

Research Article

Automatic Rock Classification Algorithm Based on Ensemble Residual Network and Merged Region Extraction

Chenlu Guo  and Zhiyuan Li 

College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China

Correspondence should be addressed to Chenlu Guo; gcl0329@126.com

Received 19 November 2021; Revised 13 February 2022; Accepted 1 March 2022; Published 20 March 2022

Academic Editor: Rajesh Kaluri

Copyright © 2022 Chenlu Guo and Zhiyuan Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Lithology identification of rocks is an important part in the field of oil and gas exploration, mineral exploration, and geological analysis. How to accomplish rock classification is a key issue for the further development of the geology industry. The current main method for classifying rock pictures containing background is to select sample points or disregard the disturbance of the background. For more accurate classification, the rock part extraction method for rock images containing boundaries is designed to eliminate the influence of background. First, the rock parts are extracted based on the image gradient information and color information, respectively. Then, the two images are intersected to realize the refinement of pixel-level information to obtain a pure rock image. Ensemble ResNet18 (ERN18) is designed as an image classification model. It contains basic blocks to reduce the loss of features during the training process. The method breaks the neglect of most previous studies on background interference. The effect of misclassification in certain regions on the results is eliminated by ensemble learning based on the voting method. The classification results are further improved. Compared with the effects of LeNet, AlexNet, and ResNet, ERN18 has achieved significant results.

1. Introduction

Rock is the main component of the earth's crust, which is one of the elements of the earth's lithosphere. The classification of rock samples is an important field in geology, and it is an essential link in oil and gas exploration, mineral resource exploration, solid metal mineral resource exploration, and geological analysis.

At present, the methods of rock classification can be divided into three categories. The first is the mathematical-statistical analysis method. Chu Shuwu et al. used the detrital components of sandstone for rock classification [1], and Guo and Sun used the statistical analysis of chemical characteristics and ore-bearing (CR) properties for ultrabasic bodies from the Sartuhai of Xinjiang [2]. The second is physical test methods. Wang et al. used the wave of seismic reflection to construct a rock classifier [3], and Fan et al. used hyperspectral remote sensing technology for rock classification. The third is the vision-based neural network method.

Patel and Chatterjee proposed a visual and probabilistic neural network-based rock classification method to classify different types of limestone by histogram features [4].

The traditional methods of rock identification are a series of sampling, experiments, and analysis by specialists, which are highly dependent on human experts and require a high level of professional knowledge. However, if relying on experts' operation manually, it is time-consuming and labor-consuming, and work errors are unavoidable. At present, the main methods of sample classification depend on gravity and magnetism, seismic activity, electromagnetism, hand specimen, and slices, but they have the disadvantages of high cost and strong subjectivity. Therefore, how to classify rocks efficiently and accurately is essential for the further development of the mining industry. Although the image recognition method is not as accurate as of the traditional method, it can implement a rough classification efficiently, which greatly reduces the prework by experts.

Computer vision refers to the use of computers or electronic device terminals to perform operations such as recognition and measurement of target objects through digital methods instead of the eyes, making it a form of image and device detection more suitable for human observation [5]. Currently, computer vision is the most important and active research area in artificial intelligence and is widely used in smart transportation, smart health care, etc. At the level of a large number of individual vision tasks, machine vision capabilities have surpassed the accuracy of human vision. There are already some applications of computer vision in the rock field. Li et al. used BP neural network for collapsing coal-rock identification [6], and Han et al. combined decision tree and SVM for lithology classification [7]. However, the above algorithms have some shortcomings; for example, SVM is a black-box model, so the middle process cannot be controlled, and the fitting effect of the decision tree cannot be guaranteed with too many samples and so on.

In recent years, with the development of artificial intelligence algorithms and the computing power of GPU, the accuracy of deep learning algorithms has been improved with the high operating speed. At present, there are many mature image classification models, such as LeNet [8], AlexNet [9], ResNet [8], and VGG [10]. These networks already have many applications in the rock field. Li et al. proposed a classification model for Mars rock images based on the migration deep learning method and used the VGG-16 network to classify the Mars rock images [11]. Zhang et al. proposed a rock classification model based on the Inception-v3 deep learning model for recognition of granite, kilometre, and breccia images [12]. Ran et al. proposed a rock classification model based on deep convolutional networks for six types of rocks (granite, limestone, conglomerate, sandstone, shale, and vesicular rocks) [13]. Rock classification can be influenced by the background. However, the deep learning research methods mentioned above are simpler for rock images with backgrounds, either by directly ignoring the influence of the background or by directly selecting sample points in the rock region. Since the misjudgement of sample points has a large impact on the final result, this study adopts two methods to reduce this impact and improve the accuracy of classification. Based on the above analysis, a novel rock classification method based on ensemble residual network and merged region extraction is proposed in this study.

2. The Principles of Rock Classification

In this study, a novel rock classifying system based on the convolutional neural network is proposed, which can be regarded as an intelligent classification model of rock lithology.

In some datasets of rock images, images with backgrounds may exist. To avoid background disturbance, for some rock images with boundaries, the edge detection and extraction are supposed to be firstly performed on these image data. For pure rock images (rock image without background), this step can be skipped and the classification can be done directly. The two types of rock images are shown in Figure 1.

The Gaussian denoising is used to reduce abnormal image noise. Grayscale transformation and binarization and the Canny operator are used for the intermediate result image. Multiple terms of dilation and corrosion and median filtering processing are used to obtain more independent and clear edges. Finally, to identify the edge of the rock accurately, the images after binarization and the Canny operator are dilated, respectively, and the images are intersected to obtain a pure rock image without background.

In this study, the ERN18 convolutional neural network image recognition model is established. Residual units are added to reduce the loss of feature information during deep training. The testing results are compared and analyzed with the ones of LeNet, AlexNet, ResNet18, and ResNet50 models to obtain the optimal effect, and the ROC curve is drawn to visually display various recognition effects to verify the accuracy and reliability of the model.

3. Merged Region Extraction

Because the image data in the data set can be divided into rock images with/without background, it is necessary to detect the edge of the image data with background. Based on the Canny operator and morphological image processing algorithm, the pixel-level rock region in the image is extracted.

3.1. Processing for the Rock Image with Boundary. Since some of the images in the data set contain background and the rock area is centered in the middle part, the following algorithm is designed to remove the background around the rock region.

3.1.1. Rough Extraction of Rock Areas. In this study, two methods are used for the rough extraction of rock regions, based on the brightness and gradient information. Since the color of the rock area is dark and the color of the background area is bright, one way is to use adaptive threshold binarization to locate the rock roughly, and because the rock is granular with large gradient variations, the other way is to extract the edge information by the Canny operator. Then, the pixel-level information is refined based on the extracted regional information of the two.

(1) Rough Extraction of Rock Areas Based on Color Information. It used Figure 2 as an example to perform the following operations. Firstly, the Gaussian denoising is carried out on the image. Generally, the image noise is concentrated on the high-frequency signal, which will be mistaken as the rock edge during segmentation. Therefore, it is necessary to reduce the noise to improve the accuracy of segmentation. In the Gaussian filtering, the closer the pixel is to the center, the greater the weight is. The specific weight value is determined by the Gaussian formula matrix.

Secondly, according to the different gray values of the grayscale images of the rock and the background, the adaptive threshold binarization is used to roughly locate the

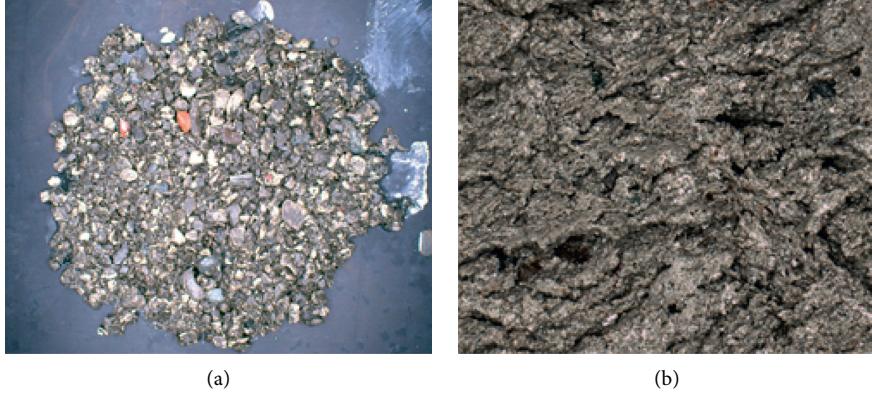


FIGURE 1: Two kinds of rock images in the dataset. (a) Rock image with boundary. (b) Pure rock images.

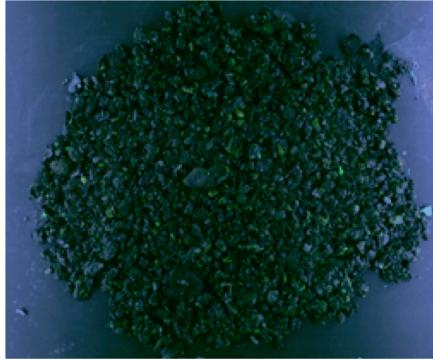


FIGURE 2: Original picture of camera obscura. This image is taken as an example to show the effect of each operation of boundary extraction.

area of the rock particles after denoising the image, and the extraction results are shown in Figure 3.

(2) *Rough Extraction of Rock Areas Based on Gradient Information.* The Canny operator edge gradient detection is an algorithm based on image gradient changes to detect and extract useful information from it. The double-threshold method is adopted when the Canny operator processes image edges, which has a good effect on noise suppression and edge connection. The obtained edge information is of high accuracy. It is a standard multilevel detection algorithm and has been widely used in practical applications [14]. Its algorithm flow chart is shown in Figure 4.

In this algorithm, the original input image is set as $f(x, y)$, which is processed with the Gaussian filter, and the gradient of the processed image $g(x, y)$ is as follows:

$$\Delta g(x, y) =$$

$$\begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{dg}{dx} \\ \frac{dg}{dy} \end{bmatrix}. \quad (1)$$

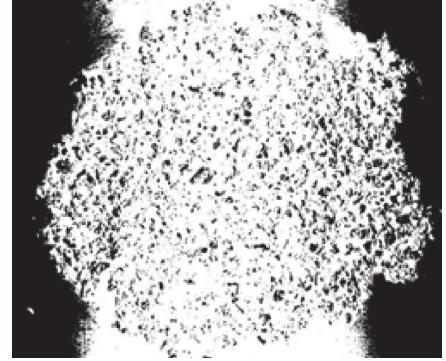


FIGURE 3: Image binarization diagram.

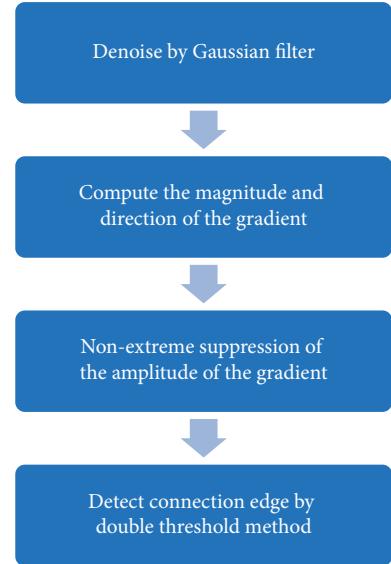


FIGURE 4: Flow chart of edge gradient detection algorithm.

According to the characteristics of convolution operation,

$$\begin{aligned} \nabla g(x, y) &= \nabla f(x, y) \otimes \nabla G(x, y) \\ &= f(x, y) \otimes \nabla G(x, y), \end{aligned} \quad (2)$$

where $G(x, y)$ is the Gaussian function and $\nabla G(x, y)$ is the gradient vector.

Since the blur and width of the image edge will increase after Gaussian processing, the Canny operator is used to sharpen the image through non-maximum suppression.

Suppose the binary Gaussian filter function as follows:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \quad (3)$$

where (x, y) is the point coordinate and σ is the standard deviation.

The algorithm decomposes the two filtering convolution templates of the gradient vector into two one-dimensional row and column filters. Then, the two filters convolve with $f(x, y)$, respectively. The non-maximum suppression technique is to find the local maximum value in the normal direction of the image edge and retain the points with the significant variation of local gradient amplitude [15].

Due to the interference of nonrock parts such as water in the rock image, there are many pseudo-edges in the background. The double-threshold method of the Canny operator can remove the pseudo-edges and solve the problem effectively. The double-threshold method selects thresholds t_1 and t_2 , and $t_2 = 3t_1$ or $t_2 = 2t_1$. Through these two thresholds, the obtained images are thresholding, and two double-threshold edge images $g_1(i, j)$ and $g_2(i, j)$ can be obtained. The edge of the image is discontinuous after the high threshold processing, making that all points higher than the high threshold are edge points. After low threshold processing, the image has many noises and pseudo-edges, so that all points below the lower threshold are not edge points. Then, the discontinuous edge is reconnected by the double-threshold method. If the intermediate pixel is connected with a value higher than the high threshold, it is an edge point [16]. The double-threshold method is used to continuously collect the edges of g_1 image until the edges in g_2 are connected, to detect image edges.

The schematic diagram of image gradient information extracted by the Canny operator is shown in Figure 5. It can be seen that after the Canny operator gradient detection processing, the pseudo-edges of the image can be removed effectively and relatively accurate edge images can be obtained.

Because the rock region obtained by the Canny operator is fuzzy and the edge discrimination is not obvious, to further refine the image edge, the image is processed by dilation, erosion, and median filtering. The dilation and erosion algorithm processing in morphology can segment more independent image elements and connect adjacent elements, and the brightness contrast makes the required rock part in the image more prominent [17]. Median filtering can replace the value of a point in the digital image or digital sequence with the median value of all points in the neighbourhood of the point, to eliminate isolated noise points [18]. After gradient detection, the image obtained by three-term dilation and erosion and median filtering treatment is shown in Figure 6.

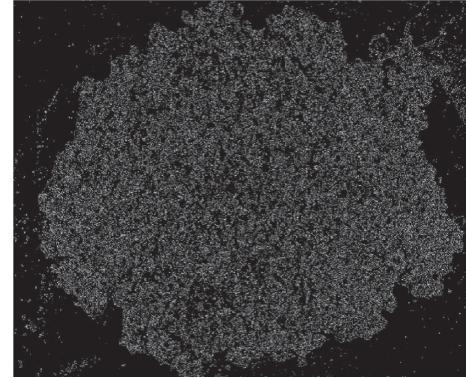


FIGURE 5: Image processed by the Canny operator.

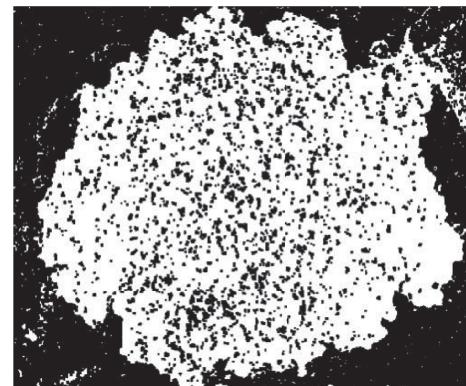


FIGURE 6: Image processed by dilation, erosion, and median filtering.

According to Figure 6, after further processing, more independent image elements are segmented and adjacent elements are connected to make the edge more obvious and the effect of edge detection clearer and more accurate.

3.1.2. The Pixel-Level Fine Positioning of the Rock Areas. The rough location of rock particle area is obtained through color information and gradient information, respectively, but both of them contain a lot of voids and different types of noise information. In this study, morphological expansion operators are used to fill cavities, and the effect picture after filling is shown in Figure 7. The left image is the dilation effect after binarization (Figure 3), and the right image is the dilation effect after extraction by gradient information (Figure 6).

To further reduce the image noise outside the rock area, this study takes the intersection of two types of roughly located rock areas in the same original image to obtain a more accurate edge detection effect. The final edge obtained by the intersection of the two images in Figure 7 is shown in Figure 8. As can be seen from the figure, after the intersection processing, the noise in the original image is effectively eliminated, the rock edge is clear and accurate, and the image detection effect is good, to obtain the pixel-level fine positioning of the rock area.

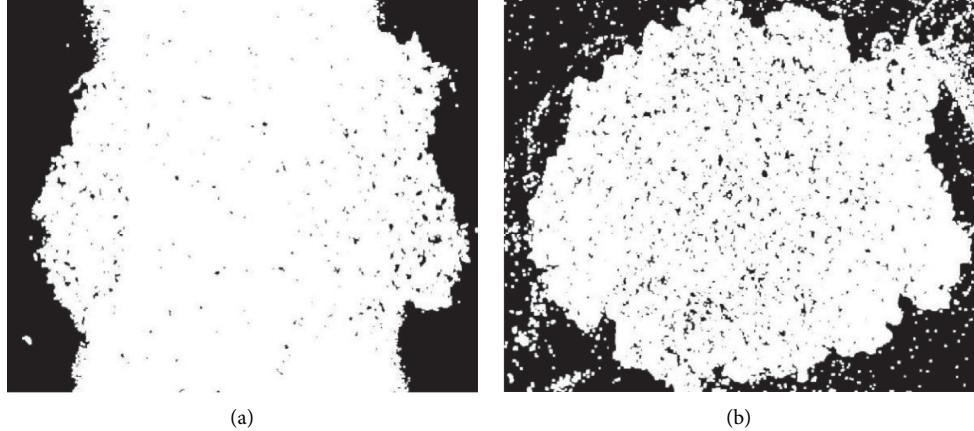


FIGURE 7: (a) Figure 3 processed by expansion. (b) Figure 6 processed by expansion.

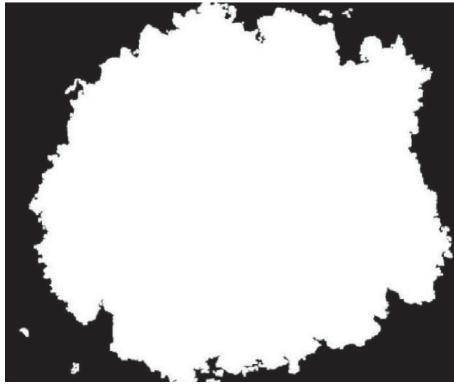


FIGURE 8: Image processed by intersection. The extraction of boundary is accurate, and the pseudo-edge removal is effective.

4. Ensemble Learning Based on the ResNet18 Network

4.1. Introduction to ResNet18. Convolutional neural networks are widely used in computer vision. The traditional feedforward neural network only has fully connected layers, while the convolutional neural network mainly uses the convolutional layers and pooling layers as the core layers. Compared with the traditional neural network, the convolutional neural network has a stronger ability in feature extraction and generalization and has better recognition characteristics and prediction performance [19]. In the development of CNN structure, its basic idea is to increase the diversity by increasing the number of convolutional network layers, to improve the accuracy of classification layer by layer.

However, simply stacking network layers does not improve the training ability of the model and may even lead to gradient explosion or gradient disappearance. Therefore, a residual convolutional neural network model is proposed.

Residual convolution network (ResNet) uses the underlying mapping to fit a certain residual map, and the core idea is behind a high accuracy of shallow network, plus an identity mapping, which means the output is equal to the input. This

makes it possible to increase the depth of the network without increasing the error, to improve the model accuracy [20].

Two residual blocks are commonly used in the structure of ResNet. One is the serial connection of two 3×3 convolutional networks as a residual block, and the other is the serial connection of three convolutional networks as a residual block (with the size of 1×1 , 3×3 , and 1×1). The schematic diagram of the core idea is shown in Figure 9.

The residual mapping expression is as follows:

$$F(x) = H(x) - x, \quad (4)$$

where x is the output of the shallow layer, $H(x)$ is the output of the deep layer, and $F(x)$ is the two layers sandwiched between them, representing transformation. The output of a shallow neural network is added to the output of the deep layer. When the network feature reaches a certain value, the task of the deeper identity mapping is released from the original stack layer to the new identity mapping relationship, and all tasks in the original layer will be close to 0 from the identity mapping. That is, when the x feature is mature enough, any change in it will make the error larger. At this point, $F(x)$ will automatically tend to be 0, and x will continue to be transmitted along the identity mapping path so that the layer behind the deep layers can realize the role of identity mapping and promote network optimization by changing the way of forward and backward information transmission [21].

For the output of L layers, it can be regarded as the sum of the output of the previous $L - 1$ layers and the output of the residual block in the middle. This allows the network to treat any layer and the layer before it as residual blocks, which ensures smooth forward propagation of the network, and the model does not need to stack too many residual blocks, to avoid overfitting.

4.2. Introduction to Ensemble Learning. For small sample data, this study designs an ensemble learning algorithm based on the ResNet18 network. Because the image of rock sample with a smaller grain size has the characteristic of regional consistency, the subgraphs extracted from the high-definition image can be used as independent samples for

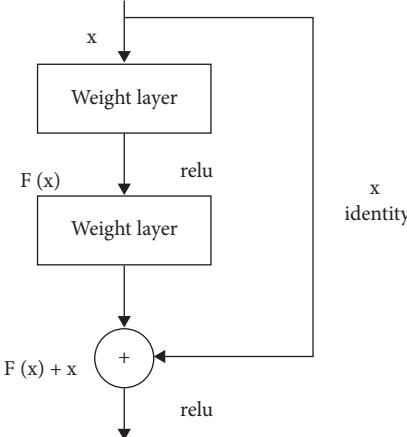


FIGURE 9: Building block in ResNet.

rock classification. In this study, the method of segmenting high-definition images to obtain sub-images can reduce the loss in processing large images and obtain more detailed information of rock particles. Ensemble learning strategy is used to comprehensively evaluate the classification results of subgraphs to improve the accuracy of classification.

Ensemble learning refers to the use of one or more learning algorithms to generate a set of different basic classifiers and then combine them in series or parallel so that a more appropriate decision can be calculated using different algorithms.

Voting is a common technique in ensemble learning. The main idea of voting is to synthesize the choices of various voters to get a general result. Through the integration of multiple models to reduce variance, the robustness of the model can be improved, so the accuracy of the results obtained by voting is higher.

Suppose that $f_n^{(i)}(x)$ is the output of the decision whether voter n predicts that the category of x is i . Then, the voting result $f_{\text{vote}}^{(i)}(x)$ of x belonging to category i can be expressed as follows:

$$f_{\text{vote}}^{(i)}(x) = \omega^T f^{(i)}(x), \quad (5)$$

where $f^{(i)}(x) = [f_1^{(i)}(x), \dots, f_N^{(i)}(x)]^T$ and ω is the weight vector.

$f_n^{(i)}(x)$ is 1 if the category x of voter n 's prediction is i , and 0 otherwise, and then, $f_{\text{vote}}^{(i)}(x)$ can be expressed as follows:

$$f_{\text{vote}}^{(i)}(x) = \sum_{n=1}^N f_n^{(i)}(x), f_n^{(i)}(x) \in \{0, 1\}. \quad (6)$$

According to the voting rule as shown in formula (6), the final classification category is as follows:

$$\max_{i \in \text{accept, reject}} f_{\text{vote}}^{(i)}(x), y \longrightarrow i. \quad (7)$$

5. Results and Discussion

To verify the effectiveness of the proposed rock classification method, tests were carried out on the dataset of question B of

the 9th “TipDM Cup” Data Mining Challenge (<https://www.tipdm.org:10010/#/competition/1354705811842195456/question>), and comparisons were made with existing deep learning methods.

The dataset contains multiple image data and a table of rock types for each sample. The dataset consists of 7 types of rocks and 350 photographs taken by white light. In this study, classes 1–7 are used to represent in Figure 10, respectively.

According to the data requirements of the ensemble learning method designed in this study, each image is divided into 3×3 subgraphs, as shown in Figure 11. Data augmentation methods such as random horizontal flip and random vertical flip are adopted to conduct mass training on the segmented data set.

5.1. Parameter Setting. The images in the dataset have been preprocessed and divided into seven classes of rocks. The image data are divided into the training set and the test set in the ratio of 5:1, and some data augmentation methods are used to expand the dataset and avoid the loss of information during training effectively.

In the method proposed in this study, each decision branch (ResNet18) adopts two 3×3 residual blocks of convolution operations as the basic operation unit to reduce parameters, which is composed of 17 convolution layers and 1 fully connected layer.

Under Ubuntu 18.04.5 LTS, GeForce RTX 2080 Ti GPU, and PyTorch framework environment, ERN18 designed in this study is compared with ResNet18 [8], ResNet50 [8], AlexNet [9], and LeNet [8] on the mentioned dataset for rock recognition experiments. ResNet18 has 18 layers, and ResNet50 has 50 layers. These two networks add residual units through a shortcut connection for solving the degradation problem. AlexNet has 8 layers, including 5 convolutional layers and 3 fully connected layers. Each convolutional layer contains an excitation function (RELU) and local response normalization (LRN) processing. LeNet has 3 convolutional layers, 2 subsampling layers, 1 fully connected layer, and 1 Gaussian connected layer, for a total of 7 layers. The input sizes of the images are all 224×224 . In the experimental training period, each model is trained for 300 epochs with cross-entropy loss function and Adam optimizer, the hidden layer activation function is ReLu function, the batch size is 8, and the learning rate is set to 2×10^{-4} . The training set is the same for each model. The best training weights of the models are kept during the training period to test the test set.

5.2. Experimental Results. The results of the validation set and the test set were compared and verified. The accuracy of the final classification result of the model is 82.87%, and the effect is good. Compared with ResNet50, AlexNet, and LeNet, ERN18 has achieved significant results. Some examples accurately classified by the model as light gray fine sandstone are shown in Figure 12.

Some examples accurately classified by the model as dark gray fine sandstone are shown in Figure 13.

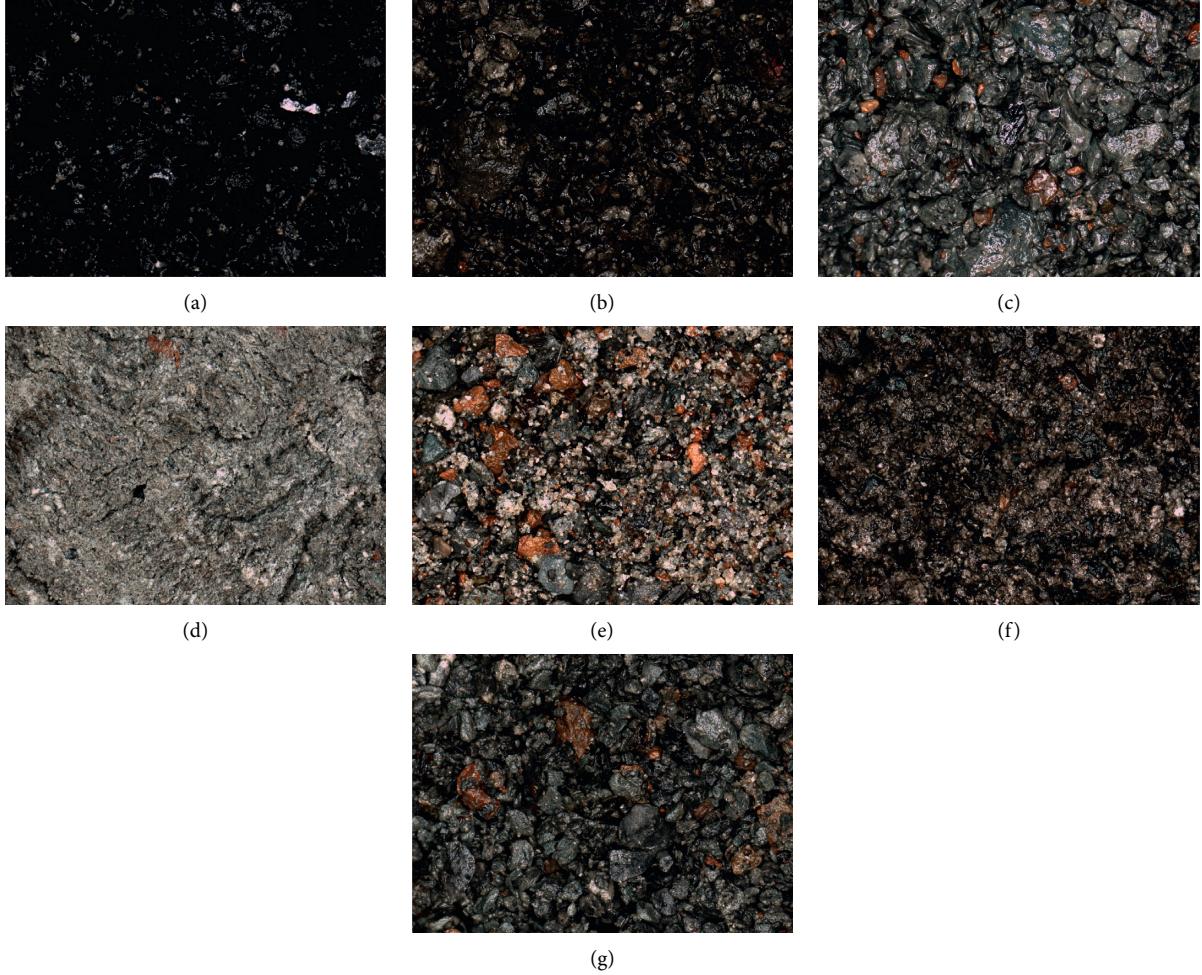


FIGURE 10: Seven types of rock diagrams in the dataset. (a) Black coal. (b) Gray-black mudstone. (c) Gray muddy siltstone. (d) Gray fine sandstone. (e) Light gray fine sandstone. (f) Dark gray siltstone. (g) Dark gray mudstone.

5.3. Discussion. The classification effects of ERN18, ResNet50, ResNet18, AlexNet, and LeNet models are compared and analyzed in Table 1.

The training accuracy of ResNet50 has reached more than 90%, while the testing accuracy is less than 80%, which shows the phenomenon of overfitting. ERN18 model has the best training effect, which is higher than other models to a large extent. It is suitable for rock sample classification. ERN18 has a recall rate of 82.76%, which shows that it can classify rocks more accurately. This is due to the fact that ERN18 builds on ResNet18, and the residual blocks in ResNet18 eliminate the problem of gradient explosion, which allows for deeper network layers and a more robust network. The ensemble learning approach divides the images into subgraphs, and the subgraphs are classified separately without affecting each other, avoiding the influence of a certain region leading to inaccurate results further and improving its accuracy.

ROC curve can accurately reflect the relationship between the specificity and the sensitivity of an analysis method, demonstrate the accuracy of the experimental results, and visually display the classification effect of each

model, as shown in Figure 14. When the false-positive rate is 0.2, the true rate of ERNT has reached about 0.7, and AUC is 0.81. The classification ability of ERN18 is much stronger than LeNet and other networks shown in this figure.

As can be seen from Figure 15, class 7 has the best classification effect with an accurate classification rate of 99%, followed by class 1 and class 3 with the rate of 96%, and then followed by class 5, class 4, and class 6 with the rate of 95%, 94%, and 86%, respectively. Class 2 has the lowest effect with the rate of 74%. However, it can also complete the task of image classification. In summary, ERN18 has a good effect on rock classification and it is suitable for intelligent lithology classification of rock samples.

The confusion matrix derived from the classification results of the ERN18 network for the training set data is shown in Figure 16. In the dataset, the number of images for each class of rock is different, but it can still be seen that the distribution of the dataset is on the diagonal, from which it can be seen that an accurate classification can be achieved.

The F1 score presentation of these seven rocks is shown in Figure 17. It can be seen that class 4 and class 7 have the best result.



FIGURE 11: Image segmentation diagram. The image is divided into 3×3 subgraphs on ensemble learning using the voting method.

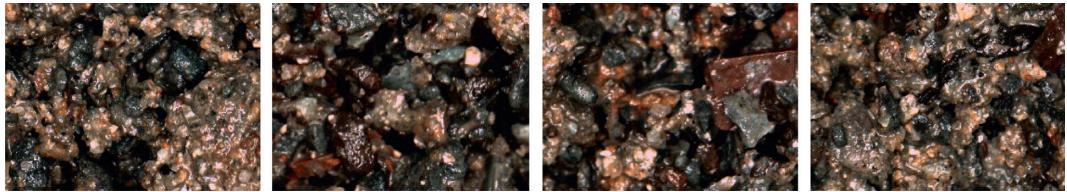


FIGURE 12: Partial results of test set data classified as light gray fine sandstone.



FIGURE 13: Partial results of test set data classified as dark gray fine sandstone.

TABLE 1: Training and testing accuracy table for each model.

Model	Train loss	Train accuracy (%)	Test loss	Test accuracy (%)	Recall (%)
LeNet [8]	1.5684	68.54	1.2148	62.31	72.54
AlexNet [9]	1.3451	65.85	1.1908	64.76	75.19
ResNet50 [8]	0.1583	95.94	0.8861	75.68	81.74
ResNet18 [8]	0.1926	93.84	0.9256	72.72	79.63
ERN18	0.0193	99.38	0.0976	82.87	82.76

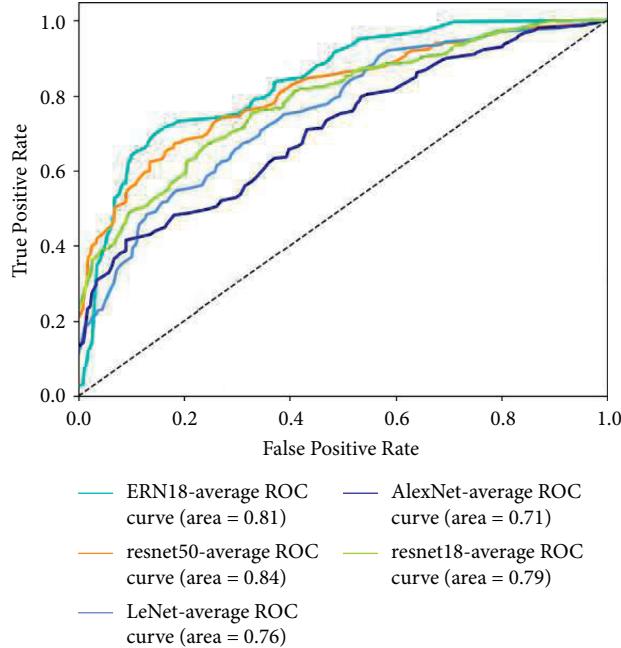


FIGURE 14: Comparison of ROC curves of each model.

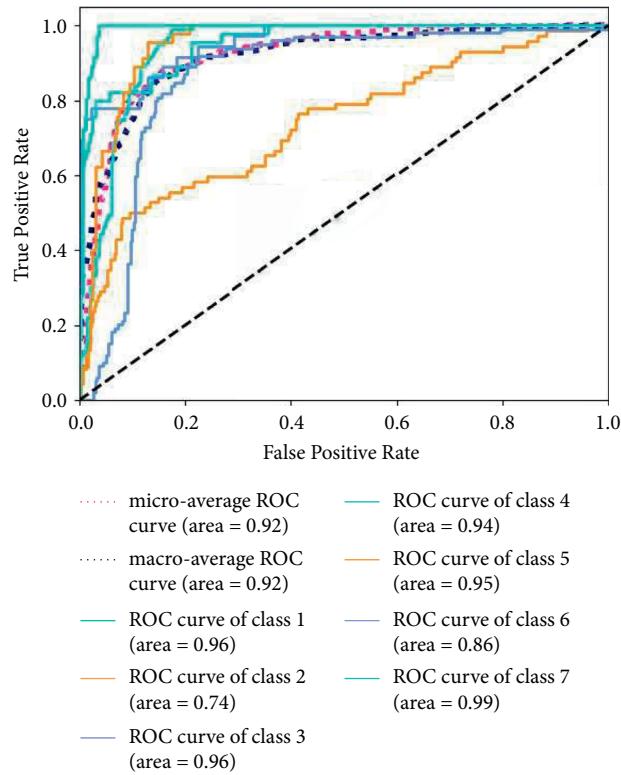


FIGURE 15: ROC curves of 7 classes of samples on ERN18.

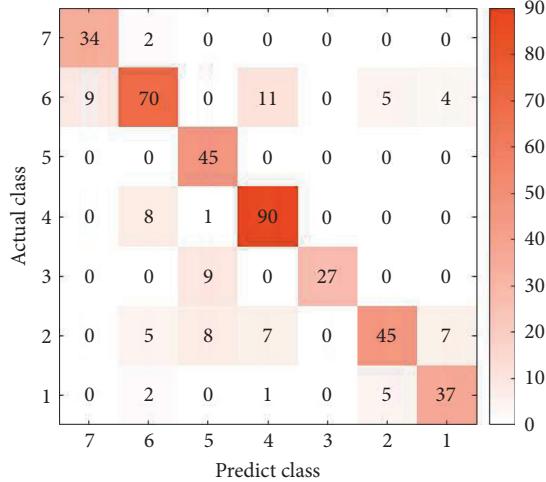


FIGURE 16: Confusion matrix of ERN18 for the training set.

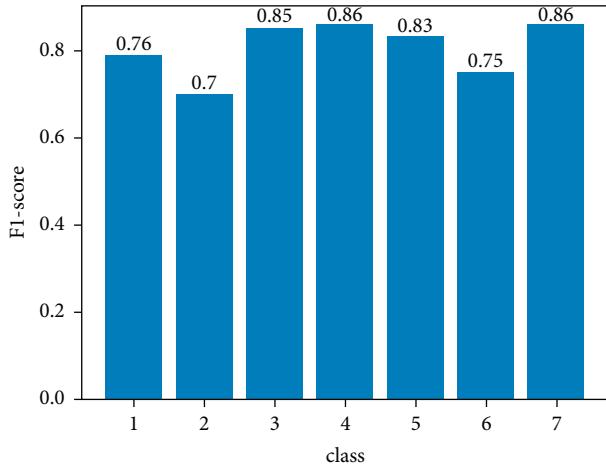


FIGURE 17: F1 score of 7 classes of samples on ERN18.

6. Conclusions

In this study, in the field of rock recognition, a method based on binarization, Canny operator, and morphological methods is established to extract rocks from images containing rocky backgrounds. A network named ERN18 is created to improve the accuracy of rock classification by combining ResNet18 and integrated learning. The model created in this study has good practicality, changing the method of not dealing with background or sampling in previous models, and using the voting method to avoid wrong results due to misidentification of a region and improve the accuracy. The network can be further applied to rocks that are not easily accessible to humans directly, such as Martian rocks and lunar rocks.

The classification of rock is achieved by ERN18. The accuracy of the test set reaches 82.87%, far more than ResNet18, LeNet, AlexNet, and other neural networks. It can accurately classify the type of rock it belongs to. ERN18 has wide adaptability, which can be extended to other industries such as material identification research and has a good universality and application prospect.

On this basis, some optimizations can be made. Although deep network models have achieved excellent performance in rock recognition tasks, the current research mainly classifies rocks with obvious color and texture features and cannot effectively extract rock detail features in terms of particle size features. For similar types of rocks that have similar color and texture features but slight differences in grain size, the above methods cannot effectively extract the grain size information to accurately identify and classify them. In addition, in the field of rock and mineral recognition, most scholars study the relevant content of rock slices, which cannot be well applied to specific practice. Therefore, the study of rock classification still needs to be further explored.

Data Availability

A publicly archived dataset is published at <https://www.tipdm.org:10010/#/competition/1354705811842195456/question>. The dataset can be downloaded at https://drive.google.com/file/d/1CWGSbvIlOliAzwUzPhbSCK_7gigWEI6T/view?usp=sharing.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

References

- [1] S. W. Chu and S. Y. Li, "Comparison between two methods for statistical study of detrital composition of sandstones," *Journal of Hefei University of Technology*, vol. 30, no. 6, pp. 668–671, 2007.
- [2] Y. S. Guo and S. R. Sun, "Petrochemistry characteristics and statistical analysis of ore (Cr)bearing nature of ultrabasic bodies from the Sartuohai of Xinjiang," *Journal of Lanzhou University*, vol. 24, no. S1, pp. 21–28, 1988.
- [3] Y. J. Wang, D. C. Li, N. Yin et al., "Application of decision tree algorithm in lithologic identification of geological modeling," *Urban Geotechnical Investigation & Surveying*, no. 1, pp. 198–202, 2020.
- [4] A. K. Patel and S. Chatterjee, "Computer vision-based limestone rock-type classification using probabilistic neural network," *Geoscience Frontiers*, vol. 7, no. 1, pp. 53–60, 2016.
- [5] Z. N. Sun, Q. Li, and Y. F. Liu, "Research progress of computer vision and pattern recognition," *E-science Technology & Application*, vol. 10, no. 4, pp. 3–18, 2019.
- [6] Y. M. Li, S. C. Fu, Y. B. Jiao, and M. Wu, "Collapsing coal-rock identification based on fractal box dimension and wavelet packet energy moment," *Journal of China Coal Society*, vol. 42, no. 3, pp. 803–808, 2017.
- [7] Q. D. Han, X. T. Zhang, and W. Shen, "Application of support vector machine based on decision tree feature extraction in lithology classification," *Journal of Jilin University (Earth Science Edition)*, vol. 49, no. 2, pp. 611–620, 2019.
- [8] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [9] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, no. 2, 2012.

- [10] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, 2016.
- [11] J. L. Li, Z. Li, Z. C. Wu et al., “Autonomous Martian rock image classification based on transfer deep learning methods,” *Earth Science Inform Atics*, vol. 13, no. 10, pp. 951–963, 2020.
- [12] Y. Zhang, M. C. Li, and S. Han, “Automatic identification and classification in lithology based on deep learning in rock images,” *Acta Petrologica Sinica*, vol. 34, no. 2, pp. 333–342, 2018.
- [13] X. Ran, L. Xue, Y. Zhang, Z. Liu, X. Sang, and J. He, “Rock classification from field image patches analyzed using a deep convolutional neural network,” *Mathematics*, vol. 7, no. 8, p. 755, 2019.
- [14] W. M. Wells, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, “Adaptive segmentation of MRI data,” *IEEE Transactions on Medical Imaging*, vol. 15, no. 4, pp. 429–442, 1996.
- [15] H. R. Li, J. Gao, T. Wu, and Z. Y. Shu, “Crack detection method of insulators based on improved canny operator,” *Power Grid Analysis & Study*, vol. 49, no. 2, pp. 91–98, 2021.
- [16] Z. R. Zhao, B. L. Gao, Y. Y. Guo, and Li Tian, “Edge detection of noise image based on improved algorithm,” *Computer Measurement & Control*, vol. 28, no. 12, pp. 202–206+212, 2020.
- [17] H. M. Du, B. B. Jiang, L. B. Chang, C. Y. Guo, and K. B. Ji, “Improvement and parallel implementation of dilation and erosion algorithms,” *Journal of Posts and Telecommunications*, vol. 22, no. 1, pp. 88–93, 2017.
- [18] B. Hao and X. P. Lu, “Research on image edge denoising algorithms based on median filtering,” *Modern Computer*, no. 20, pp. 38–41+49, 2019.
- [19] Y. B. Zhang, M. H. Chen, S. G. Yang, and H. W. Chen, “Optoelectronic convolutional neural networks based on time-stretch method,” *Science China (Information Sciences)*, vol. 64, no. 2, pp. 180–191, 2021.
- [20] H. Wang, “Waste Bottle Classification system based on depth residuals network ResNet,” *Science and Technology & Innovation*, no. 14, pp. 71–72, 2020.
- [21] S. X. Guan and W. C. Ge, “Recognition and environmental influence of speed bump based on ResNet18,” *Communications Technology*, vol. 54, no. 3, pp. 597–603, 2021.