

Design of Rock Sample Classification and Recognition Technology Based on Deep Learning (Post-print)

Authors: Ruifang Zhang, Yongan Li, Ruifang Zhang

Date: 2022-11-17T00:00:00+00:00

Abstract

In oil and gas exploration and mineral resource exploration, rock sample identification plays an invaluable role due to its strong operability. In addition to conventional methods such as gravity, magnetic, remote sensing, and electromagnetic techniques for identifying rock samples, a novel approach involves establishing an automatic identification and classification model for rock samples using deep learning-based image analysis methods. This paper employs transfer learning to construct a deep convolutional neural network model for automatic rock sample classification and recognition. First, the original data is preprocessed to ensure optimal usability for model training. Then, using the TensorFlow deep learning framework, this study applies transfer learning to adapt the Inception-V3 deep convolutional neural network model to the preprocessed target dataset. This improves the learning effectiveness of the target task, enabling extraction of more reasonable image features until the model is fully trained for the target domain, thereby achieving better rock sample classification.

Full Text

Preamble

Design of Rock Sample Classification and Recognition Technology Based on Deep Learning

Ruifang Zhang, *Yongan Li*

Guilin University of Technology, Guilin, China

Corresponding author, e-mail: 631600961@qq.com

Abstract: In oil and gas exploration and mineral resource exploration, rock sample identification plays an invaluable role due to its strong operability. In ad-

dition to conventional methods such as gravity, magnetic, remote sensing, and electromagnetic techniques for identifying rock samples, a novel approach involves establishing an automatic identification and classification model for rock samples using deep learning-based image analysis methods. This paper employs transfer learning to construct a deep convolutional neural network model for automatic rock sample classification and recognition. First, the original data is preprocessed to ensure optimal usability for model training. Then, using the TensorFlow deep learning framework, this study applies transfer learning to adapt the Inception-V3 deep convolutional neural network model to the preprocessed target dataset. This improves the learning effectiveness of the target task, enabling extraction of more reasonable image features until the model is fully trained for the target domain, thereby achieving better rock sample classification.

Keywords: deep learning, rock image, inception-v3, classification recognition

1 Introduction

The identification and classification of rock samples are crucial for oil and gas exploration, mineral resource exploration, and geological analysis. Currently, rock sample identification primarily encompasses methods such as gravity and magnetic surveys, well logging, seismic analysis, remote sensing, electromagnetic techniques, geochemistry, hand specimen examination, and thin section analysis (Anonymous, 2017). The application of deep learning-based image analysis to establish automatic identification and classification models for rock samples holds significant development potential. This approach can reduce reliance on experimental equipment and specialized expertise while achieving automatic identification and classification through image recognition, thereby offering substantial practical value for research, exploration, and development in oil and gas fields and mineral resources.

Numerous scholars have conducted in-depth research on rock image classification using image recognition technology. Hossein Izadi et al. (2017) developed an ore flake recognition system that achieved promising practical results through neural network classification based on color and texture feature analysis of ore flakes. Hong et al. (2017) proposed a method combining image processing, fractal theory, and artificial neural networks to quantitatively determine geological strength indicators from rock joint surface images. Yadigar (2019) introduced an effective well geological facies classification model based on deep learning. In the domain of rock image recognition, many domestic researchers have also made significant progress. Liu Juexian et al. (2016) utilized texture, shape, and spatial features as parameters for support vector machine classification, achieving favorable results. Cheng Guojian et al. (2015) similarly analyzed color, texture, shape, and spatial characteristics of rock slice images for SVM-based classification with good outcomes. Xu Shuteng (2018), leveraging the TensorFlow deep learning framework, designed a specialized U-Net convolutional neural network model to automatically extract deep feature information of ore minerals under a

mineralogical microscope, enabling intelligent identification and classification of microscopic ore minerals. With current breakthroughs in deep learning for image processing, combining these advantages will significantly enhance cuttings image recognition.

Based on deep convolutional networks and transfer learning methods, this study employs image sample sets of black coal, gray-black mudstone, and gray fine sandstone to establish a corresponding automatic recognition and classification model. Testing results demonstrate that the model possesses strong recognition capabilities.

2 Introduction of Related Theories and Models

In rock image classification tasks, experts must first manually label rock types in sample images and create network training files based on color, texture, and particle size characteristics for discrimination. This process is difficult, time-consuming, and labor-intensive, resulting in very limited labeled training sets. Therefore, for rock image classification with small sample datasets, selecting a network structure with strong generalization capability and fewer training parameters is particularly important.

The GoogLeNet network improves computing performance by clustering sparse matrices into denser sub-matrices. The “Inception” module serves as the basic building block of GoogLeNet, increasing both network width and depth. Multiple Inception modules combined with other structures form a 22-layer network characterized by high sparsity and computational efficiency. Furthermore, GoogLeNet achieved over 93% top-5 accuracy on the large-scale ImageNet dataset. Considering these advantages—fewer training parameters, good versatility, and high accuracy—this paper proposes using GoogLeNet Inception-V3 for rock sample classification based on transfer learning of training parameters (Hu, Yan & Xia, 2017; Szegedy, Vanhoucke, Ioffe, Shlens & Wojna, 2016). The schematic diagram of the GoogLeNet network structure is shown in Figure 1 [Figure 1: see original paper], with the basic Inception module illustrated in Figure 2 [Figure 2: see original paper].

The Inception-V3 network expands the network architecture without increasing computational cost, extracting more subtle features under equivalent computing power and improving training effectiveness. Its design theory primarily builds upon the Hebbian principle and multi-scale processing intuition to increase network depth and width (Wang & Fujimoto, 2018). In the Inception basic module, 1×1 convolution is employed to process input image dimensions and reduce computational cost. The Inception-V3 structure replaces 5×5 convolution with two consecutive 3×3 convolutions, further reducing computational expense. Rather than continuously stacking convolutional layers, it utilizes convolution kernels of different sizes to maintain a constant receptive field while enlarging it, ultimately achieving concatenation for multi-scale feature fusion. The most notable characteristic of Inception-V3 is the expanded convolutional computation between layers.

3.1 Sample Data and Data Enhancement

The rock image samples used in this experiment are sourced from the official dataset of the 9th “Teddy Cup” Data Mining Challenge, focusing primarily on rock images of black coal, gray fine sandstone, and gray-black mudstone for training and recognition. All image data is divided into training, validation, and test sets in a 60:20:20 ratio. The training set is used to train the deep convolutional neural network model after data augmentation; the validation set is used to select the optimal model; and the test set is used to evaluate the accuracy of the final model.

Since deep convolutional neural networks for image classification require large numbers of training samples to obtain optimal model parameters, and the samples collected in this study are far fewer than the hundreds of thousands typical in other fields, image processing techniques such as flipping, rotation, color contrast and saturation transformation, and random noise addition are employed to augment the original dataset. This serves dual purposes: obtaining a larger dataset and preventing the neural network from learning irrelevant patterns, thereby continuously improving network structure and fundamentally enhancing overall network performance. However, some of these techniques involve complex image processing that may alter relevant characteristics. To improve modeling efficiency, this study utilizes image flipping and translation techniques to rapidly expand the original data without changing the primary characteristics of rock images (Wu & Xiao, 2019). The augmentation operations are illustrated in Figure 3 [Figure 3: see original paper].

As shown in Figure 3, due to inherent limitations in rock texture characteristics, rock shapes vary in size and appear anywhere within images, making uniform network training difficult. By translating samples at different distances and rotating them at various angles, the rock image training sample volume can be expanded to approximately three times the original size. This not only improves model generalization capability but also enhances the accuracy of the rock sample classification network.

3.2 Realization of Rock Classification

The experimental environment uses Windows 10 as the operating system with TensorFlow 2.3.0. The training and testing processes based on Inception-V3 within the TensorFlow environment are illustrated in Figure 4 [Figure 4: see original paper] and Figure 5 [Figure 5: see original paper], respectively.

The data flow for transfer learning training with the Inception-V3 model proceeds as follows: First, a rock image is input and processed through the feature extraction model, which comprises the convolutional and pooling layers of Inception-V3. These layers are transferred and utilized to compute image features, representing each rock image as a 2048-dimensional feature vector stored in a buffer.

The complete network structure begins at the input side, starting with three convolutional layers connected to a pooling layer, followed by two convolutional layers connected to another pooling layer, and finally connecting to 11 mixed layers. The original model also includes a Dropout layer, fully connected layer, and Softmax layer. For rock image recognition in this study, these three final layers require retraining.

3.2.1 Feature Extraction

This step freezes the training weights of the model after pre-training in the source domain and transfers them to the target dataset. This approach enables training on small sample images by modifying only the final classifier of the network (Zhang, Li & Han, 2018). The purpose of feature extraction in transfer learning is to apply features extracted from the source domain to small sample data in the target domain, simplifying feature extraction while obtaining a better-performing model. The detailed feature extraction steps are as follows:

- (1) To prevent modification of basic weight information in the pre-trained model during training, weights pre-trained by Inception-V3 on the ImageNet dataset are used as input for the target dataset, with the top layer frozen to establish the base model.
- (2) Online data augmentation is performed on the small sample dataset to obtain more target data details and enhance model generalization capability.
- (3) The image input size is set to (299, 299, 3), with pixel values normalized from [0, 255] to [-1, 1] to build a feature extractor.
- (4) The defined base model and feature extractor are connected to construct the model, with global average pooling applied to the acquired feature vectors.
- (5) To further reduce computational load, Dropout is set to 0.3, the classification layer is established, rock sample images are classified, and the image classification model is constructed.

The final model parameters for feature extraction are: learning rate of 0.001; batch size of 32; number of iterations of 50; optimizer is Adam, which calculates adaptive learning rates for each parameter and offers easy implementation, low memory requirements, and high computational efficiency; classifier is Softmax.

The evolution process of various rock feature extractions across some layers during model training is shown in Figure 6 [Figure 6: see original paper].

3.2.2 Fine Tuning

After the network model completes training on the training and validation sets and their data reaches a state of convergence, fine-tuning is required to better adapt the target data to the target domain by unfreezing all or part of the base

model. Generally, for image classification tasks, the bottom layers of a model contain the most fundamental image features applicable to almost any image type.

Therefore, the bottom layers require no adjustment and remain frozen. The top portion of the model must be adjusted by continuing training with the top layers and the added classification layer on the small-sample target domain dataset. This adjusts the pre-trained network weights to better suit the small-sample target data, ultimately improving model classification accuracy.

3.3 Training Results and Model Evaluation

Since deep learning models are end-to-end architectures, manual feature selection is unnecessary; the dataset simply needs to include as many different rock image types as possible. When raw data is input, the model automatically extracts features for each category. During training, rock images from the validation set are randomly selected at each step for model prediction and evaluation. Their feature vectors are input for prediction, results are compared with actual category labels, and model weight parameters are updated through backpropagation. Model accuracy continues to increase with each iteration step. Although the feature selection process cannot be directly observed, feature selection quality and model results can be evaluated through training accuracy, test accuracy, and cross-entropy values.

By analyzing log files from the rock sample classification network training process, the accuracy and loss function curves of the Inception-V3 network for the three rock sample image categories are shown in Figure 7 [Figure 7: see original paper] and Figure 8 [Figure 8: see original paper], where the blue curve represents the training set and the orange curve represents the validation set.

Training accuracy refers to the percentage of correctly classified images in the current training batch, while test accuracy refers to the percentage of correctly classified images in randomly selected test samples. Cross-entropy indicates learning effectiveness during model training, with smaller values representing better learning outcomes. Each training prediction is compared with actual values, and the final layer weights are updated through backpropagation.

As shown in Figure 7 and Figure 8, the accuracy of both training and validation sets improves rapidly and eventually stabilizes above 80%, while cross-entropy decreases significantly and stabilizes at a low value, indicating good model training performance.

3.4 Model Test

Test set images not used in training are employed for identification and analysis to verify model generalization capability—specifically, whether the model can achieve effective recognition and classification of unseen images.

The test set comprises 5 black coal images, 7 gray fine sandstone images, and 7 gray-black mudstone images. Recognition and classification results are presented as probability values, with each image receiving three probabilities corresponding to the three rock types. The rock type with the highest probability is considered the predicted class (Hu & Wu, 2021). Results for black coal are shown in Table 1 , gray fine sandstone in Table 2 , and gray-black mudstone in Table 3 .

Table 1 shows that the model correctly recognizes and classifies all five black coal test images with probabilities above 90%, with most classification probabilities exceeding 95%. This indicates effective extraction of black coal characteristics and accurate identification of this category.

Table 2 demonstrates that the model correctly classifies all seven gray fine sandstone test images with probabilities above 83%, with most classification probabilities above 90%, showing accurate identification of gray fine sandstone.

Table 3 reveals that while the model correctly classifies all seven gray-black mudstone test images, the recognition probabilities are less stable. The model can accurately identify gray-black mudstone, but the recognition probability for this category needs improvement. This may be due to fewer samples in the gray-black mudstone training set and less distinctive sample features, affecting the model's recognition probability for this class.

Since results are presented probabilistically, with the highest probability determining the predicted rock type, the model correctly recognizes and classifies all three rock types. However, some images exhibit low recognition probabilities, likely due to limited training data and few or no images similar to the test set, resulting in inadequate feature extraction. Overall, the model correctly identifies and classifies images of all three rock types, providing valuable reference for future research and application in rock sample classification.

4 Summary

This paper establishes a deep learning transfer classification model for rock images based on Inception-V3, achieving effective recognition of black coal, gray fine sandstone, and gray-black mudstone. The model independently learns features without manual intervention, reducing subjective influence. Moreover, the training process has minimal requirements regarding rock image size and brightness. Testing with separate test set images yielded no errors, demonstrating strong model robustness and generalization capability, along with effective identification of rock characteristics. However, due to currently limited data sources, the model does not yet meet standards for practical engineering applications. Future work will involve collecting more data to improve model accuracy and meet real-world engineering needs.

References

- Anonymous. (2017). Status and progress of lithology identification technology. *Progress in Geophysics*, 32(141), 32-46.
- Izadi, H., Sadri, J., & Bayati, M. (2017). An Intelligent System for Mineral Identification in Thin Sections Based on a Cascade Approach. *Computers & Geosciences*, 99, 37-49.
- Hong, K., Han, E., & Kang, K. (2017). Determination of geological strength index of jointed rock mass based on image processing. *Journal of Rock Mechanics and Geotechnical Engineering*, 9(4), 702-708.
- Imamverdiyev, Y., & Sukhostat, L. (2019). Lithological facies classification using deep convolutional neural network. *Journal of Petroleum Science and Engineering*, 174, 216-228.
- Liu, J. X., Teng, Q. Z., & Wang, Z. Y. (2016). Multi-feature fusion rock classification based on collaborative representation. *Computer Applications*, 36(3), 854-858.
- Cheng, G. J., & Yin, J. J. (2015). Rock slice image classification based on SVM. *Science and Technology Innovation and Application*, (01), 38.
- Xu, S. T., & Zhou, Y. Z. (2018). Experimental research on intelligent recognition of minerals under mirror based on deep learning. *Acta Petrologica Sinica*, 34(11), 3244-3252.
- Hu, K., Yan, H., & Xia, M. (2017). Satellite cloud image cloud classification based on transfer learning. *Journal of Atmospheric Sciences*, 40(6), 856-863.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).
- Wang, Y., & Fujimoto, Y. (2018, September). Fast and High-performance Multi-convolution Deep Neural Network Structure with Residuals. In *2018 12th France-Japan and 10th Europe-Asia Congress on Mechatronics* (pp. 288-293). IEEE.
- Wu, R. X., & Xiao, Q. K. (2019). Multi-object image recognition based on deep network and data enhancement. *Foreign Electronic Measurement Technology*, 38(5), 86-90.
- Zhang, Y., Li, M. C., & Han, S. (2018). Automatic lithology recognition and classification method based on deep learning of rock images. *Acta Petrologica Sinica*, 34(02), 81-90.
- Hu, S. L., & Wu, J. (2021). A small sample image classification method based on transfer learning. *Journal of Hubei Institute of Technology*, 37(02), 27-32.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.