

Article

Research on Coal and Rock Recognition in Coal Mining Based on Artificial Neural Network Models

Yiping Sui ¹, Lei Zhang ^{1,2,*}, Zhipeng Sun ¹, Weixun Yi ¹ and Meng Wang ¹

¹ College of Coal Engineering, Shanxi Datong University, Datong 037003, China; 210857002126@sxdtdx.edu.cn (Y.S.); 220857002150@sxdtdx.edu.cn (Z.S.); 230857002147@sxdtdx.edu.cn (W.Y.); 230857002119@sxdtdx.edu.cn (M.W.)

² Key Laboratory of Deep Coal Mining of the Ministry of Education, School of Mines, China University of Mining and Technology, Xuzhou 221116, China

* Correspondence: dtblack84@cumt.edu.cn

Abstract: In the process of coal mining, one of the main reasons for the high labor intensity of workers and the frequent occurrence of casualties is the low level of intelligence of coal mining equipment. As the core equipment in the process of coal mining, the intelligence level of shearers directly affects the safety production and mining efficiency of coal mines. Coal and rock recognition technology is the core technology used to realize the intelligentization of shearers, which is an urgent technical problem to be solved in the field of coal mining. In this paper, coal seam images, rock stratum images, and coal–rock mixed-layer images of a coal mining area are taken as the research object, and key technologies such as the construction of a sample image library, classification and recognition, and semantic segmentation are studied by using the relevant theoretical knowledge of artificial neural network models. Firstly, the BP neural network is used to classify and identify coal seam images, rock stratum images, and coal–rock mixed-layer images, so as to distinguish which of the current mining targets of a shearer is the coal seam, rock stratum, or coal–rock mixed layer. Because different mining objectives will lead to different working modes of a shearer, it is necessary to maintain normal power to cut coal when encountering a coal seam, to stop working when encountering rock stratum, and to cut coal along the boundary between a coal seam and rock stratum when encountering a coal–rock mixed stratum. Secondly, the DeepLabv3+ model is used to perform semantic segmentation experiments on the coal–rock mixed-layer images. The purpose is to find out the distribution of coal and rocks in the coal–rock mixed layer in the coal mining area, so as to provide technical support for the automatic adjustment height of the shearer. Finally, the research in this paper achieved a 97.16% recognition rate in the classification and recognition experiment of the coal seam images, rock stratum images, and coal–rock mixed-layer images and a 91.2% accuracy in the semantic segmentation experiment of the coal–rock mixed-layer images. The research results of the two experiments provide key technical support for improving the intelligence level of shearers.

Keywords: coal mining area; coal and rock identification; wavelet decomposition; neural network; semantic segmentation



Citation: Sui, Y.; Zhang, L.; Sun, Z.; Yi, W.; Wang, M. Research on Coal and Rock Recognition in Coal Mining Based on Artificial Neural Network Models. *Appl. Sci.* **2024**, *14*, 864. <https://doi.org/10.3390/app14020864>

Academic Editor: Nikolaos Koukouras

Received: 6 December 2023

Revised: 8 January 2024

Accepted: 11 January 2024

Published: 19 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Safety is the most important concern in coal mine production, and one of the main reasons for the frequent occurrence of casualties in coal mines is the low level of automation and intelligence of coal mining equipment [1]. Therefore, it is necessary to improve the automation level of coal mining equipment, so as to reduce the labor intensity of coal miners and to protect the life safety of workers.

A shearer is the core equipment for coal mining, and its continuous and safe operation is an important guarantee for efficient coal production [2]. The coal and rock recognition technology for the mining area of a shearer is the key to realizing the intelligent control of the shearer, and it is also an important basis for realizing the intelligent mining of

coal [3,4]. The existing coal–rock recognition technology includes image recognition, multi-sensor fusion recognition, process signal monitoring and recognition, electromagnetic wave recognition, and ultrasonic detection and recognition. There are many kinds of coal–rock recognition technologies, but they are rarely used in actual mining [5]. The main reasons are as follows. First, the underground environment of coal mines is complex, including high-density dust, electromagnetic fields generated by high-power equipment, changes in light intensity, and strong vibrations and noise generated by various equipment [6]. Second, the geological conditions are complex and changeable, including faults, cracks, folds, and magma intrusion, which will lead to changes in the thickness and hardness of coal seams [7]. These all bring challenges to the recognition of coal and rock characteristics. The use of deep learning for image recognition has the characteristics of a low cost and real-time and strong applicability, which is also the focus of this paper.

The coal and rocks (including waste rocks) cut by a shearer will enter the mine conveyor together and are transported to the ground by a transmission belt after passing through a mine crusher for sorting (machine sorting or manual sorting). The separation process will separate the coal and rocks and will screen the rocks. Those rocks that still have industrial uses will be separated for further treatment, while waste rocks such as coal gangue will be stacked in places with little environmental pollution or used to backfill coal mining subsidence areas. Coal gangue is close to coal in color and is a black-gray rock with a greater hardness than coal. However, the image recognition technology based on visual features proposed in this paper can still distinguish the two sheets. The neural network model designed in this paper relies on extracting and comparing the gray features and texture features of the target to identify coal and rocks. Although the color of coal gangue is close to that of coal, there are great differences in texture features. Compared with gangue, the surface of coal is more glossy and easier to reflect, while the surface of gangue has a more diffuse reflection of light and thus almost no reflection. In this paper, in the study of coal and rock identification technology in a fully mechanized coal mining face, coal gangue is classified as a kind of rock, and coal is distinguished.

More and more experts and scholars use image recognition technology to solve the problem of coal and rock recognition in coal mine production. Si Lei et al. proposed a method of a convolutional neural network to identify the coal–rock interface. Aiming at the problem of over-fitting caused by few training samples, three regularization methods, a dropout, L2 regularization, and batch normalization, were used to construct the coal–rock image data set, and the coal–rock image data set was enhanced by adding noise, image scaling, and image rotation. The experimental results show that the designed network has a better coal–rock image recognition performance [8]. A new coal–rock image segmentation algorithm model, the Coal–Rock Pyramid Network, was proposed by Gao et al., which was trained by collecting three types of images of coal–rock integrity, crack shadows, and dark light. Compared with U-net and Segnet image segmentation algorithms, the recognition accuracy of this algorithm was 2.46% and 7.64% higher than that of the U-net algorithm and Segnet algorithm, respectively, reaching 91.54% [9]. Sun et al. used deep separable convolution to improve the YOLOv3 algorithm. In view of the uncountable and continuous characteristics of the coal–rock interface, the ratio of the total projection length of the prediction frame in the x and y directions to the total projection length of the coal–rock interface curve is used as the evaluation standard of target detection accuracy. The results show that the recognition accuracy of the improved YOLOv3 model in the x direction is 5.85% higher than that of the original model, which is 89.89%. The recognition accuracy in the y direction is 16.99% higher than that of the original model, which is 73.30% [10].

The coal and rock recognition technology for coal mining areas mainly includes two aspects. (1) A judgment of the cutting target of the shearer: There are three situations in the mining area when the shearer works [11]. First, it is necessary to maintain the current operating power to cut coal when encountering coal seams. The second is to stop the operation of the shearer when it encounters the rock layer and to start cutting coal again after selecting the appropriate way to complete the rock breaking. Thirdly, when the

coal–rock mixed layer is encountered, that is, the mining area contains both a coal seam and rock stratum, it is necessary to adjust the cutting height of the shearer drum to ensure that the drum can cut coal along the boundary between the coal seam and rock stratum, so as to avoid encountering rocks in the cutting process. (2) The determination of the distribution state of the coal seam and rock stratum in coal–rock mixed layers: the distribution state of a coal seam and rock stratum in a coal–rock mixed interface is accurately identified by coal and rock recognition technology, so that the shearer can cut coal along the boundary of the coal–rock mixed layer [12].

2. Coal and Rock Image Acquisition and Image Feature Extraction

2.1. Collection of Coal Image and Rock Image Samples in Coal Mining Area

The application of deep learning in image recognition requires a large number of image samples as support [13]. This paper designs an image acquisition system for a coal mining area, as shown in Figure 1.

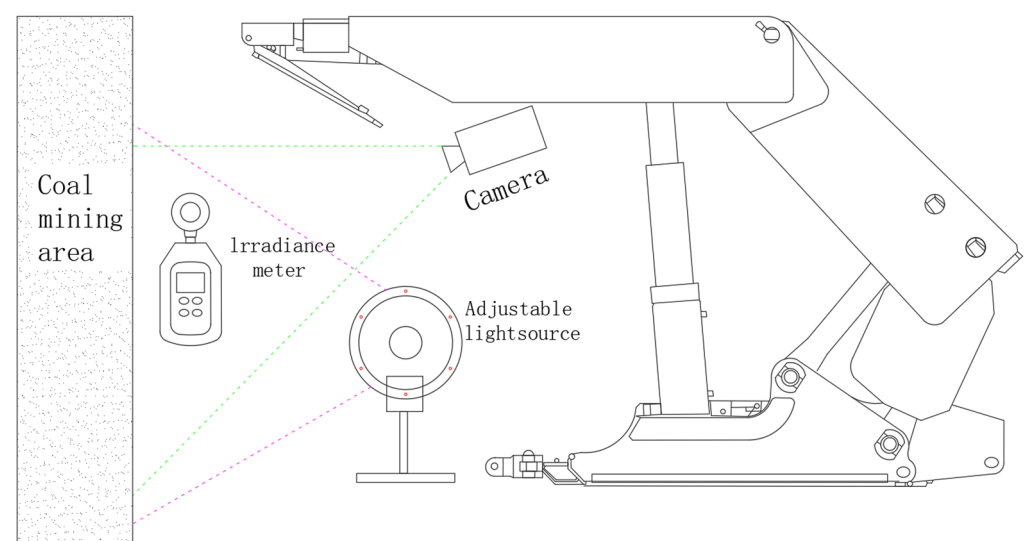


Figure 1. Coal mining area image acquisition system.

Illumination is an important factor affecting image quality and parameters. Even for the same piece of coal or rock, the characteristic parameters extracted under different illumination conditions are not the same [14]. The illumination conditions in coal mines are poor and uneven, so an adjustable light source is designed to fill the mining area, and the illuminance value of the surface of the mining area is measured in real time using an illuminance meter. In this paper, 7000 Lux is set as the standard illumination value when the sample image is collected, and the adjustable light source is used to ensure the unity of the experimental background conditions. After ensuring that the illumination value is unchanged, a mine explosion-proof camera is used to take pictures of the coal mining area to obtain the sample pictures.

The experimental site is the 52603 fully mechanized mining face of the Tashan Coal Mine in Shanxi Province, with a total length of 135 m. The mine explosion-proof camera is used to video capture the mining area of the whole fully mechanized mining face. The captured video is intercepted every 1 s, and a total of 1860 pictures are obtained. Then, the sample maps that meet the experimental requirements are selected, and the number of samples is expanded by image cropping. A total of 270 sample maps that meet the experimental requirements are obtained. The acquired sample images are classified into coal seam images, rock stratum images, and coal–rock mixed-layer images. At the same time, these pictures are numbered. Some samples are shown in Figure 2. Finally, the sample images are randomly divided into training sets and test sets in a ratio of 2:1 for deep learning training.

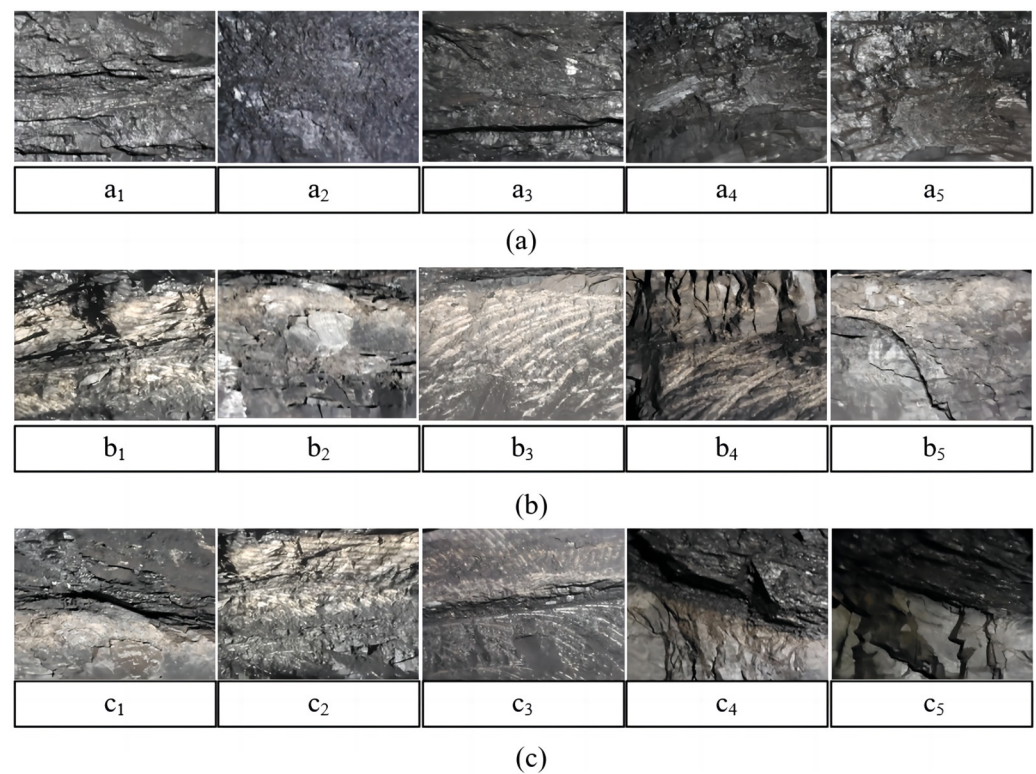


Figure 2. Coal mining area image acquisition system. (a) Coal seam sample image; (b) Rock stratum sample image; (c) Coal-rock mixed-layer sample image.

2.2. Image Feature Extraction Based on Wavelet Three-Level Decomposition

The experimental site is the 52603 fully mechanized mining face of the Tashan Coal Mine in Shanxi Province, with a total length of 135 m. A mine explosion-proof camera is used to video capture the mining area of the whole fully mechanized mining face. The captured video is intercepted every 1 s, and a total of 1860 pictures are obtained. Then, the sample maps that meet the experimental requirements are selected, and the number of samples is expanded by image cropping. A total of 270 sample maps that meet the experimental requirements are obtained. The acquired sample images are classified into coal seam images, rock stratum images, and coal-rock mixed-layer images and numbered. Some samples are shown in Figure 2. Finally, the sample images are randomly divided into training sets and test sets in a ratio of 2:1 for deep learning training. When using the neural network model for coal and rock recognition, it is preferred to determine the type of image features and to extract the eigenvalues. In this paper, a feature extraction method using wavelet three-level decomposition to extract texture features of images is proposed. The three-level wavelet decomposition is to extract image features by the multi-scale and multi-resolution decomposition of the image. It can fully reflect the difference of coal seam image, rock stratum image, and coal-rock mixed-layer image features while avoiding the redundancy and over-fitting problems caused by too many features and can thus ensure the high-precision and fast recognition of coal images and rock images by the network model [15,16].

The principle of the wavelet decomposition of an image is to perform convolutional filtering on the image through a low-pass filter and a high-pass filter and then to perform a two-to-one downsampling. Each decomposition continuously decomposes the low-frequency part of each level, and the high-frequency part does not move [17,18]. The principle of wavelet three-level decomposition is shown in Figure 3. In the figure, L and H represent low-frequency and high-frequency information, respectively. LL is a low-frequency sub-band, which is an approximate representation of the image. HL, LH,

and HH are all high-frequency sub-bands, which represent the details of the image in all directions. The subscript represents the level of wavelet decomposition.

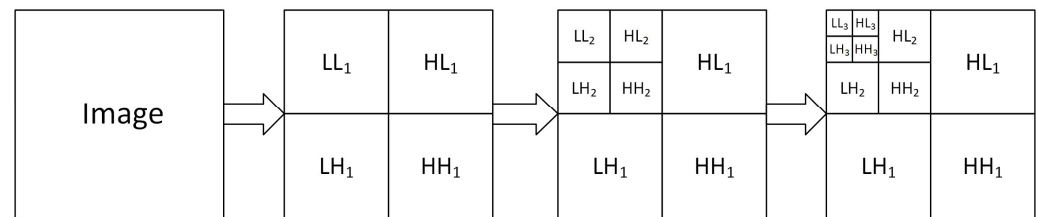


Figure 3. Wavelet three-level decomposition schematic diagram.

Three kinds of images (coal seam images, rock images, and coal–rock mixed-layer images) are decomposed by wavelet three-level decomposition, and the wavelet coefficients of their low-frequency sub-bands (LL sub-bands) are calculated as the eigenvalues of image texture features. In this section, the wavelet three-level decomposition results of coal seam images are taken as an example to demonstrate the texture feature extraction results of images based on wavelet three-level decomposition, as shown in Figure 4a. In the experiment, the wavelet coefficients calculated by the LL_1 sub-band, that is, the texture features of the coal seam image extracted in the LL_1 sub-band, are drawn into a network diagram, as shown in Figure 4b. The wavelet coefficients calculated from the LL_2 sub-band, that is, the texture features extracted from the LL_2 sub-band, are drawn into a network, as shown in Figure 4c. The wavelet coefficients calculated from the LL_3 sub-band, that is, the texture features extracted from the LL_3 sub-band, are drawn into a network, as shown in Figure 4d.

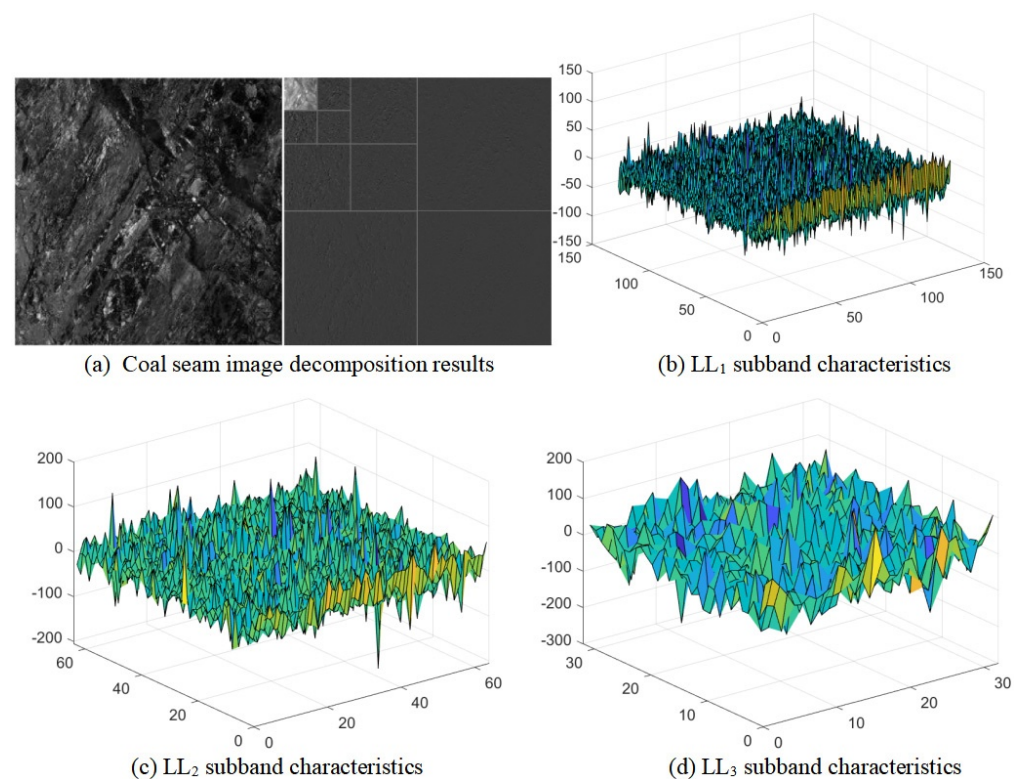


Figure 4. Coal seam image decomposition results and sub-band characteristic network diagram. (a) Three-level wavelet decomposition results of coal seam image; (b) Characteristics of LL_1 sub-band in wavelet first-order decomposition; (c) Characteristics of LL_2 sub-band in wavelet second-order decomposition; (d) Characteristics of LL_3 sub-band in wavelet third-order decomposition.

As the decomposition level increases, the area of the decomposed image will decrease exponentially, that is, the extracted feature information will decrease exponentially. The LL_1 sub-band is $1/4$ of the area of the original image, which contains the most feature information, such as Figure 4b. The LL_2 sub-band is $1/16$ of the area of the original image, which contains less feature information but more than the LL_3 sub-band, such as Figure 4c. The LL_3 sub-band is $1/64$ of the area of the original image, which contains the least feature information, such as Figure 4d.

The wavelet coefficient of the decomposed coal and rock images is calculated, and the final calculation result is the extracted image texture feature value. The low-frequency part coefficient and the high-frequency part coefficient of the wavelet decomposition can be calculated by calculating the neighborhood pixel value of the image. The calculation method is shown in Formulas (1) and (2), where L_i is the low-frequency Wavedec2 two-dimensional wavelet coefficient, H_i is the high-frequency Wavedec2 two-dimensional wavelet coefficient, i is the Wavedec2 wavelet coefficient index number, and P is the digital image pixel value.

$$L_i = (P_{2i} + P_{2i+1}) \frac{1}{\sqrt{2}} \quad (1)$$

$$H_i = (P_{2i} - P_{2i+1}) \frac{1}{\sqrt{2}} \quad (2)$$

The coal seam image is imported into the two-dimensional wavelet decomposition program based on Wavedec2, written by Matlab 2021b, and then the image is decomposed by wavelet three-level decomposition, resulting in a low-frequency sub-band and nine high-frequency sub-bands. In the experiment, the total amount of eigenvalues extracted after each decomposition is defined as $Show_i$ ($i = 1, 2, 3$), where i represents the decomposition level, and the statistical method is shown in Formula (3).

$$\begin{cases} Show_1 = LL_1 + LH_1 + HL_1 + HH_1 \\ Show_2 = Show_1 + LH_2 + HL_2 + HH_2 \\ Show_3 = Show_2 + LH_3 + HL_3 + HH_3 \end{cases} \quad (3)$$

The total amount of eigenvalues extracted from each layer of wavelet decomposition is counted and drawn into a network diagram, as shown in Figure 5. The experimental results show that with the increase of the number of decomposition layers, Figure 5a to Figure 5c, the total amount of texture feature value data contained in the mesh graph increases. This proves that as the number of wavelet decomposition layers increases, the total amount of texture feature values extracted from images also increases.

In order to improve the accuracy of the network model for image classification and recognition, some statistics of texture features are added in the experiment to increase the difference of eigenvalues. The coal and rock images are decomposed by wavelet three-level decomposition. At each level of the decomposition scale, the angular second moment, contrast, correlation, mean, and variance of the low-frequency coefficient image are calculated.

The angular second moment of the low-frequency coefficient image is calculated by Formula (4), where $\beta_{LL}(i)$ is the contrast of the LL sub-band when the sub-band scale value is i ($i = 1, 2, 3$); i is the decomposition level; LL is the low-frequency sub-band; L is the number of gray levels; and $P(m, n)$ is the probability of gray level n , starting from gray level m .

$$\beta_{LL}(i) = \sum_{m=0}^{L-1} \sum_{n=0}^{L-1} P(m, n)^2 \quad (4)$$

The contrast of the low-frequency coefficient image is calculated by Formula (5), where $\alpha_{LL}(i)$ is the contrast of the LL sub-band when the sub-band scale value is i ($i = 1, 2, 3$); and t is the gray level.

$$\alpha_{LL}(i) = \sum_{n=0}^{L-1} t^2 \left\{ \sum_{m=0}^{L-1} \sum_{n=0}^{L-1} P(m, n) \right\} \quad (5)$$

The correlation of low-frequency coefficient images is calculated by Formula (6), where $G_{LL}(i)$ is the correlation of the LL sub-bands when the sub-band scale value is i ($i = 1, 2, 3$); μ_1 and μ_2 are the mean values of $P(m, n)$; and σ_1 and σ_2 are the variances of $P(m, n)$.

$$G_{LL}(i) = \sum_{m=0}^{L-1} \sum_{n=0}^{L-1} \frac{mnP(m, n) - \mu_1\mu_2}{\sigma_1^2\sigma_2^2} \quad (6)$$

The mean value of the low-frequency coefficient image is calculated by Formula (7), where $E_{LL}(i)$ is the mean value of the LL sub-band when the sub-band scale value is i ($i = 1, 2, 3$).

$$E_{LL}(i) = \sum_{m=0}^{L-1} \sum_{n=0}^{L-1} mP(m, n) \quad (7)$$

The variance of the low-frequency coefficient image is calculated by Formula (8), where $\sigma_{LL}(i)$ is the variance of the LL sub-band when the sub-band scale value is i ($i = 1, 2, 3$); and μ is the mean value of $P(m, n)$.

$$\sigma_{LL}(i) = \sum_{m=0}^{L-1} \sum_{n=0}^{L-1} (m - \mu)^2 P(m, n) \quad (8)$$

A total of 15 texture features were extracted based on wavelet three-level decomposition and were applied to the training of network models. The eigenvalues of the 15 texture features extracted from some sample images are shown in Table 1. Take $\beta_{LL(1)}$ as an example, where $\beta_{LL(1)}$ represents the first-order angular second-order moment of the low-frequency LL sub-band, and the number 1 in the bracket represents the number of wavelet decompositions.

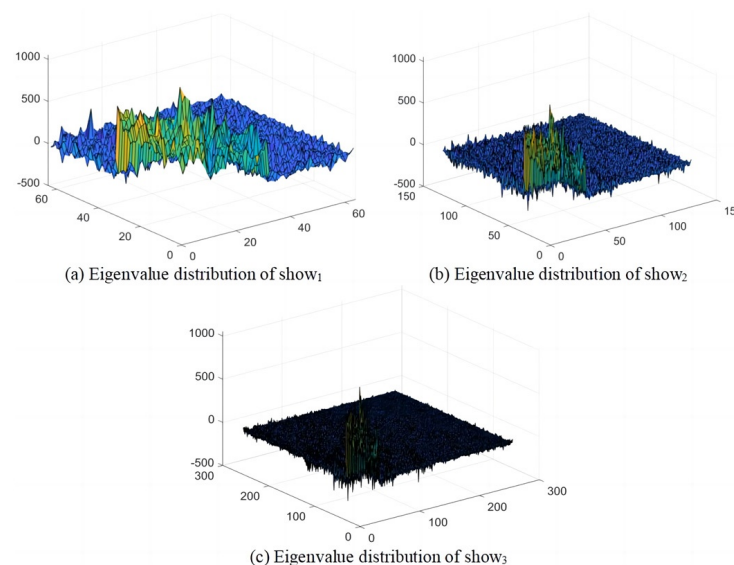


Figure 5. Sub-band characteristics of wavelet decomposition in each layer of a coal seam image. (a) Wavelet first-order decomposition eigenvalue extraction results (Show₁); (b) Wavelet second-order decomposition eigenvalue extraction results (Show₂); (c) Wavelet third-order decomposition eigenvalue extraction results (Show₃).

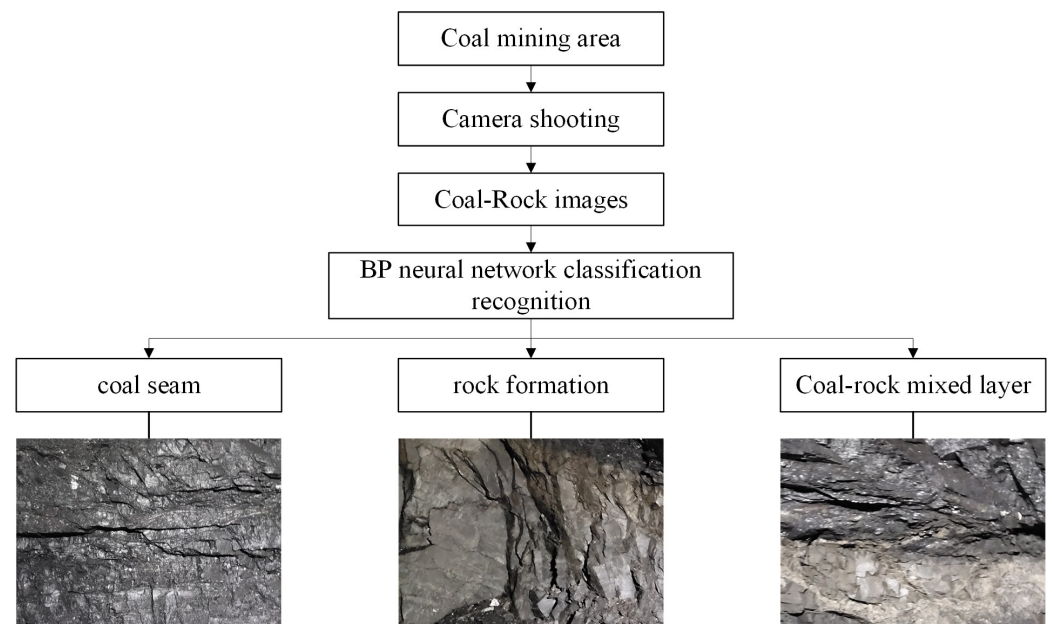
Table 1. Eigenvalues extracted from some sample images.

Feature Number	Coal Seam Characteristic Value	Rock Characteristic Value	Characteristic Value of Coal–Rock Mixed Layer
β_{LL} (1)	0.00983	0.0057	0.0038
α_{LL} (1)	46.496	79.853	133.133
G_{LL} (1)	0.0293	0.0368	0.0192
E_{LL} (1)	119.345	144.28	219.836
σ_{LL} (1)	25.6	41.696	14.42
β_{LL} (2)	0.014	0.00766	0.00375
α_{LL} (2)	29.361	74.28	202.18
G_{LL} (2)	0.0215	0.0193	0.0414
E_{LL} (2)	221.5	255.691	301.225
σ_{LL} (2)	41.491	65.2	61.579
β_{LL} (3)	0.0184	0.00977	0.0057
α_{LL} (3)	101.577	139.641	369.682
G_{LL} (3)	0.00681	0.00694	0.00607
E_{LL} (3)	477.1	426.608	786.49
σ_{LL} (3)	70.662	89.390	59.581

3. Coal and Rock Recognition Experiment Based on the BP Neural Network

3.1. BP Neural Network Construction and Matlab Simulation Experiment

There are three situations that a shearer will encounter when working in a mining area, which are coal seams, rock layers, and coal–rock mixed layers. In order to improve the intelligence level of the shearer, it is necessary to realize the coal and rock recognition technology, so that the shearer can automatically change the working mode according to the coal and rock recognition results of the computer. The purpose of this chapter is shown in Figure 6.

**Figure 6.** Purpose of the experiment in this chapter.

In this section, the BP neural network is designed to carry out a coal and rock recognition experiment. The BP neural network is a multi-layer feedforward network trained by error back propagation. Based on the gradient descent method, the mean square error of the actual output value and the expected output value of the network is minimized. The algorithm includes two processes: forward propagation of the signal and back propagation of the error. Based on the previous experimental results, a BP neural network with an input

layer neuron number of 15, a hidden layer neuron number of 50, an output layer neuron number of 3, an initial weight of 0.5, a learning rate of 0.1, an expected error of 10^{-1} , and a transfer function of the Sigmoid function is built. The BP neural network architecture of the experimental design is shown in Figure 7.

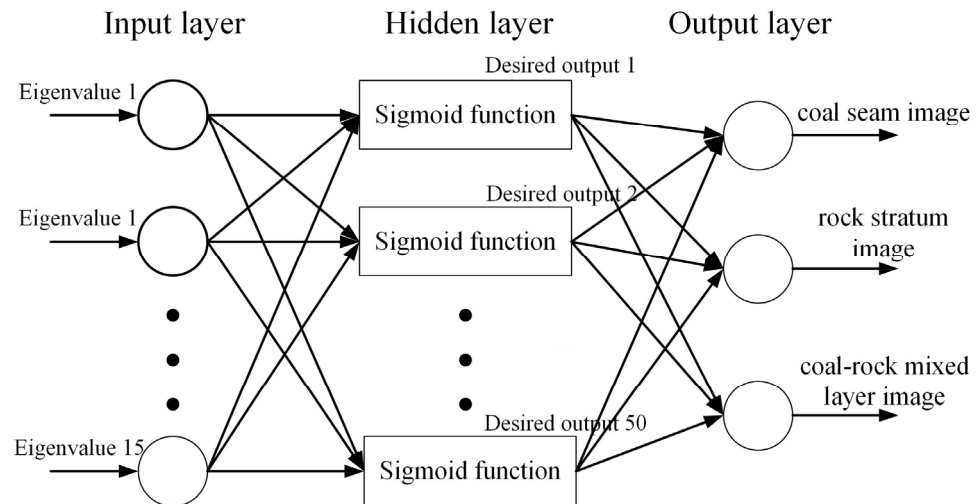


Figure 7. BP neural network architecture for experimental design.

In order to expand the scale of the data set and to enhance the characteristics of the data set, the data enhancement method is used to expand the image data set of the coal and rock samples. The common methods of data enhancement used in this training set include image size scaling, rotation, pixel translation, changing contrast, perspective transformation, adding noise, and so on. Through data enhancement, the number of sample images in the experimental training set increased from 270 to 1347. Data enhancement is performed on the sample images, as shown in Figure 8.

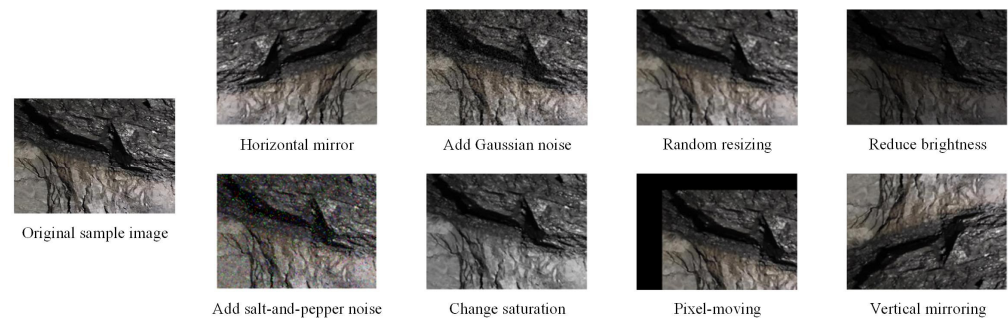


Figure 8. Data enhancement.

The Matlab simulation experiment of coal and rock image recognition is carried out with the trained BP neural network. A total of 1347 images of experimental samples are randomly divided into a training set and test set, according to the ratio of 2:1. The training set is used to determine the network threshold and weight, and the test set is used to test the classification performance of the network. The Matlab simulation results of the coal and rock recognition of the BP neural network model are shown in Figure 9. In this figure, ordinate 1 represents that the model recognition target is a coal seam, ordinate 2 represents that the model recognition target is a rock stratum, and ordinate 3 represents that the model recognition target is a coal–rock mixed layer. In Figure 9, the red part represents the real type of image samples, and the blue part represents the image type predicted by the BP neural network model. The higher the degree of overlap between the two representations, the higher the accuracy of the model prediction. Observing the experimental results in

Figure 10, it is found that the recognition rate of the model to the training set is 98.76%, and the recognition rate of the test set is 97.16%. The recognition rate of the test set represents the classification performance of the network model, that is, the recognition rate of the BP neural network model to the coal and rocks is 97.16%.

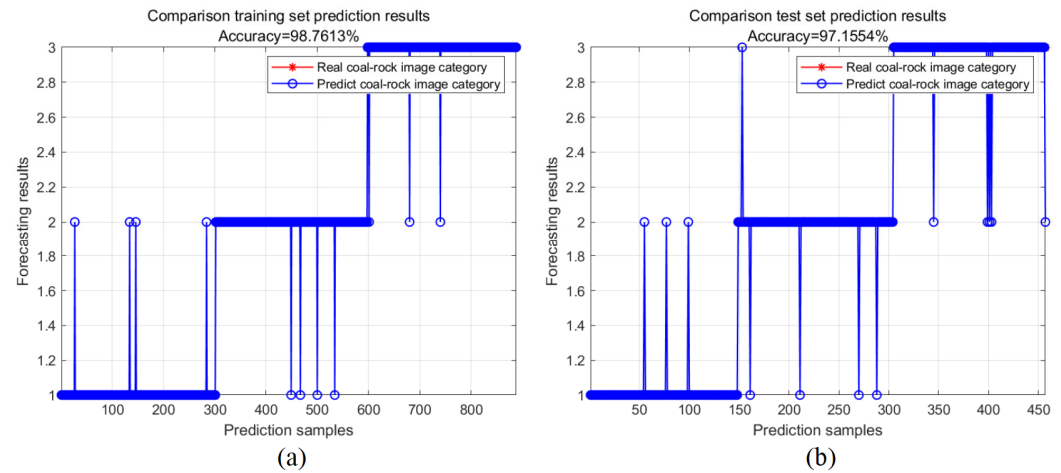


Figure 9. Matlab simulation results of BP neural network. (a) Training set recognition rate; (b) Test set recognition rate.

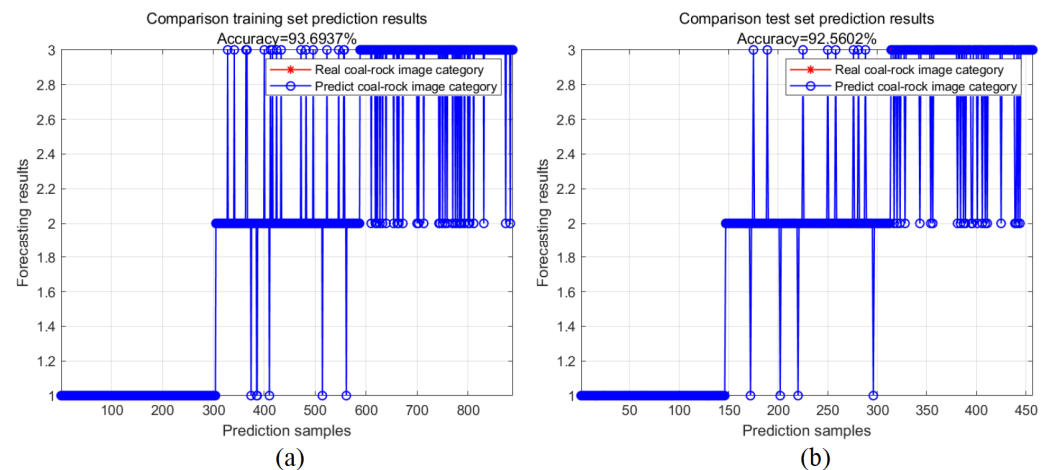


Figure 10. Matlab simulation results of VGG16 network model. (a) Training set recognition rate; (b) Test set recognition rate.

3.2. Control Group Experiment Based on Convolutional Neural Network

In order to verify that the BP neural network performs better than other models in coal and rock recognition experiments, this section also uses two convolutional neural networks for image classification and recognition experiments. They are the VGG16 network model and LeNet-5 network model. The Matlab simulation experiment of coal and rock recognition in the control group used a total of 540 sample images, and the sample images were randomly divided into a training set and a test set, according to the ratio of 2:1. The structure of these two convolutional neural networks and the results of the Matlab simulation experiments are as follows:

(1) VGG16 network model

VGG16 is an earlier network model for image recognition based on visual features. It is characterized by the use of small convolutions to improve performance by increasing network depth. At the same time, it also uses smaller convolution kernels to increase the nonlinear expression ability of the model. Therefore, the model obtains a larger receptive

field and also makes the parameters of the network model smaller. The multi-layer activation function makes the network more capable of learning features [19]. The Matlab simulation results of coal-rock image recognition based on the VGG16 network model are shown in Figure 10. It is found, in Figure 10, that the recognition rate of the model to the training set is 93.69%, and the recognition rate to the test set is 92.56%. The recognition rate of the test set represents the classification performance of the network model, that is, the recognition rate of the VGG16 network model to the coal and rocks is 92.56%.

(2) LeNet-5 network model

Each layer of the LeNet-5 network model contains trainable parameters, and each layer has multiple feature maps. Each feature map extracts a feature through a convolutional filter, and each feature map has multiple neurons, so the model has good learning and recognition capabilities [20]. The Matlab simulation results of coal and rock recognition based on the LeNet-5 network model are shown in Figure 11. It is found, in Figure 11, that the recognition rate of the model to the training set is 90.77%, and the recognition rate to the test set is 89.28%. The recognition rate of the test set represents the classification performance of the network model, that is, the recognition rate of the volume LeNet-5 network model to the coal and rocks is 89.28%.

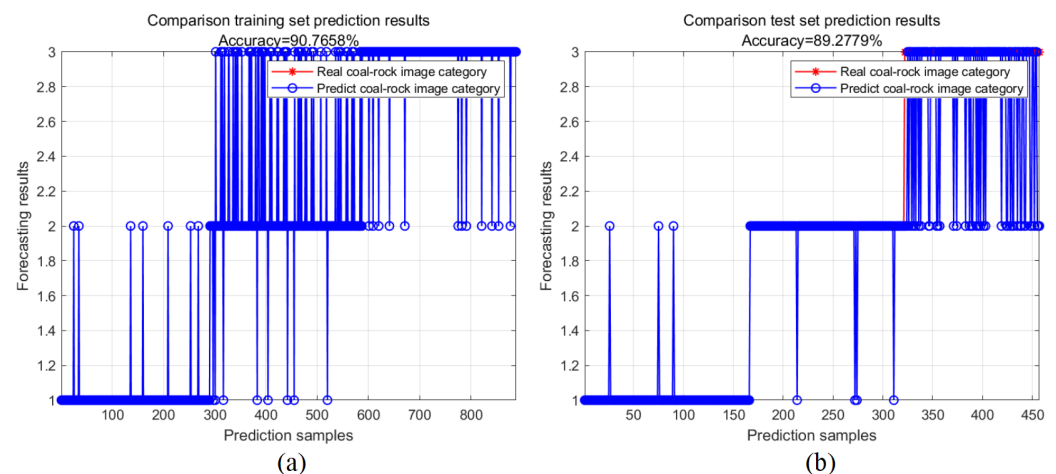


Figure 11. Matlab simulation results of LeNet-5 network model. (a) Training set recognition rate; (b) Test set recognition rate.

3.3. Performance Comparison of Network Models

The performance of the network model is not only determined by the recognition rate but also by the accuracy, precision, recall, specificity, and F1-score value of the five network model evaluation indicators, and these evaluation indicators are calculated by the confusion matrix of the network model test set [21,22]. In order to compare which of the four network models in this chapter has the best performance and to select the best coal and rock recognition network model, it is necessary to calculate the confusion matrix of the test set of the simulation experiment results. The confusion matrix principle is shown in Figure 12, where TP represents the true class, FN represents the false negative class, FP represents the false positive class, and TN represents the true negative class [23].

The concept of the evaluation indicators of these five types of network models is as follows, and the calculation method is shown in Table 2. The higher their values are, the better the effect of model prediction is. The recall rate represents the ratio of the number of samples that the model correctly identifies as positive to the total number of positive samples. The accuracy rate represents the ratio of the number of samples recognized by the model as a positive class to the total number of samples. The precision rate represents the ratio of the number of samples identified as positive classes to the total number of positive class samples. The specificity represents the ratio of the number of samples identified as negative classes by the model to the total number of negative classes. The F1-score is a

new evaluation index obtained by a comprehensive consideration of the recall rate and precision rate [24].

Confusion matrix		Predicted label	
		Positive	Negative
Real category	Positive	TP	FN
	Negative	FP	TN

Figure 12. Confusion matrix structure diagram.

Table 2. The calculation method of the network performance evaluation index.

Network Performance Evaluation Index	Computing Formula
Recall rate	$Recall = \frac{TP}{TP+FN}$
Accuracy rate	$Accuracy = \frac{TP+TN}{TP+FN+FP+TN}$
Precision rate	$Precision = \frac{TP}{TP+FP}$
Specificity	$Specificity = \frac{TN}{TN+FP}$
F1-score	$F1-score = 2 \frac{Precision \cdot Recall}{Precision+Recall}$

After the training of the network model, the performance evaluation indexes of the above five network models are calculated by the confusion matrix of the test set to compare and analyze the performance of the three coal–rock image-recognition models: the BP neural network model, the VGG16 network model, and the LeNet-5 network model. The test set confusion matrix of the simulation results of the above three network models is shown in Figure 13.

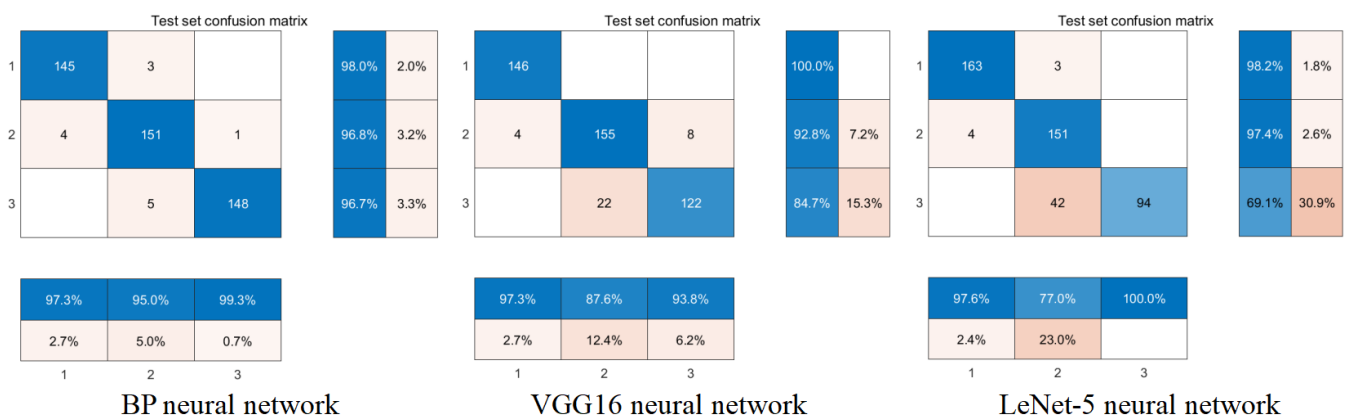


Figure 13. Confusion matrix of three network models.

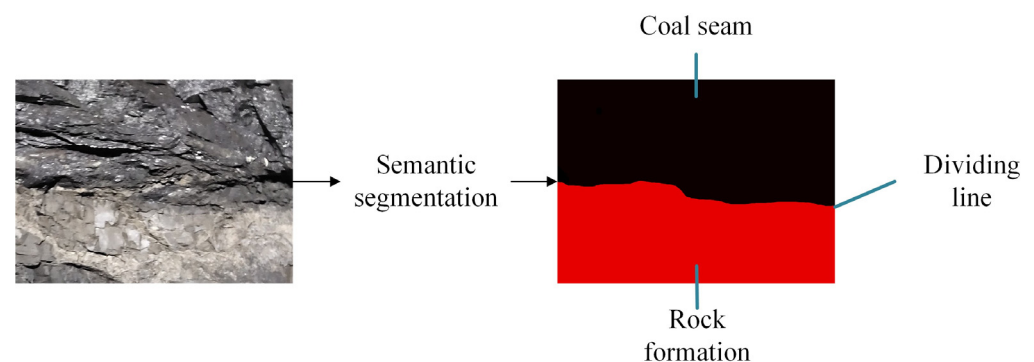
The calculation results of the performance evaluation indexes of each network model are shown in Table 3. Through calculations, the recall rate, accuracy rate, precision rate, specificity, and F1-score value of the BP neural network are the highest, which are 96.7%, 97.16%, 95%, 96.32%, and 95.84%, respectively. Therefore, it is concluded that the BP neural network has the best performance in solving the problem of coal and rock classification and recognition.

Table 3. The calculation results of network model performance evaluation index.

Network Model Name	Accuracy Rate	Precision Rate	Recall Rate	Specificity	F1-Score
BP neural network	97.16%	95%	96.7%	96.32%	95.84%
VGG16 network	92.56%	87.6%	84.7%	91%	86.12%
LeNet-5 network	89.28%	77%	69.91%	89.27%	73.29%

4. The Image Semantic Segmentation Experiment of the Coal–Rock Mixed Layer

There is a coal–rock mixed layer in the mining area of the fully mechanized coal mining face. In order to realize the automatic height adjustment of the shearers, it is necessary to determine the distribution state of coal and rocks in the coal–rock mixed layer. This paper proposes a method to determine the distribution state of coal and rocks in the coal–rock mixed layer in the coal mining area by using machine semantic segmentation technology, that is, the coal and rocks in the coal–rock mixed images are segmented with different colors. The principle of image semantic segmentation is that the machine marks the category of the object to which each pixel belongs in the image to achieve pixel-level image recognition. For the mixed image of coal and rocks, semantic segmentation can distinguish coal and rocks with different colors, according to different pixels, so as to determine the distribution state of coal and rocks in the mixed layer of coal and rocks in the mining area and to divide the boundary between a coal seam and rock stratum. The purpose of this experiment is shown in Figure 14.

**Figure 14.** The purpose of semantic segmentation experiment.

In order to determine the distribution state of coal and rocks in the coal–rock mixed layer in the coal mining area, this chapter uses the Deeplabv3+ network model to perform semantic segmentation experiments on coal–rock mixed images. Deeplabv3+ adopts the encoder–decoder structure. In the encoder, the initial effective feature layer compressed four times is extracted with different sizes of dilated convolutions, then merged, and then compressed with 1×1 convolutions. In the decoder, the initial effective feature layer compressed twice is used to adjust the number of channels by 1×1 convolutions and then stacked with the results of the effective feature layer upsampling after the hole convolution [25,26]. After the stack is completed, two depthwise separable convolutions are performed to obtain a final effective feature layer, which is the feature concentration of the entire picture. Finally, the model uses a 1×1 convolution to adjust the final effective feature layer and uses the resize function to upsample so that the width and height of the final output layer are consistent with the input image [27,28]. The overall architecture of the DeepLabv3+ model is shown in Figure 15.

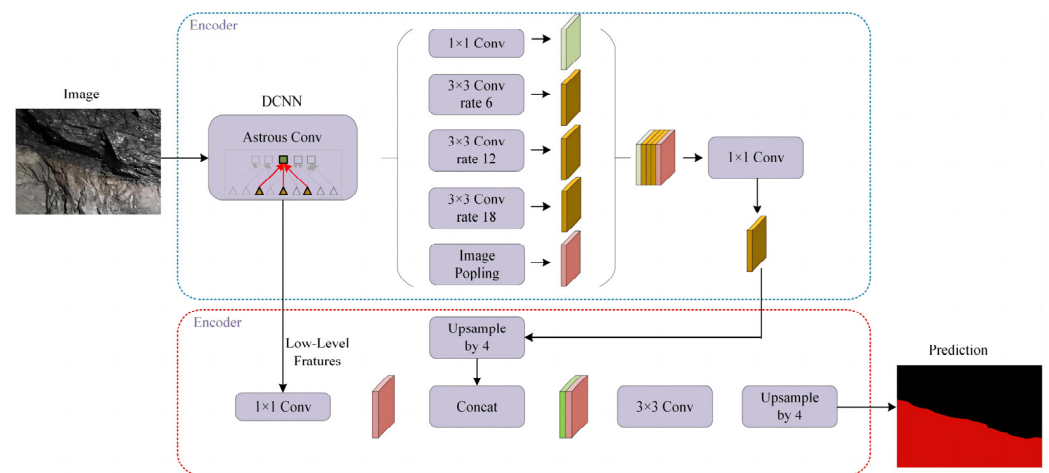


Figure 15. Deeplabv3+ network structure diagram.

Figure 16 shows the changes of accuracy and loss function in the training process of the Deeplabv3+ network model under the Adam adaptive algorithm. The loss function is used to represent the gap between the predicted results and the actual data. The smaller the loss function, the better the robustness of the model.

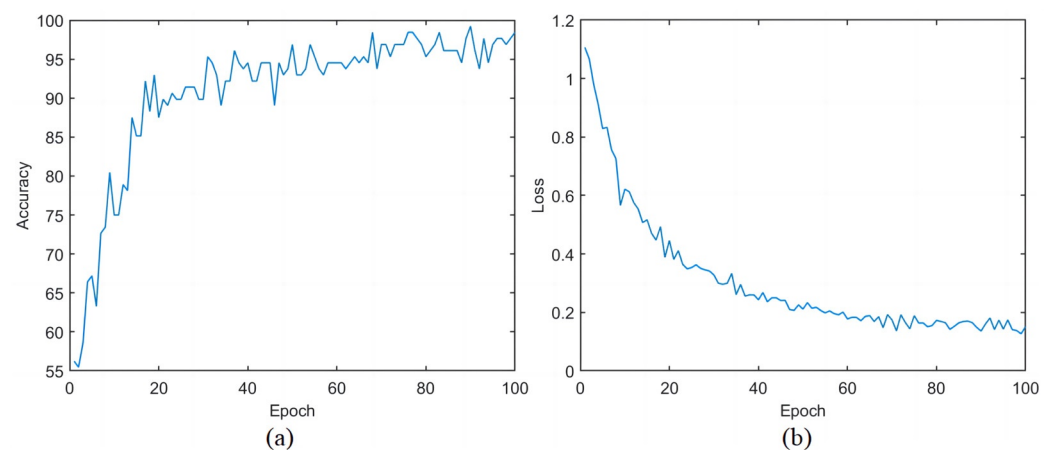


Figure 16. Deeplabv3+ network model training curve. (a) Variation curve of accuracy; (b) Loss function change curve.

In order to verify that the Deeplabv3+ network model performs better than other models in the semantic segmentation experiment of coal–rock mixed-layer images, this section also uses two other types of semantic segmentation models for experiments. They are the FCN-8s network model and U-Net network model.

The size of the coal–rock mixed-layer image data set after cutting is 600×400 , and there are 100 images in the test set. This semantic segmentation experiment only contains two objects, so the segmented image is processed into an image containing only red and black colors in order to distinguish different types of objects, in which the black part represents coal, and the red part represents rock. The results of the semantic segmentation of five groups of sample images are selected as the experimental results for demonstration, as shown in Figure 17. From left to right are the following: the input image, the Deeplabv3+ semantic segmentation result, the FCN-8s semantic segmentation result, and the U-Net semantic segmentation result.

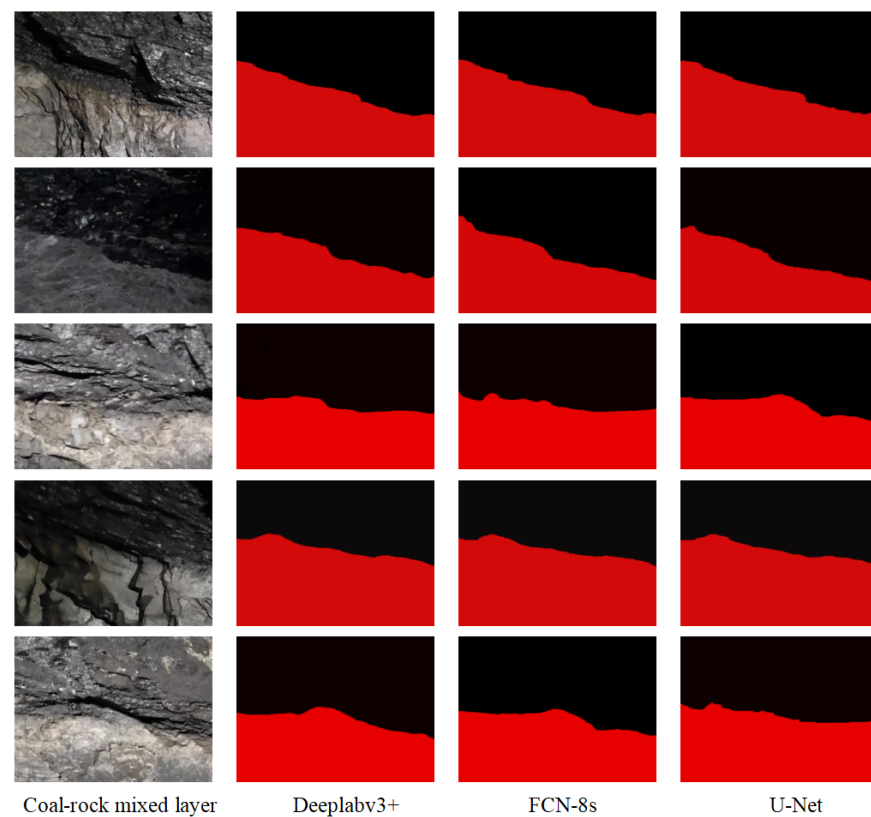


Figure 17. Semantic segmentation results of three types of network models.

Semantic segmentation is a pixel-level classification, which commonly uses the concept of intersection over union (IOU) in target monitoring as an evaluation index [29]. The intersection and union ratio calculates the overlap rate of the ‘predicted border’ and the ‘real border’, that is, the ratio of their intersection and union [30,31]. The higher the value is, the better the prediction effect is, and the ideal situation is to completely overlap, that is, the ratio is 1. The calculation method of the intersection and union ratio is shown in Formula (9), where K represents the number of pixel types in the image, t_i represents the number of pixels of type i , n_{ii} represents the number of pixels of actual type i and predicted type i , and n_{ji} represents the number of pixels of actual type i and predicted type j . The results of the calculation of the intersection over union after various model experiments are shown in Table 4. According to the calculation results, the intersection over union of the Deeplabv3+ model is the highest among the three models, which proves that its semantic segmentation effect is the best.

$$IOU = \sum_{i=1}^K \frac{n_{ii}}{t_i + \sum_{i=1}^K (n_{ji} - n_{ii})} \quad (9)$$

Table 4. Intersection over union evaluation results.

IOU	Deeplabv3+	FCN-8s	U-Net
Sample 1	0.893	0.827	0.914
Sample 2	0.945	0.893	0.961
Sample 3	0.899	0.833	0.856
Sample 4	0.914	0.876	0.910
Sample 5	0.905	0.910	0.933
.....			
Sample 100	0.940	0.851	0.855
Mean value	0.912	0.860	0.891

5. Conclusions

This paper designs an image acquisition system for a coal mining area. With the help of an adjustable light source and illuminance meter, the coal seam image, rock stratum image, and coal–rock mixed-layer image shooting under a 7000-lux illuminance value is completed, and the image sample library in the coal mining area is successfully constructed. At the same time, in order to increase the type of image feature extraction, an image feature extraction method based on wavelet three-level decomposition is proposed. Through the calculation of five statistics, the eigenvalue extraction of the images is completed.

After constructing the sample image library, the BP neural network is built to carry out the simulation experiment of coal and rock classification and recognition. Through the continuous training of the network model, the recognition rate of coal and rocks is 97.16%. At the same time, in order to verify the superiority of the BP neural network in solving the problem of coal and rock recognition, two control groups were set up for experiments, and the VGG16 model and LeNet-5 model were used to carry out simulation experiments of coal and rock classification and recognition, respectively. Finally, it was concluded that the BP neural network was the best choice to solve the problem of coal and rock recognition.

Finally, this paper builds a Deeplabv3+ network model for semantic segmentation experiments and performs a pixel-level segmentation on the coal–rock mixed-layer image, which solves the problem of determining the distribution state of coal and rocks in the coal–rock mixed layer and achieves 91.2% accuracy. The research results of this paper realize the coal and rock recognition technology by means of image recognition, which provides key technical support for improving the intelligence level of shearers, which is of great significance to improve the production efficiency of coal mines and to reduce the occurrence of coal mine safety accidents.

Author Contributions: Writing—original draft preparation, Y.S.; writing—review and editing, Y.S. and L.Z.; software, L.Z. and Z.S.; investigation, W.Y.; funding acquisition, M.W.; data curation, Y.S., L.Z., Z.S., W.Y. and M.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the 2022 industry-university-research project of Shanxi Datong University (grant number: 2022CXY14); the 2022 Datong Science and Technology Plan Project (grant number: 2022005); the Datong University Education Innovation Project (grant number: 21GJ02); the Postgraduate Education Innovation Project of Shanxi Province (grant number: 2022Y766); and the Datong University 2022, 2023 graduate student innovation project (grant number: 22CX07, 22CX41).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The experimental data used to support the findings of this study are included within the article.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Wu, X.; Li, H.; Wang, B.; Zhu, M. Review on Improvements to the Safety Level of Coal Mines by Applying Intelligent Coal Mining. *Sustainability* **2022**, *14*, 16400. [\[CrossRef\]](#)
2. Chen, W.; Jie, Z. New advances in automatic shearer cutting technology for thin seams in Chinese underground coal mines. *Energy Explor. Exploit.* **2021**, *40*, 3–16. [\[CrossRef\]](#)
3. Wang, G.; Zhang, D. Innovation practice and development prospect of intelligent fully mechanized technology for coal mining. *J. China Univ. Min. Technol.* **2018**, *47*, 459–467.
4. Wang, G. New development of longwall mining equipment based on automation and intelligent technology for thin seam coal. *J. Coal Sci. Eng.* **2013**, *19*, 97–103. [\[CrossRef\]](#)
5. Sun, C.; Li, X.; Chen, J.; Wu, Z.; Li, Y. Coal-Rock Image Recognition Method for Complex and Harsh Environment in Coal Mine Using Deep Learning Models. *IEEE Access* **2023**, *11*, 80794–80805.
6. Liu, C.; Jiang, J.; Jiang, J.; Zhou, Z.; Ye, S. Automatic Coal-Rock Recognition by Laser-Induced Breakdown Spectroscopy Combined with an Artificial Neural Network. *Spectroscopy* **2023**, *38*, 25–32.
7. Gorai, A.K.; Raval, S.; Patel, A.K.; Chatterjee, S.; Gautam, T. Design and development of a machine vision system using artificial neural network-based algorithm for automated coal characterization. *Int. J. Coal Sci. Technol.* **2021**, *8*, 737–755. [\[CrossRef\]](#)

8. Si, L.; Xiong, X.; Wang, Z.; Tan, C. A deep convolutional neural network model for intelligent discrimination between coal and rocks in coal mining face. *Math. Probl. Eng.* **2020**, 2020, 2616510. [\[CrossRef\]](#)
9. Gao, F.; Yin, X.; Liu, Q.; Huang, X.; Bo, Y.; Zhang, Q.; Wang, X. Coal-rock image recognition method for mining and heading face based on spatial pyramid pooling structure. *J. China Coal Soc.* **2021**, 46, 4088–4102.
10. Sun, C.M.; Wang, Y.P.; Wang, C.; Xu, R.J.; Li, X.E. Coal-rock interface identification method based on improved YOLOv3 and cubic spline interpolation. *J. Min. Strat. Control. Eng.* **2022**, 4, 81–90.
11. Fang, X.Q.; He, J.; Zhang, B.; Guo, M.J. Self-positioning system of the shearer in unmanned workplace. *J. Xi'an Univ. Sci. Technol.* **2008**, 28, 349–353.
12. Chad, O.H.; James, C.A.; Ralston, J.C. Infrastructure-based localization of automated coal mining equipment. *Int. J. Coal Sci. Technol.* **2017**, 4, 252–264.
13. Yaghoobi, H.; Mansouri, H.; Farsangi, M.A.E.; Nezamabadi-Pour, H. Determining the fragmented rock size distribution using textural feature extraction of images. *Powder Technol.* **2019**, 342, 630–641. [\[CrossRef\]](#)
14. Wang, J.C.; Li, L.H.; Yang, S.L. Experimental study on gray and texture feature extraction of coal gangue image under different illumination. *J. Coal Sci. Eng.* **2018**, 43, 3051–3061.
15. Zhang, K.; Kang, L.; Chen, X.; He, M.; Zhu, C.; Li, D. A Review of Intelligent Unmanned Mining Current Situation and Development Trend. *Energies* **2022**, 15, 513. [\[CrossRef\]](#)
16. Wu, Q. Research on deep learning image processing technology of second-order partial differential equations. *Neural Comput. Appl.* **2023**, 35, 2183–2195. [\[CrossRef\]](#)
17. Daubechies, I. Ten Lectures on Wavelets. In *CBMS-NSF Regional Conference Series in Applied Mathematics*; Department of Mathematics, University of Lowell: Lowell, MA, USA; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 2016.
18. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, SMC-3, 610–621.
19. Ha, V.K.; Ren, J.C.; Xu, X.Y.; Zhao, S.; Xie, G.; Masero, V.; Hussain, A. Deep Learning Based Single Image Super-resolution: A Survey. *Int. J. Autom. Comput.* **2019**, 16, 413–426.
20. Ying, C.; Huang, Z.; Ying, C. Accelerating the image processing by the optimization strategy for deep learning algorithm DBN. *EURASIP J. Wirel. Commun. Netw.* **2018**, 2018, 232. [\[CrossRef\]](#)
21. Luque, A.; Carrasco, A.; Martín, A.; de las Heras, A. The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognit.* **2019**, 91, 216–231.
22. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **2020**, 21, 6.
23. Grandini, M.; Bagli, E.; Visani, G. Metrics for Multi-Class Classification: An Overview. *arXiv* **2020**, arXiv:2008.05756.
24. Tharwat, A. Classification assessment methods. *Appl. Comput. Inform.* **2021**, 17, 168–192.
25. Aouat, S.; Ait-hammi, I.; Hamouchene, I. A new approach for texture segmentation based on the Gray Level Co-occurrence Matrix. *Multimed. Tools Appl.* **2021**, 80, 24027–24052. [\[CrossRef\]](#)
26. Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.; Villena-Martinez, V.; Rodríguez, J.G. A Review on Deep Learning Techniques Applied to Semantic Segmentation. *arXiv* **2017**, arXiv:1704.06857.
27. Wang, X.; Liu, F.; Ma, X. Mixed distortion image enhancement method based on joint of deep residuals learning and reinforcement learning. *Signal Image Video Process.* **2021**, 15, 995–1002. [\[CrossRef\]](#)
28. Reibman, A.R.; Bell, R.M.; Gray, S. Quality assessment for super-resolution image enhancement. In *Proceedings of the 2006 International Conference on Image Processing*, Atlanta, GA, USA, 8–11 October 2006; pp. 2017–2020.
29. Zhou, H.; Hou, J.; Wu, W.; Zhang, Y.; Wu, Y.; Ma, J. Infrared and visible image fusion based on semantic segmentation. *J. Comput. Res. Dev.* **2021**, 58, 436–443.
30. Rahman, M.A.; Wang, Y. Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation. In *Advances in Visual Computing*; Bebis, G., Boyle, R., Parvin, B., Koracin, D., Porikli, F., Skaaf, S., Entezari, A., Min, J., Iwai, D., Sadagic, A., et al., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 234–244.
31. Shanbhag, A.G. Utilization of Information Measure as a Means of Image Thresholding. *CVGIP Graph. Model. Image Process.* **1994**, 56, 414–419. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.