2025:10(3):232-244

Korshunov D. M. et al. From visual diagnostics to deep learning..

GEOLOGY OF MINERAL DEPOSITS

Research paper

https://doi.org//10.17073/2500-0632-2025-05-416 UDC 550.8



From visual diagnostics to deep learning: automatic mineral identification in polished section images

¹ Geological Institute of the Russian Academy of Sciences (GIN RAS), Moscow, Russian Federation ² Lomonosov Moscow State University, Moscow, Russian Federation

⊠ dmit0korsh@gmail.com

Abstract

Studying mineralogical composition of ores is a fundamental step in the exploration of new deposits, as it allows determining the forms in which useful components are found, the processes of ore formation, and the potential recoverability of valuable elements. The mineral associations, textures, and structures of ores not only provide key information about the geology of a deposit, but also determine the choice of beneficiation methods. Despite the development of modern analytical tools and existing solutions for automatic mineral diagnosis, such as those based on the SEM-EDS method, optical microscopy remains the most accessible means of quantitative mineralogical analysis. However, it remains labor-intensive and requires highly skilled specialists. In addition, its visual nature limits the accuracy and reproducibility of results, creating a need for more effective approaches. One promising area is the automation of ore mineral identification based on images of polished sections. The aim of the work was to develop and validate a universal segmentation model based on deep learning. In the course of the research, related problems were also solved, including the creation of an open LumenStone dataset, the development of color adaptation methods, joint analysis of PPL and XPL images, panorama construction, and the development of a fast annotation method. The work applied convolutional neural network architectures, color correction and joint image processing algorithms, as well as an original sampling method that compensates class imbalance. The proposed segmentation model demonstrated high accuracy (IoU up to 0.88, PA up to 0.96) for nine minerals. The obtained results confirmed the effectiveness of integrating deep learning and modern image processing algorithms in mineralogical analysis systems and laid the foundation for further development of digital methods in automated petrography.

Keywords

mineralogy, mineragraphy, digital petrography, automatic image analysis methods, segmentation, deep learning, color adaptation, panoramic images

Financing

This research was supported by the Russian Science Foundation (project no. 24-21-00061).

For citation

Korshunov D.M., Khvostikov A.V., Nikolaev G.V., Sorokin D.V., Indychko O.I., Boguslavsky M.A., Krylov A.S. From visual diagnostics to deep learning: automatic identification of minerals in polished section photographs. *Mining Science and Technology* (*Russia*). 2025;10(3):232–244. https://doi.org//10.17073/2500-0632-2025-05-416

ГЕОЛОГИЯ МЕСТОРОЖДЕНИЙ ПОЛЕЗНЫХ ИСКОПАЕМЫХ

Научная статья

От визуальной диагностики к глубокому обучению: автоматическое определение минералов на фотографиях аншлифов

Д.М. Коршунов¹ [р ⊠, А.В. Хвостиков² [р SC, Г.В. Николаев² [р, Д.В. Сорокин² [р, О.И. Индычко² [р, М.А. Богуславский² [р, А.С. Крылов² [р SC]

 1 Геологический институт РАН, г. Москва, Российская Φ едерация

Аннотация

Изучение минерального состава руд является основополагающим этапом при разведке новых месторождений, поскольку именно оно позволяет определить формы нахождения полезных компонентов, процессы рудообразования и потенциальную извлекаемость ценных элементов. Минеральная ассоциация, текстуры и структуры руд не только дают ключевые сведения о геологии месторождения, но и определяют выбор методов обогащения. Несмотря на развитие современной аналитической базы и существующие решения автоматической диагностики минералов, например, на основе СЭМ-EDS метода, оптическая микроскопия является самым доступным средством количественного минералогического анализа. Однако она остаётся трудоемкой и требует высокой квалификации специалиста. А её визуальный характер ограничивает точность и воспроизводимость результатов, что создает необходимость в разработке более эффективных подходов. Одним из перспективных направлений является автоматизация идентификации рудных минералов по фотоизображениям аншлифов. Целью работы являлась разработка и валидация универсальной сегментационной модели на основе глубокого обучения. В процессе исследования также были решены сопутствующие задачи, включая формирование открытого набора данных LumenStone, разработку методов цветовой адаптации, совместного анализа PPL- и XPL-изображений, построения панорам и разработки метода быстрой разметки. В работе были применены свёрточные нейросетевые архитектуры, алгоритмы коррекции цвета и совместной обработки изображений, а также оригинальный метод семплирования, компенсирующий дисбаланс классов. Предложенная модель сегментации продемонстрировала высокую точность (IoU до 0,88, PA до 0,96) по девяти минералам. Полученные результаты подтвердили эффективность интеграции глубокого обучения и современных алгоритмов обработки изображений для задач минералогического анализа и заложили основу для дальнейшего развития цифровых методов в автоматизированной петрографии.

Ключевые слова

минералогия, минераграфия, цифровая петрография, автоматические методы анализа изображений, сегментация, глубокое обучение, цветовая адаптация, панорамные изображения

Финансирование

Исследование выполнено за счет гранта Российского научного фонда (проект № 24-21-00061).

Для цитирования

Korshunov D.M., Khvostikov A.V., Nikolaev G.V., Sorokin D.V., Indychko O.I., Boguslavsky M.A., Krylov A.S. From visual diagnostics to deep learning: automatic identification of minerals in polished section photographs. *Mining Science and Technology (Russia)*. 2025;10(3):232–244. https://doi.org//10.17073/2500-0632-2025-05-416

Introduction

Studying mineralogical composition of ores is a fundamental step in the exploration of new deposits, as it allows determining the forms in which useful components are found, the processes of ore formation, and the potential recoverability of valuable elements. The mineral associations, textures, and structures of ores not only provide key information about the geology of a deposit, but also determine the choice of beneficiation methods.

Despite the development of modern analytical tools and existing solutions for automatic mineral diagnosis, such as those based on the SEM-EDS method [1, 2], optical microscopy remains the most accessible means of quantitative mineralogical analysis. However, it remains labor-intensive and requires highly skilled specialists. In addition, its visual nature limits the accuracy and reproducibility of results, creating a need for more effective approaches.

A promising area is the automation of ore mineral identification based on photographs of polished sections. This approach not only reduces span time, but also minimizes subjective errors associated with visual diagnostics and enables the implementation of accurate statistical analysis methods. The aim of this work is to describe our experience in developing a segmentation model for automatic detection of minerals in photographs of polished sections and

solving a number of related problems that arose during the research. The paper systematically outlines the main problems encountered by the authors and the solutions they propose.

Current state of the problem

The first attempts to create tools for the automatic diagnostics of ore minerals under a microscope were made in the second half of the 20th century [3, 4]. At that time, spectrophotometers were used to measure the color of minerals, in particular, a mineral type was interpreted based on the absorption spectra of light in the visible range. Due to its low accuracy, this method was not widely used. More advanced methods of automatic mineral identification were developed in the second half of the 1990s and were based on the analysis of photographs of polished sections under a microscope [5, 6].

Attempts were made to automatically analyze mineral associations using cluster analysis in order to find patterns between different objects in photographs [7]. Special mention should be made of the attempt by the authors [8] to compile a digital atlas of all minerals and to identify the minerals themselves using a dendrogram based on a digital questionnaire.

To date, existing classical solutions (without the use of deep learning) for automatic mineral identification can be divided into two main groups:

2025;10(3):232-244

Korshunov D. M. et al. From visual diagnostics to deep learning..

- 1. Based on reflected light intensity in conjunction with color characteristics expressed in the RGB or LAB color space [9];
- 2. Based on statistical principles of color palette separation to identify minerals in a specific sample [10–12].

Both approaches have significant limitations. Methods that use color and reflectivity are unable to distinguish between minerals with similar optical properties. Statistical methods, in turn, require recalibration for each new geological object, making their application "situational" and limited. This is well illustrated in [12], which shows the features of applying this principle to separate copper ore into three minerals and three lithological types at a specific deposit.

It is worth noting that there are also highly specialized solutions available in the form of extensions for popular image analysis software packages such as Fiji/ImageJ. For example, [13] describes a method for automatic determining hematite grade, size, and intergrowth types in ore using this software. The problem with such solutions is that they solve a narrow, specific problem and lack the necessary level of versatility.

The most effective way to overcome the shortcomings of the classical methods and achieve fundamentally better results in the automatic analysis of such images is to use trainable deep models (e.g., convolutional neural networks) that are capable of extracting complex hierarchical features from images, taking into account not only local textures and shapes, but also global relationships between image fragments. Instead of manually selecting color and statistical characteristics, such models whether traditional convolutional neural networks (CNN) [13–15], modern transformers with a self-attention mechanism [16], or hybrid architectures (e.g., Mamba [17]) — learn to identify the distinctive morphological and structural-textural features of each mineral.

For instance, convolutional neural networks were used to detect surface defects and examine polishing quality of metal products [18, 19] and to analyze carbon distribution in cast iron based on microphotographs of rough workpiece surfaces [20]. In [21], a method is presented for separating hematite and quartz in iron ore polished sections, with the determining their size classes to optimize the feed for a processing plant. It is also worth noting a number of works devoted to the assessment and classification of the dimensionality of individual mineral individuals [22, 23], as well as the analysis and typification of the morphology of intergrowths in a system with known mineral associations [24, 25]. The segmen-

tation model proposed in these works achieved 98% accuracy in predicting iron ore quality and hematite recoverability, highlighting the potential of deep models in solving industrial problems.

In [14], the effectiveness of deep convolutional networks for three-dimensional mineral identification and free grain analysis was demonstrated, and in [26], the authors showed that combined analysis of optical micrographs using CNN improves the accuracy of mineral content estimation in charge. In [15], the authors improved the methods of feature downsampling, classifying rocks more accurately based on polished section images.

It is worth noting that applying modern deep learning approaches facilitates transition from image fragment classification [27] to full semantic segmentation, which allows for accurate pixel-level segmentation of images by mineral, see [17, 26] and [14, 28]. At the same time, works [14, 29] demonstrated the fundamental feasibility of creating high-quality ore mineral segmentation models with high identification accuracy (> 0.8 by to the IoU metric).

The main advantage of using deep trainable neural networks when working with images of ore samples is their ability to take into account the context of an image and adapt to the variability of mineral associations. Most importantly, it allows reliable differentiation even between minerals with very similar characteristics (pyrite-marcasite, covellite-chalcocite, etc.) without the need for permanent recalibration of the algorithm for new samples, unlike other computer vision methods. However, there are still relatively few studies devoted directly to the diagnostics of mineral species using such approaches. Deep learning models can also be used in conjunction with domain adaptation methods, which allow the segmentation model to be retrained on "new" images - taken with different equipment or under different lighting conditions, and thus maintain high performance even with significant variations in input data. Extensive reviews on domain adaptation [14] and examples of successful application in semantic segmentation of geological and satellite images [30, 31] confirm that this approach provides versatility and stability in a wide variety of conditions. The fundamental feature of most deep learning methods is the need for complete image annotation for training. This is often a very labor-intensive process, but the use of specialized weak supervision methods, which appear to the user as annotation the image with rough strokes (ScribbleSup [32], Scribble-Seg [33]) or clicks [34], allows in many cases to significantly speed up the collection and preparation of training data.

To build a reliable system based on deep learning, the following fundamental problems must be solved, which are discussed in detail in this paper:

- 1. Development of neural network methods for mineral segmentation.
- 2. Development of adaptive methods for image calibration and preprocessing.
- 3. Development of methods for joint processing of heterogeneous images.
- 4. Development of a method for creating panoramic images.
- 5. Development of auxiliary methods for processing and analyzing images of polished sections.

Research Materials and Techniques

This study used a collection of polished sections provided by the Department of Geology, Geochemistry, and Mineral Resources of the Faculty of Geology at Lomonosov Moscow State University. A Carl Zeiss AxioScope 40 polarizing microscope with a Canon PowerShot G10 camera was used to obtain images of the polished sections. All photos were taken with a magnification of ×50 and have a resolution of 3396×2547 pixels.

The main drawback of existing solutions that use deep neural network models in the considered problems of analyzing photos of polished sections [23, 35], according to the authors, are the proprietary image sets used and the proprietary code base, which makes it impossible to compare the methods being developed. Therefore, all annotated (indexed) image sets created as part of the work are presented as a single open dataset LumenStone¹, and the software implementation of all developed methods is published as a petroscope² library with open source code for the Python 3 programming language.

The LumenStone image dataset contains several subsets focused on solving various image analysis problems for polished sections. The main subsets are S1, S2, and S3, which are aimed at the problem of mineral segmentation (automatic identification) and are formed taking into account mineral associations and mineral properties:

- LumenStone S1 (84 images): complex ores (galena, sphalerite, chalcopyrite, bornite, fahlore);
- LumenStone S2 (39 images): sulfide copper-nickel ores (pyrrhotite, pentlandite, chalcopyrite);
- LumenStone S3 (35 images): minerals with strong anisotropic properties (arsenopyrite, covellite).

Pixel masks of the corresponding minerals were created for all images of the datasets using Super-

visely and Adobe Photoshop software. The masks are necessary for training and testing deep learning models.

It should be noted that due to natural reasons (frequency of occurrence in nature), the collected set of images has a significant minerals imbalance (the percentage ratio is given in Table 1). This fact is an additional complication for the development of methods for automatic mineral segmentation and must be taken into account.

The authors also collected additional subsets of images necessary for solving related problems:

- 1. LumenStone V1: a special dataset of images of the same 10 specimens (sections) with different shooting conditions, designed for developing and testing color adaptation methods. The images were obtained using the same equipment with blue and yellow light filters, as well as using a LOMO Microsystems PLM-215 microscope with a Canon EOS 40D camera.
- 2. LumenStone P1: 875 images obtained for 35 polished sections. For each polished section, 25 photographs were taken with 20–30% overlap, intended for creating panoramic microscopic images.

To solve the problem of simultaneous analysis of anisotropic mineral photographs in PPL and XPL, "rotated" photographs of a single field of view were taken with the microscope stage rotation increment of 5 and 15° and additionally included in LumenStone S3.

Problems and their solutions (discussion)

The descriptions of the problems in the field of image processing and analysis that are to be considered the complex problem of automatic mineral identification in microscopic images of polished sections and the approaches proposed by the authors to solve these problems are given below.

1. Neural network methods for mineral segmentation. In this work, we consider convolutional neural networks to solve segmentation problems. Transformer-based alternatives, although promising, remain excessively resource-intensive for standard laboratory conditions [36]. Despite the good generalization ability of convolutional neural networks, they are quite sensitive to class imbalance in training dataset [37, 38], which is characteristic of the collected data (Table 1). Furthermore, neural network methods cannot be directly applied to high-resolution images due to hardware limitations. To mitigate these shortcomings, we proposed a specialized method for sampling the training dataset during the training process, which extracts small fragments from images (patches) and acts as a data balancer.

¹ LumenStone Dataset. URL: https://imaging.cs.msu.ru/en/research/geology/lumenstone

² GitHub. URL: https://github.com/xubiker/petroscope

Table 1 Distribution of minerals on labeled photographs of polished sections in the LumenStone S1, S2, and S3 sets for solving segmentation problems (the distribution when divided into training and test datasets is Provided in square brackets)

Mineral	Percentage in set S1 [training, test], %	Percentage in set S2 [training, test], %	Percentage in the S3 set [training, test], %	Total percentage (S1 + S2 + S3), %
Nonmetallic minerals	16.4 [12.6. 3.8]	9.8 [8.0. 1.8]	11.4 [8.8. 2.6]	37.6
Chalcopyrite	2.0 [1.1. 0.9]	3.1 [2.7. 0.4]	0.9 [0.6. 0.3]	6
Galena	3.9 [3.2. 0.8]	_	1.1 [0.9. 0.3]	5
Magnetite	-	0.4 [0.4. 0.1]	0.1 [0.1. < 0.1]	0.5
Bornite	2.0 [1.7. 0.3]	_	0.5 [0.4. 0.1]	2.5
Pyrrhotite	-	8.9 [6.2. 2.7]	-	8.9
Pyrite	12.9 [9.5. 3.4]	_	1.9 [1.5. 0.4]	14.8
Pentlandite	-	2.4 [1.6. 0.8]	-	2.4
Sphalerite	13.8 [10.9. 2.9]	_	0.5 [0.3. 0.2]	14.3
Arsenopyrite	-	_	3.9 [3.0. 1.0]	3.9
Tennantite	2.1 [1.6. 0.5]	_	-	2.1
Covellite	-	_	1.8 [1.4. 0.3]	1.8
Other (not used)	_	0.1	0.1	0.2

The objective of the developed sampling method is to equalize the distribution of mineral classes fed into a neural network during training. For each pair "training image – mineral type", a matrix containing the extracted area of a selected mineral at each point in the case of selecting a patch centered at that point is calculated. The resulting set of matrices is used as probability maps when selecting patches for training. For instance, at each sampling iteration, for the mineral that is currently the least represented, 1) an image from the training sample collection is selected (proportional to the content of this mineral), 2) the center of the patch is selected in accordance with the previously calculated probability maps, 3) the patch is extracted, and 4) the information about the representation of minerals in the used data is updated. With moderate patch sizes (256–384 px), this method allows for a significant equalization of the distribution of minerals in the LumenStone S1, S2, and S3 sets that has a positive effect on the training speed of segmentation models and on the final segmentation quality metrics.

When developing neural network models for mineral segmentation, we reviewed and investigated a number of convolutional architectures, ranging from the traditional UNet [40] and its modification ResUNet [29] to more modern PSPNet [41] and UPerNet [42]. The advantage of the latter lies in the ability to analyze images at different scales, correctly identify both small and very large objects simultaneously, and take

into account local and global context that significantly improves the quality of segmentation based on the available data.

To evaluate the quality of segmentation in this work, IoU (Intersection over Union) metric was used [43]. This is one of the simplest and most common methods of geometric evaluation of segmentation when reference labeling is available. The metric takes values from the range [0, 1], where 1 corresponds to a complete match between the predicted and reference labelings (ideal case), and 0 corresponds to no intersections between the predicted and reference segmentation annotation. An IoU value greater than 0.7 is usually considered satisfactory, although this depends, of course, on the subject area.

In our case, training the PSPNet neural network with the ResNet18 encoder on the LumenStone S1 and S2 datasets, together with the class-balanced sampling method described above, allowed segmenting nine minerals and a generalized class of nonmetallic minerals with very high quality (the average IoU value on the test set was 0.88). The training uses a cross-entropy loss function, random augmentations (rotation, slight changes in scale, brightness, and color), an Adam optimizer with an initial learning rate of 0.001 and a decrease upon reaching a plateau. The training took approximately 3 hours using a Nvidia A6000 GPU. An example of applying the trained mineral segmentation model to an image from a test set is shown in Fig. 1.

2025;10(3):232-244 Korshunov D. M. et al. From visual diagnostics to deep learning.

2. Adaptive methods for image calibration and preprocessing. One of the main problems encountered by the authors when working with primary data is high sensitivity of segmentation models to the color palette of images. Differences in color characteristics between training images and real images lead to a significant deterioration in the quality of mineral identification. The color and brightness characteristics of images are determined by many factors: microscope parameters, camera settings, lighting, etc.

A solution to this problem is to use automatic color correction based on the color difference between the received image and a known reference (e.g., [44]).

We proposed the method for correcting color distortions in [45]. The main idea is to construct a tran-

sition matrix (Color Correction Matrix, CCM) [46] between the color spaces of a distorted and reference images (images from the training set are taken as reference).

The process includes extracting the averaged colors of minerals and the background using partial labeling, linearizing colors through gamma correction (γ = 2.2), and calculating the affine transformation. The minimization problem is solved in LAB space, using the sum of the squares of color differences calculated using the CIEDE2000 formula [47] as a loss function. The work uses a 4×3 matrix with initial approximation initialization based on the "white balance" method [46]. The final step is to transform the distorted image through matrix multiplication by the previously calculated color correction matrix.

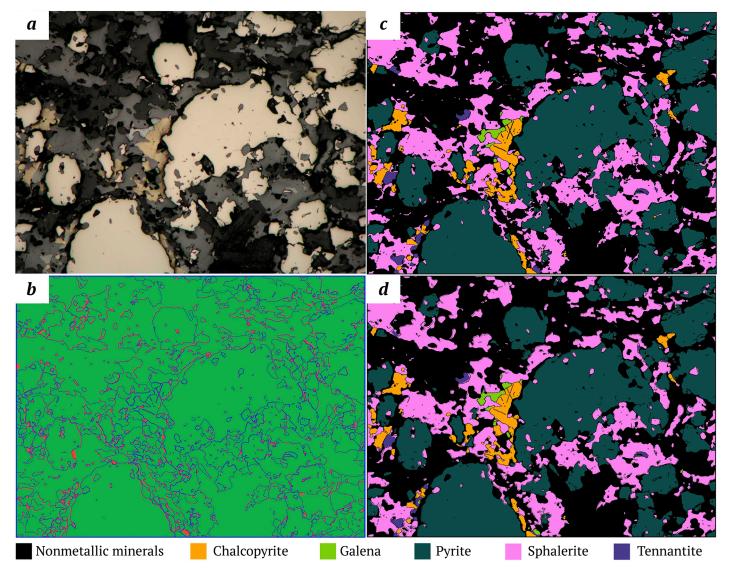


Fig. 1. Example of a polished section image segmentation with the trained PSPNet model: a – image; b – error map (correctly recognized areas are highlighted in green, segmentation errors are highlighted in red); c – mineral mask (expert annotation); d – model prediction

The proposed method allows preserving color differences that are critical for mineral identification (Fig. 2), while minimizing the influence of lighting changes and equipment settings. The algorithm supports two operating modes: an individual correction for each image and a "calibration" mode for a series of images, where the correction matrix is calculated once and applied to the entire group. The method does not require any prior training, and processing a single image takes less than 10 seconds on an Intel Xeon Gold 6226R CPU.

3. Methods for joint processing of heterogeneous images. Many minerals are identified not only by their color and reflectivity, but also by the presence or absence of anisotropic properties. Anisotropy manifests itself in the ability of minerals to "fade" in doubly polarized light (crossed nicols) when the optical axes of a mineral coincide with the direction of the microscope polarizers. This property is a key to distinguishing minerals with similar reflectance and color parameters. For example, pyrite (isotropic) and marcasite (anisotropic) have similar optical characteristics but differ in the manifestation

of anisotropy. Similarly, pyrite and arsenopyrite, although they have slightly different reflectivity and color, can also be reliably separated based on the manifestation of anisotropy by arsenopyrite.

We developed a neural network segmentation method that uses XPL and PPL images as additional input data for the segmentation neural network to improve the accuracy of mineral segmentation [48]. The key step in this method is to align images taken at different angles of rotation with the reference PPL image. For this purpose, SIFT algorithms [49] were used to detect stable key points in images, and RANSAC algorithm [50] was used to calculate the affine transformation between images based on the found matches. Thus, all images were referred to a single coordinate system (Fig. 3). Then XPL images referred to a single coordinate system were used as additional input channels for a neural network based on the architecture proposed earlier by the authors [29]. The used hyperparameters are described in [29], and the model training time is approximately 6 hours with the use of a NVidia A6000 GPU.

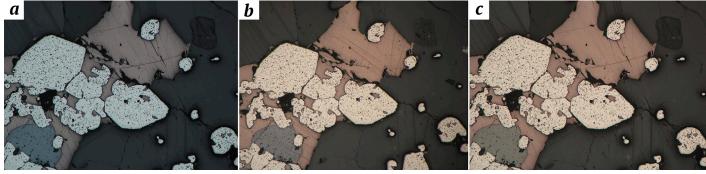


Fig. 2. An example of how the proposed color calibration method works: a – original image taken with alternative equipment; b – reference image; c – initial image after applying the method

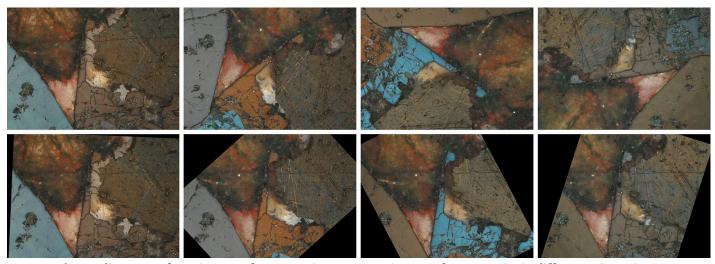


Fig. 3. Alignment of XPL images of arsenopyrite: top row – images of arsenopyrite in different orientations; bottom row – images of arsenopyrite in different orientations after alignment. Four different orientations out of 24 are presented for each image with anisotropic minerals

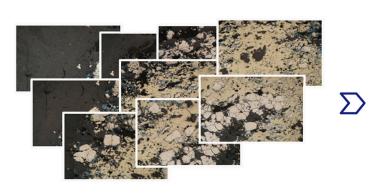




Fig. 4. Illustration of the developed method for constructing panoramas: on the left, several images of the same polished section, taken with overlap; on the right, the constructed panorama

4. Methods for creating panoramic images.

The average polished section area is several square centimeters, with typical x50 magnification. Under such conditions, only a small part of a polished section, measuring a few square millimeters, is visible in each photograph. The use of photographs covering a large area of a sample would allow more accurate information to be obtained about the distribution of minerals in a sample and their relative positions, which would have a positive effect on the quality of the analysis.

Scanning electron microscopes (SEM) can be used to obtain large images in geology, but such equipment is very expensive, structural and textural features may be lost due to the nature of the method, and the identification of mineral phases requires additional effort. Therefore, like other researchers [51], we have opted for software stitching series of overlapping images into a single panorama.

Currently, there are many examples of software for automatic stitching disparate photos into a single panoramic image. These include Adobe Photoshop, Fiji/ImageJ, and many others. However, using third-party software has a number of disadvantages. Powerful tools such as Adobe Photoshop can overly transform a panorama (unnaturally change colors, remove important details, mistaking them for artifacts of stitching). Integrating a third-party implementation into own system is tricky and it also makes it impossible to make the changes needed to the algorithm to fit the specifics of the problem you're trying to solve.

We developed our own algorithm for stitching photographs into a panoramic image of the surface of a polished section [52] (Fig. 4). The algorithm con-

sists of two main stages: image alignment and further post-processing to improve visual perception. At the first stage, with the use of calibration images, geometric distortions of images are corrected using the Brown-Conrady model [53], and photometric distortions are corrected using flat field compensation [54]. Then, using the LoFTR neural network [55], common key points are found in images that have overlapping areas. These are used to calculate perspective transformations (homographies) for pairs of adjacent images using RANSAC [50], after which all images are transformed into the coordinates of a single image (reference image). Finally, global panorama optimization is performed to minimize the alignment errors. The result of this stage is a preliminary panorama, a collage. The second stage involves improving the initial panorama. The differences in exposure between images are compensated. The seams between images are masked by constructing the least noticeable seam using the graphcut method [56], taking into account the differences in color and gradients of neighboring pixels. The final step is to blend the images near the joints of the panorama tiles to remove any remaining stitching artifacts. The LumenStone P1 dataset, compiled for the panorama construction, was used to test the algorithm. The method does not require prior training, and the processing speed for a single panorama consisting of 25 images at an Intel Xeon Gold 6226R CPU is approximately 5 minutes.

5. Additional methods for processing and analyzing images of polished sections. The application of deep learning methods for mineral segmentation in images requires accurate annotation of a large number of images, which is a labor-intensive process. To sim-

plify the annotation process and create a segmentation model capable of recognizing the main ore minerals, the authors are developing a method of accelerated interactive annotation using superpixel clustering based on the SLIC [57] and Felzenswalb [58] methods. A geologist roughly annotates minerals with strokes, labeling entire areas of an image with the label of a particular mineral based on the scribble data and the superpixel map. The user adjusts the method's predictions until the final annotation is obtained. A distinctive feature of this approach is multi-scale clustering, which allows to quickly label both large homogeneous areas and small fragments, automatically breaking large clusters into smaller ones as needed.

One can also reduce the labor costs of data annotating by extending the training set with partially annotated data. The main idea behind this approach is to highlight areas of uncertainty (lack of confidence) in the trained segmentation model on images. The authors suggest highlighting areas of uncertainty in images [59] using a hyperbolic radius [60] that reduces the scope of annotation to 5–10% of the original image (Fig. 5).

The final stage after recognizing and segmenting all minerals in the images is the statistical analysis of an image. It is responsible for conducting quantitative analysis to assess the areal ratio of mineral phases and their particle size analysis with separating fractions by size class for each mineral. This stage is currently under development.

Findings

The result of the authors' research into the automatic analysis of microscopic images of geological polished sections to determine mineral composition was the creation of an open image dataset called LumenStone and a number of algorithms and methods that solve the main problems encountered:

1. Neural network methods for mineral segmentation. A convolutional neural network model for mineral segmentation and a special method for sampling training data have been developed, allowing the existing class imbalance to be neutralized. The accuracy of mineral segmentation according to the IoU metric was as follows: non-metallic minerals – 0.912, bornite – 0.938, chalcopyrite – 0.899, galena – 0.905, magnetite – 0.650, pentlandite – 0.790, pyrrhotite – 0.928, pyrite – 0.964, sphalerite – 0.922, tennantite – 0.882. The overall pixel accuracy (PA) of the segmentation was 0.96. The differences in mineral identification results can be explained by the difference in the size of the training sets used for LumenStone S1 and LumenStone S2.

- **2. Adaptive methods for image calibration and preprocessing.** An algorithm has been developed for adapting images of polished sections obtained under different shooting conditions using partial user annotation. Pixel segmentation accuracy for distorted images increased from 0.29 (before) to 0.87 after adaptation using annotation covering approximately 30–35% of the image.
- 3. Methods for joint processing of heterogeneous images. The developed algorithm for segmenting anisotropic minerals using additional rotated XPL images improved the quality of anisotropic mineral segmentation by 3–12%. It has been shown that the best results can be achieved by using 6 additional rotated images.
- **4. Methods for creating panoramic images.** A method for constructing panoramic microscopic images of polished sections has been developed. The root mean square error of alignment of panorama tiles from 25 images was 0.5–0.6 px. The resulting panoramas have a resolution of 12000×8000 pixels and can be used for automatic mineral segmentation. The implemented method does not have the disadvantages of less specialized solutions such as Adobe Photoshop, Fiji, and Panorama Studio.
- **5. Additional methods for processing and analyzing images of polished sections.** A prototype method for interactive annotation of polished section images has been developed, which significantly speeds

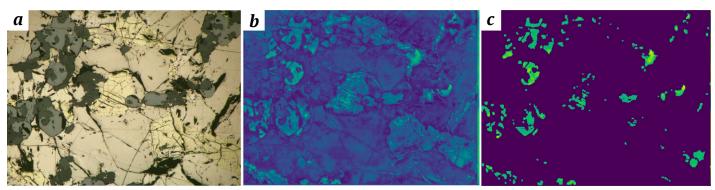


Fig. 5. The result of uncertainty area assessment method:

a – original image; b – prediction of segmentation model uncertainty areas; c – areas for manual labeling

MINING SCIENCE AND TECHNOLOGY (RUSSIA)

Korshunov D. M. et al. From visual diagnostics to deep learning.

2025;10(3):232-244

up the process of preparing data for training segmentation models. A method for automatically searching for areas of uncertainty has also been developed, allowing image annotation to be prioritized and significantly reducing the scope of annotations required.

Conclusion

This paper presents the authors' experience in developing a set of methods for the automatic analysis of images of polished sections for the identification of ore minerals. The developed segmentation method based on a convolutional neural network is capable of identifying nine ore minerals (with correct distinguishing between ore minerals and non-metallic phases) with an IoU accuracy = 0.88 and PA = 0.96. The potential of using additional information from XPL images to increase the accuracy of anisotropic mineral identification was demonstrated.

The developed methods of interactive annotation and image adaptation significantly accelerate and improve the training and use of segmentation models on new data. It is worth noting the method developed by the authors for obtaining panoramic images of polished sections, which allows detailed images of the entire surface of polished sections to be obtained in high resolution without expensive equipment. Unlike existing software solutions, this method does not distort the final panorama that has a positive effect on the segmentation results. Working with large-format images opens up new possibilities for the automatic analysis of images of polished sections.

The results obtained justify the need for further development of the area under consideration and form the basis for the creation of an intelligent quantitative assessment system capable not only of identifying minerals and calculating their areal fractions and performing particle-size analysis by size class, but also determining the types of minerals intergrowths. The implementation of this methodology will open up new possibilities in digital petrography, enabling fast, economical, and reproducible mineralogical analysis on optical microscopes in reflected light. Ultimately, this will enable the formation of unified criteria for analyzing the structural and textural characteristics of mineral associations for genesis comparison of different deposits.

Currently, the authors are integrating most of the methods and algorithms described in this paper into their PathScribe software platform [61]. This platform is a cloud-based client-server solution with cross-platform clients for working with ultra-high-resolution images and is designed for universal use in both scientific and educational applications. The authors hope that the ability to work with panoramic images of polished sections using convenient tools for annotation and fully automatic analysis will be useful for geologists of various specializations.

References

- De Castro B., Benzaazoua M., Chopard A., et al. Automated mineralogical characterization using optical microscopy: Review and recommendations. Minerals Engineering, 2022;189:107896. https:// doi.org/10.1016/j.mineng.2022.107896
- Duncan P., Gavyn K.R. Unlocking the applications of automated mineral analysis. *Geology Today*. 2011;27(6):226-235. https://doi.org/10.1111/j.1365-2451.2011.00818.x
- Yushko S.A. Methods of laboratory ore research. Moscow: Nedra; 1984. 389 p. (In Russ.)
- Craig J.R., Vaughan D.J. Ore microscopy and ore petrography. Manchester: A Wiley-interscience Publication; 1994. 446 p.
- Bonifazi G. Digital multispectral techniques and automated image analysis procedures for industrial ore modelling. Minerals Engineering. 1995;8(7):779-794. https://doi.org/10.1016/0892-
- Marschallinger R. Automatic mineral classification in the macroscopic scale. Computers & Geosciences. 1997;23(1):119–126. https://doi.org/10.1016/S0098-3004(96)00074-X
- Berry R., Walters S.G., McMahon C. Automated mineral identification by optical microscopy. In: Ninth International Congress for Applied Mineralogy. Brisbane, Australia, 8-10 September 2008. Brisbane: QLD; 2008. Pp. 91-94.
- Shoji T., Keneda H. An interactive system to assist mineral identification in ore microscopy. Mathematical Geology. 1994;26:961-972. https://doi.org/10.1007/BF02083424
- López-Benito A., Catalina J.C., Alarcón D., et al. Automated ore microscopy based on multispectral measurements of specular reflectance. I-A comparative study of some supervised classification techniques. Minerals Engineering. 2020;146:106136. https://doi.org/10.1016/j.mineng.2019.106136
- 10. Berrezueta E., Ordóñez-Casado B., Bonilla W., Banda R., Castroviejo R., Carrión P., Puglla S. Ore petrography using optical image analysis: application to Zaruma-Portovelo deposit (Ecuador). Geosciences. 2016;6(2):30. https://doi.org/10.3390/geosciences6020030

Korshunov D. M. et al. From visual diagnostics to deep learning.

- 11. Köse C., Alp I., İkibas C. Statistical methods for segmentation and quantification of minerals in ore microscopy. Minerals Engineering. 2012;30:19–32. https://doi.org/10.1016/j.mineng.2012.01.008
- 12. Krawczykowska A., Trybalski K., Krawczykowski D. The application of modern techniques and measurement devices for identification of copper ore types and their properties. Archives of Mining Sciences. 2013;58(2):433–448. https://doi.org/10.2478/amsc-2013-0029
- 13. Iglesias J.C.A, Augusto K.S., Gomes O.D.F.M., et al. Automatic characterization of iron ore by digital microscopy and image analysis. Journal of Materials Research and Technology. 2018;7(3):376-380. https://doi.org/10.1016/j.jmrt.2018.06.014
- 14. Tang K., Chen J., Zhou H., Liu J. Deep convolutional neural network for 3D mineral identification and liberation analysis. Minerals Engineering. 2022;183:107592. https://doi.org/10.1016/j. mineng.2022.107592
- 15. Zhou Z., Yuan H., Cai X. Rock Thin section image identification based on convolutional neural networks of adaptive and second-order pooling methods. *Mathematics*. 2023;11(5):1245. https://doi. org/10.3390/math11051245
- 16. Kirillov A., Mintun E., Ravi N., et al. Segment anything. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023;4015-4026. https://doi.org/10.48550/arXiv.2304.02643
- 17. Hatamizadeh A., Kautz J. Mambavision: A hybrid mamba-transformer vision backbone. arXiv preprint arXiv: 2407.08083. 2024. https://doi.org/10.48550/arXiv.2407.08083
- 18. Liu M.W., Lin Y.H., Lo Y.C., et al. Defect Detection of grinded and polished workpieces using faster R-CNN. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2021. Pp. 1290–1296. https://doi.org/10.1109/AIM46487.2021.9517664
- 19. Zhongliang L.V., Zhenyu Lu., Kewen Xia., et al. LAACNet: Lightweight adaptive activation convolution network-based defect detection on polished metal surfaces, Engineering Applications of Artificial Intelligence. 2024;133(E):108482. https://doi.org/10.1016/j.engappai.2024.108482
- 20. Sivkova T., Gusev A., Syropyatov A. Technology for cast iron microstructure analysis in SIAMS software using neural networks. In: Proceedings of the 31th International Conference on Computer Graphics and Vision. September 27–30, 2021, Nizhny Novgorod, Russia. 2021;2:772–780.
- 21. Amaral B., Soares A.K., Iglesias J.C.Á., Caldas T.D.P., Santos R.B.M., Paciornik S. Instance segmentation of quartz in iron ore optical microscopy images by deep learning. Minerals Engineering. 2024;211:108681. https://doi.org/10.1016/j.mineng.2024.108681
- 22. Maitre J., Bouchard K., Bédard L.P. Mineral grains recognition using computer vision and machine learning. Computers & Geosciences. 2019;130:84-93. https://doi.org/10.1016/j.cageo.2019.05.009
- 23. Song Y., Huang Z., Shen C., et al. Deep learning-based automated image segmentation for concrete petrographic analysis. Cement and Concrete Research. 2020;135:106118. https://doi.org/10.1016/j. cemconres.2020.106118
- 24. Donskoi E., Hapugoda S., Manuel J.R., et al. Automated optical image analysis of iron ore sinter. Minerals. 2021;11(6):562. https://doi.org/10.3390/min11060562
- 25. Donskoi E., Poliakov A. Advances in optical image analysis textural segmentation in ironmaking. Applied Sciences. 2020;10(18):6242. https://doi.org/10.3390/app10186242
- 26. Santoro L., Lezzerini M., Aquino A., et al. A novel method for evaluation of ore minerals based on optical microscopy and image analysis: preliminary results. Minerals. 2022;12(11):1348. https://doi. org/10.3390/min12111348
- 27. Su C., Wang Y., Zhu J., Zhang X.C. Rock classification in petrographic thin section images based on concatenated convolutional neural networks. Earth Science Informatics. 2020;13:1477–1484. https:// doi.org/10.1007/s12145-020-00505-1
- 28. Tang H., Wang H., Wang L., et al. An improved mineral image recognition method based on deep learning. JOM. 2023;75:2590–2602. https://doi.org/10.1007/s11837-023-05792-9
- 29. Khvostikov A. V., Korshunov D. M., Krylov A. S., Boguslavskii M. A. Automatic identification of minerals in images of polished sections. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2021;XLIV-2/W1-2021:113-118. https://doi.org/10.5194/ isprs-archives-XLIV-2-W1-2021-113-2021
- 30. Chen H., Zhang H., Yang G.A., Zhang L. Mutual information domain adaptation network for remotely sensed semantic segmentation. In: IEEE Transactions on Geoscience and Remote Sensing. 2022;60:1-16. https://doi.org/10.1109/TGRS.2022.3203910
- 31. Nasim M.K., Tannistha M., Shrivastava A., Singh T. Seismic facies analysis: a deep domain adaptation approach. In: IEEE Transactions on Geoscience and Remote Sensing. 2020;60:1-16. https://doi. org/10.1109/TGRS.2022.3151883

ГОРНЫЕ НАУКИ И ТЕХНОЛОГИИ

MINING SCIENCE AND TECHNOLOGY (RUSSIA)

2025;10(3):232-244

Korshunov D. M. et al. From visual diagnostics to deep learning...

- 32. Lin D., Dai J., Jia J., et al. Scribblesup: Scribble-supervised convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA; 2016. Pp. 3159–3167. https://doi.org/10.1109/CVPR.2016.344
- 33. Chen X., Cheung Y.S.J., Lim S.N., Zhao H. *ScribbleSeg: Scribble-based interactive image segmentation*. arXiv preprint arXiv: 2303.11320. 2023. https://doi.org/10.48550/arXiv.2303.11320
- 34. Cheng B., Parkhi O., Kirillov A. Pointly-supervised instance segmentation. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, LA, USA; 2022. Pp. 2617–2626. https://doi.org/10.1109/CVPR52688.2022.00264
- 35. Tang H., He. L., Huang B., et al. Segmentation and labeling of polished section images based on deep learning. *Mining, Metallurgy & Exploration*. 2025;42:1053–1063. https://doi.org/10.1007/s42461-025-01205-4
- 36. Tabani H., Balasubramaniam A., Marzbanet S., et al. Improving the efficiency of transformers for resource-constrained devices. In: *24th Euromicro Conference on Digital System Design (DSD)*. Palermo, Italy; 2021. Pp. 449–456. https://doi.org/10.1109/DSD53832.2021.00074
- 37. Bressan P.O., Junior J.M. Martins J.A.C., et al. Semantic segmentation with labeling uncertainty and class imbalance. *International Journal of Applied Earth Observation and Geoinformation*. 2022;108:102690. https://doi.org/10.1016/j.jag.2022.102690
- 38. Li Z., Kamnitsas K., Glocker B. Analyzing overfitting underclass imbalance in neural networks for image segmentation. In: *IEEE Transactions on Medical Imaging*. 2021;40(3):1065–1077. https://doi.org/10.1109/TMI.2020.3046692
- 39. Kochkarev A., Khvosticov A., Korshunov D., Boguslavskii M. Data balancing method for training segmentation neural networks. In: *Proceedings of the 30th International Conference on Computer Graphics and Machine Vision (GraphiCon 2020)*. Saint Petersburg, Russia, 22–25 September. Saint Petersburg: Ceur Workshop Proceedings; 2020.
- 40. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Navab N. *Medical Image Computing and Computer-Assisted Intervention MICCAI 2015*. Cham: Springer; 2015. Pp. 234–41. https://doi.org/10.1007/978-3-319-24574-4 28
- 41. Zhao H., Shi J., Qi X., et al. Pyramid scene parsing network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA; 2017. Pp. 2881–2890. https://doi.org/10.1109/CVPR.2017.660
- 42. Xiao T., Liu Y., Zhou B., et al. Unified perceptual parsing for scene understanding. In: Ferrari V., Hebert M., Sminchisescu C., Weiss Y. (eds.) *Computer Vision ECCV 2018. ECCV 2018. Lecture Notes in Computer Science. Vol 11209.* Springer, Cham; 2018. Pp. 418–434. https://doi.org/10.1007/978-3-030-01228-1 26
- 43. Rezatofighi H., Tsoi N., Gwak J., et al. Generalized intersection over union: A metric and a loss for bounding box regression. In: *Proceedings of the IEEE*. 2019. Pp. 658–666. https://doi.org/10.48550/arXiv.1902.09630
- 44. Reinhard E., Adhikhmin M., Gooch B., Shirley P. Color transfer between images. *IEEE Computer Graphics and Applications*. 2001;21(5):34-41. https://doi.org/10.1109/38.946629
- 45. Indychko O.I., Khvostikov A.V., Korshunov D.M., Boguslavskii M.A. Color adaptation in images of polished sections of geological specimens. *Computational Mathematics and Modeling*. 2022;33:487–500. https://doi.org/10.1007/s10598-023-09588-z
- 46. Wolf S. *Color correction matrix for digital still and video imaging systems*. Washington, D.C.: National Telecommunications and Information Administration; 2003. 28 p.
- 47. Sharma G., Wu W., Dalal E.N. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application*. 2005;30(1):21–30. https://doi.org/10.1002/col.20070
- 48. Razzhivina D.I., Korshunov D.M., Boguslavskiy M.A., et al. Registration and segmentation of PPL and XPL images of geological polished sections containing anisotropic minerals. *Computational Mathematics and Modeling*. 2024;34:16–26. https://doi.org/10.1007/s10598-024-09592-x
- 49. Lowe D.G. Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision*. 2004;60:91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94
- 50. Fischler M.A., Bolles R.C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*. 1981;24(6):381–395.
- 51. Ro S.-H., Kim S.-H. An image stitching algorithm for the mineralogical analysis. *Minerals Engineering*. 2021:169;106968. https://doi.org/10.1016/j.mineng.2021.106968

Korshunov D. M. et al. From visual diagnostics to deep learning

- 52. Nikolaev G., Korshunov D., Khvostikov A. Automatic stitching of panoramas for geological images of polished sections. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2024;X-2/W1-2024:39–46. https://doi.org/10.5194/isprs-annals-X-2-W1-2024-39-2024
- 53. Brown D.C. Decentering distortion of lenses. *Photogrammetric Engineering*. 1966:32(3);444–462.
- 54. Seibert J.A., Boone J.M., Lindfors K. K. Flat-field correction technique for digital detectors. In: *Proceedings of SPIE, Medical Imaging 1998: Physics of Medical Imaging.* 1998:3336;348–354. https://doi.org/10.1117/12.317034
- 55. Sun J., Shen Z., Wang Y., et al. LoFTR: Detector-free local feature matching with transformers. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021. Pp. 8922–8931. https://doi.org/10.48550/arXiv.2104.00680
- 56. Boykov Y., Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2004;26(9):1124–1137. https://doi.org/10.1109/TPAMI.2004.60
- 57. Achanta R., Shaji A., Smith K., et al. SLIC superpixels compared to state-of-the-art superpixel methods. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2012;34(11):2274–2282. https://doi.org/10.1109/TPAMI.2012.120
- 58. Felzenszwalb P.F., Huttenlocher D.P. Efficient graph-based image segmentation. *International Journal of Computer Vision*. 2004:59;167–181. https://doi.org/10.1023/B:VISI.0000022288.19776.77
- 59. Indychko O., Korshunov D., Khvostikov A. Using uncertainty to expand training sets for mineral segmentation in geological images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2025. In press.
- 60. Franco L., Mandica P., Kallidromitis K., et al. Hyperbolic Active Learning for Semantic Segmentation under Domain Shift. In: *Proceedings of the 41st International Conference on Machine Learning*. 2024. https://doi.org/10.48550/arXiv.2306.11180
- 61. Khvostikov A., Ippolitov V., Krylov A., et al. PathScribe: new software to work with whole slide histological images for education and research. In: *Proceedings of the 2023 8th International Conference on Biomedical Imaging, Signal Processing*. Singapore: ACM; 2023. Pp. 63–70. https://doi.org/10.1145/3634875.3634884

Information about the authors

Dmitrii M. Korshunov – Cand. Sci. (Geol. and Miner.), Researcher, Geological Institute of the Russian Academy of Sciences (GIN RAS), Moscow, Russian Federation; ORCID 0000-0002-8500-7193; e-mail dmit0korsh@gmail.com

Alexander V. Khvostikov – Cand. Sci. (Phys. and Math.), Senior Researcher of the Laboratory of Mathematical Methods for Image Processing, Faculty of Computational Mathematics and Cybernetics (CMC), Lomonosov Moscow State University, Moscow, Russian Federation; ORCID 0000-0002-4217-7141, Scopus ID 57188856261; e-mail khvostikov@cs.msu.ru

Gleb V. Nikolaev – M. Sc. Student, Faculty of Computational Mathematics and Cybernetics (CMC), Lomonosov Moscow State University, Moscow, Russian Federation; ORCID 0009-0003-0814-3997; e-mail nickolaev.gleb03@gmail.com

Dmitry V. Sorokin – Cand. Sci. (Phys. and Math.), Senior Researcher of the Laboratory of Mathematical Methods for Image Processing, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow, Russian Federation; ORCID 0000-0003-3299-2545; e-mail dsorokin@cs.msu.ru

Olesya I. Indychko – PhD-Student, Faculty of Computational Mathematics and Cybernetics (CMC), Lomonosov Moscow State University, Moscow, Russian Federation; ORCID 0009-0007-0936-4088; e-mail olesvaindychko@gmail.com

Mikhail A. Boguslavskii – Cand. Sci. (Geol. and Miner.), Associate Professor of the Department of Geology, Geochemistry and Mineral Economics, Faculty of Geology, Lomonosov Moscow State University, Moscow, Russian Federation; ORCID 0000-0003-0133-7185, ResearcherID V-4671-2017; e-mail mboguslavskiy@yandex.ru

Andrey S. Krylov – Dr. Sci. (Phys. and Math.), Professor, Head of the Laboratory of Mathematical Methods for Image Processing, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow, Russian Federation; ORCID 0000-0001-9910-4501, Scopus ID 7202280261, ResearcherID B-9651-2014; e-mail kryl@cs.msu.ru

 Received
 30.05.2025

 Revised
 05.07.2025

 Accepted
 07.07.2025