

Identifying Rock Thin Section Based on Convolutional Neural Networks

Ren Wei ⁺, Zhang Minghua, Zhang Sheng, Qiao Jihua, Huang Jinming

China Geological Survey Development and Research Center, Beijing 100037, China

Abstract. This paper uses the convolutional neural networks to train and identify rock thin sections. There are three innovations in this method. Firstly the convolutional neural network is used in the field of geological experiment testing and it can automatically identify and classify the rock thin sections. Secondly the original image of rock thin sections are sliced and segmented, so that the convolutional neural network could learn more details of the rock mineral texture without damaging the original resolution of the image. Thirdly using other image enhancement techniques, such as random flipping and standardization, can expand the sample data set and enhance the robustness of the model. Finally the training model achieves the desired results.

Keywords: deep learning, convolutional neural network, rock thin section, lithology identification

1. Introduction

In geology work, mineral samples are often used for identification, and the most common sample is the rock thin section. Through mineralogy, petrology, chalcography and other methods combined with microscopic observation, the rock thin section mainly uses the collected rock or ore samples to research on the mineral composition, ore structure, mineral symbiosis and generation sequence, alteration features, secondary change of ore and surrounding rock. Therefore, it can provide information basis for determining the mineral type of rock or ore, analyzing geological structure, inferring the geological conditions of deposit formation, understanding the performance of ore processing technology, and classifying the natural type of ore.

Because the light-colored main minerals in the rock thin section under the microscope are colorless and transparent, and the refractive index is similar, the gray scales are very similar and the image contrast is very low or there is almost no contrast^[1]. Therefore, the classification and identification of the traditional rock thin section is difficult, requiring observers to adjust the slice angle and magnify the power, and repeatedly confirm the mineral type and content to complete the classification. Geologists, such as Wu Jimin, have also made a lot of attempts in the automatic identification of thin sections, using the gray threshold method to segment the image of the section. However, the threshold is difficult to control, and the resolution of the minerals with similar gray scales as quartz is hard to distinguish^[2]. Researchers, such as Zhang Jie, selected the RGB component, HSI component and the gray mean value of the pixel as features. But the image segmentation is hard to infer the mineral cleavage feature^[3]. Li Zhiqiang used the support vector machine (SVM) classification method to classify the granite section images and complete the automatic division of quartz^[4]. Until now, there have been few papers on the automatic classification and identification of various types of rock thin sections.

There are three innovations in the method of using convolutional neural networks to identify rock thin sections. Firstly the convolutional neural network is used in the field of geological experiment testing and it

⁺ Corresponding author. Tel.: +86-010-58584351 ; fax: +86-010-58584351.
E-mail address:renwei@mail.cgs.gov.cn.

can automatically identify and classify the rock thin sections. Secondly the original image of rock thin sections are sliced and segmented, so that the convolutional neural network could learn more details of the rock mineral texture without damaging the original resolution of the image. Thirdly using other image enhancement techniques, such as random flipping and standardization, can expand the sample data set and enhance the robustness of the model. Finally the training model achieves the desired results.

2. Deep Learning

Deep learning is a branch of machine learning, and originates from the research of artificial neural network, which builds a multi-layer neural network structure similar to human brain, extracts the input data from the bottom to the top, and deeply mines the data origin to achieve learning. Deep network structure obtained by deep learning includes a large number of single neuron, and each neuron is connected with a large number of other neurons. The connection weight between neurons is modified during the learning process and determines the network function. The deep network structure obtained through deep learning complies with the characteristics of neural network.

In the past, multi-layer neural networks faced many difficulties in training. On the one hand, due to data set limitations, training on insufficient data led to over-fitting; on the other hand, the impact of network complexity was affecting the result. The commonly used backpropagation algorithm (BP) also has problems, such as falling into local optimum and gradient dispersion, which undoubtedly hinders further research in neural networks. In 2006, the greedy layer-by-layer training algorithm proposed by Professor Hinton made a breakthrough in deep learning^[5]. The greedy layer-by-layer training algorithm solves the difficulties of deep learning in training, and the multi-level neural network has excellent learning ability, through which the obtained features can express better, so as to achieve better visualization and classification. The wave of deep learning in academy and business fields has become an inevitable trend in the development of Internet, Big Data and artificial intelligence.

Each module in the deep learning model transforms the low-level representations into high-level representations, compounding transforms as much as possible in a layer-by-layer mode to learn complex functions. According to the structure, deep neural networks are divided into the following categories: feed-forward deep networks (FFDN), feed-back deep networks (FBDN), bi-directional deep networks (BDDN)^[6]. For deep learning-based cognition, high-level representations can enhance certain aspects of input and suppress unrelated variables. For example, in an image where array represent pixel values, whether an edge exists at particular position or direction of the image is described in the first layer of the representation, the second layer detects the block formed by the edges, and the third layer combines the blocks to form a certain portion of an object and a larger block. Finally, these large pieces are combined into one body to construct a target to be detected. The feature layer of deep learning does not need to be manually designed, and it automatically acquires the input hierarchical features through system training, and finally integrates the features learned in each sub-region into advanced features.

3. Convolutional Neural Network

In deep learning, the convolutional neural network is a deep and artificial feedforward neural network, which has become one of the popular researches in many scientific fields. The convolutional neural network is a neural network that uses the local characteristic training data of the image to construct a partially connected network, where the hidden layer uses the convolution kernel method to calculate. The core idea is the local connection, weight sharing, and time or spatial sampling. The hierarchical structure of the convolutional neural network includes a data input layer, a convolutional layer, an activation function layer, a pooling layer, and a fully connected layer. The convolutional layer is used to detect partial connections of features in the previous layer, and the pooling layer is used to fuse similar features. Two or three phases of convolution, activation, and pooling are stacked so that there are more convolution and fully connected layers, and the gradients are propagated backwards through the network, allowing the weights of all convolution kernel groups to be trained. The convolutional neural network can deal with the problem of unclear inference rules in complex environments. It has high fault tolerance, parallel processing capability, and self-learning ability. This network structure is highly invariant to translation, scaling, tilting and other

deformations. Convolutional neural networks have a good application effect in the fields of speech recognition, image recognition, face recognition, and autopilot [7-14].

The input to the convolutional neural network is usually the original image X . In this paper, H_i is used to represent the feature map of the i layer of the convolutional neural network ($H_0=X$). Assuming that H_i is a convolutional layer, the generation process of H_i can be described as:

$$H_i = f(H_{i-1} \otimes \Omega_i + b_i) \quad (1)$$

Where: W_i represents the weight vector of the i layer convolution kernel; the operation symbol “ \otimes ” represents the convolution operation of the convolution kernel and the $i-1$ layer image or feature map, and the output of the convolution and the offset of the i layer The vector b_i is added, and finally the feature map H_i of the i layer is obtained by the nonlinear excitation function $f(x)$.

In the training process, the commonly used optimization method for convolutional neural networks is the gradient descent method. The residuals are backpropagated by gradient descent, updating the trainable parameters (W and b) of the various layers of the convolutional neural network layer by layer. The learning rate parameter (η) is used to control the strength of residual back propagation:

$$W_i = W_i - \eta \frac{\partial E(W, b)}{\partial W_i} \quad (2)$$

$$b_i = b_i - \eta \frac{\partial E(W, b)}{\partial b_i} \quad (3)$$

4. Construct Convolutional Neural Network

The convolutional neural network used in this experiment was VGG16^[15], and VGG16 won the first place as the world's largest image recognition database (ImageNet) in the annual challenge of 2014. The network is characterized by the use of a smaller convolution filter (size 3×3), which helps to extract local features more carefully, reduces the number of parameters, and improves the discriminability of decision functions^[16]. The network consists of five stacked ConvNet layers, three fully connected layers and one Softmax layer of the convolutional neural network, which is a "network-in-network" architecture (Figure 1). Each ConvNet layer consists of multiple convolutional layers, followed by pooling layers. After the output goes through the convolution and pooling layers and three layers of fully connection layers, the output of the last fully connected layer is used as input to the Softmax regression model, and the regression model generates the results of the rock thin section classification. The network also adds a nonlinear ReLU activation function layer, and the output of the convolutional layers and the fully connected layers are processed by the ReLU activation function, which can greatly shorten the network training time. In addition, the network uses a regularization method, Dropout, to avoid overfitting on the fully connected layer.

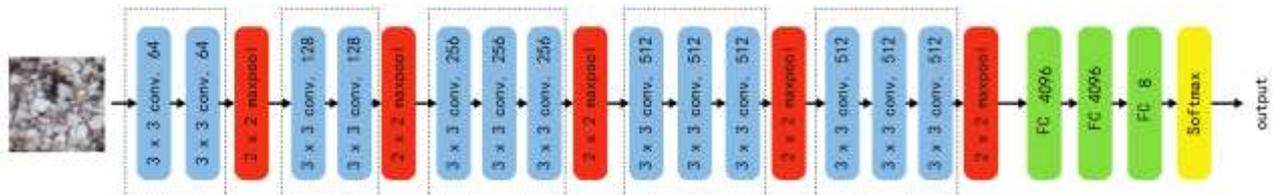


Fig.1: Data flow diagrams of VGG16 convolutional layers.

This time we built the learning and training process of deep neural network on the Tensorflow framework. TensorFlow is a Google open-source artificial intelligence learning framework developed at the end of 2015. It has rich applications in graphics classification, audio processing, recommendation systems, and natural language processing. TensorFlow can be run on CPUs and GPUs and even desktop PCs, servers, and mobile devices. It provides the most comprehensive library of algorithms in all open-source deep learning frameworks, which have basically met most of the common application requirements. These advantages make TensorFlow the most popular deep learning program.

5. Dataset Preparation for Deep Learning Training

Rock thin section micrographs have different advantages and disadvantages than traditional deep learning image recognition. The disadvantages are: First, the rock microstructure does not have obvious and specific contour features like objects, but specific texture information and mineral composition. Second, the rocks are mostly composed of light-colored minerals, and the photomicrographs are relatively single. Third, due to the formation mechanism of magmatic rocks and sedimentary rocks, the internal structure is disordered, unlike the traditional objects whose internal structure has a strict positional relationship. The advantages are also obvious. First, the photos are highly standardized, with uniform magnification and light intensity. Second, the resolution of the photos is large and the image is clear.

To perform deep learning, first build an image dataset for training neural networks. Photographs of rock thin section are often observed using both single-polarization and orthogonal polarization modes. In order to learn the original color characteristics of rocks, only single-polarized photos are selected for learning training.

The rock thin section data comes from the Natural Resources Physical Geological Data Center of China Geological Survey. It is the official geological data collection institution of the state. It undertakes the collection, management, development, research and utilization of important national geological data, and provides decisions for government authorities. Supporting business technology and providing public welfare services to the society. This time, we selected three typical types of thin section datasets in the eight major rock types, including sandstone, granite, limestone, basalt, peridotite, sulphate, gneiss, and tuff, nearly 100 sheets of images in each category .

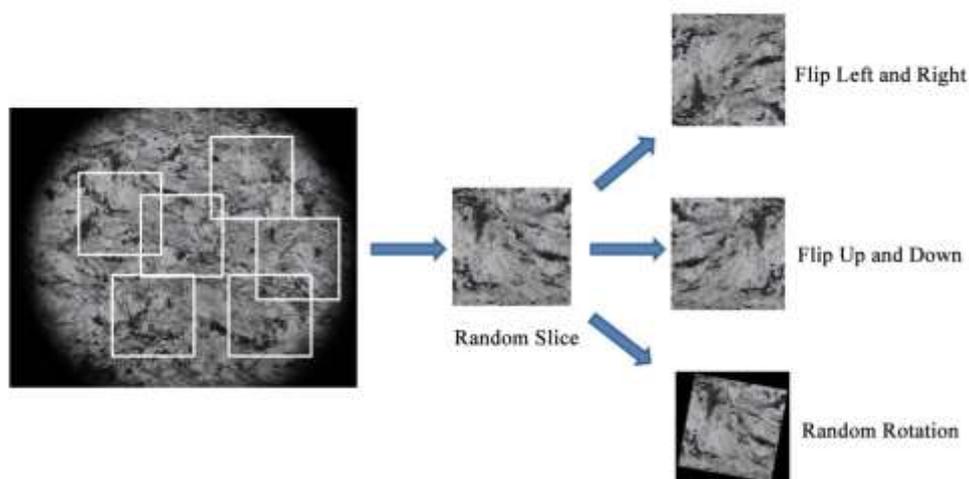


Fig.2: A method for slicing microphotograph of rock thin section.

Due to memory limitations, the pictures of the training set and test set cannot be too large. The pictures should be adjusted to 300*300 pixels. In order to preserve the details of the rock thin section to the maximum extent, the picture is not scaled and the required image size is only cut out from the original image. Because of the disorder of the rock internal structure, the slice position is randomly selected within the original picture. In order to enhance the robustness of the training results, we randomly flipped the thin section up and down, left and right, while rotating a random angle. Finally, 8 types of rock thin section photomicrograph training sets were obtained, each with 10,000 pictures. The same method was used to process other micrographs of thin sections to obtain a test set with multiple pictures for each type of rocks for verification of the experimental results.

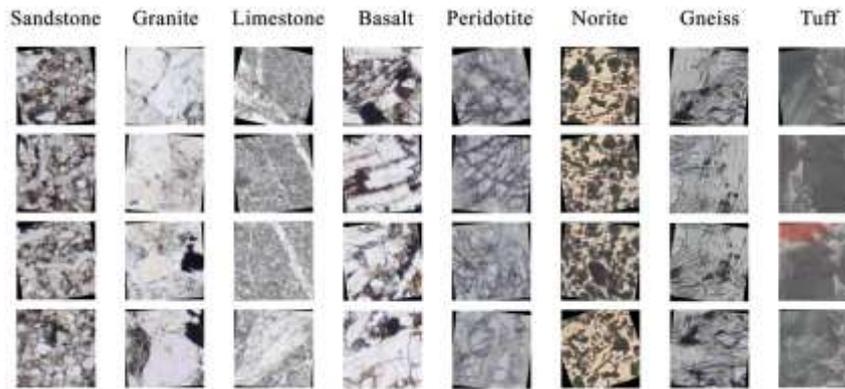


Fig.3: Training set and test set examples of rock thin section.

6. Model training and result analysis

The training process of the neural network model is run on the NVIDIA GeForce GTX 1060 graphics card with 6 GB of memory. The training uses a stochastic gradient descent method, and sets the initial learning rate to 0.01. Due to the memory limitation, 32 pictures are trained in each batch, and a total of 10,000 batch iterations were performed.

Initially, the training set selected thin sections randomly and imported them to the convolutional neural network for training. The blue curve in Fig. 4 (a) is the loss function value of the training process. The convergence of the loss function value is not as expected, and the curve fluctuates greatly. The loss function value of the model after iterating 4,000 epochs is starting to oscillate between 0.4-1.3. Then to improve the training effect, we standardized the thin sections before importing them into the model by mapping the RGB values of each pixel to the interval of 0 to 1. The training result of the green curve in (b) shows that the convergence of the loss function value is greatly improved. The loss function value is finally stable at 0.25, and the fluctuation amplitude of the curve is also reduced.

In order to further improve the training effect, the transfer learning technique is introduced, where the parameter model of VGG16 trained by ImageNet^[17] is introduced into the model of this training as the initial value. Transfer learning can solve the problem of the network's limited feature learning ability when the number of picture training samples is insufficient, and it can improve the convergence of the training loss function value^[18-20]. From the red curve of (c), we found that the training effect is surprisingly excellent. The loss function value curve converges fast during the training process. The loss function value is maintained at near 0 after about 100000 times of iteration, and the loss function value curve does not fluctuate greatly but is relatively stable.

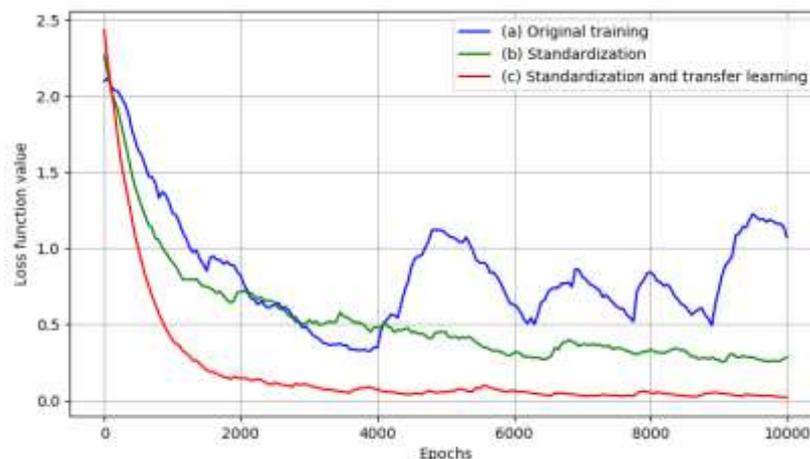


Fig.4: Model training loss function value curve.

- (a) The results of the original training;
- (b) The results of training after standardization;
- (c) The results of training after standardization and transfer learning.

Finally, the models trained before and after using the transfer learning technology are tested and verified. 992 slices of different types are randomly selected in the test set prepared in the data preparation stage, and imported for identification. The statistical results are shown in Table 1. We found that the accurate rate of the training model after transfer learning is indeed higher, reaching 95.16%. The rock thin section identification model achieves the desired results.

In order to verify the effect of the model, we used the Inception v3 convolutional neural network to verify the training of this data set. Inception v3 is a strong competitor in VGGNet in the ImageNet competition. By rethinking the inception architecture, computational efficiency and fewer parameters are realized. With fewer parameters, 42-layer deep learning network, with similar complexity as VGGNet, can be achieved. Using the test data to verify the trained Inception v3 model, we found that the accuracy rate was slightly inferior to VGG16.

Table1: Test verification of training results

Model category	Number of Tests	Correct	Accuracy
VGG16(standardization only)	992	847	85.38%
VGG16(migration learning)	992	944	95.16%
Inception V3	992	868	87.50%

7. Conclusions

In this paper, the deep learning method is used to train and identify the classification of eight types of rock thin sections. The methods of random position slicing, random flipping and rotating slicing are used to expand the number of training sets and enhance the robustness of the models. In the training phase, we used standardized and transfer learning techniques on the input pictures to accelerate the convergence of accuracy. The training model obtained by the final optimization has an accuracy rate of 95.16% for identification of the test set. Also, the Inception v3 model was used for comparison verification, and the accuracy rate was found to be slightly inferior to VGG16.

Deep learning is feasible in the field of rock thin section identification with a promising application prospect. At the same time, deep learning has a strong application scene expansion, which is very suitable for other geological fields, such as classification and identification of paleontological fossils, borehole cores and rock ore specimens. This article serves only as a trigger, expecting that other scholars will apply deep learning techniques to more geological fields and resolve more geological issues in the future.

8. References

- [1] Wu Jimin. Application of Thin Slice Microscopic Delineation and Automatic Image Analysis Technology in Quantitative Evaluation of Petrology. *Mineral Rock*, 19(2), 1999, 6: 25-30.
- [2] Wu Jimin. Image Threshold Calculation Method and Extraction of Mineral Information in Rock Flakes. *Advances in Science and Technology of Water Resources*, 18(3), 1998: 24-27.
- [3] Zhang Jie, Teng Qizhi, Wang Zhengyong. Rock Slice Image Segmentation Based on Support Vector Machine. *Journal of Chengdu University of Information Technology*, 2007,(4):186-189.
- [4] Li Zhiqiang, Chen Xiaogang. Study on Quartz Segmentation Method of Granite Slice Image Based on SVM. *Science Mosaic*, 2008,(10):108-111.
- [5] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, 18(7): 1527-1554.
- [6] YIN Baocai, WANG Wentong, WANG Lichun. Review of Deep Learning Research. *Journal of Beijing Polytechnic University*, 2015, 41(01):48-59.
- [7] Li Yandong, Hao Zongbo, Lei Hang. A review of convolutional neural networks[J]. *Journal of Computer Applications*, 2016, 36(9): 2508-2515.
- [8] IMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition // *ICLR.[S.l.]*:

[s.n.], 2015:1-14.

- [9] HAN K, WANG D L. Neural network based pitch tracking in very noisy speech. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, 2014, 22(12): 2158-2168.
- [10] LeCUN Y, BOTTOU L, BENGIO Y, et al. Gradient based learning applied to document recognition // *Proceedings of the IEEE. Berlin: IEEE*, 1998: 2278-2324.
- [11] LAWRENCE S, Lee GILES C, Ah CHUNG TSOI, et al. Face recognition: a convolutional neural network approach. *IEEE Trans on Neural Networks*, 1997, 8(1) : 98-113.
- [12] TIVIVE, FOK HING CHI, BOUZERDOUM, et al. An eye feature detector based on convolutional neural network // *Proc 8th Int Symp Signal Process Applic. Sydney, Australia: IEEE*, 2005: 90-93.
- [13] MATE S, AKIRA Y, MUNETAKA Y, et al. Pedestrian detection with convolutional neural networks // *Proceedings on IEEE Intelligent Vehicles Symposium. Madrid, Spain: IEEE*, 2005: 224-229.
- [14] Le CUN Y, MULLER U, BEN J, et al. Off-road obstacle avoidance through end-to-end learning . *Advances in Neural Information Processing Systems. Cambridge, USA: MIT Press*, 2005.
- [15] Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks . *Advances in Neural Information Processing Systems*, 2012, 25:2012.
- [16] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2016-11-20]. <http://www.robots.ox.ac.uk/~vgg/publications/2015/Simonyan15/simonyan15.Pdf>.
- [17] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet large scale visual recognition challenge . *International Journal of Computer Vision*, 2015, 115(3): 211 – 252.
- [18] DAI W, YANG Q, XUE G R, et al. Boosting for transfer learning // *Proceedings of the 24th International Conference on Machine Learning. New York: ACM*, 2007: 193 – 200.
- [19] YANG Q, CHEN Y, XUE G R, et al. Heterogeneous transfer learning for image clustering via the social Web // *Proceedings of the 47th Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the AFNLP. Stroudsburg, PA: Association for Computational Linguistics*, 2009: 1-9.
- [20] XIE M, JEAN N, BURKE M, et al. Transfer learning from deep features for remote sensing and poverty mapping // *Proceedings of the 30th AAAI Conference on Artificial Intelligence. Menlo Park, CA: AAAI Press*, 2016: 3929-3935.