size, a critical hyperparameter influencing the convergence and generalization of the model, was fine-tuned through a series of experiments. Starting with a conservative value, we gradually increased the batch size while monitoring the training dynamics performance. This process allowed us to determine an optimal batch size that strikes a balance between computational efficiency and model convergence. The value of dropout is typically in the range of 0.2 to 0.5. If the dropout rate is too small (e.g., 0.1 or lower), it might not have a strong regularizing effect. Conversely, if the dropout rate is too large (e.g., 0.7 or higher), it may hinder the network's ability to learn and generalize effectively. After conducting multiple tests, the appropriate value for Dropout can be determined. Furthermore, the parameter of learning rate, which impacts convergence speed and model stability, was rigorously tested. We employed a learning rate schedule, gradually reducing the learning rate during the training epochs and evaluating its impact on training dynamics, ensuring stable convergence without the risk of divergence.

Three comparative algorithms are utilized to compare and validate the effectiveness of the proposed method in this study.

(1) 1D-CNN model: 1D-CNN encompasses a convolutional layer, pooling layer, fully connected layer, and classification output layer. In this experiment, three convolutional layers are used, and each convolutional layer executes two convolutional calculations. The pooling layer adopts maximum pooling after each convolutional layer. Dropout is implemented before the fully connected layer, with a value of $Dropout = 0.3$, $Epochs = 2, Batchsize = 400$. The optimizer used in this study is Adam, with a learning rate set at 1×10^{-4}. The original power plant data are preprocessed and split into training and test sets in a 7:3 ratio. Following this step, the features are non-linearly combined using the fully connected layer, and the coal categories are obtained through the classification output layer. The results of model training and testing are recorded.

(2) MIC-CNN model: During the establishment of the MIC-CNN coal identification model, the data are initially subjected to feature selection using MIC, followed by data pre-processing and division into a training set and a test set at a ratio of 7:3.

Subsequently, the 1D-CNN model is trained to accomplish the coal identification task and evaluated for accuracy and generalization ability on the test set. The resulting training and test results of the model are documented.

(3) BP neural network model: In building a multilayer perceptron for coal classification, an input layer with 47 nodes, an output layer with five nodes, 32 hidden layers, Relu as the activation function, and Softmax classifier as the output layer are used to output the classification results. The accuracy and training test results of the model are recorded.

This paper compares the experimental results of the three aforementioned methods. Table 2 presents a comparison of the classification performance of MIC-CNN, CNN, and BP.

When utilizing the BP neural network for feature extraction and classification of coal types (Figures 4 and 5), the network can extract coal type features despite the heterogeneity of the data volume and the presence of redundant and highly correlated features. However, the classification effect is not ideal, as the recognition accuracy of coal type 5 is poor. The model can correctly identify five types of coal species, with an overall test set accuracy of 68.65%. In terms of memory usage and training time, the BP neural network occupies 2237.6 MB of memory during training and takes 202.02 s to complete the training process.

When passing the data through 1D-CNN for coal species recognition (Figures 6 and 7), features are extracted through four convolutional layers. Despite some high-dimensional invalid features, coal species recognition can still be achieved on the basis of the feature extraction ability of the model and sharing of convolutional kernels. During the recognition of five types of coal, the recognition accuracy of coal types 1–4 is significantly higher, and the recognition accuracy of coal type 5 is greatly improved, with an accuracy of 78.38% in its test set, which is nearly 10% higher than the accuracy achieved by the traditional BP neural network. In terms of memory usage and training time, the 1D-CNN occupies 2413.7 MB of memory during training and takes 188.58 s to complete the training process.

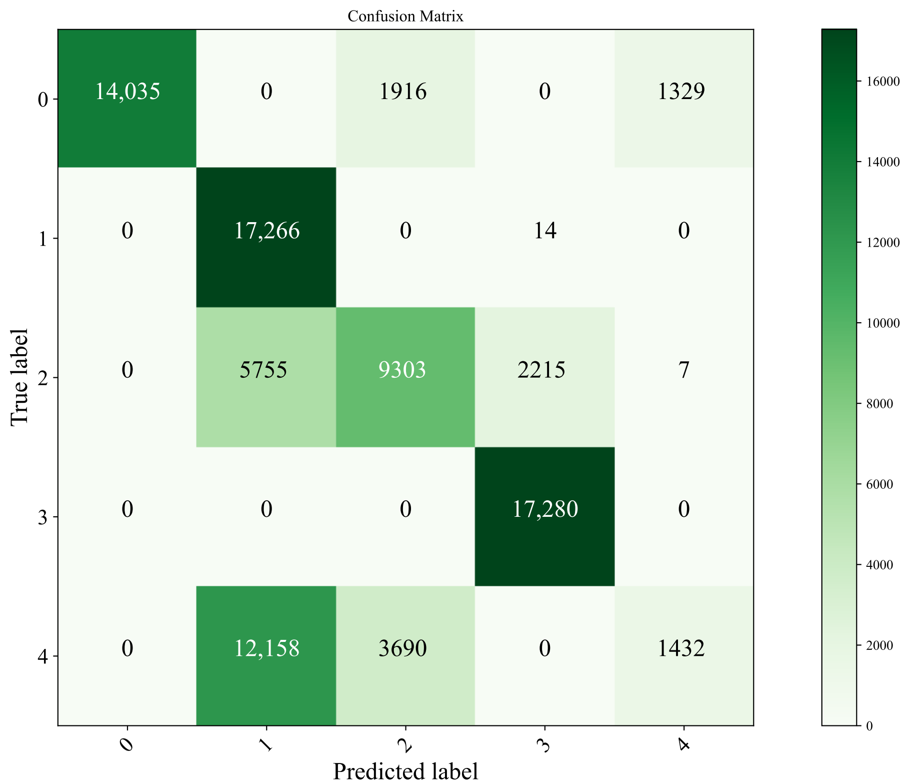

Figure 4. Confusion matrix of the BP neural network.

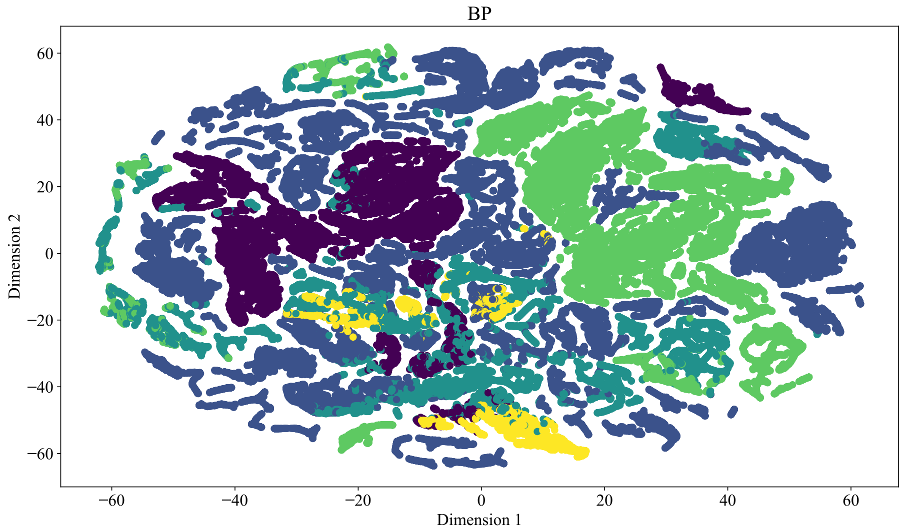

Figure 5. Clustering diagram of BP neural network classification.

Table 2. Classification performance.

Accuracy	BP	1D-CNN	MIC-CNN
Training set	78.17%	88.54%	98.23%
Test set	68.65%	78.38%	94.02%

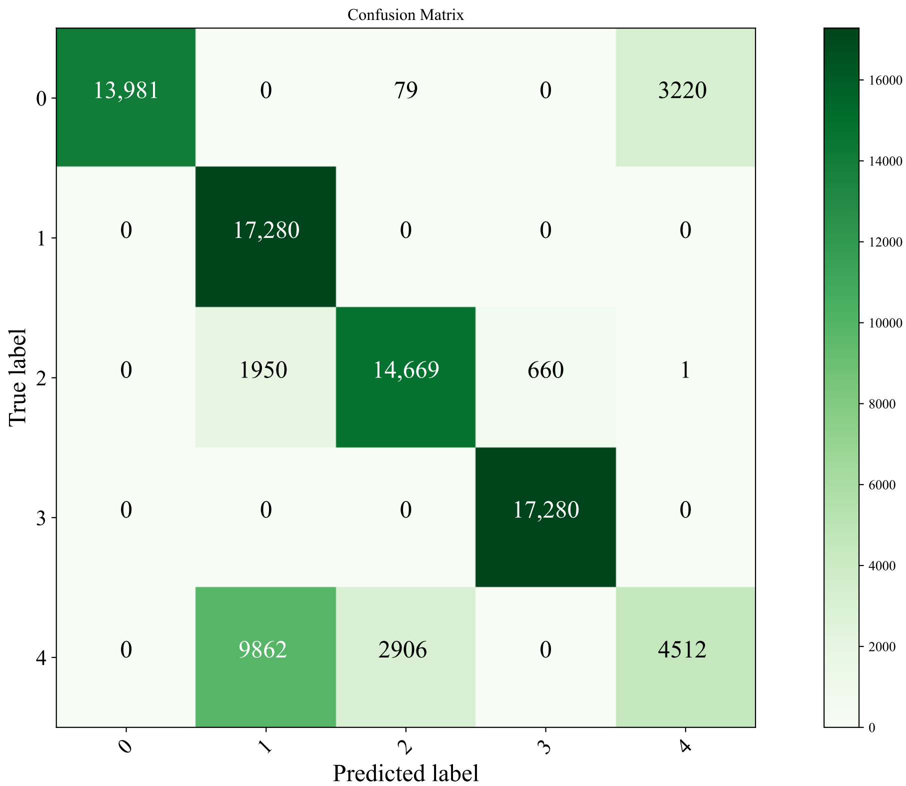

Figure 6. Confusion matrix of 1D-CNN.

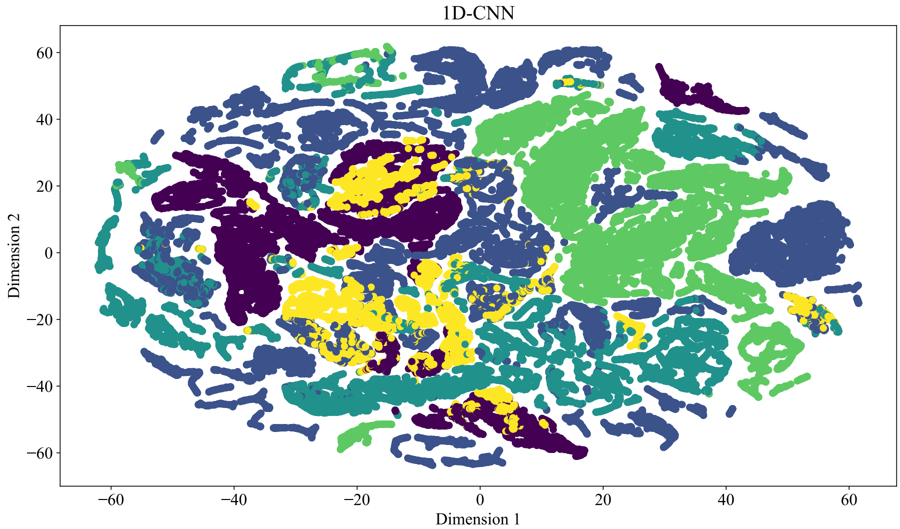

Figure 7. Clustering diagram of 1D-CNN classification.

In enhancing the accuracy of the model, considering the strong inter-correlation among features in actual industrial data, the MIC-CNN method has been adopted. Initially, the features with strong correlations are removed by using the MIC method, while retaining the independent features. The streamlined data, consisting of 30 features after MIC screening, are then fed into the 1D-CNN network. This process significantly reduces the interference of redundant features (Figures 8 and 9). The recognition accuracy of coal types in the test set processed by MIC-CNN reaches 94.02%, which is nearly 16% higher than the classification accuracy of 1D-CNN, indicating the efficacy of the proposed approach in achieving better coal species recognition. In terms of memory usage and training time, the MIC-CNN occupies 2217.5 MB of memory during training and takes 179.95 s to complete the training process.

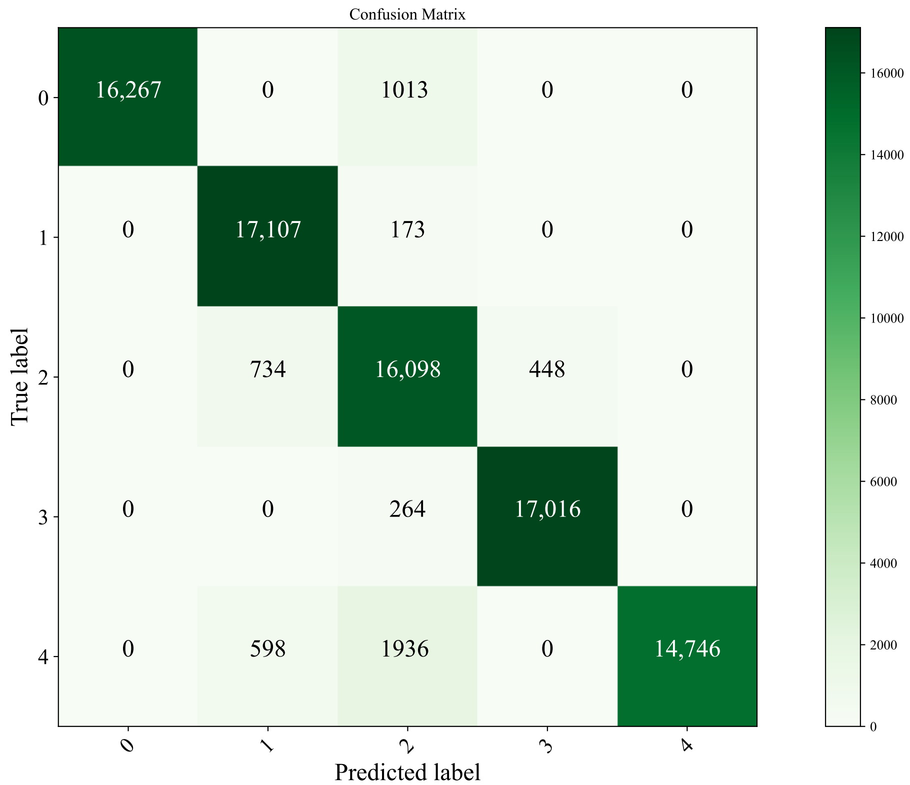

Figure 8. Confusion matrix of MIC-CNN.

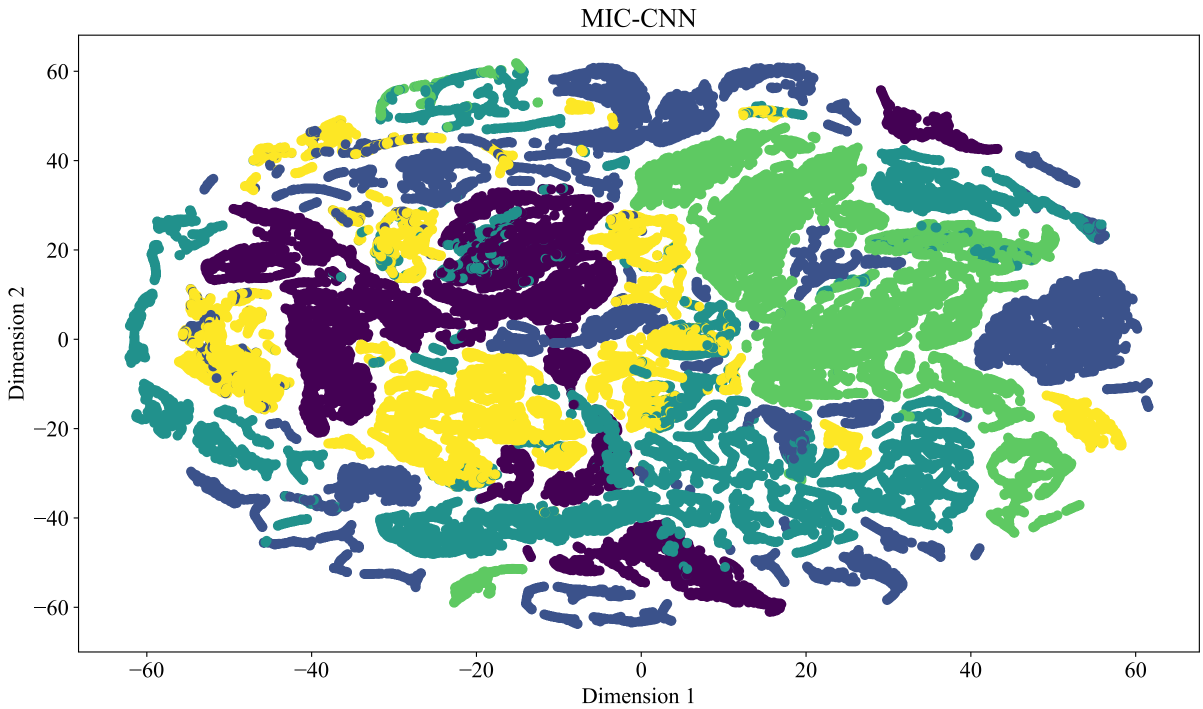

Figure 9. Clustering diagram of MIC-CNN classification.

6. Discussion

As shown in Table 2, the MIC-CNN method achieves the best recognition effect for coal types, with higher classification accuracy and faster computation speed compared with 1D-CNN and BP neural networks. The MIC method can quickly and accurately identify the correlation among data features and reduce the interference of redundant features. When combined with the powerful feature extraction and classification ability of CNN, the model accuracy can be further improved. The recognition effect of 1D-CNN follows, which is affected by the presence of redundant features that lower the classification accuracy and long computation time of the model. The BP neural network has the poorest recognition effect because of the high requirements of data quality and features, and it is prone to overfitting when handling large amounts of power plant data. These experiments demonstrate that data-driven deep learning methods have a good effect on coal identification, and combining with the MIC feature selection method can improve the identification accuracy of the deep learning model. This paper provides a new approach for coal identification in TPPs, which has potential applications. A dedicated server is established on-site to host the software components. This server acts as the central hub for data processing, storage, and distribution, which ensures data accessibility and maintains a stable connection with the Oracle database. The Oracle database, provided by Oracle Corporation, is utilized to store and manage the data generated by the system. The software establishes a connection to the Oracle database using appropriate credentials, enabling data retrieval and storage. This integration ensures data consistency, security, and efficient management. The software is configured to display the data locally, which means that the data processed and retrieved from the Oracle database is visualized on a user interface accessible within the local environment. This local display provides real-time insights into the data's behavior and characteristics. The software's functionality extends to the Distributed Control System (DCS) monitoring interface. Through integration with the DCS, the software enables real-time monitoring and control of the processes. This integration enhances operational efficiency by allowing operators to monitor the system's performance and make informed decisions. The software is deployed using Python programming language, a versatile and widely-used scripting language. Python's robust libraries and frameworks facilitate seamless communication between various components, making it suitable for interfacing with the Oracle database, local display, and DCS monitoring.

7. Conclusions

Online coal identification forms the basis of combustion optimization control and plays an important role in improving boiler efficiency. In the present work, in order to monitor the changes of coal types effectively, an online coal identification method is proposed based on 1D-CNN. First, MIC is used to extract important features from the original process variables. The selected features are then used to construct a 1D-CNN model. Based on MIC, the proposed method compensates for the influence of feature redundancy and high feature dimension on CNN model construction, improves the modeling efficiency, and reduces the influence of noise on CNN's performance by eliminating redundant features. Industrial data experiments show that the accuracy of coal identification can reach over 94.02%. The purpose of online coal identification is to detect changes in coal quality over time so as to adjust the control parameters of combustion optimization. In the future work, we will focus on approaches to combine the results of coal identification with combustion optimization control, as well as utilize intelligent algorithms for automatic parameter tuning to further enhance boiler combustion efficiency. in order to further improve the boiler combustion efficiency.

Author Contributions: S.M.: Software, data curation, Visualization, Investigation; K.H.: Conceptualization, methodology, Writing—Original draft preparation; X.P.: Writing—Reviewing and Editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partly supported by the National Natural Science Foundation of China (NSFC) under Grant (62273214, 62233012), partly by the Research Fund for Postdoctoral Innovation Talent Support Plan of Shandong Province of China (SDBX2022015), and partly by the Research Fund for the Taishan Scholar Project of Shandong Province of China (Corresponding author: Kaixun He).

Data Availability Statement: The raw/processed data required to reproduce these findings cannot be shared at this time as the data also form part of an ongoing study and are not authorized for publication.

Acknowledgments: The authors would like to thank the editors and anonymous referees for their invaluable comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

TPP	Thermal power plants
CNN	Convolutional neural network
1D-CNN	One-dimensional convolution network
SVM	Support vector machine
VGG	Visual geometry group
ADARTS	Attention for differentiable neural architecture search
MOEA-PS	Multi-objective evolutionary algorithm with a probability stack
EPC-DARTS	Efficient partial channel connection for differentiable architecture search
MIC	Maximum information coefficient
SWUP	Spiral Wound Universal Pressure Boiler
DCS	Distributed Control System

Appendix A

Table A1. Process variable information.

Number	Variables	Units
1	2 AB phase voltage of generator stator	KV
2	2A The temperature of the lower bearing of the coal mill's rotary separator	DEG C
3	A Motor speed of coal feeder	rpm/mi
4	A Instantaneous coal feeding rate for coal feeder	t/h
5	A Three-selection output of primary air temperature at the outlet of air preheater	DEG C
6–11	A Coal mill motor stator coil temperature 1–6	DEG C
12–15	A Coal mill motor bearing temperature 1–4	DEG C
16	A Current of coal mill	A
17–19	A Coal mill air-powder mixture temperature1–3	DEG C
20–21	A Coal mill planetary gearbox input bearing temperature 1–2	DEG C
22–25	A Coal mill planetary gearbox bearing temperature 1–4	DEG C
26	A Feedback on the position of the cold primary air electric adjustment damper of the coal mill	%
27	A Differential pressure between the sealing air of the coal mill and the lower part of the grinding bowl	kPa
28	A Differential pressure above and below the grinding bowl of the coal mill	kPa
29	A Position feedback of electric regulating damper for hot primary air of coal mill	%
30	A Lubricating oil return temperature of coal mill	DEG C
31	A Coal mill lubricating oil temperature 1	DEG C
32	A Temperature of lubricating oil tank of coal mill	DEG C
33	A Lubricating oil pressure of coal mill	MPa
34	A Coal mill rotary separator current	A
35	A Bearing temperature on the rotary separator of the coal mill	DEG C
36	A Coal mill rotary separator speed output	rpm/mi
37	A Lower bearing temperature of rotary separator of coal mill	DEG C
38	A Primary air pressure of coal mill	kPa
39	A Inlet air temperature of forced draft fan	DEG C
40	BTU correction command feedback	
41	B Three-selection output of primary air temperature at the outlet of air preheater	DEG C
42	Generator active power three-selection output	MW
43	Three-selection output of primary air volume of coal mill	t/h
44	Primary Air Temperature Dual Selection Output of Coal Mill	DEG C
45–47	Hot primary air header pressure 1–3	MPa

References

1. Rushdi, A.; Sharma, A.; Gupta, R. An experimental study of the effect of coal blending on ash deposition. *Fuel* **2004**, *83*, 495–506. [CrossRef]
2. Chen, G.; Xia, J.; Peng, P.; Yang, T.; Gao, W.; Yan, Z.; Jiang, J. Development and Application of the Intelligent Decision System for Entire Process of Coal Blending. *ASME Power Conf.* **2011**, *44595*, 215–220.
3. Nomura, S.; Arima, T.; Kato, K. Coal blending theory for dry coal charging process. *Fuel* **2004**, *83*, 1771–1776. [CrossRef]
4. Dai, C.; Cai, X.H.; Cai, Y.P.; Huang, G.H. A simulation-based fuzzy possibilistic programming model for coal blending management with consideration of human health risk under uncertainty. *Appl. Energy* **2014**, *133*, 1–13. [CrossRef]
5. Hu, L.; Zhang, Y.; Chen, D.; Fang, J.; Zhang, M.; Wu, Y.; Zhang, H. Experimental study on the combustion and NOx emission characteristics of a bituminous coal blended with semi-coke. *Appl. Therm. Eng.* **2019**, *160*, 113993. [CrossRef]
6. Liu, C.; Li, M.; Zhang, Y.; Han, S.; Zhu, Y. An enhanced rock mineral recognition method integrating a deep learning model and clustering algorithm. *Miner* **2019**, *9*, 516. [CrossRef]
7. Hui-ling, M.; Man, L. Characteristic analysis and recognition of coal-rock interface based on visual technology. *Int. J. Signal Process. Image Process. Pattern Recognit.* **2016**, *9*, 61–68.
8. Pu, Y.; Apel, D.B.; Szmigiel, A.; Chen, J. Image recognition of coal and coal gangue using a convolutional neural network and transfer learning. *Energies* **2019**, *12*, 1735. [CrossRef]
9. Pontil, M.; Verri, A. Properties of support vector machines. *Neural Comput.* **1998**, *10*, 955–974. [CrossRef]
10. Joachims, T. Text categorization with support vector machines: Learning with many relevant features. In Proceedings of the 10th European Conference on Machine Learning, Chemnitz, Germany, 21–23 April 1998; pp. 137–142.
11. Ng, A.; Jordan, M. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. In Proceedings of the Advances in Neural Information Processing Systems 14 (NIPS 2001), Vancouver, BC, Canada, 3–8 December 2001; Volume 14.
12. Peng, H.; Long, F.; Ding, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1226–1238. [CrossRef]
13. Xu, S. Bayesian Naïve Bayes classifiers to text classification. *J. Inf. Sci.* **2018**, *44*, 48–59. [CrossRef]
14. Quinlan, J.R. Induction of decision trees. *Mach. Learn.* **1986**, *1*, 81–106. [CrossRef]
15. Quinlan, J.R. Improved use of continuous attributes in C4.5. *J. Artif. Intel. Res.* **1996**, *4*, 77–90. [CrossRef]

16. Banfield, R.E.; Hall, L.O.; Bowyer, K.W.; Kegelmeyer, W.P. A comparison of decision tree ensemble creation techniques. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *29*, 173–180. [CrossRef] [PubMed]
17. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, 55, 119–139. [CrossRef]
18. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM*. **2017**, 60, 84–90. [CrossRef]
19. Zhang, X.; Zou, J.; He, K.; Sun, J. Accelerating very deep convolutional networks for classification and detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *38*, 1943–1955. [CrossRef] [PubMed]
20. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
21. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9.
22. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
23. Xue, Y.; Qin, J. Partial connection based on channel attention for differentiable neural architecture search. *IEEE Trans. Industr. Inform.* **2023**, 19, 6804–6813. [CrossRef]
24. Xue, Y.; Chen, C.; Słowik, A. Neural Architecture Search Based on A Multi-objective Evolutionary Algorithm with Probability Stack. *IEEE Trans. Evol. Comput.* **2023**, 27, 778–786. [CrossRef]
25. Cai, Z.; Chen, L.; Liu, H.L. EPC-DARTS: Efficient partial channel connection for differentiable architecture search. *Neural Netw.* **2023**, 166, 344–353. [CrossRef] [PubMed]
26. Li, D.; Zhang, Z.; Xu, Z.; Xu, L.; Meng, G.; Li, Z.; Chen, S. An image-based hierarchical deep learning framework for coal and gangue detection. *IEEE Access* **2019**, *7*, 184686–184699. [CrossRef]
27. Liu, Q.; Li, J.; Li, Y.; Gao, M. Recognition methods for coal and coal gangue based on deep learning. *IEEE Access* **2021**, *9*, 77599–77610. [CrossRef]
28. Lv, Z.; Wang, W.; Xu, Z.; Zhang, K.; Lv, H. Cascade network for detection of coal and gangue in the production context. *Powder Technol.* **2021**, *377*, 361–371. [CrossRef]
29. Reshef, D.N.; Reshef, Y.A.; Finucane, H.K.; Grossman, S.R.; Mcvean, G.; Turnbaugh, P.J.; Lander, E.S.; Mitzenmacher, M.; Sabeti, P.C. Detecting novel associations in large data sets. *Science* **2011**, 334, 1518–1524. [CrossRef] [PubMed]
30. Chen, H.; Chen, A.; Xu, L.; Xie, H.; Qiao, H.; Lin, Q.; Cai, K. A deep learning CNN architecture applied in smart near-infrared analysis of water pollution for agricultural irrigation resources. *Agric. Water Manag.* **2020**, *240*, 106303. [CrossRef]
31. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, 77, 354–377. [CrossRef]
32. Milletari, F.; Navab, N.; Ahmadi, S.A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 565–571.
33. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollar, P. Focal loss for dense object detection. In Proceedings of the International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2980–2988.