Automated Classification of High-resolution Rock Image Based on Residual Neural Network

Weibo Cai¹, Juncan Deng¹, Qirong Lu¹, Kengdong Lu¹ and Kaiqing Luo¹*

¹ School of Physics and Telecommunication Engineering, Guangdong Provincial Engineering Research Center for Optoelectronic Instrument, South China Normal University, Guangzhou, China

*Corresponding author email: kqluo@scnu.edu.cn

Abstract. The identification and classification of high-resolution rock images are significant for oil and gas exploration. In recent years, deep learning has been applied in various fields and achieved satisfactory results. This paper presents a rock classification method based on deep learning. Firstly, the high-resolution rock images are randomly divided into several small images as a training set. According to the characteristics of the datasets, the ResNet (Residual Neural Network) is optimized and trained. The local images obtained by random segmentation are predicted by using the model obtained by training. Finally, all probability values corresponding to each category of the local image are combined for statistics and voting. The maximum probability value and the corresponding category are taken as the final classification result of the classified image. Experimental results show that the classification accuracy of this method is 99.6%, which proves the algorithm's effectiveness in high-resolution rock images classification.

Keywords: rock image; residual neural network; automated classification.

1. Introduction

The identification and classification of rock samples are complicated problems in oil and gas exploration. In mineral resources exploration, especially in solid metal mineral resources exploration, rock image recognition also plays an immeasurable role. Recently, the methods of rock image recognition mainly include gravity and magnetism, well logging, seismic, remote sensing, electromagnetic, geochemistry, hand specimen, thin section analysis, etc. The traditional rock image recognition and classification is mainly based on naked-eye observation, which has long periods, difficult quantification and low recognizable efficiency. In recent years, deep learning-based methods have been widely applied in many fields, such as digital recognition, medical image recognition, face detection, image fusion, etc. Therefore, it is a more effective way to establish an automatic recognition and classification model of rock images by using deep learning.

Dunlop et al. [1] use Support Vector Machines to classify rock image datasets in natural scenes and achieve high accuracy of 86.3%. Zhou Guanwu et al. [2] use RBF neural network to perform calculations related to reservoir parameters and lithology identification in oilfield reservoir characterization. Some people use the K-means clustering segmentation algorithm to segment rock thin slice image into target pore and background rock, extract features, and then use a probabilistic neural network to classify images. The results show that classification and recognition accuracy reaches 95.12% after several verifications [3]. After preprocessing and featuring extraction of rock thin slice images, some scholars use the neural network to identify the rock fabric of rock thin slice images, and the accuracy reached 93.3% [4]. Zhang Jiafan et al. [5] proposed a rock CT image segmentation and quantification method
based on a clustering analysis algorithm. Zhang Cuifen et al. [6] use feature vectors of lithologic units for color synthesis of images, which significantly increased the identifiability of lithologic companies. Li et al. [7] use the transfer learning method to train microscopic sandstone images and finally obtain a high-accurate sandstone tiny image classification model. Cheng Guojian et al. [8] classified rock images based on a 4-layer convolutional neural network and proved the research value of the deep convolutional neural network in rock fabric analysis. Alexis et al. [9] use 5-layer CNN to classify images of natural scenes of rocks, achieving classification accuracy of 89.43%. Zhao et al. [10] realized automatic annotation of rock images. Liu et al. [11] use an extraction network based on simplified VGG16 and achieved the classification accuracy of 96%. Mohammadeza et al. [12] use CNN to predict pixel-level fractures in rock images, achieving an F-score of 84.0. In the submission process of this paper, Cheng et al. [13] also use ResNet to classify rock images and published the paper. The difference is that they aim at thin-section image rather than high-resolution rock image.

Previous experimental results show that the method based on deep neural network can significantly improve the efficiency of rock image recognition. It provides convenience for rock image analysis. However, their work still has the following shortcomings: (1) They did not use more advanced image classification models. (2) There is no processing for high-resolution rock images. (3) Negative samples are not considered in the training set.

The following improvements are made to overcome those problems: (1) The most popular image classification model, ResNet, is introduced for training. (2) The local image fusion prediction strategy is used to predict high-resolution images. (3) The local images that do not contain rocks are introduced as negative samples in the training set.

## 2. Our Approach

### 2.1. Fundamentals of ResNet

Traditional deep learning algorithms generally adopt data initialization and regularization [14] to solve problems such as high consumption of computing resources, easy over-fitting of models, and gradient extinction/gradient explosion. Although this method solves the gradient problem to a certain extent, network degradation and other phenomena appear. Based on this, He Kaiming et al. [15] proposed Residual Neural Network (ResNet) and invented Shortcut Connection for the degradation phenomenon, which greatly eliminated the difficulty of neural network training with excessive depth. ResNet helps significantly improve model prediction performance. While it does not significantly increase the training time of the model.

In this paper, ResNet is selected as the feature extraction network for the pre-training and fine-tuning model. To increase the number of layers of the network, the identity mapping layer is superimposed on the shallow neural network to construct residual learning units. A portion of the original input information is fed directly to the next layer via the identity mapping layer by learning the residuals. This avoids the loss of features of the convolutional layer in information transmission to a certain extent and can learn new features according to the input features, which has better performance. As shown in figure 1, Resnet 152 is selected as the network, and the input size is 224×224.

![Figure 1. Model building.](image)

### 2.2. Local Image Fusion Prediction Strategy

The rock image can only be observed in the local image. Based on this, this paper proposes a local image fusion prediction strategy.

Due to the unique structure of the convolutional neural network in the process of image classification, the image will be blurred and smoothed continuously after multiple convolutions and pooling operations. Many local details will be lost in the process. In this case, if the original image is directly sent into the
convolutional neural network for training, some small local features will not be extracted, and the classification results will be affected. Therefore, to better mine the local features of the image, we proposed the local image fusion prediction strategy.

There are many methods for image segmentation. The first method is the division that does not include overlapping areas. The second is the division of overlapping areas. The last method is to use random segmentation to cut the image into multiple partial images randomly. The partial images may or may not contain overlapping parts. These three methods can achieve image segmentation, but the choice of different methods dramatically impacts the final classification results. To improve the generalization ability and reduce the computation redundancy, we choose to use the random cutting method.

The specific method is shown in figure 2. Firstly, to fully extract local features, the images in the training set are randomly cut into the ResNet residual neural network model for training, and the classification model of local images is obtained. Then, the images to be classified in the test set are randomly cut and sent into the rock local image classification model. The first three rock categories with the highest probability and their corresponding ascribing probability values should be obtained from each local image, indicating that the image belongs to this category. The first three prediction results of all sub-blocks of each image to be classified are classified according to categories. Then all probability values corresponding to each category are combined to make a statistic. They are summed, and the maximum probability value and the corresponding category are taken as the final classification results for the classified images.

![Diagram](image)

**Figure 2.** Flowchart of our method.

2.3. ResNet Training and Optimization

2.3.1. Model Training Framework and Parameter Setting

The setting of network mode parameters will affect the performance of the network and the training speed. The ResNet model is built using the PyTorch framework. Parameter setting: the training batch size is eight, and the learning rate is $10^{-6}$. The network resolution is 256×256.

2.3.2. Model Optimization

The rock images in the experiment have complex background, fuzzy boundary, and uneven sample. In order to obtain more high-resolution information, the image is optimized and adjusted.

1) After each convolution calculation, Batch Normalization is carried out to avoid gradient dispersion, improve the model's generalization ability, and enable the network to allow a higher learning rate to accelerate convergence.

2) The Cross-Entropy Loss function is adopted to improve the stability of the loss value. The calculation formula is:

$$CE(p, q) = -\sum_{l=1}^{C} p_l\log(q_l)$$  \hspace{1cm} (1)

Where $C$ represents the number of categories, $P_l$ is the actual value, and $Q_l$ is the predicted value. From the above calculation, we can know that the forecast more accurate, the loss less.

3) By using the adaptive optimizer Adamax [16], the learning rate is adjusted automatically, and the step size is annealed to avoid falling into the local minimum value during model training. Adamax is able to observe a large and influential information gradient contributed by a few small batches of samples, reducing the effects of selection inhomogeneity.
4 The negative sample, which is equivalent to the positive sample, is introduced to improve the robustness of the model to a certain extent.

![Figure 3. The original rock image datasets.](image)

![Figure 4. Random cutting demonstration.](image)

3. Analysis of Experimental Results

3.1. Experimental Dataset
The data in this paper came from the 9th Teddy Cup National College Student Data Mining Challenge. The data include images of seven rock types: black coal, gray-black mudstone, gray argillaceous siltstone, fine gray sandstone, light gray fine sandstone, dark gray silty mudstone, and dark gray mudstone. As is shown in figure 3. The number of images in each category are 22, 31, 47, 19, 86, 41, and 76. Most of the images are 4096x3000, and a few are 2448x2048. Eighty percent are training sets, 10 percent are validation sets, and 10 percent are testing sets. 

The above images are called the original image. For each original image, we divided it into 200 local images by random cutting. A total of 51,600 experimental sample data are obtained, including 50,000 positive samples and 1,600 negative samples (images of rocky areas are omitted). The segmentation schematic diagram is shown in figure 4.

3.2. Results and Analysis

3.2.1. Rock image classification results based on optimized ResNet
When the training parameters of the basic model are the same, the training process before and after model optimization is shown in figure 5 and figure 6. After the algorithm is improved and enhanced, the dataset's training loss curve oscillates less and converges better. To evaluate the performance of the optimized ResNet model, we compared the accuracy of the ResNet model, the GoogLeNet model, and the optimized ResNet model on the images (including some dirty data) in the test sets, as shown in Table 1.
In the rock image data set, the accuracy of the ResNet model is higher than that of the GoogLeNet model, while the accuracy rate of the ResNet model optimized according to the data set is as high as 0.996 on the test set, which is 0.092 higher than that before optimization. And the accuracy of the optimized ResNet model on the test set is very close to that on the verification set, which indicates the robustness of our method.

**Figure 5.** ResNet training loss and verification set accuracy curve before optimization.

**Figure 6.** ResNet training loss and verification set accuracy curve after optimization.

**Table 1.** Comparison of model performance.

| Net           | Accuracy |
|---------------|----------|
| ResNet        | 0.904    |
| GoogLeNet     | 0.887    |
| Optimized ResNet | 0.996   |

**Table 2.** Classification results of our method on high-resolution rock images.

| Rock type                  | Number of images | Number of images correctly predicted |
|----------------------------|-------------------|--------------------------------------|
| black coal                 | 2                 | 2                                    |
| gray black mudstone        | 3                 | 3                                    |
| gray argillaceous siltstone| 5                 | 5                                    |
| gray fine sandstone        | 2                 | 2                                    |
| light gray fine sandstone  | 9                 | 9                                    |
| dark gray silty mudstone   | 4                 | 4                                    |
| dark gray mudstone         | 8                 | 8                                    |

3.2.2. **Test results of the original image**

There are seven types of original high-resolution images that are tested: black coal, gray-black mudstone, gray argillaceous siltstone, fine gray sandstone, light gray fine sandstone, dark gray silty mudstone, and dark gray mudstone. The number of images in each category is 2, 3, 5, 2, 9, 4, 8. We test these original images using an optimized ResNet model and the local image predictive fusion strategy. Table 2 shows the prediction results of our method for authentic high-resolution images. Our process correctly
predicted seven types of high-resolution rock images, proving the reliability and robustness of our approach.

4. Conclusion
A high-resolution rock images classification algorithm is proposed based on a deep learning model. The method includes: (1) High-resolution images are predicted using the local image fusion prediction strategy. (2) The Resnet model is optimized for the characteristics of the dataset. The results show that this method's classification accuracy and model robustness is good, and the model is reliable and practical when verified using the test data set of the Teddy Cup National Student Data Mining Challenge. The method provides a reference to realize the intelligence of rock image classification and improve the efficiency of rock image classification and has high application value.

References
[1] H. Dunlop. Automatic Rock Detection and Classification in Natural Scenes. Thesis(Carnegie Mellon University), 2006.
[2] Guanwu Zhou, Guojian Cheng. BRF Neural Networks for Petroleum Reservoir Characterization. Computer Engineering and Applications, 2007, pp. 174-176+224.
[3] Guojian Cheng, Jing Yang, Quanzhou Huang. Rock Image Classification Recognition Based on Probabilistic Neural Networks, Science Technology and Engineering, 2013, pp. 9231-9235.
[4] Guojian Cheng, Wei Ma. Research of rock texture identification based on image processing and neural network, Journal of Xi'an Shiyou University (Natural Science Edition), 2013, pp. 105-110+1.
[5] Jiafan Zhang, Xuejia Zhang. A Method of Rock CT Image Segmentation and Quantification based on Clustering Algorithm. Journal of Xi'an University of Science and Technology, 2016, 36(02): 171-175.
[6] Cuifen Zhang, Lina Hao. An Image Enhancement and Lithology Identification Method based on Landsat8 OLI Data. Geology and Prospecting, 2017, pp. 325-333.
[7] Na Li, Huizhen Hao, Qing Gu. A transfer learning method for automatic identification of sandstone microscopic images, Computers & Geosciences, Volume 103, 2017, pp. 111-121.
[8] Guojian Cheng, Wenhui Guo, Pengzhao Fan. Study on Rock Image Classification Based on Convolution Neural Network. Journal of Xi'an Shiyou University (Natural Science Edition), 2017, pp. 116-122.
[9] A. D. P. Pascual, L. Shu, J. Szoke-Sieswerda, K. McIsaac and G. Osinski, "Towards Natural Scene Rock Image Classification with Convolutional Neural Networks," 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), 2019, pp. 1-4.
[10] H. Zhao, X. Zhao, W. Liu, J. Wei, G. Cheng and X. Qiang, "Rock Image Automatic Annotation based on Deep Neural Network," 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), 2020, pp. 1824-1828.
[11] X. Liu, H. Wang, H. Jing, A. Shao and L. Wang, "Research on Intelligent Identification of Rock Types Based on Faster R-CNN Method," in IEEE Access, vol. 8, pp. 21804-21812, 2020.
[12] M. Asadi, M. T. Sadeghi and A. Yarahmadi Bafghi, "A Multi-Classifier System for Rock Mass Crack Segmentation Based on Convolutional Neural Networks," 2021 26th International Computer Conference, Computer Society of Iran (CSICC), 2021, pp. 1-6.
[13] C. Guojian and L. Peisong, "Rock Thin-Section Image Classification based on Residual Neural Network," 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP), 2021, pp. 521-524.
[14] S. IOFFE, C. SZEGEDY. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, Proceedings of the 32nd International Conference on International Conference on Machine Learning (ICML), 2015, pp. 448-456.
[15] K. He, X. Zhang, S. Ren and J. Sun. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778.
[16] D. Kingma, J. Ba. Adam: A method for stochastic optimization. Proceedings of the 3rd International Conference for Learning Representations, 2015.