Rotary Kiln Combustion State Recognition Based on Convolutional Neural Network

Tao Li, Tao Peng and Hua Chen

Department of Computer Science and Electronic Engineering, Hunan University, Changsha, Hunan, 410082, China.
Email: 450720262@qq.com

Abstract. The accurate recognition of combustion state is an important part of combustion control system. The rotary kiln is nonlinear, large lag, multi-disturbance and multivariable in the combustion process, so it is difficult to realize the intelligent control of the rotary kiln. Moreover, the traditional image recognition method is easily affected by the external environment, and the accuracy of recognition depends on the extraction of artificial features, so the satisfactory results can not be obtained. In order to solve these problems, a method of flame image recognition of combustion state in rotary kiln based on convolution neural network is proposed in this paper. In this method, convolution neural network vgg-16 model is used for feature migration, and the flame images of different combustion states in rotary kiln are trained and tested by network in order to achieve the purpose of automatic recognition of combustion state. The experimental results show that the combustion control system based on convolution neural network image recognition is effective, robust and has high accuracy. Compared with the traditional image recognition method, this method can achieve more effective accurate and more reliable automatic classification effect.

1. Introduction

Rotary kilns are widely used in chemical industries such as metallurgy, building materials, cement, aluminum firing, and power plant construction. The temperature of the rotary kiln firing zone is a key factor in determining the quality of clinker, so the current research on the rotary kiln intelligent control system is roughly selected. The control variable is the temperature value. The literature [1-2] proposes a pid controller to control the temperature of the power station boiler. At present, the fuzzy neural network [3-4] is widely used in the rotary kiln temperature control system. These methods are based on the temperature in the kiln as the controlled variable. These controllers achieve automatic control by directly measuring the temperature of the rotary kiln. Due to the influence of dust and smoke in the kiln, traditional temperature measurement systems, such as infrared systems, thermal coupling systems, and colorimeter systems, will not be used in this case. In recent years, through the analysis of the thermal signals collected by the sensors, a technology for automatic monitoring and control of the rotary kiln [5-6] has been proposed. Due to the complex physical and chemical reactions in the kiln, heat The signal identification system has a lag problem.

Digital image processing methods have become one of the research hotspots in the field of rotary kiln combustion status recognition to determine the combustion status and coal feed control in a more timely manner [7-11]. The main research steps of these methods include image preprocessing, image segmentation and Extract the image features, then input the extracted image features to the pattern classifier, and finally perform pattern recognition on the target to determine the flame combustion state. The flame is formed by the mixed combustion of coal powder and air, and the geometry and intensity of the flame are the combustion The quality and performance of the process provide instant..
information. In determining the combustion state, the image-based method is faster and more direct than the thermal signal-based method. At present, the recognition method of the combustion state based on digital image technology has been studied by many scholars. These research methods have various algorithms and high recognition, and they are quite mature, but due to the complex and changeable environment in the kiln, it is impossible to monitor the flame status in the rotary kiln online in real time, and it is still necessary to manually monitor the combustion status of the rotary kiln. The subjective influence of the subject is very large, such as technical level, responsibility and patience. The internal combustion state fluctuates, which is easy to cause under-combustion or over-combustion of abnormal combustion state, which is ultimately not conducive to the improvement of clinker quality and output. And the recognition method based on digital image processing technology usually depends on the extraction of artificial features, the selection of artificial features To obtain reasonable and effective, the accuracy of recognition is high, the method of artificial feature extraction is unreasonable, and the accuracy of recognition cannot obtain good results, but good feature extraction usually requires rich experience and solid professional knowledge.

Combustion state recognition based on deep learning can better overcome the above deficiencies. In recent years, deep learning based on convolutional neural network (CNN) has become a research hotspot, widely used in image recognition [12], pedestrian detection [13], speech Recognition [14] and other fields. This paper proposes a method for identifying the combustion state of a rotary kiln flame image based on a convolutional neural network. The main idea is to remove the fully connected layer of the VGG-16 network trained on the ImageNet dataset and retain the features. The layer is transferred to the flame image data set in this paper to extract the characteristics of the flame image data. The feature layer usually contains complex and diverse features such as edge features, texture features, etc. and deeper features. The flame image is automatically extracted by the model. Features, no need to manually extract image features, overcome many inherent shortcomings of traditional image processing algorithms and can effectively improve the accuracy of image classification and recognition.

Recent studies have shown that convolutional neural networks have great advantages in feature extraction and have certain invariance to operations. Existing neural network models have the characteristics of large calculations and high computing resources, while deep convolutional neural networks. The network model is prone to local optimization problems, making transfer learning an ideal choice. In order to improve the overall recognition performance, this paper adds a transfer learning method, compares the parameter initialization model with the transfer learning model, and compares it with other traditional experimental methods. Comparison. The experimental results show that this method can significantly improve the accuracy of the rotary kiln combustion state identification, and has good robustness and generalization ability.

2. Research Background

2.1. Convolutional Neural Network

The convolutional neural network is essentially a multi-layer perceptron, which was designed inspired by the neural mechanism of vision. In 1962, Hubel and Wiesel [15] studied the electrophysiology of the visual cortex of cats and proposed the concept of receptive field. 1980, Fukushima [16] introduced the concept of CNN for the first time on the basis of the receptive field. The difference between the convolutional neural network and the ordinary neural network is that the convolutional neural network contains a feature extraction consisting of a convolutional layer and a sub-sampling layer. In the convolutional layer of the convolutional neural network, a neuron is only connected to a part of the neurons in the adjacent layer. In a convolutional layer of CNN, it usually contains several feature planes, and each feature plane is arranged by some rectangles. The neurons of the same feature plane share the weights, and the shared weights are also called convolution kernels. The direct benefit of weight sharing is to reduce the connection between the various layers of the network, while reducing the over-simulation Risk.

The five main structures of the convolutional neural network structure are the input layer, the convolutional layer (convolutional layer), and the pooling layer (pooling layer), also known as the
sampling layer, the fully connected layer, and the output layer. The layer is usually composed of multiple convolutional layers and pooling layers stacked alternately, and the fully connected layer corresponds to the hidden layer and logistic regression classifier of the traditional multi-layer perceptron. The features extracted by the convolutional layer and the sub-sampling layer are usually As the input of the first fully connected layer. The last output layer is a classifier. The classic convolutional neural network classifier used for image classification usually uses softmax regression, so sometimes the output layer is also called softmax layer. Use logistic regression, or support vector machines and other classifiers.

The input layer is the input of the entire neural network. In the convolutional neural network that processes images, it generally represents the pixel matrix of a picture. The length and width of the matrix represent the size of the image, and the depth of the three-dimensional matrix represents the image. Color channel (channel). For example, the depth of a black and white picture is 1, while in the RGB color mode, the depth of the image is 3. Starting from the input layer, the convolutional neural network uses a different neural network structure to convert the three-dimensional matrix of the previous layer. The three-dimensional matrix converted to the next layer is converted to the three-dimensional matrix of the next layer until the last fully connected layer.

The convolutional layer is the most important part of a convolutional neural network. Unlike the traditional fully connected layer, the input of each node in the convolutional layer is only a small block in the previous layer of the neural network, and the size of this small block is 3 * 3 or 5 * 5, which is the convolution kernel. The convolution layer attempts to analyze each small block in the neural network more deeply to obtain higher abstraction features. In general, it is processed by the convolution layer. The node matrix of will become deeper. Through the convolution operation, the convolutional layer of cnn extracts different features of the input layer, the convolutional layer extracts low-level features such as edge features, texture features, etc. at the lower layer, and the convolutional layer of the higher layer extracts more advanced features. Cnn weight sharing can reduce model complexity and make the network easier to train.

The pooling layer in the neural network usually does not change the depth of the three-dimensional matrix, but it can reduce the size of the matrix. The pooling operation can be considered as converting a higher-resolution picture into a lower-resolution picture. The layer can further reduce the number of nodes in the final fully connected layer, so as to achieve the purpose of reducing the parameters in the entire neural network. The pooling layer immediately follows the convolution layer and is composed of multiple feature surfaces like the convolution layer. , Each feature surface corresponds to a feature surface of the previous layer, and does not change the number of feature surfaces. The convolutional layer is the input layer of the pooling layer, and a feature surface of the convolutional layer uniquely corresponds to the feature surface, the neurons of the pooling layer are connected to the local acceptance domain of the input layer, and the local acceptance domains of different neurons do not overlap. The purpose of the pooling layer is to obtain features with spatial invariance by reducing the resolution of the feature surface. The pooling layer is mainly used for the extraction of secondary features. Each neuron performs a pooling operation on the local acceptance field. The window where the pooling layer slides on the previous layer is also called the pooling kernel. The convolution kernel in CNN Equivalent to pooled core Hubel-WieselType receptive field in engineering implementation, the convolution layer cells used to simulate simple Hubel-Wiese theory, pooling the cell to simulate the complex layer theory. The size of each output feature surface (number of neurons) of each pooling layer in CNN $\text{DoMapN}$ for

$$\text{DoMapN} = \frac{o\text{MapN}}{D\text{Window}}$$

Among them, the size of the pooled core is $D\text{Window}$. The size is 2 * 2. The pooling layer reduces the number of connections between the convolutional layers. The pooling operation reduces the number of neurons and the calculation amount of the network model.
In the CNN structure, after multiple convolutional layers and pooling layers, one or more fully connected layers are connected (VGG-16 is connected to three fully connected layers). Each neuron in the fully connected layer is fully connected with all the neurons in the previous layer. The fully connected layer obtains local information with class distinction in the convolutional layer or the pooling layer. The excitation of each neuron in the fully connected layer the function generally uses the ReLU function. The ReLU function is defined as

\[
f(x) = \max(0, x)
\]

The output value of the last fully connected layer is passed to an output layer, which is usually the result of the output classification. If softmax is used for classification, this layer can also be called the softmax layer. In order to reduce the problem of training overfitting, the regularization method is usually used in the fully connected layer, that is, the dropout technology. Due to the randomness of the dropout technology, the samples input to the network each time correspond to different network structures. However, all network structures share weights. Since a neuron cannot exist depending on other specific neurons, this technique reduces the complexity of neuron adaptation and strengthens the robustness of neuron learning. Classic convolutional neural networks (such as VGGNET, Alexnet, etc.) basically use ReLU + dropout technology, and also achieved good classification performance.

2.2. Vgg-16 network
VGGNet [17] was jointly proposed by the researchers of the Visual Geometry Group (Oxford University) and Google DeepMind Company. In the ILSVRC-2014 competition project, the positioning task won the first place and the classification task won the second place. The image classification data set of the ImageNet 1000 class is very expandable, and it has a good generalization ability when it is migrated to other data sets. VGGNet optimizes multiple logistic regression objective functions for model training by using mini-batch gradient descent with momentum, uses L2 regularization and uses Dropout in the first two fully connected layers. Heuristic method is used for learning The adjustment of the rate, that is, when the verification error rate stops decreasing, the current learning rate is divided by 10, and the learning rate is reduced by 3 times during the entire training process. In the random initialization weights, the normal value from the mean is 0 and the variance is 0.01. The sampling weight in the distribution, the offset is initialized to 0. When the crop size is fixed at 224 × 224, the method of random cropping from the rescaling training example is adopted, and each SGD iteration is cropped one time. In order to further expand the training set, the cropped picture is randomly flipped horizontally and randomly transferred.
Figure 1. VGGNet-16 network structure diagram

VGG has six network structures. In this paper, a 16-layer CNN (VGG-16) network is used for transfer learning. All the previous 13 layers of VGG-16 are migrated to the research of this paper, removing the last three layers of full connection. Image feature extraction. As shown in Figure 1, the VGG-16 network structure, the input of the network uses a 224 × 224 RGB color image, the preprocessing is only to subtract the average value of RGB from each pixel value in the data set. VGGNet-16 has 13 convolutional layers, followed by 3 fully connected layers, the first two fully connected layers all have 4096 channels, the third layer has a total of 1000 channels representing 1000 label categories, and the last layer is the softmax layer. There are 5 max-pooling layer, the window is 2*2, and the step size is 2. Each convolution layer is connected to an activation layer of the ReLU activation function, and then connected to the pooling layer. The VGG-16 network is characterized by the use of a smaller size convolution filter (size 3 × 3), which is conducive to local feature extraction, reduces the number of parameters, and improves the differentiation of decision functions.

3. Rotary Kiln Combustion State Recognition Method Based on CNN Transfer Learning
The rotary kiln combustion control system is shown in Figure 2. The system is mainly composed of a ccd camera, a computer with an image processing program, and a control unit. The resolution of the camera on the RGB channel is 704 * 576. The image is captured by the camera and sent to the computer for processing. The main part of the image processing program is the deep transfer learning model for combustion state recognition. The control unit can make decisions based on the combustion state.

Figure 2. Combustion control system
3.1. Analysis of the Combustion Status of the Flame Image of the Rotary Kiln

Taking the alumina rotary kiln as an example, in the production process of sintered alumina, the slurry passes through the drying belt, the pre-heating zone, the decomposition zone, the firing zone and the cooling zone [18], which are different in these five stages. The physicochemical reaction of the final clinker is obtained. In these five processes, the firing reaction occurring in the firing zone plays a decisive role in the quality of the clinker, of which the flame combustion state of the firing zone is one of the key parameters. Commonly used methods for identifying combustion status are manual identification and temperature identification.

![Flame image of the rotary kiln burning state](image)

(a) Under-combustion (b) Normal combustion (c) Over-combustion

**Figure 3. Flame image of the rotary kiln burning state**

In recent years, the method of judging the burning state of the firing band by flame image combustion state recognition has been paid more and more attention. Rotary kiln operation experts generally judge the combustion state according to the temperature of the firing zone and the flame shape. The combustion state of the rotary kiln is generally divided into "normal state" and "abnormal state". Among them, the normal state is the state of normal combustion, and the abnormal state has two states of under-burning and over-burning. "Normal burning", "over-burning" and "under-burning" are the three typical burning states of the rotary kiln. The flame shape is shown in Figure 3.

3.2. Transfer Learning Based on CNN Model

Convolutional neural network-based transfer learning [19] refers to learning and training model parameters on a specific data set using a new target data set. Transfer learning can not only learn low-level features such as color pre-training data sets and textures, but also help learn to perform advanced semantic classification on the target data set, so that the classification performance of the model is better. The model in this article pre-trains the model on an ImageNet dataset containing more than 1.2 million natural images and more than 1,000 different categories. Then transfer the model parameters to the rotary kiln flame image data set for fine-tuning training. In this experiment, the weights of specific layers of the network were migrated, training and learning were performed on the original data, and the model was fine-tuned. Migration learning based on the CNN model The method is shown in Figure 4.

![CNN-based transfer learning](image)

**Figure 4. CNN-based transfer learning**
3.3. Model Architecture Based on Vgg-16 Network Transfer Learning to Identify Combustion State

The model architecture used in this article is shown in Figure 5. The migration learning model based on the vgg-16 network identifies the combustion status of the rotary kiln. The left side of the figure is the vgg-16 network structure model, and the right side of the figure is the migration network on the left. The recognition model consists of two parts: one part is the convolution layer and the pooling layer, this part migrates the vgg-16 network structure and pre-trained parameters, and the other part is the fully connected layer, which is composed of the flame data set Trained fully connected layer.

4. Experiment and Result Analysis

4.1 Experimental Data Set
The experimental data in this paper comes from two 10-minute kiln flame monitoring videos (25 frames / second) in front of the No. 6 rotary kiln of a large alumina plant of Aluminum Corporation of China, and extracts the picture as a training data set. The image contains as shown in Figure 4. According to the on-site experience, this series of flame images are divided into three combustion states: under-combustion, normal combustion, and over-combustion by experienced kiln workers. This article extracts 3,000 pictures from the video, and Mark classification. One video data set is used as the training set. The number of images in the normal burning state of the training set is 1500, and the number of images in the under-burning state and the over-burning state is 750. Select 2000 pictures from the other video data set as the test set , The normal burning, over-burning and under-burning data sets of the test set are 1000, 500 and 500. The image size extracted from the video is 704 × 576 (rgb).

4.2 Experimental Setup
The experimental hardware devices are all carried out on CPU: Intel (R) i7-7700HQ, GPU: NVIDIA Ge-Force GTX 1080, memory 32G, using the Pytorch deep learning framework. The SGD optimizer is selected, and the parameters in the SGD method are set to : Learning rate is $10^{-4}$, The impulse factor is 0.9 , The weight attenuation factor is $10^{-6}$ The number of network training iterations is 50. In order to improve the efficiency of training, the data is preprocessed, all image data is adjusted to the same size (3 × 256 × 256), and at the same time, in order to improve the model, the image is different. The fitting ability of angle, size, position and noise adopts data enhancement techniques, including image
displacement, flipping, rotation and scaling. In order to prove that the trained model can be applied to data in new fields with no intersection between data sets, it is also verified. The effect of transfer learning was compared with two training methods, random initialization of parameters and fine-tuning of transfer.

### 4.3 Analysis of Experimental Results

![Feature map of the second convolution layer](image)

![Feature map of the third convolution layer](image)

**Figure 6.** Feature map of some layers

Figure 6 is a feature map extracted using the recognition model based on the convolutional neural network mentioned in this article, showing the hidden layer feature map of the single picture part of the migration under the ImageNet dataset using the VGG-16 network. It can be seen in this method that this method can propose some high-level abstract features, such as texture, edges, shape features, etc. Because of the large amount of data in ImageNet, the trained VGG-16 network can learn the abstract features of the samples in more detail and precision, and migrate to the flame image recognition problem to extract the abstract features of the image well, thereby improving the generalization ability of the model.

This paper implements the experiment based on vgg-16 transfer learning model. As shown in Figure 7, the convolutional neural network is used to identify the combustion state of the rotary kiln, with an accuracy rate of 96%, and the literature [21] uses traditional digital image classification methods to manually extract features. Using elm and svm classifiers, the classification results are shown in Table 1, and the accuracy is only 85%. As shown in Table 1, this method for identifying flame images based on convolutional neural networks is more accurate than traditional image recognition methods. The rate has been increased by nearly 10%. The experimental results show that the method proposed in this paper has a higher accuracy in identifying the combustion state of the rotary kiln.

| Test set (%) | Training set (%) |
|--------------|------------------|
| Manual feature extraction + elm classification [20] | 87.25 | 85.43 |
| Manual feature extraction + svm classification [20] | 87.43 | 85.65 |
| Migrating the CNN architecture of vgg-16 vgg-16 CNN architecture | 96.78 | 96.23 |
| vgg-16 CNN architecture | 82.58 | 81.13 |
The results show that compared with the initial network model, the transfer learning network model has better robustness and stronger generalization ability. The initial network model is affected by its own deep structure and training parameters, and the network training speed is slow and easy Local optimization and overfitting problems occur. The specific network layer of the transfer learning model freezes the weight parameters, and the semantic level parameters of the model are fully trained and adjusted, and the extracted target features are more clear.

5. Conclusion
The correct recognition of the combustion state is the key to the sintering quality control of the rotary kiln. Research shows that the image of the combustion zone can be used to determine the temperature state of the rotary kiln. This paper proposes a method based on the convolution neural network to identify the flame image of the rotary kiln to determine the rotary kiln combustion state. This method overcomes the problem of traditional digital image manual extraction of features and automatically extracts features. Experimental results show that this method and traditional digital image processing have higher accuracy and recognition accuracy of 96%. Therefore, the method proposed in this paper is effectively, it is of great significance to study the intelligent control system of the rotary kiln in the future to improve the combustion efficiency and reduce the emission of polluted gas.

6. References
[1] LI S W, LI B, I I C R. T he study of temperature control of cement decomposition furnace on the basis of fuzzy immune self-adaptive PID control method[C]. 2013 International Conference on Information Science and Technology. Y anzhou : IEEE .2013: 86 –90.
[2] Wang Dongfeng, Han Pu, Guo Qigang. Drnn network self-tuning PID of boiler superheated steam temperature system control. Chinese Journal of Electrical Engineering, 2004, 24 (8): 196-200.
[3] Jiang Fangle, Wang Mingchun. Fuzzy neural network control of boiler main steam temperature. Boiler Manufacturing, 2004, (4): 5-7.
[4] Cheng Qiming. Design of fuzzy neural network controller based on chaos optimization algorithm of simulated annealing strategy. Measurement and Control Technology, 2004, 23 (8): 31-33.
[5] I. Makaremi, A. Fatehi, B. N. Araabi, and M. Azizi, “Abnormal condition detection in a cement rotary kiln with system identification methods,” J. Process Control., vol. 19, no. 9, pp. 1538–1545, Jun. 2009.
[6] M. Sadeghian and A. Fatehi, “Identification, prediction and detection of the process fault in a cement rotary kiln by locally linear neuro-fuzzy technique,” J. Process Control., vol. 21, no. 2, pp. 302 –308, Feb. 2011.
[7] W. T. Li, D. Wang, and T. Y. Chai, “Flame image-based burning state recognition for sintering process of rotary kiln using heterogeneous features and fuzzy integral”, IEEE Trans. Ind. Informat., vol. 8, no. 4, pp. 780 – 790, Nov. 2012.
[8] B. Lin and S. B. Jgensen, “Soft sensor design by multivariate fusion of image features and process measurements,” J. Process Control., vol. 21, no. 4, pp. 547 – 553, Apr. 2011.
[9] M. He, J. Zhang, and M. Yan, “Texture analysis and classification for clinker in rotary kiln,” in Proc. Int. Conf. Opt. Photon. Energy Eng., May 2010, pp. 211 – 214.
[10] H. Chen, J. Zhang, X. G. Zhang, and H. P. Hu, “Recognition of sintering state in rotary kiln using a robust extreme learning machine,” in Proc. Int. Joint Conf. Neu. Netw., Beijing, China, Jul. 2014, pp. 2563 – 2570.
[11] J. Matthes, P. Waibel, and H. B. Keller, “A new infrared camera-based technology for the optimization of the Waelz process for zinc recycling.” Miner. Eng., vol. 24, no. 8, pp. 944 – 949, Jul. 2011.
[12] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.
[13] Szarvas M,Yoshizawa A,Yamamoto M.et al,Pedestrian detection with convolutional neural networks[C] proceedings of 2005 IEEE Intelligent Vehicles Symposium.2005.224-229.
[14] Abdel-Hamid O,Mohamed A,Jiang H.et al.Convolutional Neural Networks for Speech Recognition[J].IEEE/ACM Transactons on Audio, Speech, and Language Processing, 2014,22(10):1533-1545
[15] Hubel D H, Wiesel T N.Receptive fields, binocular interaction and functional architecture in the cat's visual cortes[J].The Journal of physiology, 1962,160(1):106-154
[16] Fukushima K. Neocognition :A self-organization neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological Cybernetics,1980,36(4):193-202.
[17] Simonyan K, Zisserman A, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint. 2014, 1409.1566.
[18] Fang Binhao.Research and development of temperature setting control software for alumina clinker firing rotary kiln [d].Shenyang: Northeastern University, 2006.
[19] Angelova A, Zhu S. Efficient object detection and segmentation for fine-grained recognition. In Computer Vision and Pattern Recognition(CVPR), 2013 IEEE Conference on, pp.811-818, 2013.
[20] H. Chen, X.G. Zhang, P.Y. Hong, H.P. Hu, X. Yin, Recognition of the temperature condition of a rotary kiln using dynamic features of a series of blurry flame images, IEEE Trans. Ind. Infor. 12 (1) (2016) 148-157.