Applied Mathematics, Modeling and Computer Simulation
C. Chen (Ed.)

doi:10.3233/ATDE220044

Research on Rock Image Recognition Based on Deep Learning

Mingmin GONG[a,1], Lei ZHENG[b], Sheng ZHOU[a], Quan HAO[a] and Letao WANG[a]
[a] *School of Information Engineering, Wuhan College, Wuhan 430212, China*
[b] *Shen Zhen Institute of Technology, Shen Zhen 518116, China*

Abstract. The classification and identification of rock samples is an important part of geological analysis. In order to reduce the cost of identification and prevent human subjective judgment from affecting the classification results, the use of deep learning to establish intelligent classification is a new way. Based on the deep convolutional neural network model of the AlexNet network structure, a deep learning migration model for rock image set analysis is established, and the automatic recognition and classification of rock lithology is realized by using the migration learning method. The traditional method of artificial image recognition of rock samples has big drawbacks. Based on the deep learning method, this paper establishes an image recognition model suitable for cuttings and core samples using industrial cameras at the logging site. Due to the small number of samples, the data is enhanced first. Then, feature extraction is performed on the rock image data based on a network structure such as a convolutional neural network. Finally, compare and analyze the experimental results of each model, and obtain the best model data for automatic identification of rock lithology. Thereby reducing the cost of rock image recognition and improving the efficiency of geological work.

Keywords. AlexNet network, image recognition, image binarization.

1. Introduction

In oil and gas exploration, rock sample identification is a basic and important link; in mineral resource exploration, especially in solid metal mineral resource exploration, rock sample identification also plays an immeasurable role; rock sample identification and Classification is extremely important for geological analysis. At present, the methods of rock sample identification mainly include gravity and magnetic, logging, seismic, remote sensing, electromagnetic, geochemistry, hand specimens and thin section analysis methods, and the use of image deep learning methods to establish automatic identification and classification models of rock samples is a new way. The rock sample image recognition program based on deep learning can automatically extract the characteristics of image features, so that the drawbacks of manual feature extraction in traditional machine learning image classification have been solved, and the adverse effects of subjective recognition on experimental results can be effectively avoided. In addition, deep learning models can extract deeper and more complex abstract image features.

[1] Corresponding Author, Mingmin GONG, School of Information Engineering, Wuhan College, China; Email: zhangyingyi@bucea.edu.cn.

Thanks to this advantage, neural networks can achieve more types and numbers of classifications. Based on the deep convolutional neural network model of the AlexNet network structure, a deep learning migration model for rock image set analysis is established, and the automatic recognition and classification of rock lithology is realized by using the migration learning method. Using this method, 315 rock images were divided into training set and test set according to a certain ratio for learning and classification training. Through multiple training and model optimization, the test accuracy of the model reached 92% in a short number of training times. The image recognition rate after stretching is not good, and it may be solved by further supplementing the training set. According to the manual test results of the single-picture recognition program we designed, it has a relatively good effect on the 7 types of rocks included in the training set, and it has strong practicability.

1.1. Convolutional Neural Network

Convolutional neural network (convolutional neural network) is a neural network that contains a convolutional layer. The convolutional neural networks introduced here all use the most common two-dimensional convolutional layer. It has two spatial dimensions, height and width, and is commonly used to process image data. Although the convolutional layer is named after the convolution operation, we usually use the more intuitive cross-correlation operation in the convolutional layer. In a two-dimensional convolutional layer, a two-dimensional input array and a two-dimensional kernel (kernel) array output a two-dimensional array through a cross-correlation operation.

1.2. Alexnet

In 2012, AlexNet turned out(structure as shown in figure 1). The name of this model comes from the name of the first author of the paper, Alex Krizhevsky [1]. AlexNet (figure 1) uses an 8-layer convolutional neural network and won the ImageNet 2012 Image Recognition Challenge with a great advantage. It proved for the first time that the learned features can surpass the manually designed features, thus breaking the previous state of computer vision research in one fell swoop.

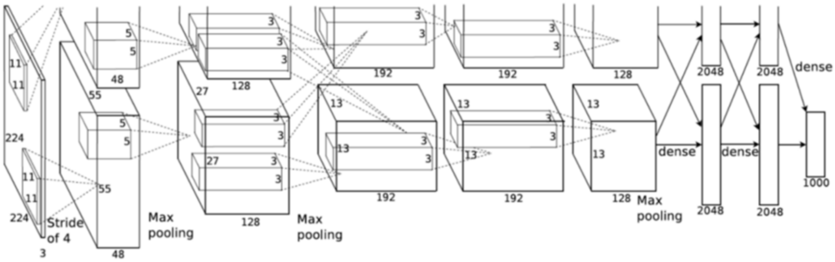


Figure 1. AlexNet network structure

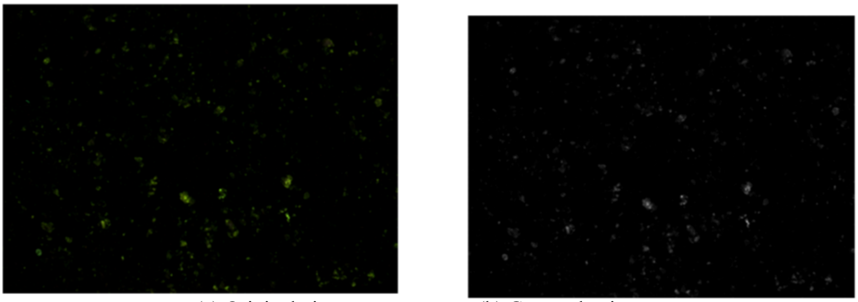
1.3. Constructing an Intelligent Recognition Model of Rock Sample Lithology

Traditional metthods of artificial image recognition of rock samples have major drawbacks. Based on the deep learning method, this paper establishes an image recognition model suitable for cuttings and core samples using industrial cameras at the logging site.[2] Due to the small number of samples, the data is enhanced first. Then, feature extraction is performed on the rock image data based on a network structure such as a convolutional neural network. [3] Finally, analyze the experimental results of each model, and obtain the best model data for automatic identification of rock lithology. Thereby reducing the cost of rock image recognition and improving the efficiency of geological work.

2. Data Preprocessing

2.1. Image Binarization

Binarization is the simplest method of image segmentation. Binarization can convert grayscale images into binary images. Set the grayscale of the pixel greater than a certain critical grayscale value as the maximum grayscale value, and set the grayscale of the pixel less than this value as the minimum grayscale value, so as to realize the binarization. In the binarization of images with low brightness can achieve good results. First, binarize such images, as shown in figure 2 below.

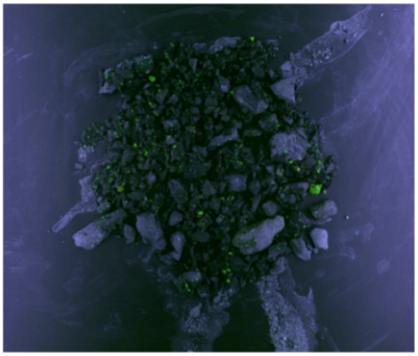


(a) Original picture (b) Grayscale picture

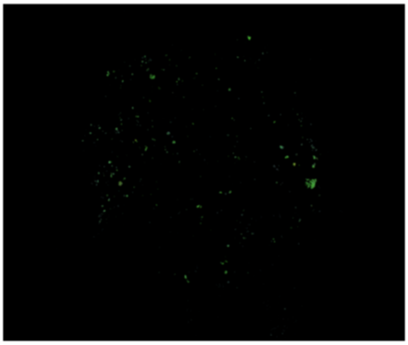
Figure 2. Image binarization

2.2. Image Division

In the task, the brightness of some pictures is not low, so the direct binarization effect of such pictures is not good. For this type of picture, directly use the RGB value to limit the upper and lower intervals, and get the green part of the picture, as shown in figure 3 below.



(a) Original Picture



(b) Green Part Extraction

Figure 3. Image division

2.3. Data Enhancement

The original data set has 315 pictures, which is far from enough for the neural network, so the data set is expanded through data enhancement.

Because the number of pictures of various types of rocks in the original data is different, the number of pictures is balanced during the enhancement. Each type of rock is 4000 pictures, for a total of 28,000 pictures. All pictures are standardized and written into Pytorch's Tensor data structure.

2.4. Data Set Division

In the data set, pictures with -1 suffix are used for training, and pictures with -2 suffix are used to detect the percentage of oil and gas. The training set is divided into a training set and a test set. The training set is composed of the original data and the enhanced data, and contains 28,000 pictures, each with 4,000 pictures. The test set consists of 2100 pictures with additional data enhancement, 300 pictures in each category. 3 Research on rock image recognition method based on deep learning

3. Identification Method Design

3.1. Network Structure Design

In order to avoid over-complex network model leading to over-fitting and long calculation time, compared with the original AlexNet, the network structure in this paper simplifies the number of network layers, from the 8-layer network of AlexNet to 4-layer network, as shown in figure 4 below .

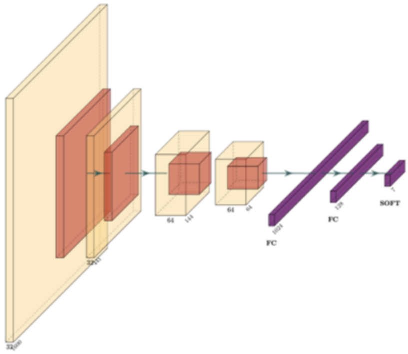


Figure 4. Visualization of network structure

In addition, each layer of the network uses dropout [4] to deal with the overfitting problem. Batch normalization [5] is used in the network to normalize the input of the layer. During model training, batch normalization uses the mean and standard deviation of the small batch to continuously adjust the intermediate output of the neural network, so that the entire The output value of the neural network in the middle of each layer is more stable.

3.2. Loss Function

In this paper, the cross entropy function is used as the loss function,[6] and the expression of the loss function is as follows (see Equation 1):

$$L = -[ylog\hat{y} + (1 - \mathrm{y})\log(1 - \hat{y})] \tag{1}$$

Defined in Pytorch as:

criterion = nn.CrossEntropyLoss()

3.3. Optimizer

In this paper, Adam is used as the optimizer. The Adam algorithm also does an exponentially weighted moving average of small batch stochastic gradients on the basis of the RMSProp [7] algorithm. This algorithm can be regarded as a combination of the RMSProp algorithm and the momentum method. Through the optimizer instance named "Adam", we can use the Adam algorithm provided by PyTorch:

optimizer = torch.optim.Adam(model.parameters(), lr=0.001)

3.4. Evaluation Index

This article evaluates the training effect of a network from the training loss (Train Loss), test accuracy (Test Accuracy) and other indicators. The training loss is the calculation result of the loss function, which can be expressed as:

$$L_{train} = criterion(output, label) \tag{2}$$

Test accuracy refers to the ratio of the model outputting correct results on the test set. The calculation formula of test accuracy is as follows (see Equation 2):

$$P_{test} = \frac{n_{testCorrect}}{n_{testSet}} \tag{3}$$

4. Test Process

4.1. Model Establishment

This article uses the Pytorch deep learning framework to train models and tests. PyTorch is an open source machine learning library based on the Torch library, which is mainly used for applications developed by Facebook's AI Research Laboratory (FAIR), such as computer vision and natural language processing.

The model in this paper uses a convolutional neural network with 4 convolutional layers, in which the number of pooling layers is 4, and the number of fully connected layers is 2. After the first layer of the network convolution operation is over, the convolution window shape is 40 * 40, and then a maximum pooling is performed. The core size of each pooling layer is 2 and the number of steps is 2. After pooling, batch normalization is performed once, and the final window shape is 20 * 20. After passing through the second pooling layer, the convolution window is 10*10, after the third layer pooling, the volume machine window is 6*6, after the last layer pooling, the window size is 4*4, and the number of windows is 64. After completing two full connections, the result is mapped to the category number 7.

4.2. Experimental Results and Comparative Analysis

Table 1. Hyperparameter settings and test accuracy

Number	LearningRate	BatchSize	epochs	dropout	Accuracy
1	0.0001	128	100	0	58.12%
2	0.0001	128	200	0	64.81%
3	0.0001	128	200	0.2	86.23%
4	0.0001	128	200	0.2	86.80%
5	0.005	128	200	0.2	92.57%
6	0.0005	256	200	0.2	90.42%
7	0.001	512	200	0.2	91%
8	0.0005	1024	200	0.2	90%

The model hyperparameter settings and test accuracy are shown in table 1.

It can be observed that before using the dropout method, the model has overfitting. Overfitting means that the gap between the training error and the test error is too large. In other words, the complexity of the model is higher than the actual problem. The model performs well on the training set, but it performs poorly on the test set. Such a model has poor generalization ability. After using the discarding method, the occurrence of over-fitting is well suppressed, and the performance on the test set is good. The learning rate as a hyperparameter controls the magnitude of weight update, as well as the speed and accuracy of training. If the learning rate is too large, the target (cost) function will fluctuate greatly, making it difficult to find the optimum. If the learning rate is too small, the convergence will be too slow and time-consuming. After testing, the performance is best when the learning rate is 0.005, and the final test accuracy is stable at about 92%. The change of the number 5 hyperparameter test accuracy with the increase of the number of training is shown in figure 5 below, and the change of Loss is shown in figure 5 and figure 6 below.

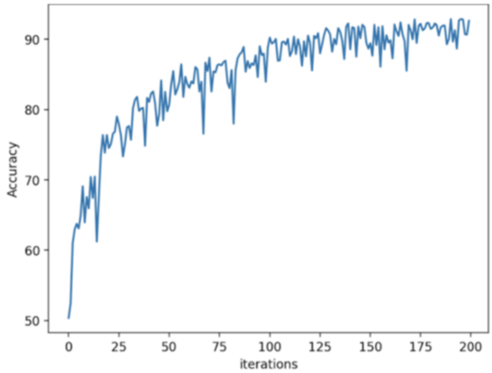


Figure 5. Test accuracy change chart

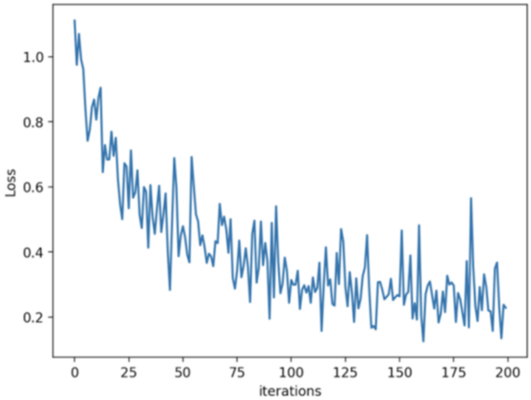


Figure 6. Loss change diagram

5. Summary and Outlook

In the process of training the model, the discarding method effectively suppressed the occurrence of over-fitting and achieved good results. At the same time, we studied the impact of different learning rates and batch size settings on the experimental results, and selected the best performing parameters for training, and we got a training model with a test accuracy of up to 92%. The image division algorithm successfully eliminates the non-oil part of the picture by setting appropriate parameters, and the calculated oil occupation is more accurate. Although the accuracy of the model test in this article is good, the effect in actual production still needs to be investigated, and the limitations are more obvious. Mainly the following three questions:

(1) It is difficult to recognize the image after stretching and deformation. In the single-picture prediction program of the subsequent test, the center of the image must be cropped to ensure that the image is not deformed. This is the result of the lack of stretching and deformation in the training set, which leads to the general use of the model Sexual decline.

(2) There are too few samples of a certain type of rock in the data set, and the cropping window is fixed, which makes it easier to generate exactly the same pictures when the data is enhanced. This will have a certain impact on the training results and reduce the versatility of the model.

(3) The accuracy of the image division algorithm designed when calculating the oil content is difficult to guarantee, and the green and yellow in the picture cannot be perfectly extracted. In addition, it is difficult to accurately identify the part of the picture that is not a rock, and it may be better to use other methods. Therefore, designing a more versatile image recognition network structure and image partitioning algorithm is an important research direction in the future.

Acknowledgement

Fund Project: Supported by the Industry-University-Research Innovation Fund of the Science and Technology Development Center of the Ministry of Education (2018A02016) Introduction to the author: Gong Mingmin, female, born in 1977, master degree, associate professor, research interests in image processing, machine learning; Zhou Sheng , Male, born in 1978, PhD, research direction is mathematical theory, data analysis.

References

[1] Aston Zhang, hands-on deep learning, Beijing: People's Posts and Telecommunications Press, 2019.
[2] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.
[3] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).
[4] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. JMLR
[5] "Glossary of Deep Learning: Batch Normalisation". medium.com. 27 June 2017. Retrieved 24 April 2018.
[6] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
[7] Tieleman, T., & Hinton, G. (2012). Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. COURSERA: Neural networks for machine learning, 4(2), 26-31.