Core Photo Lithological Interpretation Based on Computer Analyses

Alexander V. Ivchenko
Department of Radio Engineering and Cybernetics
Moscow Institute of Physics and Technology
Dolgoprudny, Russia
ivchenko.a.v@phystech.edu

Evgeny E. Baraboshkin[1], Leyla S. Ismailova[2], Denis M. Orlov[3], Dmitry A. Koroteev[4]
Department of Petroleum Engineering
Skolkovo Institute of Science and Technology
Moscow, Russia
[1]Evgenii.Baraboshkin@skoltech.ru,
[2]L.Ismailova@skoltech.ru, [3]D.Orlov@skoltech.ru,
[4]d.koroteev@skoltech.ru

Evgeny Yu. Baraboshkin
Department of Regional Geology and Earth History
Moscow State University
Moscow, Russia
ejbaraboshkin@mail.ru

***Abstract*— Demands for speed which geological data should be processed is rising. The problem doubles when you don't have any unite Government or production standards for the processing of it. In addition to some geological tasks, there are no criteria for subjectivism reduction or hard to distinguish in boundary rock task. On the basis of machine learning and neural networks, we demonstrate the automatization of geological information analysis based on core images. We show the ability of the algorithm to distinguish six lithotypes with highest precision 98 % on several classes.**

Keywords— machine learning; neural networks; sedimentology; core; images; description; segmentation; lithology; automatization

I. Introduction

Geophysical, geochemical, geological data are continuously collected by different methods. Those methods give a large amount of information which are continuously growth through technology development. At the same time, companies want all this data to be processed with minimal costs [9]–[11], [18], [20]–[21]. Such challenges can be solved with machine learning and neural networks application. Additional advantages of such method are lack of subjectivity in those questions which will be solved by computer and a unified scheme of data preparation and processing. Here, we describe methods for automated core sample analyses.

A. State of the art

Several works address the problem of lithotypes recognition. In 2012 A. K. Patel and others [16] made classification of rocks with SVR classifier. Six rock types were identified with high precision. Results were obtained by unified sampling at special conditions. Each sample was taken using the same camera in daylight. Authors didn't show any results on distinguishing geological samples with different structures and textures.

I. S. Khasanov made several works [1] devoted to the half-automated core image processing. Images are converted to RGB or HSL format. An operator gives petrophysic properties (porosity, permeability, etc.) for each color class and corresponding lithotypes. Afterwards, the program analyses distribution of color in the core sample and plots histograms. As the input data, photos taken both in daylight and in UV. The disadvantage of this method is that you need an expert to give properties for each color class. This tool can be used to improve the data processing speed.

W. E. Ellington's [14] analyzed color and type of rock using the same approach. The program determines rock color parameters for each sample. Subsequently, it analyzes the whole well. Finally, threshold by color margin is set for each lythotype. By using this threshold, synthesized well logs of lythotypes are built for the whole well.

Recently, M. M. Mezghani [15] performed automated core sample analysis using high definition core photos taken in daylight. Additionally, the author used fullbore microimager images (FMI) in his dataset. Missing parts of core (for example, cylindrical samples from full core (plugs), for terminology refer to [1]) have been reconstructed from FMI and surrounding structure and texture by multi-point statistic (MPS). The disadvantage of this approach that it reduces the precision of the method. Main gap of the work is comparison of all images with image samples library to estimate lithology by contrast and pixel intensity. These parameters are not quite good for description of rock, since they strongly depend on the

978-1-5386-5824-6/18/$31.00 ©2018 IEEE

quality of photos. Also the author did not use photos taken in UV.

B. Goals and objectives

The main goal of our research is to make an algorithm which will correctly detect rock type from large amount of images for short amount of time (from minutes to hours). Such an algorithm is highly demanded in case of large dataset processing.

We use the following strategy: (1) determine and classify parameters, which will be used for the research; (2) gather data for algorithm training; (3) process the data for training purposes; (4) train the algorithm.

II. METHODS AND MATERIALS

A. Data preparation

For easier data manipulation core properties tree was made. It increased accuracy of analyses and reduced computational complexity.

An essential part of the project is preliminary core images processing data preparation. It includes segmentation of the image into separate fragments and preliminary extraction of images from standard formats for storage of geological reports XLS and Corel (Fig. 1 and Fig. 3).

B. Analizator work flow

The analyzer takes as an input a color or grayscale image and converts it into an n-dimensional color distribution matrix and normalizes these values in the range from 0 to 1. Afterwards, it includes segmentation, reduction in the dimension of the matrix (convolution), and other transformations.

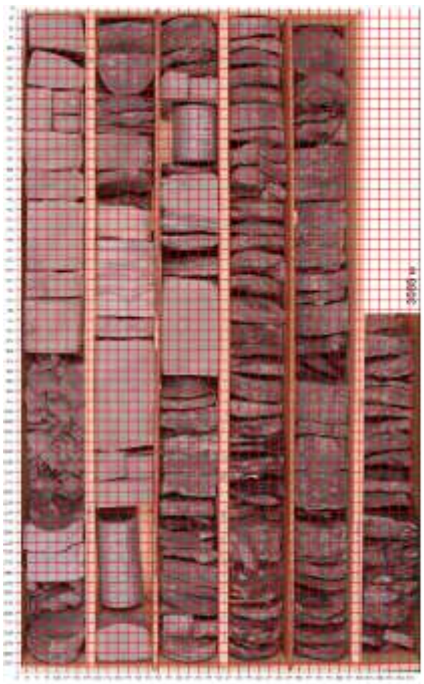

Fig. 1. Data for segmentation example

Finally, the deep neural network segments the entire image into separate fragments according to the specified step. Each segment is given the probability of belonging to a particular type of rock or other determined characteristic (porosity, permeability, etc.). At this stage, the most likely characteristic is assigned to the studied sample. As an output, the results are visualized in the form of an image with a marked map of the

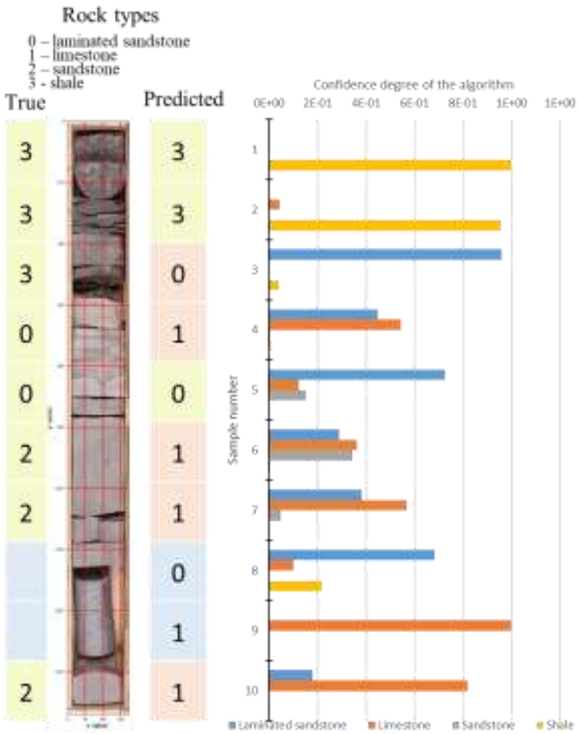

Fig. 2. An example of results given by analizator

We choose the neuro net core the open software library for machine learning Google TensorFlow [23]. A number of experiments with shells for the core led to the selection of the Keras library [8]. Preliminary image segmentation is produced in Python using the PIL library (Fig. 1) [17]. The processing of data arrays is carried out using the Numpy library [12], a number of machine learning algorithms are taken from the library Scikit-learn [19]. The work directly with the images is produced using the OpenCV library [13]. To improve the accuracy of the neural network and reduce the risk of retraining, the Dropout algorithm [22] is applied.

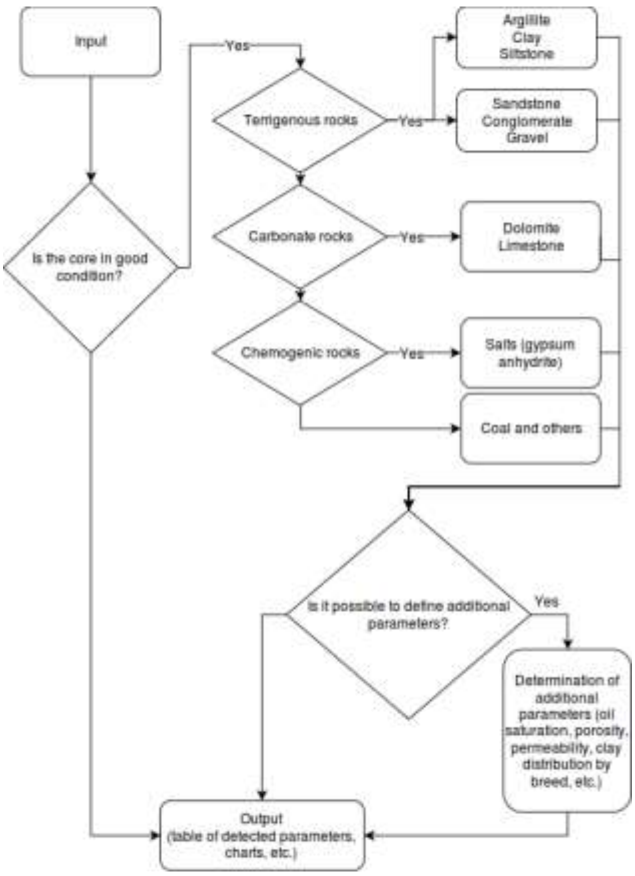

Fig. 3. Schematic diagram of the operation of a separate unit of the system

Based on these libraries deep convolutional neural network was build [10]. Parameters which were tuned are: number of neurons in layers, number of layers, backpropagation functions and image sizes. The network got 84% precision inside each class with size of image 128x128 px within equally distributed quantity of samples. For training and test of the network different quality data were used. These data represents six types of rocks (Fig. 4):

- laminated sandstone;
- limestone;
- non-cutted;
- crushed;
- massive sandstone;
- shale.

Images were taken in daylight and UV light under different conditions (light intensity, angle, resolution etc.).

Fig. 4. An example of initial data from left to right: layered sandstone, limestone, broken core, massive sandstone, unsolved core, argillite

III. Results

For training of neural network 800 images were used. They were uniformly resized to 512x512 px, containing six types of rock. For further research the number of classes reduced to four as uncutted and crushed core classes defined correct from the geologist point of view (the type of the rock determined in exact way) otherwise wrong for the computer (the type of crushed core determined as a rock). Quantity of materials is distributed irregular. The biggest type is shale (335 samples), smallest – massive sandstone (43). 60% of materials where used as training set and 40% as a test set. Gained precision is up to 96% and recall up to 98% (Fig. 5). In sime types this numbers falls to 42% of precision (limestones, 124 samples).

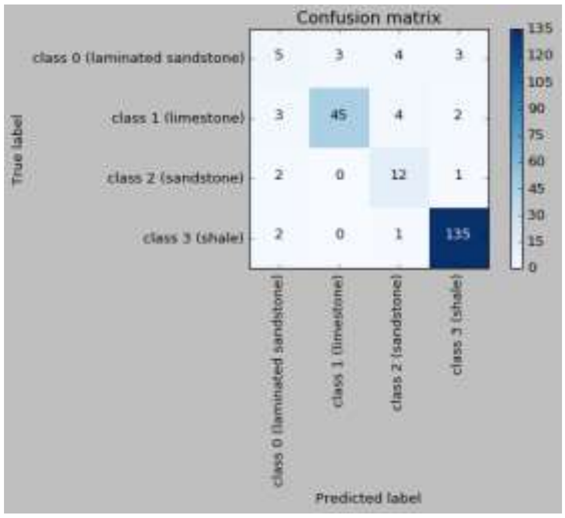

Fig. 5. Confusion Matrix for 4 classes

Conclusions

Our results demonstrate the possibility of applying convolutional neural networks for automatic analysis of core images Further work will be aimed to increase the geological data base for the network training, optimize the network operation speed. The database with core photos will also be expanded. Neural network recognition ability will be tuned for those lithotypes where the system shows lowest results. Modules for search and analysis of objects in the core will be added, particularly – grain size analysis, core samples depth alignment.

Acknowledgment

Authors thanks Elena Anatolyevna Zhukovskaya for the consultation and Marcel Kharitonov Khayrullin for the materials provided.

References

[1] Andersen M.A., Brent D., McLin R. Core Truth in Formation Evaluation. Oilfield Review. 2013. Vol. 25. No. 2. Pp. 16–25.

[2] Postnikov A.V., Khasanov I.I. Computer modeling of structural and texture features of rocks. Proceedings of the Baltic School-Seminar Petrophysical Modeling of Sedimentary Rocks. Petromodel 2015. St. Petersburg. September 14 - 18. 2015. (In Russian).

[3] Postnikov A.V., Postnikova O.V., Olenova K.Yu., Sivalneva O.V., Khasanov I.I., Osintseva N.A., Ganaeva M.R. New Methodological Aspects of Lithological Studies of Bazhenov Formations. Oil Industry. 2015. No. 10. Pp. 23 - 27. (In Russian).

[4] Khasanov I.I. Analysis of chromaticity of rocks from digital cores images. Geology of oil and gas. 2014. No. 5. Pp. 33-39. (In Russian).

[5] Khasanov I.I. Method of quantitative evaluation of chromaticity of core material and its practical application. Materials of the Baltic School-Seminar "Petrophysical modeling of sedimentary rocks, Petromodel 2013". 2013. (In Russian).

[6] Khasanov I.I. Application of computer analysis of digital images for the study of core material (abstracts). 17th Conference on Exploration and Development of Oil and Gas Fields EAGE "Geomodel 2015". Gelendzhik. 7-10 September. 2015. (In Russian).

[7] Khasanov I.I., Ponomarev I.A., Postnikov A.V., Osintseva N.A. Methodology for quantitative estimation of capacitive parameters of reservoir rocks using digital core photo processing (abstracts). 18th EAGE Geomodel 2016 Exploration and Developmental Oil and Gas Development Conference. Gelendzhik. 12-15 September. 2016. (In Russian).

[8] Chollet F. Google, Microsoft, others. Keras. GitHub repository. 2015. URL: https://keras.io

[9] Gurevich I.B., Yashina V.V. Descriptive Image Analysis: Genesis and Current Trends. Pattern Recognit. Image Anal. 2017. No. 27. Pp. 653–674.

[10] Lecun Y., Boser B., Denker J., Henderson D., Howard R.E., Hubbard W., Jackel, L. Handwritten Digit Recognition with a Back-Propagation Network. Neural Information Processing Systems. 1997. No. 2.

[11] Lu B., Cui M., Liu Q., Wang Y. Automated grain boundary detection using the level set method. Computers & Geosciences. 2009. No. 35(2). Pp. 267-275.

[12] NumPy. Package for scientific computing with Python. Version 1.14.2. URL: http://numpy.org/

[13] OpenCV. Open Source Computer Vision Library. Version 3.4.0. URL: https://opencv.org/

[14] Pat. US 2013/0156270 A1. Ellington W.E.J., Moore J.C., Smith M.A., Dubinsky G.L. Products and methods for identifying rock samples. 2013.

[15] Pat. US2017286802. Mezghani M.M., Shammari S.H., Anifowose F.A. Automated core descripton. 2017.

[16] Patel A.K., Chatterjee S., Gorai A.K. Development of online machine vision system using support vector regression (SVR) algorithm for grade prediction of iron ores. 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA). May 8-12. 2017. IEEE. 2017. Pp. 149–152.

[17] Pillow. Python Imaging Library. Version 5.0.0. URL: http://pillow.readthedocs.io/

[18] Prince C.M., Chitale J. Core Image Analysis: Reliable Pay Estimation in Thin-Bedded Reservoir Units. Soc. of Core Analysts Annual Meeting. 2008. Pp. 1-6.

[19] Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay E. Scikit-learn: Machine Learning in Python. JMLR 12. 2011. Pp. 2825−2830. Version 0.19.1. http://scikit-learn.org

[20] Rabbani A., Ayatollahi S. Comparing three image processing algorithms to estimate the grain-size distribution of porous rocks from binary 2D images and sensitivity analysis of the grain overlapping degree. Special Topics and Reviews in Porous Media. 2015. No. 6. Pp. 71-89.

[21] Richa R., Mukerji T., Mavko G., Keehm Y. Image analysis and pattern recognition for porosity estimation from thin sections. Seg Technical Program Expanded Abstracts. 2006. Pp. 1968-1972.

[22] Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research. 2014. No. 15. Pp. 1929-1958.

[23] TensorFlow. An open source machine learning framework for everyone. Google. Version 1.6.0. URL: https://tensorflow.org/