

Computers and Geosciences
**Graph-based Deep Learning Segmentation of EDS Spectral Images for Automated
Mineralogy Analysis**
--Manuscript Draft--

Manuscript Number:	CAGEO-D-21-00142R2
Article Type:	Research Paper
Section/Category:	Algorithms
Keywords:	electron microscopy; Mineralogy; Segmentation; Energy-dispersive X-ray spectroscopy; Deep learning
Corresponding Author:	Martin Kolar Vysoké učení technické v Brně Fakulta Informačních Technologií Brno, CZECH REPUBLIC
First Author:	Martin Kolar
Order of Authors:	Martin Kolar Roman Juránek, PhD Jakub Výravský, PhD David Motl, Ing Pevel Zemčík, Prof. Dr. Ing
Abstract:	We introduce a novel method for graph-based segmentation of spectral images obtained using a Scanning Electron Microscope (SEM) equipped with an Energy Dispersive X-ray spectroscopy (EDS) detector. The method exploits deep learning along with fusion of rasterized electron microscopy images with sparse EDS samples to obtain accurate mineralogy segmentation with high efficiency. Improvements over previous methods are with respect to the goal of an improved quantitative and qualitative assessment of segmentation, so that volumetric composition is indirectly addressed. We describe the principles of the novel method, show experimental results on real samples and demonstrate its advantages in comparison to the state of the art. The new method performs unsupervised clustering on sparsely measured EDS spectra, allowing for classification of unseen mineralogical compounds. Then, the processed spectra are combined with single channel SEM measurements through an optimised lattice, where a Markov Field is used to perform spatial segmentation in image. The benefit of this material-agnostic method is that clusters can then be (separately) classified, analyzed, and small grains with distinct EDS measurements are more accurately separated than in previous methods. These improved results are evaluated quantitatively on ground-truth electron microscope measurements with dense high-count EDS data, as well as visually through analysis by a mineralogist.
Suggested Reviewers:	
Response to Reviewers:	reviewer responses are attached separately

We present a novel and improved method for the segmentation of minerals, when scanned under an electron microscope. Our method optimally combines Backscattered Electron (BSE) and Energy Dispersive X-Ray (EDS) measurements to improve segmentation prior to classification, despite differing modalities and locations of each type of measurement.

This is a third re-submission of the work here at Computers and Geosciences, as it was rejected twice at the pre-review stage. The reason for the first rejection was a series of formatting issues and lack of shared code, and this has been remedied in the second submission. The second re-submission was rejected because of citation formatting, which was resolved by changing \usepackage[numbers]{natbib} to \usepackage{natbib} in the LaTex code.

We would like to apologise for the series of formal and formatting errors, and sincerely hope that the paper in its current form will satisfy the requirements of the Computers and Geosciences journal.

Conflict of Interest

We authors declare that we have no financial/personal interest or belief that could affect our objectivity with regard to the presented work.

Dear Reviewers. We value your feedback, and have included all recommendations into a new, revised edition.

See each point addressed below.

Reviewer 1: Dear authors

The main aim of this paper is to propose image segmentation algorithm especially for EDS spectral images. In this respect, a graph-based method is proposed which uses deep learner as clustering technique. The current manuscript has enough novelty in proposed algorithm. It can be accepted after moderate revision in terms of technical details and paper organization. So, consider following comments in the final revised manuscript:

1. The abstract is too short for a scientific paper. Add more details about proposed algorithm in the abstract. It is suggested to add quantitative results too.

Abstract has been extended, and quantitative results have been mentioned

2. Discuss about the number of bins EDS measurement. Does 3000 bins affect the final accuracy of the proposed system? Discuss about it briefly.

The necessary information has been added, explaining the quantization tradeoff this pertains to. See section 3

3. CNN is used in this paper as a projection technique to convert original spectra to a latent space with 64 dimensions. Why did you use 64 nodes as output layer of CNN? Is it possible to change number of output dimensions? Just discuss about it for eager researchers. It is not needed to report new results.

Exploration is future work, this has been mentioned in the revised text

4. Image segmentation is the main aim of this paper. So, it is necessary to review related works. I find a paper titled "Multi-Resolution and Noise- Resistant Surface Defect Detection Approach Using New Version of Local Binary Patterns" that has enough relation with this paper. Cite this paper and some other.

This paper and others have been added to citations

5. It is better to add more visual results of segmentation process.

Means more results in the paper - added images

6. What are your reasons to use average pooling? Did you evaluate using max pooling? Discuss about it.

Explanation added into text, see section 3.5

Reviewer 2: Hello! Authors

Appreciate putting together this type of work that automates the mineral phase identification in samples with complex compositions. Overall flow is good. I enjoyed reading it in the first part but from the section 2.1 it become more difficult to track because almost similar words are repeated for different methods.

The text has been revised and phrasing has been improved throughout

It would have been really helpful if the authors could have used a set of images to show the difference between the existing techniques. The alternatively, each of these techniques must have been discussed to show the relative shortcomings with respect to the proposed novelty. In general, the sentence structure was simple but in places, it loses the audience. Some specific comments are mentioned below.

Section: Title

I suggest that the title be changed to reflect 'Automated Mineral Phase Analysis' instead of 'Automated Mineralogy Analysis' because this is what is intended in the paper. The authors

have not tried to argue the mineral quantification, its validation between different techniques, and the upscaling methods with associated uncertainties.

Title has been changed

Section: Highlights

Line 9: In works related to geosciences, the words like 'Accurate' and 'Precise' needs to be dealt very carefully. What is the reference to justify the word "accurate" in your work?

Increased average segmentation accuracy over previous methods, as explained in the revised text

Line 11: What is better than the State of the Art? You could simply say that this work establishes new state of the art segmentation process. But with world moving away from threshold-based segmentation to clustering base, I wonder as to what the state of the art is for image processing?

Claim removed

Section: Abstract

Quoted Text: "The method exploits deep learning along with fusion of rasterized electron microscopy images with sparse EDS samples to obtain accurate mineralogy segmentation with high efficiency."

The above text followed by the introduction does not specify clearly if the objective is to clearly identify different mineral phases (a qualitative work) or the volumetric composition of the sample (a quantitative work). Your abstract should make a clear distinction for the purpose this work was carried out. Simply saying segmentation better than state of the art does not cut much ice.

The goal is an improved quantitative and qualitative assessment of segmentation, and volumetric composition is to be analyzed only indirectly. This has been explained in the text

Section: Introduction

Something important and particularly true for geological sample is that the strength of the current used by the probes is the first thing that should be adjusted for not missing out on many elements and therefore missing out the mineral that they are part of. You have not mentioned this limitation in any place in your introduction. I am sure the method suggested here cannot compensate for the missed-out minerals in EDS spectra.

This has been addressed in the text on lines 60 to 63

Line 50: Until unless you have left it purposefully, put the word 'mean' before the 'Z'.

Correct. This has been updated in the manuscript

Line 50: the start of new sentence with 'On the other hand' kind of breaks the flow here. You can merge it with the previous sentence by putting 'however' and make the following sentence a little shorter.

Change made

Line 54: put ',' before the word 'one'.

Change made

Line 65 and 68: put the word 'algorithm OR segmentation' after watershed to keep the good flow. Repeat in other places also for similar occurrences.

Change made throughout text

Line 76-77: Could you please quantify about the relative difference between the segment size that can be handled effectively w.r.t the conventional method that was making it time

consuming and you suggested coarser grid at Line 54-55. What is the maximum size one can work with and not miss any finer details?

The quantization tradeoff has been addressed extensively in the new manuscript on lines 185 to 225.

Section: Prior Work

Line 110-112: Re-write the sentence in proper voice.

Change made

Line 153-159: Referencing is all wrong. Keep the format consistent with the remaining paper.

Referencing style unified

Line 160-166: Reduction of dimensionality is an obvious step to make the calculations of phases easy but at the same time, it would be useful for the audience to first look at the evidence/type of dimensionality inherent with such multispectral data.

because of the inherent high-dimensionality, we focus on reducing information to a lower dimensional space. This has been explained in the updated manuscript here

Line 167: Change 'In' with 'It'. It is also a good idea to put a reference in this sentence or show a picture of the kind of noise introduced because of derivation.

Grammatical change made, and figure added to Figure 10

Line 218 - Dropped polymorphs have been mentioned

The following reviewer comments do not pertain to our submission. This issue has been raised with editors Pauline Collon <pauline.collon@univ-lorraine.fr>, Renguang <zrguang@cug.edu.cn>, and Perumal, Dhilip Kumar (ELS-CHN) <D.Perumal@elsevier.com>. No other reviews have been found. They are included below for completeness.

Page 7, Line 118-119: How do you check the uncertainty in the measured velocity, given that there could be uncertainty in the measured length and first break pick. You have also not

discussed anything on the mineralogical composition of the samples to constrain the uncertainty in grain density and the therefore bulk density. It would be a good idea to present a table with values of property for a few selected samples and calculate uncertainty in velocity values.

Page 7, Line 125-126: Why not the velocity value picked or scaled up from the measured values at ultrasonic frequency in these samples. This looks like a deterministic case of making it a metric resolution (also not clear/mentioned in terms of Absolute, Tuning or Limit of resolvability) by choosing to have a central frequency of 500 Hz with assumed velocity values of 4300 m/s.

Page 8, Line 150-151: Why not decide on which kind of processes would prevail between a micritic matrix versus sparitic cement, as both are capable of generating a seismic reflector?

Page 9, Line 197-198: The anti-correlation is not holding on all scales. Only in the upscaled (seismic scale) version do they look anti-correlated but this somehow weakens the work performed to describe the extent of neomorphism and diagenesis.

Page 9, Line 199-204: I believe, it is still talking of Figure2 but I am not able to follow up at all. I compared AI, RC with red and blue peaks but there is no consistency. As I said previously, scale is important. Suddenly you talk of a very high frequency variation of AI without a trend. Not able to find this as well. On the contrary, the AI track has relatively a low frequency variation in the top part, and I do not know what trend to talk here as there is none. Also could not follow, downward increasing of red hard-kicks and downward decrease of Blue soft-kicks.

Page 11, Line 220-223: This is good. A good geological observation to be linked with seismic observation.

Page 11, Line 235: Why bring dispersion into explaining of velocity acquired at single frequency. There is no change in the velocity values as a function of frequency here.

Page 12, Line 246-254: This is good. The binding acoustic velocities with geological processes is a welcome sight. Seismic otherwise recognize porosity in its equivalent form and not as distribution of pore sizes, geometries and arrangements in a matrix.

These final reviewer comments refer to our manuscript, and are addressed accordingly:

Section: Discussion and Conclusions

I did not see any discussion on the issues of making a wrong detection at the boundary transition. Both in Figure 8 and Figure 9, I can clearly see regions (black) that would easily make up some 10% of the total area (assuming volume too) still undecided.

Unknown classifications are not the subject of this work, and this has been added to the revised manuscript.

Moreover, I recommend that your ground truth should be made up of some commercially established industry method for identification of phase and quantity of minerals.

We are not aware of an existing standard, and thus created our own ground truth. The updated text explains that this is the "commercially established industry method"

This is particularly important during the analysis of softer material such shales with kerogen that is usually challenged in terms of accurate estimation of mineral and pores. I also recommend that you present a time-sensitivity analysis of the proposed algorithm with the conventional ones in terms of time and system utilization.

Optimisation is not the goal of the proposed work. Computational time and system requirements are higher than previous approaches, but the algorithm still runs in a number of minutes. Timings are reported in the new Table 3.

We would like to thank the reviewers and editors for their time and effort with this work.

⁵ Highlights

⁶ **Graph-based Deep Learning Segmentation of EDS Spectral Images for Automated Mineralogy**
⁷ **Analysis**

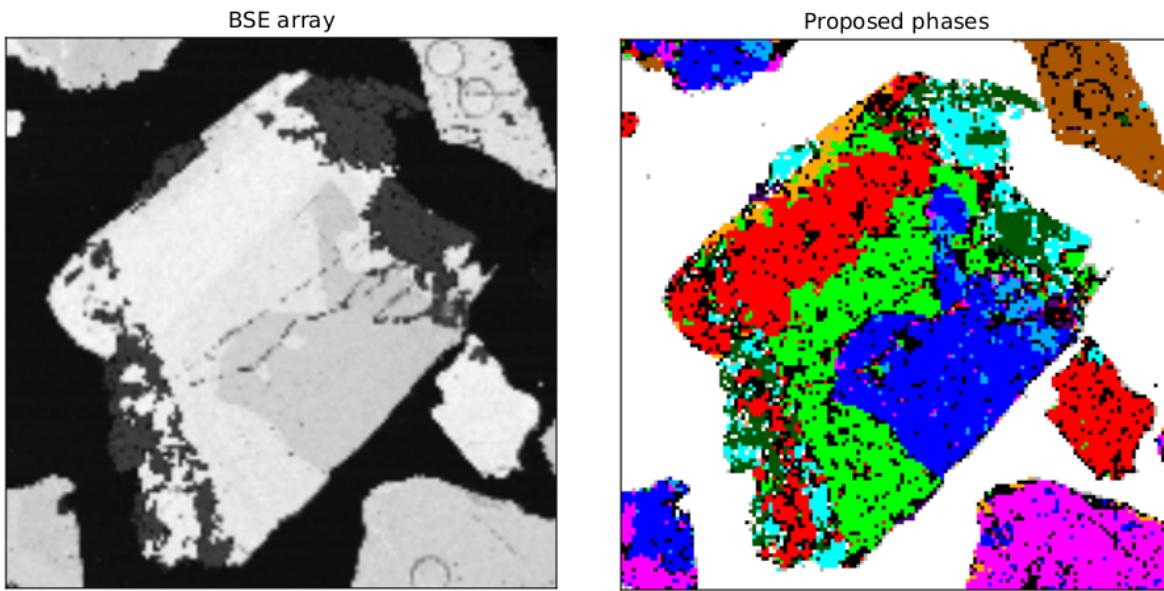
⁸ Roman Juránek, Jakub Výravský, Martin Kolář, David Motl, Pavel Zemčík

- ⁹ • Accurate segmentation using SEM image and sparse EDS spectral samples
¹⁰ • Novel approach to fusion of image and EDS spectral samples
¹¹ • Quality of segmentation is better than the state of the art
¹² • Improved discrimination of chemically similar phases

¹ Graphical Abstract

² **Graph-based Deep Learning Segmentation of EDS Spectral Images for Automated Mineralogy Analysis**

³ Roman Juránek, Jakub Výravský, Martin Kolář, David Motl, Pavel Zemčík





Click here to access/download
LaTeX Source File
paper.tex

Graph-based Deep Learning Segmentation of EDS Spectral Images for Automated Mineral Phase Analysis

Roman Juránek^a, Jakub Výravský^{b,*}, Martin Kolář^a, David Motl^b and Pavel Zemčík^a

^aBrno University of Technology, Faculty of Information Technology, Božetěchova 1/2, Brno 612 66, Czech Republic

^bTESCAN Brno, s.r.o., Libušina tř. 1, Brno 623 00, Czech Republic

ARTICLE INFO

Keywords:
segmentation
deep learning
EDS spectra
automated mineralogy

ABSTRACT

We introduce a novel method for graph-based segmentation of spectral images obtained using a Scanning Electron Microscope (SEM) equipped with an Energy Dispersive X-ray spectroscopy (EDS) detector. The method exploits deep learning along with fusion of rasterized electron microscopy images with sparse EDS samples to obtain accurate mineralogy segmentation with high efficiency. Improvements over previous methods are with respect to the goal of an improved quantitative and qualitative assessment of segmentation, so that volumetric composition is indirectly addressed. We describe the principles of the novel method, show experimental results on real samples and demonstrate its advantages in comparison to the state of the art. The new method performs unsupervised clustering on sparsely measured EDS spectra, allowing for classification of unseen mineralogical compounds. Then, the processed spectra are combined with single channel SEM measurements through an optimised lattice, where a Markov Field is used to perform spatial segmentation in image. The benefit of this material-agnostic method is that clusters can then be (separately) classified, analyzed, and small grains with distinct EDS measurements are more accurately separated than in previous methods. These improved results are evaluated quantitatively on ground-truth electron microscope measurements with dense high-count EDS data, as well as visually through analysis by a mineralogist.

1. Introduction

A Scanning Electron Microscope (SEM) equipped with an Energy Dispersive X-ray spectroscopy (EDS) detector is a well-established tool to analyze chemical composition of samples with a spatial resolution of the order of μm (Goldstein et al., 2018) and is widely used especially in material science and geological/mineralogical research. Automated mineralogy instruments are based on the capabilities of SEM with Backscattered Electron (BSE) (Goldstein et al., 2018) detector and up to 4 EDS detectors to rapidly acquire data from a statistically significant area of the subject material/mineral surface. Such automated mineralogy instruments are also equipped with dedicated software capable of distinguishing individual mineral grains, classifying them based on a user-defined list of classification rules, and deriving their morphological and textural characteristics, as well as the relative abundance of individual phases, associations, degree of liberation etc. (Fandrich et al., 2007)

The primary application for automated mineralogy is in the mineral processing industry, where epoxy-mounted polished samples of crushed mineral particles from various stages of mineral processing (feed, concentrate, tailings) are analysed in order to investigate the optimal degree of comminu-

tion, to track potential losses and in general to optimise the efficiency of the processing plant. (Gottlieb et al., 2000)

Automated mineralogy is also widely used in basic mineralogical/petrological research, where typically, glass-mounted, polished thin-sections of rocks are investigated. The required outputs include modal mineralogy, elemental deportment, mineral associations, location of minerals of interest for further analysis (EPMA, LA-ICP-MS) and many others. (Výravský et al., 2019; Hrstka et al., 2018; Breiter et al., 2017; Haluzová et al., 2015; Xu et al., 2018; Sandmann and Gutzmer, 2015) Other less common applications include e.g. oil & Gas, coal and environmental and material sciences (see (Sandmann, 2015) for details and references).

Naturally, proper recognition of phase boundaries is essential for meaningful and trustworthy analytical results and interpretation. The phases are generally recognized 1) based on their BSE brightness value only, or 2) using a combination of BSE and EDS data. The first approach is based on the observation that the number of backscattered electrons is proportional to the mean atomic number (Z) of the investigated material (Goldstein et al., 2018). This method is relatively simple but fails to distinguish phases with different composition but similar mean Z , which is rather common in geological samples (Hrstka et al., 2018), and as such, is suitable only for specific samples. The second approach makes it possible to distinguish phases with different chemistry but similar mean Z , however, acquisition of EDS data incomparably slower than acquisition of BSE data (milliseconds to microseconds per pixel, respectively), thus slowing analysis. To keep the analysis reasonably short, one could either acquire the EDS data in coarser grid than the BSE (Hrstka et al., 2018; Motl and Filip, 2013), or acquire only a relatively small number of X-ray events from each pixel (e.g.

*Corresponding author

✉ ijuranek@fit.vutbr.cz, Designed and implemented the method, trained the embedding models and evaluated the results (R. Juránek); jakub.vyravsky@tescan.com, Data acquisition, Ground truth preparation, visual evaluation (J. Výravský); kolar martin@fit.vutbr.cz, MRF segmentation, implementation and coordination (M. Kolář); david.motl@tescan.com, Software implementation and coordination (D. Motl); zemcik@fit.vutbr.cz, Coordinator (P. Zemčík)

ORCID(s): 0000-0003-0589-0172 (R. Juránek); 0000-0002-0049-4018 (J. Výravský); 0000-0001-6933-2909 (M. Kolář); 0000-0002-6583-1927 (D. Motl); 0000-0001-7969-5877 (P. Zemčík)

1
2
3 57 1,000) or combine both. The disadvantage is that the low-₁₂
4 58 count EDS spectra are prone to statistical noise, which is₁₃
5 59 typically higher than spectral differences of the phases of in-₁₄
6 60 terest. *Another disadvantage lies in the fact, that if you ac-*₁₅
7 61 *quire EDS in coarser grid, you may miss grains, which are*₁₆
8 62 *smaller than the EDS grid in case that they are not distinct in*₁₇
9 63 *BSE image* (Hrstka et al., 2018) Classifying such data pixel-₁₈
10 64 by-pixel inevitably leads to artifacts such as misclassified or₁₉
11 65 unclassified pixels. The current state of the art solution —₂₀
12 66 TESCAN Integrated Mineral Analyzer (TIMA) uses a mod-₂₁
13 67 elified multispectral watershed-based segmentation (see sec-₂₂
14 68 tion 2) to distinguish chemically coherent areas (segments)₂₃
15 69 and spectra-summing algorithms to produce a high-count,₂₄
16 70 noise-free, representative spectrum for each segment (Hrstka₂₅
17 71 et al., 2018). However, the watershed-based segmentation is₂₆
18 72 not able to distinguish chemically similar phases beyond a₂₇
19 73 certain point. This leads either to phases being mixed to-₂₈
20 74 gether or too many segments being produced with relatively₂₉
21 75 noisy sum spectra (depending on threshold settings). ₃₀

23 76 We have developed a Graph-based Deep Learning Seg-₃₁
24 77 mentation method (GDLS), with the capacity to correctly₃₂
25 78 distinguish chemically similar phases in a BSE image, com-₃₃
26 79 bined with noisy and/or sparsely occurring EDS data, while₃₄
27 80 still maintaining a large average segment size. This method₃₅
28 81 is therefore of great benefit with regard to robustness of re-₃₆
29 82 sults and speed of analysis. In this contribution, we present₃₇
30 83 this new segmentation method using a combination of deep₃₈
31 84 neural embedding and Markov Random Fields and compare₃₉
32 85 the results with the currently available method implemented₄₀
33 86 in TIMA (Hrstka et al., 2018). ₄₁

36 87 2. Prior Work

37 88 The first commercially available automated mineralogy₄₅
38 89 instruments were developed in 1980s primarily for the min-₄₆
39 90 eral processing industry. These include the Mineral Liber-₄₇
40 91 ation Analyzer (MLA) (Fandrich et al., 2007), QEM*SEM₄₈
41 92 and its successor QEMSCAN (Ayling et al., 2012; Miller₄₉
42 93 et al., 1983). Over the last decade, several other automated₅₀
43 94 mineralogy solutions have became available e.g. TIMA (TES-₅₁
44 95 CAN Integrated Mineral Analyzer (Hrstka et al., 2018)), Min-₅₂
45 96 eralogic Mining (Keulen et al., 2020) or AMICS (Advanced₅₃
46 97 Mineral Identification and Classification System (Gloy, 2016)).₅₄
47 98 Each of these solutions uses specific implementations of BSE/EDS
48 99 acquisition modes and image segmentation. ₁₀₀

51 100 2.1. Overview of existing solutions

52 101 The MLA analyzer starts with acquiring a BSE image₁₀₂
53 102 This is followed by particle de-agglomeration and phase seg-₁₀₃
54 103 mentation, where regions of homogeneous gray levels within₁₀₄
55 104 the BSE image are found. No details on methods employed₁₀₅
56 105 for phase segmentation have been published. Depending₁₀₆
57 106 on the selected acquisition method, spectroscopic data are₁₀₇
58 107 collected either in a single measurement point in the cen-₁₀₈
59 108 ter of each grain, or continually, as the electron beam scans₁₀₉
60 109 over the entire grain; alternatively, individual spectra are col-₁₁₀
61 110 lected from each point of a grid-pattern of points covering₁₁₁
62 111 the entire particle. The spectroscopic data are then used to₁₁₂
63 112 ₁₁₃

identify the mineral species (Fandrich et al., 2007). Other methods segmenting BSE only exist, using autoencoder networks (Karimpouli and Tahmasebi, 2019) and the U-Net architecture (Jasiewicz et al., 2018).

In the QEM*SEM analyser and its successor QEMSCAN, an X-ray spectrum along with the BSE signal is collected for each measurement point in an equidistant rectangular grid. The acquisition step is followed by identification of mineral species, for each measurement point individually utilizing both BSE and EDS data. Image segmentation is basically identification of continuous horizontal line segments of measurement points classified as the same mineral. Feature extraction is done by finding connected line segments from adjacent horizontal lines (Grant and Reid, 1981). A correction for boundary phases (misidentification of mineral species due to mixed spectrum at grain boundaries) is also employed, but details are not publicly known (Reid et al., 1985; Sutherland and Gottlieb, 1991).

Zeiss mineralogic uses BSE thresholding to select areas to be analyzed as well as determining grains based on BSE intensity to extract morphological characteristics. Grains are then classified based on EDS spectra obtained in several ways (one point in center of the grain, continual scan over grain, coarse grid of points, dense equidistant grid). Also, some segmentation methods are available for image postprocessing. However, specific details are not public knowledge (Keulen et al., 2020).

The AMICS automated mineralogy software from Bruker uses a proprietary BSE segmentation algorithm to determine mineral phases. It contains two adjustable interrelated parameters, which control sensitivity to gray level variation and size of the segmented region. Each segment is then analysed with EDS to classify the phase (Gloy, 2016).

TIMA achieves grain identification using multi-channel image analysis and combining image data from a BSE detector as well as EDS data. The method starts with the multi-channel gradient, followed by the watershed algorithm with adjustable threshold. The grain identification phase is followed by accumulation of spectroscopic data that are acquired from measurement points in each individual grain and determining the mean signal from the BSE detector, the accumulated data are used to identify mineral species (Motl and Filip, 2013; Hrstka et al., 2018).

155 2.2. Watershed-based Segmentation

To apply the watershed-based segmentation algorithm to multichannel data, the multichannel image gradient must be computed. Computation of the tensor gradient was introduced by DiZzenzo et al. (Di Zzenzo, 1986) for color images, Drewniok (Drewniok, 1994) extended this work to multispectral images. One application of this algorithm is also mentioned in (Karvelis et al., 2008).

To reduce dimensionality and suppress noise in low-count EDS spectra, elemental maps are produced by summing stored X-ray events from all spectrum channels corresponding to energy regions of important X-ray emission lines of the individual elements. These elemental maps are individual chan-

1
2
3 ¹⁶⁸ nels in the multichannel image and are used to calculate the
4 ¹⁶⁹ combined gradient.
5 ¹⁷⁰ (Whole paragraph has been removed.)
6
7
8 ¹⁷¹ **3. Graph-based Deep Learning Segmentation**

9 ¹⁷² The proposed segmentation method takes one input field
10 ¹⁷³ composed of a densely sampled BSE array and sparsely sam-
11 ¹⁷⁴ pled EDS data with known measurement locations.

12 ¹⁷⁵ The spectra data associated with each measurement lo-
13 ¹⁷⁶ cation has form of an 1D array of integer values. The index
14 ¹⁷⁷ to the array is an ordinal number of a spectrum channel,
15 ¹⁷⁸ the values stored in the array are frequencies of detection
16 ¹⁷⁹ of X-ray photons in the location with certain energy. We
17 ¹⁸⁰ have chosen spectrum channels to be uniformly distributed,
18 ¹⁸¹ each channel 10 eV wide. This choice of spectrum division
19 ¹⁸² has been directed by finding compromise between error that
20 ¹⁸³ arises from the quantization of the photon energy and mem-
21 ¹⁸⁴ ory footprint of the spectral data.

22 ¹⁸⁵ The spectra data associated with each measurement lo-
23 ¹⁸⁶ cation has form of an 1D array of integer values. The in-
24 ¹⁸⁷ dex to the array is an ordinal number of a spectrum chan-
25 ¹⁸⁸ nel, the values stored in the array are numbers of detection
26 ¹⁸⁹ of X-ray photons in the particular channels. We chose the
27 ¹⁹⁰ energy continuum to be divided into spectrum channels uni-
28 ¹⁹¹ formally distributed with channel width to be 10 eV, since it
29 ¹⁹² provides good compromise between quantization error and
30 ¹⁹³ memory footprint of the spectral data. Peaks corresponding
31 ¹⁹⁴ to characteristic lines as detected by the EDS detector can
32 ¹⁹⁵ be approximated by Gaussian functions superimposed on a
33 ¹⁹⁶ continuum background. Although peak width, conven-
34 ¹⁹⁷ tionally expressed as Full Width at Half Maximum (FWHM) is a
35 ¹⁹⁸ function of several factors, the most important of them to be
36 ¹⁹⁹ statistics of generating electron-hole pair in silicon (which is
37 ²⁰⁰ a function of photon energy), electronic noise in the detector
38 ²⁰¹ and preamplifier, and width of an energy measurement cir-
39 ²⁰² circuit.(Reed, 2005) We decided to set up the spectrometer to
40 ²⁰³ conditions optimized for high throughput, because such con-
41 ²⁰⁴ ditions are typical for practical sample analysis by means
42 ²⁰⁵ of an automated mineral analyzer. In our sample data, the
43 ²⁰⁶ FWHM of light element peaks, e.g. carbon, is > 70 eV. Us-
44 ²⁰⁷ ing the Gaussian model of peaks of characteristic lines, the
45 ²⁰⁸ FWHM value can be converted to standard deviation σ_{res}
46 ²⁰⁹ using formula ($\text{FWHM} = 2\sqrt{2 \ln 2} \sigma \approx 2.355 \sigma$), therefor
47 ²¹⁰ $\sigma_{\text{res}} \approx 1/2.355 \cdot 70 \text{ eV} \approx 30 \text{ eV}$.

48 ²¹¹ The additional error introduced by the quantization of
49 ²¹² energy continuum into uniformly distributed spectrum chan-
50 ²¹³ nels is $1/\sqrt{12}$ LSB (standard deviation of the distribution).(Gray
51 ²¹⁴ and Neuhoff, 1998) In our case $\Delta = 10 \text{ eV}$ and the corre-
52 ²¹⁵ sponding standard deviation of the quantization error $\sigma_{\text{quant}} =$
53 ²¹⁶ $1/\sqrt{12} \cdot 10 \text{ eV} \approx 2.9 \text{ eV}$. The standard deviation of quan-
54 ²¹⁷ tization error σ_{quant} is more than one order of magnitude
55 ²¹⁸ lower than the standard deviation that corresponds to the
56 ²¹⁹ width of the narrowest (light element) peaks σ_{res} in our sam-
57 ²²⁰ ple data. Using higher spectrum channel width, we start to
58 ²²¹ lose information, using lower width, we increase the memory
59 ²²² footprint of the spectral data unjustifiably.

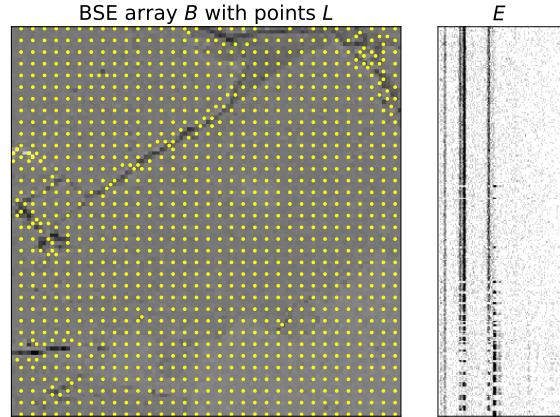


Figure 1: The field is a hyperspectral image with sparsely sampled data. The locations L of EDS measurements are superimposed on the BSE array B . The matrix E shows the actual spectra in rows. Details in section 3.1

The energy continuum of detectable X-ray photons (0 to 30,000 eV) is uniformly divided into spectrum channels each of which 10 eV wide. Therefore, the spectral data have 3,000 channels per pixel. The values in the BSE array are calibrated, with a maximal value of 2^{16} . An example field is seen in Figure 1.

The output of the segmentation process is a dense array with pixel-wise labels (integer values) where pixels with the same value belong to the same segment. The core of the proposed method is to determine if spatially close EDS measurements capture the same material, in which case they belong to the same segment and should be labeled with a common label. We construct a graph with EDS measurements represented as vertices. Edges connect spatially close measurements in a planar graph obtained by Voronoi analysis (Bowyer, 1981). A deep learning model assigns a value to each edge. This value corresponds to differences in BSE and distances between hyperspace representations of EDS spectra. Edges with high differences are removed and the resulting graph components constitute separate segments — labels. Final dense labels are obtained by a dense Markov Random Field operating over the BSE array with labels initialized from the graph. An example of graph and segmentation is in Figure 2. The individual steps of the algorithm which are detailed in the following text are:

- Projection of spectra to latent space.
- Graph construction and calculation of edge values.
- Graph component analysis.
- Markov Random Field segmentation.

3.1. Field representation

A field F (see Figure 1, left) is represented as a triplet $F = (B, E, L)$. B is $H \times W$ array with BSE measurements

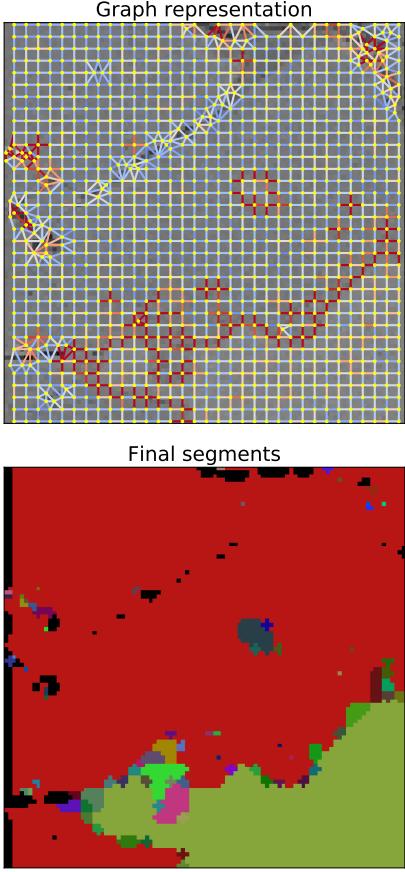


Figure 2: The proposed method constructs a graph representation of the field (edge color encodes δ_z , see Section 3.6) and decomposes the graph to obtain compact segments. Details in the text.

(an image of width W and height H). E and L are matrices containing N rows with EDS measurements

$$E = \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix} \quad L = \begin{pmatrix} l_1 \\ \vdots \\ l_N \end{pmatrix}$$

where E is $N \times D$ matrix with N spectra with dimensionality D , and L is a $N \times 2$ matrix with locations of EDS measurements in the BSE array. The spectrum s_i corresponds to location l_i in the BSE array.

The proposed method assumes full coverage by B , and sparse coverage by EDS measurements, whose locations are determined at scan-time by a separate mechanism internal to the electron microscope. Although the proposed method poses no restrictions on the locations of the sparse EDS measurements, they are best placed at regular intervals on discrete materials, and ideally do not miss small separate grains.

3.2. Mineral dataset

For training the embedding model we use a dataset of 4,507 distinct materials relevant for the application. Each material is characterized by its ideal spectrum. *In case of*

polymorphs (materials with identical chemistry, but different crystal structure e.g. graphite and diamond), which are indistinguishable by EDS, only one of the polymorphs was used in the training dataset. During the training, we generate spectra from the ideal profiles by random sampling with a random number of events to simulate data from a microscope. These spectra are close to what is captured by the microscope, but it is not perfect since the ideal profiles do not include recurring artefacts of pulse processing electronics like sum peaks or escape peaks, and other unmodeled effects (Ritchie, 2009). These differences influence the overall performance of the proposed method on the real data especially in situations where multiple materials are mixed (e.g. on boundaries between two materials) since we train on pure spectra with no “deformations”.

3.3. Spectral decomposition with deep learning

The key part of the algorithm is to determine if two spectra capture the same material, which are composed from the same elements in the same ratio. While this can be solved with relatively low error by, e.g., mean absolute difference or a two-sample K-S test on high count spectral data, in presence of noise due to low number of X-ray events per pixel the task becomes very difficult. In our solution, we process the spectra with a deep learning model to provide a compact descriptor which can differentiate between spectra with different compositions. The descriptor is obtained by embedding to a low dimensional space – latent space \mathcal{Z} as shown in Figure 3. Transforming EDS measurements into this latent space serves as a preprocessing step to separate image-space segmentation from material composition and characteristic line differentiation, but this model is made on resampled ideal spectra, which cannot contain all noise and variations of real measurements. By training the mapping to differentiate thousands of randomly selected materials, the network learns to generalize even unseen materials. Therefore, during segmentation, which is not trained for specific materials, the presented approach is not limited to previously seen minerals and solid solutions. This is demonstrated in Figure 9, where chemically-distinct domains with different Nb/Ta ratio in grains of columbite-tantalite solid solution are properly segmented despite the fact, that no spectrum corresponding to such chemical composition was in the training dataset.

3.4. Neural network design

The convolutional neural network (CNN, see Figure 4) comprises three parts — *normalization*, *feature extraction* and *embedding*. The input of the network is the vector of the spectrum s with dimensionality D followed by a normalization with mean μ and standard deviation σ calculated from ground truth spectra. We use the 200 eV to 15,560 eV spectral band with 10 eV bins (therefore the input to the network are spectra with $D = 1,536$). There are two reasons: first, coefficients outside this band are mostly uninformative and second, reducing the number of coefficients reduces the number of network parameters and this positively affects the execution speed of the network. Nothing prevents using the whole spectrum, though. Feature extraction contains 3

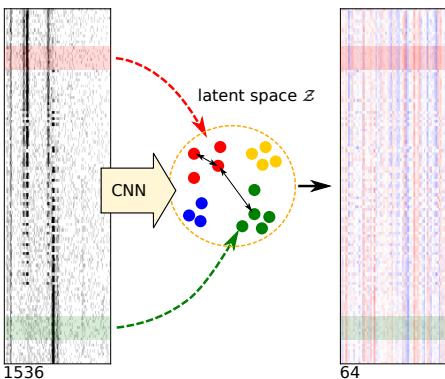


Figure 3: The original spectra are projected by CNN to a latent space \mathcal{Z} with 64 dimensions, where similar materials are mapped to similar locations

blocks with a sequence of convolution, batch normalization and ReLU activation (Nair and Hinton, 2010) followed by average pooling. Embedding processes the features with a sequence of dense layers and normalizes the output to unit length, producing a latent vector z .

Although almost any network design can be used here, we chose a simple network inspired by the VGG architecture (Simonyan and Zisserman, 2014) mainly for performance reasons. We changed the input shape, number of layers and blocks, and use *average pooling* instead of *max pooling* for downsampling since it better utilizes input data and has the ability to smooth the noise (inherent to spectral data). We did not observe any severe effect on the final results regarding the choice of downsampling strategy in preliminary experiments. Inference speed of this network is sufficient. On a single field, usually tens of thousands of measurements must be processed (in extreme cases it can be hundreds of thousands). This network can produce its output in a matter of seconds on low-end hardware which is still acceptable.

3.5. Embedding to low dimensional space

We train a neural network model M which produces a unit vector $z = M(s)$ for a spectrum s . The main purpose of this network is to make a compact, low dimensional representation of spectra. We reduce the input dimensionality from 3 000 to 64. Setting the dimensionality too low would lead to inefficient embedding which would not capture the differences in analyzed materials, setting it too high would increase memory and time requirements. We fixed the dimensionality to 64 in our experiments as a tradeoff between embedding accuracy and its size. We use the spectral information as an input and material class (i.e. its integer ID) as an output and optimize the network with ADAM (Kingma and Ba, 2014) optimizer with Semi-hard Triplet loss (Schroff et al., 2015).

The main property of the latent space \mathcal{Z} is that the difference $\|M(s_1) - M(s_2)\|_2$ is close to zero if s_1 and s_2 are from the same material and it is larger if the materials are different, whatever the materials are. By multiplying matrices M and E , individual EDS measurements s are transformed into

the spectrum decomposition z .

3.6. Graph construction and labeling

From the known locations of EDS measurements in the field, we build a planar graph $G = (V, E)$ by Voronoi analysis (Bowyer, 1981). Vertices $V = (l_i, b_i, z_i)$ are locations of the measurements, each assigned with its location l_i , BSE value b_i and spectrum decomposition z_i obtained by the neural network. Edges are formed between spatially close measurements. The edges between measurements with distance larger than the $2\times$ sampling distance are not considered. Each edge $E_k = (i, j)$ is assigned with values δ_k^b and δ_k^z capturing difference in BSE and latent space representation of the spectrum.

$$\delta_k^b = |b_i - b_j|/2^{16} \quad (1)$$

$$\delta_k^z = \|z_i - z_j\|_2 \quad (2)$$

Where the value of δ^b is in the $(0, 1)$ interval and $\delta^z >= 0$. We remove the edges with $\delta_k^b > \tau_b$ or $\delta_k^z > \tau_z$ which disconnects measurements with high difference in BSE or spectrum. The parameters τ_b and τ_z can be tuned by the user to obtain different segmentations of the input data. The components of the resulting graph then constitute sets of measurements with the same material. As a result, each vertex is assigned a value $c_i \in N$ which is the label of the component it belongs to. These labels are the starting point for the Markov Field which generates the final pixel-wise label map.

3.7. Pixel labelling with dense MRF

Markov Random Field segmentation (Li, 1994) is a flexible optimisation method to create unlabeled segments given neighbouring pixel similarities and a sparse marker initialization. The field of pixel-wise BSE measurements is set on a fully connected grid, and we initialize the segmentation map with known labels from EDS graph analysis. By using *a priori* probability of sampling different measurements from the same distribution (class), and the probability of adjacency of different classes, the entire system becomes a joint distribution whose *a posteriori* distribution can be maximised by computing conjugate gradients for a given segmentation estimate. The distribution is iteratively optimized until convergence, where classes correspond to individual class probabilities. The maximum class at each pixel is taken to be its class, producing full segmentation.

4. Evaluation

Qualitative and quantitative evaluation is performed on selected representative fields. For these, two scans are performed — a fast evaluation scan and a detailed ground-truth scan (Section 4.1).

The evaluation scans were processed using the standard industrial solution — modified multispectral watershed-based segmentation and the proposed method (GDLS). Parameters of the methods were set by an expert who ensured that all

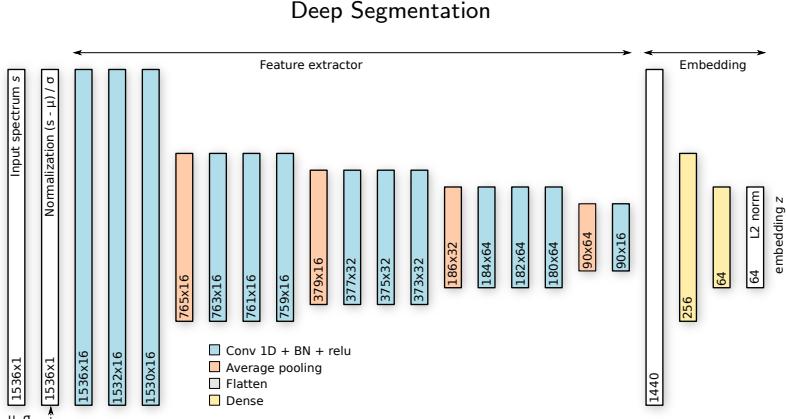


Figure 4: The neural network for mapping spectral information to latent space \mathcal{Z} . It processes the input s with normalization, feature extraction and embedding and produces the vector z .

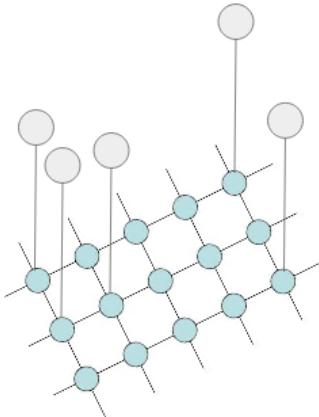


Figure 5: Dense per-pixel measurements (blue), sparse markers (gray), and relationships in a Markov Random Field

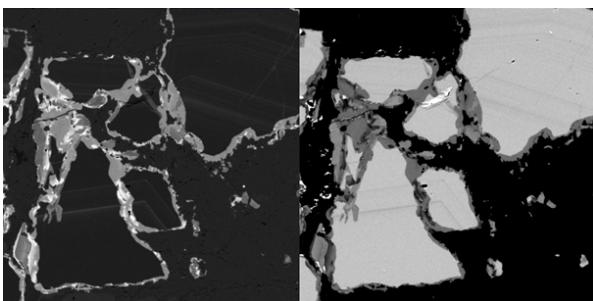


Figure 6: Low-noise, high-resolution element distribution maps from ground-truth data. Left: Al-K line intensity, Right: Fe-K line intensity. Note primary growth zoning and younger replacement zones around the edges.

relevant information is present in the resulting segmentation maps.

We evaluated the methods using standard metrics (Section 4.2)

4.1. Data acquisition

The performance of both algorithms was evaluated on a set of 4 “Ground Truth” fields representing different types of geologically relevant samples (different kinds of ores and

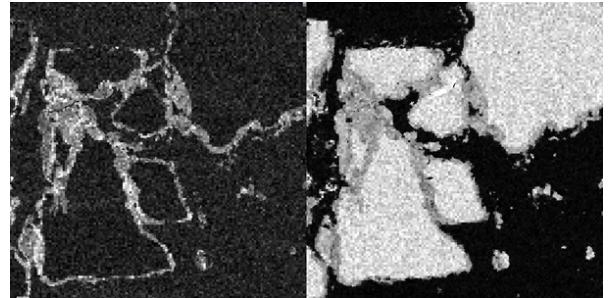


Figure 7: Element distribution maps from evaluated dot-mapping data. Left: Al-K line intensity, Right: Fe-K line intensity. Note the high level of noise and the lack of some details (e.g. the primary growth zoning).

rocks). Since, to the best of our knowledge, there does not exist any industry standard or community-recognized benchmark, which could be used to evaluate the correctness of mineral identification and quantification from BSE and EDS data, we had to create the ground truth internally. The ground truth data were always acquired in High resolution mapping mode (Hrstka et al., 2018) with 10,000–100,000 EDS events per pixel to reduce statistical noise. The definitions of individual phases (sets of ranges for normalized EDS peak intensities and/or BSE intensity for each phase) were prepared in the TIMA SW based on inspection BSE image, elemental maps, and spectra of manually selected areas of the high-count, noise-free data. The set of individual phase definitions comprise the so-called classification scheme (Hrstka et al., 2018). The High-count noise-free data were subsequently classified pixel-by-pixel using the classification scheme. The resulting phase image represented ideal segmentation (Ground Truth). At the same time, standard data with 1,000 X-ray events per pixel in high-resolution mapping were acquired from the same area and were used to compare the performance of both algorithms. For 12 fields, 1,000 X-ray events per pixel data were also acquired using the Dot mapping mode (Hrstka et al., 2018) with a 3 times coarser EDS grid compared to BSE. The following microscope settings were used: accelerating voltage 25 kV, beam current 10 nA, working distance 15 mm. The acceleration voltage of 25 kV was

Field	Ground Truth				Evaluation			
	BSE [μm]	EDS [μm]	X-ray events	Time [s]	BSE [μm]	EDS [μm]	X-ray events	Time [s]
A	3	3	20,000	8,577	3	9	1,000	55
B	3	3	20,000	7,233	3	9	1,000	49
C	1	1	20,000	20,269	2	6	1,000	42
D	10	10	100,000	2,074	10	10	1,000	26

Table 1

Overview of acquisition parameters and acquisition time for individual fields evaluated in table 2, where μm distances denote measurement spacing for each configuration. The X-ray events field specify number of X-ray events that were recorded in each measurement point and the Time field denotes the total acquisition time per field.

chosen in order to properly excite all relevant X-ray lines including e.g. Mo-K, Cu-K, Fe-K to be able to properly identify the minerals from their EDS spectra. Measurements performed on the four fields used are described in table 1, and the difference in accuracy of EDS measurements can be seen in figures 6 and 7.

4.2. Evaluation metrics

We use standard metrics for superpixel segmentation evaluation (Stutz et al., 2018). We modify the measures to take "invalid" pixels into account. Ground truth pixels can be invalid because even an expert on high-count data could not decide to which segment the pixels belong, or they were *a priori* marked as background by the device. The pixels in the segmented image can be invalid for two reasons — 1) the instrument marked them as background or 2) the method could not mark them. We do not want these pixels to participate in the evaluation. Let $S = \{S_j\}_1^K$ and $G = \{G_i\}$ be partitions of the same image where S represents the segmentation and G is ground truth segmentation.

Boundary precision (P) and recall (R) - measures adherence of the method to the edges and reflects false positive and false negative edges. Let $TP(G, S)$, $FN(G, S)$ and $FP(G, S)$ be a number of true positive, false positive and false negative edges. The P and R are defined as

$$R(G, S) = \frac{TP(G, S)}{TP(G, S) + FN(G, S)} \quad (3)$$

$$P(G, S) = \frac{TP(G, S)}{TP(G, S) + FP(G, S)} \quad (4)$$

In other words, the value of Precision reflects the ratio of real edges in the detected edges, and the value of Recall reflects the ratio of the real edges that were detected. Higher values means better adherence to edges, and consequently better segmentation.

Undersegmentation error (UE) - Measures leakage of detected segments with respect to G . It is defined as

$$UE(G, S) = \frac{1}{N} \sum_{G_i} \sum_{S_j \cap G_i \neq \emptyset} \min \{|S_j \cap G_i|, |S_j - G_i|\} \quad (5)$$

Achievable segmentation accuracy (ASA) - Measures the ratio of area that falls to the segment with highest intersection.

$$ASA(G, S) = \frac{1}{N} \sum_{S_j} \max \{|S_j \cap G_i|\} \quad (6)$$

4.3. Quantitative Evaluation

Evaluation results of the 4 fields are summarized in Table 2. Figure 8 show visualizations of segmentation for fields C and D along with metrics for different parameters settings.

The new method achieves better *average segmentation accuracy* and *undersegmentation* which means that the segments better correspond to the structures in the fields. This is also supported by the values of *precision* and *recall* metrics (as visible in Figure 8). Watershed-based segmentation can achieve similar results for the price of oversegmentation. GDLS generates lower number of segments which are more compact. The exception is the field D which was scanned in the high resolution mapping mode, and contains a large number of single-pixel segments which is impossible to achieve with watershed algorithm since it makes 1 pixel borders around segments. Despite the higher number of segments on field D, GDLS better captures homogeneous areas in the field.

4.4. Visual Evaluation

Field C (see Figure 8) consists of grains of andradite garnet ($\text{Ca}_3\text{Fe}_2(\text{SiO}_4)_3$ — red color) containing chemically distinct zones enriched in grossularite component ($\text{Ca}_3\text{Al}_2(\text{SiO}_4)_3$ — dark red color), mostly along their rims. According to EDS quantitative analysis, the composition difference of the dark red and red zones is about 3–4 wt. % of Al and Fe. The garnet grains are enclosed in wollastonite — yellow color). From the High-count Ground Truth data (Figure 6), it is seen that the grossularite-enriched zones (darker red) are younger, and they cross-cut the primary oscillatory growth zoning of the andradite crystals.

The watershed and GDLS were compared on a dot mapping dataset acquired with lower resolution and lower number of X-ray events, summarized in Table 1, therefore, lots of features and small details, especially on the elemental maps (e.g. the brightest areas on the Al distribution map), are not visible due to lower resolution and noise (compare maps in

Deep Segmentation

Field	Watershed					GDLS				
	ASA	UE	P	R	K	ASA	UE	P	R	K
A	0.97	0.059	0.61	0.97	4,080	0.98	0.047	0.81	0.98	3,427
B	0.89	0.22	0.8	0.95	12,036	0.89	0.22	0.85	0.93	8,247
C	0.93	0.13	0.3	0.99	21,822	0.93	0.15	0.69	0.83	5,102
D	0.89	0.22	0.44	1.0	2,829	0.94	0.12	0.52	1.0	5,093

Table 2

Metrics on the four testing fields evaluated for segmentation with parameters set by an expert. On average, the proposed method gives better results with lower number of segments.

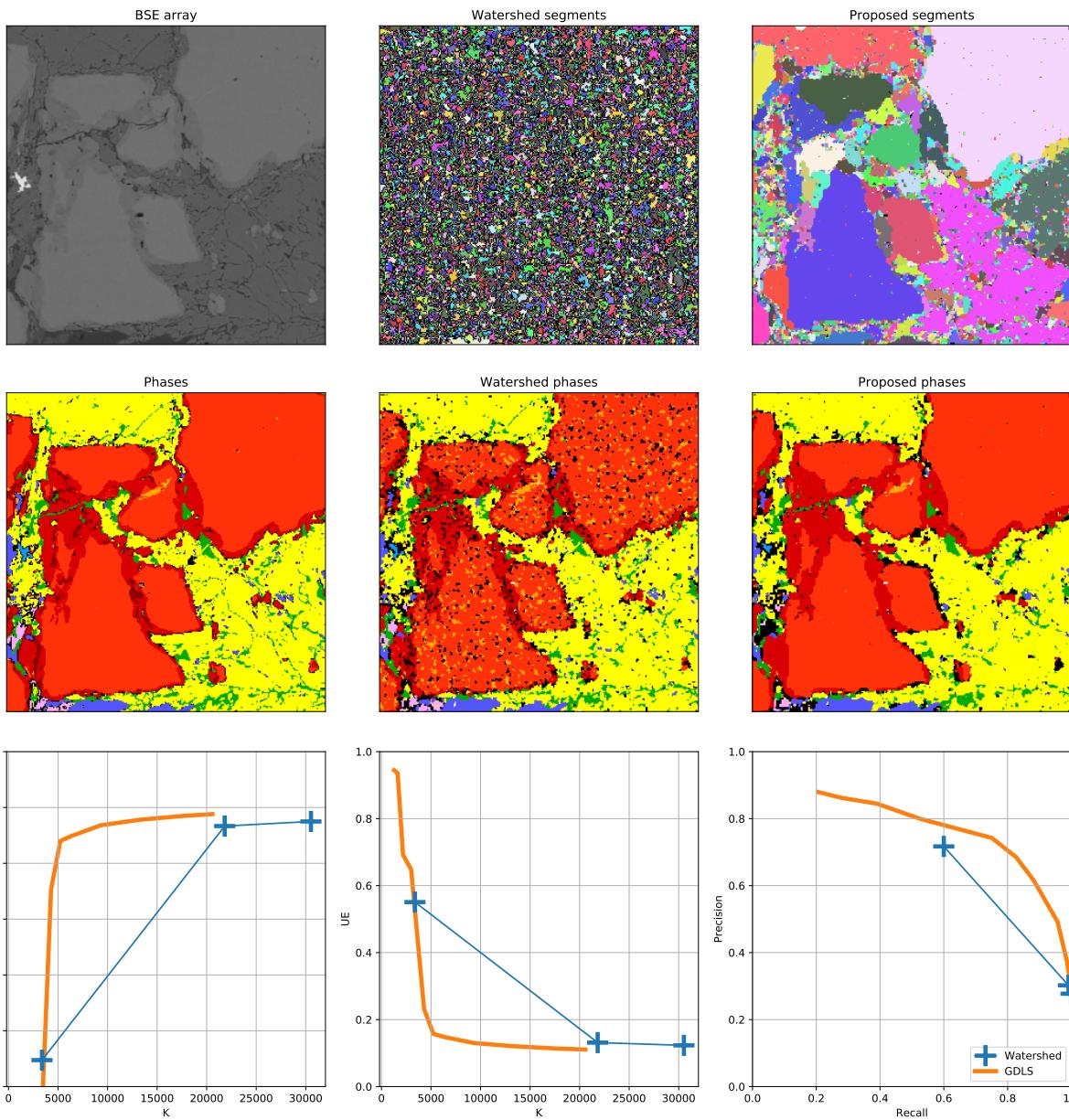
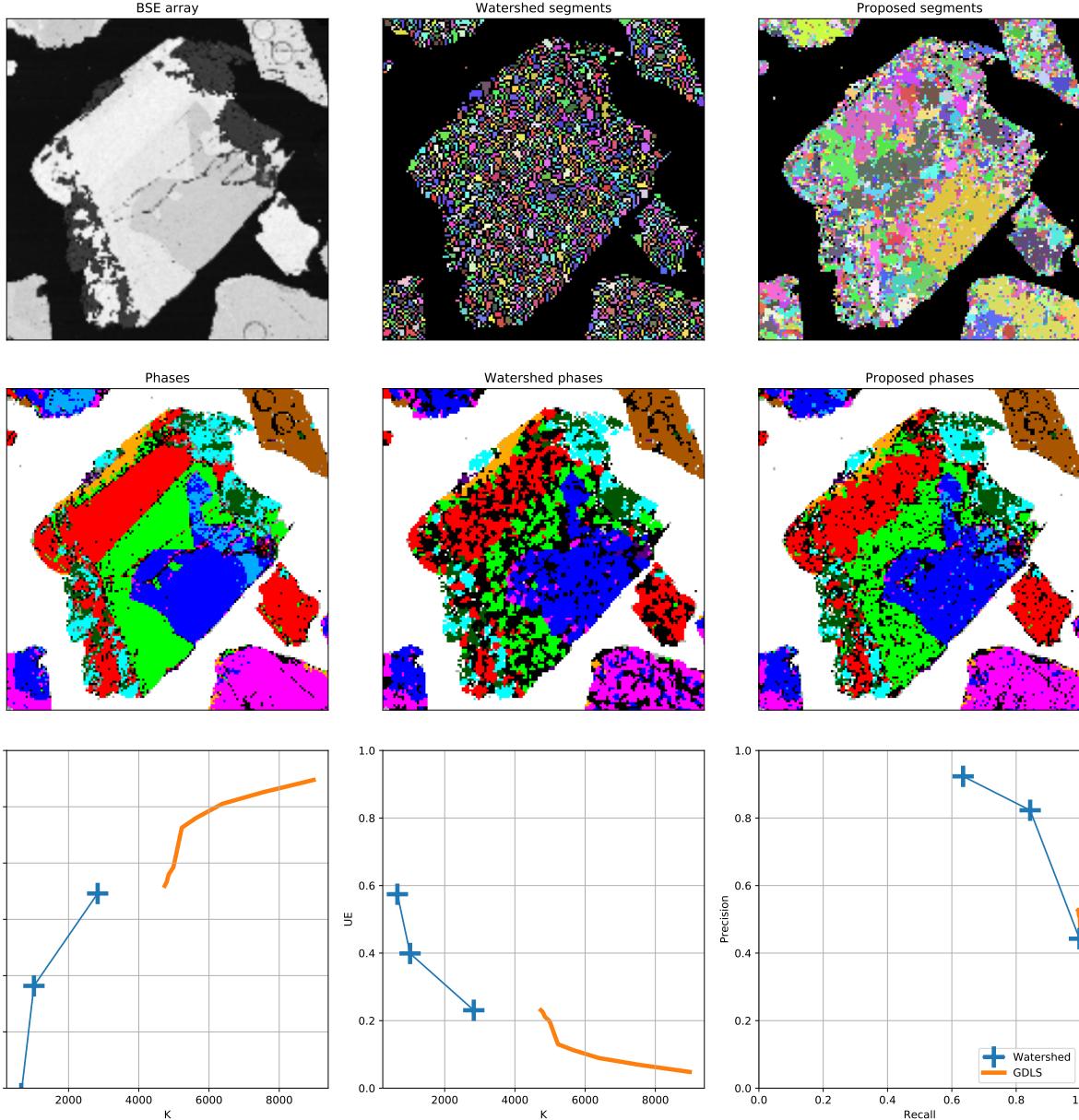


Figure 8: Example of segmentation and evaluated metrics of field C. For each field, first two rows (from left to right) show the BSE array of the field and ground truth phases, segments and phases obtained from Watershed-based segmentation, and segments and phases obtained by GDLS. The last row show the values of metrics for the different settings of watershed (blue) and GDLS (orange). The field contains andradite garnet (red) with Al-enriched, Fe-depleted (grossularite substitution) rims (dark red). The GDLS is able to distinguish these rims without significant oversegmentation and produces much clearer final phase map with less artifacts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Deep Segmentation



44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
Figure 9: Example of segmentation and evaluated metrics of field D. Subfigures are divided as in Figure 8. The field contains minerals of columbite-(Mn)-tantalite-(Mn) solid solution with distinct domains with different Nb/Ta ratio (red, green, blue and pink domains). Although GDLS produces more segments than watershed in this case, the final GDLS phase map contains less artifacts because most of the image area consist of comparatively larger segments.

516 figures 6 and 7). The criterion for setting the parameters₅₁₈ was proper recognition of the grossularite-rich zones (light₅₂₀ on the Al and darker on the Fe elemental map, respectively)₅₂₂. For watershed algorithm, separation of phases 18 had to be₅₂₄ chosen, resulting in serious oversegmentation (figure 8).₅₂₆

527 In the case of GDLS, the grossularite-rich zones were₅₂₉ recognized properly without any need for decreasing the thresh-₅₃₁ old so much, so the segments are large, and their boundaries₅₃₃ follow real phase boundaries. Unlike watershed, GDLS pro-₅₃₅ duces a much clearer phase map after classification, see Fig-₅₃₇ ure 8.

539 The field D consists of grains of Nb-Ta oxides (columbite₅₃₉

(Mn) to tantalite-(Mn)) containing zones with different Nb/Ta ratio (red, green, blue and pink colours on Figure 9). In this case, both the Ground truth and the evaluation data were acquired with dense EDS array (see table 1). The criterion for parameter settings was proper recognition of red and green zone, which differ by about 4 wt. % of Ta substituting for Nb. For watershed algorithm, the threshold needs to be lowered until it produces random small segments without any relation to phase boundaries (Figure 9) resulting in lot of artifacts in the final phase map. Compared to the dot mapping measurements, the GDLS tend to produce more segments, especially on phase boundaries and surface inhomogeneities. Despite

Deep Segmentation

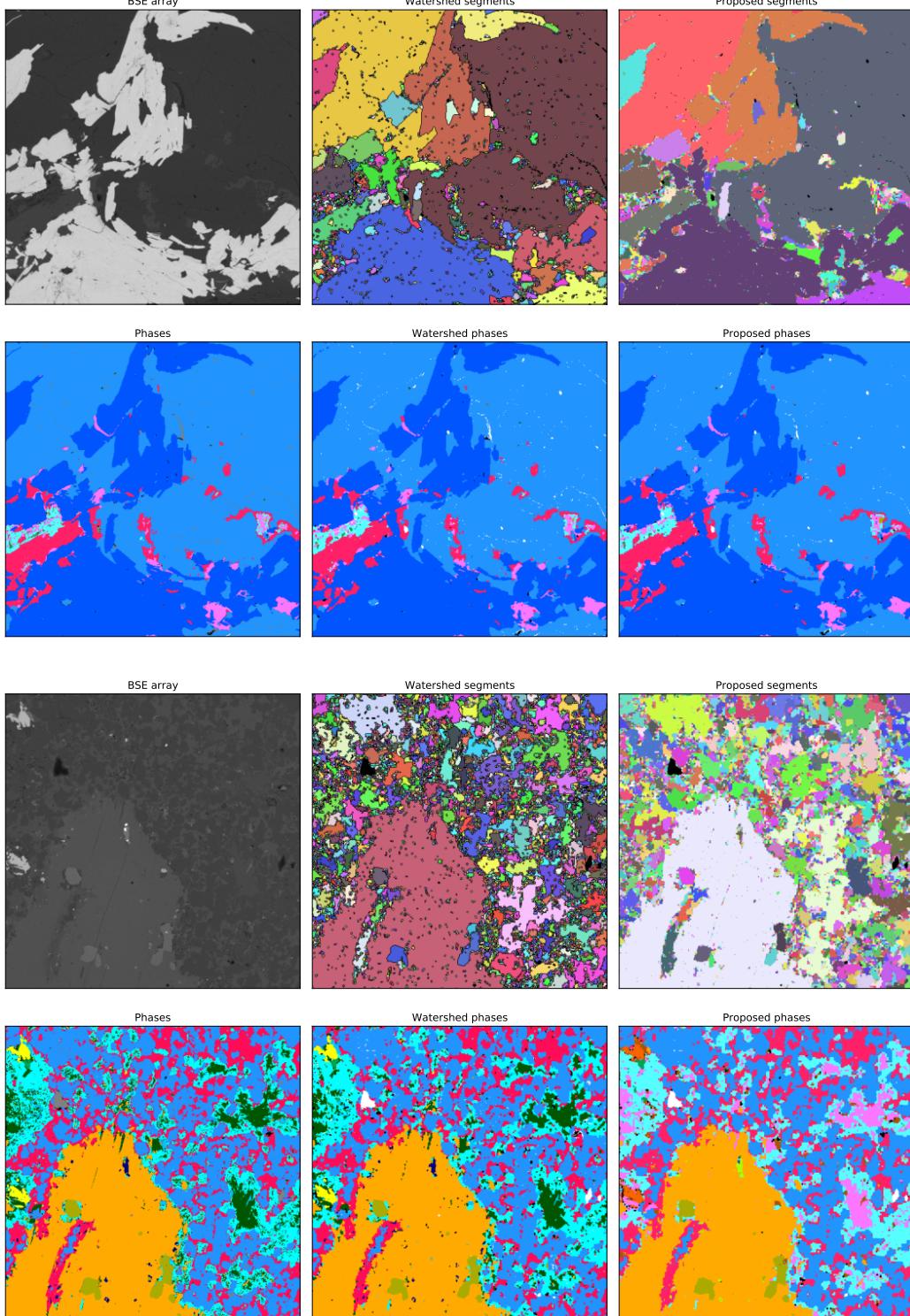


Figure 10: Comparison of the new method to Watershed of fields A and B. The evaluation metrics can be found in Table 2.

that, the GDLS segments most of the image with much larger⁵⁴⁵ segment size and produces a phase map with significantly⁵⁴⁶ less artifacts (Figure 9). Therefore, the GDLS surpasses the⁵⁴⁷ capabilities of watershed also on the dense EDS data array⁵⁴⁸ in practical evaluation.⁵⁴⁴

Figs 10 and 11 show fields A and B, respectively, where the GDLS performs comparably with watershed-based segmentation from the visual point of view.

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

4.5. Discussion

As can be gathered from the review of existing methods (Section 2), all existing commercial solutions rely heavily on BSE images to separate mineral grains. EDS data are used to classify phases pixel by pixel. In such cases, it is not possible to reliably distinguish mineral phases/individual zones in zoned mineral grains, which have similar but distinct composition. Depending on the actual composition differences, one needs to distinguish, it is necessary to acquire 1–2 orders of magnitude more X-ray events from each pixel (table 1) to correctly label all pixels and reduce the occurrence of artifacts (misclassified or unclassified pixels, which are classified incorrectly or not classified at all, as they do not fulfill the classification rules for any class due to statistical noise arising due to low number of X-ray events in the spectrum). But the 1–2 orders of magnitude more X-ray events per pixel means 1–2 orders slower data acquisition. Unlike the other solutions, the Mineralogic mining uses results of quantitative analysis of the EDS spectra (calculated wt. % of individual elements) and not the normalized peak intensities to classify the individual spectra (Keulen et al., 2020). However, this approach inherently suffers from the same limitations, because if the ratios of the peak intensities on individual pixels representing the same material are different due to statistical noise, inevitably, the calculated chemical composition from that spectra must be different as well. Applying any fine classification rules on such data to correctly label chemically similar areas and subsequently extract their morphological characteristics would not work.

TIMA is able to overcome the issues described above by producing cumulative spectra of segments and applying the classification rules on the cumulative spectra but not on every pixel, effectively eliminating statistical noise. On high-count cumulative spectra, it is possible to distinguish fine differences in chemical composition without misclassification artifacts. On the other hand, for this approach to work, it is essential that the segment boundaries follow real boundaries of the chemically coherent areas and that distinct phases are not merged together. In such cases, the cumulative spectrum does not represent any of the merged areas and gives the wrong result. For segmentation to avoid merging real phases together, one can lower the threshold (both on watershed and GDLS), which typically leads to oversegmentation and therefore fewer X-ray events in the cumulative spectra and the same problems described above.

It is seen from images and graphs in Figure 8, that up to a certain point, both methods produce comparable results, but once the threshold is lowered, GDLS will maintain large segments much longer, achieving much better overall correct classification and a clearer phase map.

The overall performance of the watershed-base segmentation and GDLS was evaluated on a desktop PC (hardware: Intel® Core™ i7-4790 @ 3.60GHz, 16 GB RAM and NVIDIA GeForce GTX 1660 SUPER, software: Windows 10 64-bit, Python 3.7.1 and TensorFlow 2.4.1). The duration of analysis for the test cases A, B, C and D (see table 1) are summarized in the table 3. The timing of GDLS was measured with

GPU support enabled and disabled. In the current state, GDLS is about one order of magnitude slower than the watershed-based segmentation (WSS). The optimum implementation of the inference phase was not the goal of this study, there is a room for optimization in future.

One of the still unresolved weak points of both the state-of-the-art and the GDLS algorithms is the behavior on the phase boundaries, where the spectra are mixed due to the overlapping EDS excitation volumes. Since these spectra does not correspond to any material, they are recognized as separate segment and after classification, they are left as unclassified since they do not fulfill classification rules for any material. Such cases need to be specifically approached in later steps e.g. by adjusting the classification scheme or performing some kind of automated classification based on similarity with neighboring segments during data post-processing.

5. Conclusion

The goal of the study we present in this paper was to develop novel graph-based segmentation of BSE and EDS hyperspectral images using deep learning. We were able to go beyond the state of the art in terms of the quality of mineralogy analysis with the same input data compared to previous approaches. We have also demonstrated that the proposed approach is readily acceptable for mineralogy experts and improves the quality of their outputs.

We evaluated the output both numerically and visually with convincing results. The new method yields superior results especially when natural solid solutions (e.g. Nb-Ta oxides, garnets etc.) containing domains with similar chemistry are present in the sample, a situation that is very common. In such cases, the new method provides larger average segment size while still correctly recognizing the boundaries of chemically similar domains. Classification of the cumulative spectra from larger segments reduces the probability of erroneous classification resulting from low X-ray events in the spectrum, which manifests in fewer artifacts in the final phase map. The new method enables better chemical/mineralogical resolution of similar phases while retaining the same speed of analysis.

Future work will focus on improvements of spectra simulation for NN training, improvements of the NN performing the mapping to energy regions, element weight concentrations or stoichiometry, with custom weights for individual elements and noise suppression. Denoising the BSE image, element map reconstruction from sparsely sampled EDS data, and better handling of mixed spectra on phase boundaries would also be beneficial.

6. Computer Code Availability

The algorithm described in this work has been tested using Python and the Tensorflow framework. Visualizations are prepared with Matlab. The code necessary to replicate these results is freely available at

<https://github.com/RomanJuranek/graph-segmentation>

Field	Field size	EDS points	WSS CPU only	GDLS	
				CPU only	GPU support
A	500 × 500	31763	1.1	23.0	12.2
B	500 × 500	29917	1.1	23.4	13.2
C	400 × 400	21717	0.7	14.2	6.8
D	150 × 150	13772	0.2	7.5	2.7

Table 3

Comparison of timing of the segmentation methods on the testing fields A, B, C and D. *Field size* = size of field (width × height), in pixels; *EDS points* = number of EDS measurement points; *WSS* = duration of the watershed-based segmentation, in seconds, *GDLS* = duration of GDLS, in seconds.

7. Acknowledgements

to be completed after review

References

- Ayling, B., Rose, P., Zemach, E., Drakos, P., Petty, S., 2012. QEM⁷¹⁵ SCAN (Quantitative Evaluation of Minerals by Scanning Electron Mi⁷¹⁶croscopy): capability and application to fracture characterization in⁷¹⁷ geothermal systems, in: AGU Fall Meeting Abstracts, pp. 1158–.⁷¹⁸
- Bowyer, A., 1981. Computing dirichlet tessellations. *The computer journal*⁷¹⁹ 24, 162–166.⁷²⁰
- Breiter, K., Ďurišová, J., Hrstka, T., Korbelová, Z., Hložková Vaňková, M.,⁷²¹ Vašinová Galiová, M., Kanický, V., Rambousek, P., Kněsl, I., Doběš,⁷²² P., Dosbaba, M., 2017. Assessment of magmatic vs. metasomatic pro-⁷²³cesses in rare-metal granites: A case study of the Cínovec/Zinnwald⁷²⁴ Sn–W–Li deposit, Central Europe. *Lithos* 292–293, 198–217. doi:⁷²⁵10.1016/j.lithos.2017.08.015.⁷²⁶
- Di Zenzo, S., 1986. A note on the gradient of a multi-image. *Computer⁷²⁷ Vision, Graphics, and Image Processing*⁷²⁸ 33, 116 – 125. doi:^{10.1016/0734-189X(86)90223-9.}⁷²⁹
- Drewniok, C., 1994. Multi-spectral edge detection. some experiments on⁷³⁰ data from Landsat-TM. *International Journal of Remote Sensing*⁷³¹ 15,⁷³² 3743–3765.⁷³³
- Fandrich, R., Gu, Y., Burrows, D., Moeller, K., 2007. Modern SEM-based⁷³⁴ mineral liberation analysis. *International Journal of Mineral Processing*⁷³⁵ 84, 310–320.⁷³⁶
- Gloy, G., 2016. High speed differentiation of mineral phases of similar BSE⁷³⁶ intensity with AMICS. https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/X-rayDiffraction_ElementalAnalysis/Microanalysis_EBSD/LabReports/App_eds_16_amics_mineral_phases_lores.pdf. Accessed⁷³⁹ 2020-07-02.⁷⁴⁰
- Goldstein, J., Newbury, D., Michael, J., Ritchie, N., Scott, J., Joy, D., 2018.⁷⁴¹ Scanning Electron Microscopy and X-ray Microanalysis. doi:^{10.1007/978-1-4939-6676-9.}⁷⁴²
- Gottlieb, P., Wilkie, G., Sutherland, D., Ho-Tun, E., Suthers, S., Perera, K.,⁷⁴⁴ Jenkins, B., Spencer, S., Butcher, A., Rayner, J., 2000. Using quantita-⁷⁴⁵tive electron microscopy for process mineralogy applications. *Jom* 52,⁷⁴⁶ 24–25.⁷⁴⁷
- Grant, G., Reid, A.F., 1981. An efficient algorithm for boundary tracing⁷⁴⁸ and feature extraction. *Computer Graphics and Image Processing*⁷⁴⁹ 17,⁷⁵⁰ 225–237. doi:^{10.1016/0146-664X(81)90003-4.}⁷⁵¹
- Gray, R., Neuhoff, D., 1998. Quantization. *IEEE Transactions on Informa-⁷⁵¹tion Theory* 44, 2325–2383. doi:^{10.1109/18.720541.}⁷⁵²
- Haluzová, E., Ackerman, L., Pašava, J., Jonášová, v., Svojtka, M., Hrstka,⁷⁵³ T., Veselovský, F., 2015. Geochronology and characteristics of Ni-Cu⁷⁵⁴ (PGE) mineralization at Rožany, Lusatian granitoid complex, Czech Re-⁷⁵⁵public. *Journal of Geosciences* 60, 219–236.⁷⁵⁶
- Hrstka, T., Gottlieb, P., Skála, R., Breiter, K., Motl, D., 2018. Automated⁷⁵⁷ mineralogy and petrology-applications of TESCAN Integrated Mineral⁷⁵⁸ Analyzer (TIMA). *Journal of Geosciences* 63, 47–63.⁷⁵⁹
- Jasiewicz, J., Stepinski, T., Niesterowicz, J., 2018. Multi-scale seg-⁷⁶⁰mentation algorithm for pattern-based partitioning of large categorica-⁷⁶¹l rasters. *Computers Geosciences* 118, 122 – 130. URL: <http://www.sciencedirect.com/science/article/pii/S0098300417305204>, doi:<https://doi.org/10.1016/j.cageo.2018.06.003>.
- Karimpouli, S., Tahmasebi, P., 2019. Segmentation of digital rock im-⁷¹²ages using deep convolutional autoencoder networks. *Computers Geosciences* 126, 142 – 150. URL: <http://www.sciencedirect.com/science/article/pii/S0098300418303911>, doi:<https://doi.org/10.1016/j.cageo.2019.02.003>.
- Karvelis, P.S., Tzallas, A.T., Fotiadis, D.I., Georgiou, I., 2008. A multichannel watershed-based segmentation method for multispectral chromo-⁷¹³some classification. *IEEE Transactions on Medical Imaging* 27, 697–⁷¹⁴708.
- Keulen, N., Malkki, S.N., Graham, S., 2020. Automated quantitative min-⁷¹⁵eralogy applied to metamorphic rocks. *Minerals* 10, 47.
- Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- Li, S.Z., 1994. Markov random field models in computer vision, in: European conference on computer vision, Springer. pp. 361–370.
- Miller, P.R., Zuiderwyk, M.A., Reid, A.F., 1983. QEM*SEM image analy-⁷¹⁶sis in the determination of modal assays, mineral associations and min-⁷¹⁷eral liberation, in: XIV International Mineral Processing Congress, pp. 1–20.
- Motl, D., Filip, V., 2013. Method and apparatus for material analysis by⁷¹⁸ a focused electron beam using characteristic X-rays and back-scattered electrons. US Patent App. 13/398,114.
- Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines, in: Proceedings of the 27th International Conference on International Conference on Machine Learning, Omnipress, Madison, WI, USA. p. 807–814.
- Reed, S.J.B., 2005. Electron microprobe analysis and scanning electron microscopy in geology. Cambridge university press.
- Reid, A., Gottlieb, P., Macdonald, K., Miller, P., 1985. QEM*SEM image analysis of ore minerals: volume fraction, liberation, and observational variances. *Liberation and Observational Variances. Applied Mineralogy*, 191–204.
- Ritchie, N.W., 2009. Spectrum simulation in DTSA-II. *Microscopy and Microanalysis* 15, 454–468. doi:^{10.1017/S1431927609990407.}
- Sandmann, D., 2015. Method development in automated mineralogy. Ph.D. thesis. Technische Universität Bergakademie Freiberg, Deutschland.
- Sandmann, D., Gutzmer, J., 2015. Nature and distribution of PGE mineralisation in gabbroic rocks of the Lusatian Block, Saxony, Germany Art und verbreitung der PGE-Mineralisation in den gabbroiden Gesteinen des Lausitzer Blocks, Sachsen, Deutschland. *Zeitschrift der Deutschen Gesellschaft für Geowissenschaften* 166, 35–53.
- Schroff, F., Kalenichenko, D., Philbin, J., 2015. FaceNet: A unified embed-⁷¹⁸ding for face recognition and clustering, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE. pp. 815–823. doi:^{10.1109/cvpr.2015.7298682.}
- Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. [arXiv:1409.1556](https://arxiv.org/abs/1409.1556).
- Stutz, D., Hermans, A., Leibe, B., 2018. Superpixels: An evaluation of the state-of-the-art. *Computer Vision and Image Understanding* 166, 1–27.
- Sutherland, D.N., Gottlieb, P., 1991. Application of automated quantitative

1
2
3 ⁷⁶² mineralogy in mineral processing. Minerals Engineering 4, 753–762.
4 ⁷⁶³ Výravský, J., Hreus, S., Cempírek, J., Vašinová Galiová, M., 2019. Assess-
5 ⁷⁶⁴ ment of scandium deportment and recovery potential in Li-Sn-W greisen
6 ⁷⁶⁵ deposit Cínovec/Zinnwald using EPMA, LA-ICP-MS and TIMA auto-
7 ⁷⁶⁶ mated mineralogy, in: Proceedings of the 15th SGA Biennial Meeting,
8 ⁷⁶⁷ 27-30 August 2019, Glasgow, Scotland, pp. 1638–1641.
9 ⁷⁶⁸ Xu, C., Kynický, J., Song, W., Tao, R., Lü, Z., Li, Y., Yang, Y., Pohanka,
10 ⁷⁶⁹ M., Vašinová Galiová, M., Zhang, L., et al., 2018. Cold deep subduction
11 ⁷⁷⁰ recorded by remnants of a Paleoproterozoic carbonated slab. Nature
12 communications 9, 1–8.
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Graph-based Deep Learning Segmentation of EDS Spectral Images for Automated Mineral Phase Analysis

Roman Juránek^a, Jakub Výravský^{b,*}, Martin Kolář^a, David Motl^b and Pavel Zemčík^a

^aBrno University of Technology, Faculty of Information Technology, Božetěchova 1/2, Brno 612 66, Czech Republic

^bTESCAN Brno, s.r.o., Libušina tř. 1, Brno 623 00, Czech Republic

ARTICLE INFO

Keywords:
segmentation
deep learning
EDS spectra
automated mineralogy

ABSTRACT

We introduce a novel method for graph-based segmentation of spectral images obtained using a Scanning Electron Microscope (SEM) equipped with an Energy Dispersive X-ray spectroscopy (EDS) detector. The method exploits deep learning along with fusion of rasterized electron microscopy images with sparse EDS samples to obtain accurate mineralogy segmentation with high efficiency. Improvements over previous methods are with respect to the goal of an improved quantitative and qualitative assessment of segmentation, so that volumetric composition is indirectly addressed. We describe the principles of the novel method, show experimental results on real samples and demonstrate its advantages in comparison to the state of the art. The new method performs unsupervised clustering on sparsely measured EDS spectra, allowing for classification of unseen mineralogical compounds. Then, the processed spectra are combined with single channel SEM measurements through an optimised lattice, where a Markov Field is used to perform spatial segmentation in image. The benefit of this material-agnostic method is that clusters can then be (separately) classified, analyzed, and small grains with distinct EDS measurements are more accurately separated than in previous methods. These improved results are evaluated quantitatively on ground-truth electron microscope measurements with dense high-count EDS data, as well as visually through analysis by a mineralogist.

1. Introduction

A Scanning Electron Microscope (SEM) equipped with an Energy Dispersive X-ray spectroscopy (EDS) detector is a well-established tool to analyze chemical composition of samples with a spatial resolution of the order of μm (Goldstein et al., 2018) and is widely used especially in material science and geological/mineralogical research. Automated mineralogy instruments are based on the capabilities of SEM with Backscattered Electron (BSE) (Goldstein et al., 2018) detector and up to 4 EDS detectors to rapidly acquire data from a statistically significant area of the subject material/mineral surface. Such automated mineralogy instruments are also equipped with dedicated software capable of distinguishing individual mineral grains, classifying them based on a user-defined list of classification rules, and deriving their morphological and textural characteristics, as well as the relative abundance of individual phases, associations, degree of liberation etc. (Fandrich et al., 2007)

The primary application for automated mineralogy is in the mineral processing industry, where epoxy-mounted polished samples of crushed mineral particles from various stages of mineral processing (feed, concentrate, tailings) are analysed in order to investigate the optimal degree of comminu-

tion, to track potential losses and in general to optimise the efficiency of the processing plant. (Gottlieb et al., 2000)

Automated mineralogy is also widely used in basic mineralogical/petrological research, where typically, glass-mounted, polished thin-sections of rocks are investigated. The required outputs include modal mineralogy, elemental deportment, mineral associations, location of minerals of interest for further analysis (EPMA, LA-ICP-MS) and many others. (Výravský et al., 2019; Hrstka et al., 2018; Breiter et al., 2017; Haluzová et al., 2015; Xu et al., 2018; Sandmann and Gutzmer, 2015) Other less common applications include e.g. oil & Gas, coal and environmental and material sciences (see (Sandmann, 2015) for details and references).

Naturally, proper recognition of phase boundaries is essential for meaningful and trustworthy analytical results and interpretation. The phases are generally recognized 1) based on their BSE brightness value only, or 2) using a combination of BSE and EDS data. The first approach is based on the observation that the number of backscattered electrons is proportional to the mean atomic number (Z) of the investigated material (Goldstein et al., 2018). This method is relatively simple but fails to distinguish phases with different composition but similar mean Z , which is rather common in geological samples (Hrstka et al., 2018), and as such, is suitable only for specific samples. The second approach makes it possible to distinguish phases with different chemistry but similar mean Z , however, acquisition of EDS data incomparably slower than acquisition of BSE data (milliseconds to microseconds per pixel, respectively), thus slowing analysis. To keep the analysis reasonably short, one could either acquire the EDS data in coarser grid than the BSE (Hrstka et al., 2018; Motl and Filip, 2013), or acquire only a relatively small number of X-ray events from each pixel (e.g.

*Corresponding author

✉ ijuranek@fit.vutbr.cz, Designed and implemented the method, trained the embedding models and evaluated the results (R. Juránek); jakub.vyravsky@tescan.com, Data acquisition, Ground truth preparation, visual evaluation (J. Výravský); kolar martin@fit.vutbr.cz, MRF segmentation, implementation and coordination (M. Kolář); david.motl@tescan.com, Software implementation and coordination (D. Motl); zemcik@fit.vutbr.cz, Coordinator (P. Zemčík)

ORCID(s): 0000-0003-0589-0172 (R. Juránek); 0000-0002-0049-4018 (J. Výravský); 0000-0001-6933-2909 (M. Kolář); 0000-0002-6583-1927 (D. Motl); 0000-0001-7969-5877 (P. Zemčík)

1,000) or combine both. The disadvantage is that the low-count EDS spectra are prone to statistical noise, which is typically higher than spectral differences of the phases of interest. Another disadvantage lies in the fact, that if you acquire EDS in coarser grid, you may miss grains, which are smaller than the EDS grid in case that they are not distinct in BSE image (Hrstka et al., 2018). Classifying such data pixel-by-pixel inevitably leads to artifacts such as misclassified or unclassified pixels. The current state of the art solution — TESCAN Integrated Mineral Analyzer (TIMA) uses a modified multispectral watershed-based segmentation (see section 2) to distinguish chemically coherent areas (segments) and spectra-summing algorithms to produce a high-count, noise-free, representative spectrum for each segment (Hrstka et al., 2018). However, the watershed-based segmentation is not able to distinguish chemically similar phases beyond a certain point. This leads either to phases being mixed together or too many segments being produced with relatively noisy sum spectra (depending on threshold settings).

We have developed a Graph-based Deep Learning Segmentation method (GDLS), with the capacity to correctly distinguish chemically similar phases in a BSE image, combined with noisy and/or sparsely occurring EDS data, while still maintaining a large average segment size. This method is therefore of great benefit with regard to robustness of results and speed of analysis. In this contribution, we present this new segmentation method using a combination of deep neural embedding and Markov Random Fields and compare the results with the currently available method implemented in TIMA (Hrstka et al., 2018).

2. Prior Work

The first commercially available automated mineralogy instruments were developed in 1980s primarily for the mineral processing industry. These include the Mineral Liberation Analyzer (MLA) (Fandrich et al., 2007), QEM*SEM and its successor QEMSCAN (Ayling et al., 2012; Miller et al., 1983). Over the last decade, several other automated mineralogy solutions have became available e.g. TIMA (TESCAN Integrated Mineral Analyzer (Hrstka et al., 2018)), Mineralogic Mining (Keulen et al., 2020) or AMICS (Advanced Mineral Identification and Classification System (Gloy, 2016)). Each of these solutions uses specific implementations of BSE/EDS acquisition modes and image segmentation.

2.1. Overview of existing solutions

The MLA analyzer starts with acquiring a BSE image. This is followed by particle de-agglomeration and phase segmentation, where regions of homogeneous gray levels within the BSE image are found. No details on methods employed for phase segmentation have been published. Depending on the selected acquisition method, spectroscopic data are collected either in a single measurement point in the center of each grain, or continually, as the electron beam scans over the entire grain; alternatively, individual spectra are collected from each point of a grid-pattern of points covering the entire particle. The spectroscopic data are then used to

identify the mineral species (Fandrich et al., 2007). Other methods segmenting BSE only exist, using autoencoder networks (Karimpouli and Tahmasebi, 2019) and the U-Net architecture (Jasiewicz et al., 2018).

In the QEM*SEM analyser and its successor QEMSCAN, an X-ray spectrum along with the BSE signal is collected for each measurement point in an equidistant rectangular grid. The acquisition step is followed by identification of mineral species, for each measurement point individually utilizing both BSE and EDS data. Image segmentation is basically identification of continuous horizontal line segments of measurement points classified as the same mineral. Feature extraction is done by finding connected line segments from adjacent horizontal lines (Grant and Reid, 1981). A correction for boundary phases (misidentification of mineral species due to mixed spectrum at grain boundaries) is also employed, but details are not publicly known (Reid et al., 1985; Sutherland and Gottlieb, 1991).

Zeiss mineralogic uses BSE thresholding to select areas to be analyzed as well as determining grains based on BSE intensity to extract morphological characteristics. Grains are then classified based on EDS spectra obtained in several ways (one point in center of the grain, continual scan over grain, coarse grid of points, dense equidistant grid). Also, some segmentation methods are available for image postprocessing. However, specific details are not public knowledge (Keulen et al., 2020).

The AMICS automated mineralogy software from Bruker uses a proprietary BSE segmentation algorithm to determine mineral phases. It contains two adjustable interrelated parameters, which control sensitivity to gray level variation and size of the segmented region. Each segment is then analysed with EDS to classify the phase (Gloy, 2016).

TIMA achieves grain identification using multi-channel image analysis and combining image data from a BSE detector as well as EDS data. The method starts with the multi-channel gradient, followed by the watershed algorithm with adjustable threshold. The grain identification phase is followed by accumulation of spectroscopic data that are acquired from measurement points in each individual grain and determining the mean signal from the BSE detector, the accumulated data are used to identify mineral species (Motl and Filip, 2013; Hrstka et al., 2018).

2.2. Watershed-based Segmentation

To apply the watershed-based segmentation algorithm to multichannel data, the multichannel image gradient must be computed. Computation of the tensor gradient was introduced by DiZzenzo et al. (Di Zenzo, 1986) for color images, Drewniok (Drewniok, 1994) extended this work to multispectral images. One application of this algorithm is also mentioned in (Karvelis et al., 2008).

To reduce dimensionality and suppress noise in low-count EDS spectra, elemental maps are produced by summing stored X-ray events from all spectrum channels corresponding to energy regions of important X-ray emission lines of the individual elements. These elemental maps are individual chan-

1
2
3
4
5
6
7
8
9
10
nels in the multichannel image and are used to calculate the
combined gradient.

3. Graph-based Deep Learning Segmentation

11
12
13
14
15
16
17
18
19
20
The proposed segmentation method takes one input field
composed of a densely sampled BSE array and sparsely sam-
pled EDS data with known measurement locations.

21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
30,000 eV) is uniformly divided into spectrum channels each
of which 10 eV wide. Therefore, the spectral data have 3,000
channels per pixel. The values in the BSE array are cal-
ibrated, with a maximal value of 2^{16} . An example field is
seen in Figure 1.

40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
The spectra data associated with each measurement location has form of an 1D array of integer values. The index to the array is an ordinal number of a spectrum channel, the values stored in the array are frequencies of detection of X-ray photons in the location with certain energy. We have chosen spectrum channels to be uniformly distributed, each channel 10 eV wide. This choice of spectrum division has been directed by finding compromise between error that arises from the quantization of the photon energy and memory footprint of the spectral data.
The spectra data associated with each measurement location has form of an 1D array of integer values. The index to the array is an ordinal number of a spectrum channel, the values stored in the array are numbers of detection of X-ray photons in the particular channels. We chose the energy continuum to be divided into spectrum channels uniformly distributed with channel width to be 10 eV, since it provides good compromise between quantization error and memory footprint of the spectral data. Peaks corresponding to characteristic lines as detected by the EDS detector can be approximated by Gaussian functions superimposed on a continuum background. Although peak width, conventionally expressed as Full Width at Half Maximum (FWHM) is a function of several factors, the most important of them to be statistics of generating electron-hole pair in silicon (which is a function of photon energy), electronic noise in the detector and preamplifier, and width of an energy measurement circuit.(Reed, 2005) We decided to set up the spectrometer to conditions optimized for high throughput, because such conditions are typical for practical sample analysis by means of an automated mineral analyzer. In our sample data, the FWHM of light element peaks, e.g. carbon, is > 70 eV. Using the Gaussian model of peaks of characteristic lines, the FWHM value can be converted to standard deviation σ_{res} using formula ($FWHM = 2\sqrt{2 \ln 2} \sigma \approx 2.355 \sigma$), therefore $\sigma_{res} \approx 1/2.355 \cdot 70$ eV ≈ 30 eV.

The additional error introduced by the quantization of energy continuum into uniformly distributed spectrum channels is $1/\sqrt{12}$ LSB (standard deviation of the distribution).(Gray and Neuhoff, 1998) In our case $\Delta = 10$ eV and the corresponding standard deviation of the quantization error $\sigma_{quant} = 1/\sqrt{12} \cdot 10$ eV ≈ 2.9 eV. The standard deviation of quantization error σ_{quant} is more than one order of magnitude lower than the standard deviation that corresponds to the width of the narrowest (light element) peaks σ_{res} in our sample data. Using higher spectrum channel width, we start to lose information, using lower width, we increase the memory footprint of the spectral data unjustifiably.

The energy continuum of detectable X-ray photons (0 to

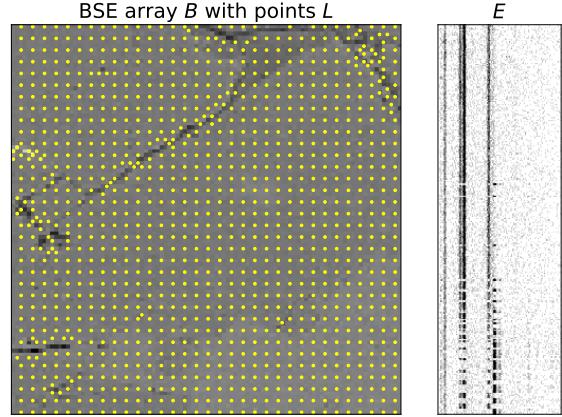


Figure 1: The field is a hyperspectral image with sparsely sampled data. The locations L of EDS measurements are superimposed on the BSE array B . The matrix E shows the actual spectra in rows. Details in section 3.1

30,000 eV) is uniformly divided into spectrum channels each of which 10 eV wide. Therefore, the spectral data have 3,000 channels per pixel. The values in the BSE array are calibrated, with a maximal value of 2^{16} . An example field is seen in Figure 1.

The output of the segmentation process is a dense array with pixel-wise labels (integer values) where pixels with the same value belong to the same segment. The core of the proposed method is to determine if spatially close EDS measurements capture the same material, in which case they belong to the same segment and should be labeled with a common label. We construct a graph with EDS measurements represented as vertices. Edges connect spatially close measurements in a planar graph obtained by Voronoi analysis (Bowyer, 1981). A deep learning model assigns a value to each edge. This value corresponds to differences in BSE and distances between hyperspace representations of EDS spectra. Edges with high differences are removed and the resulting graph components constitute separate segments — labels. Final dense labels are obtained by a dense Markov Random Field operating over the BSE array with labels initialized from the graph. An example of graph and segmentation is in Figure 2. The individual steps of the algorithm which are detailed in the following text are:

- Projection of spectra to latent space.
- Graph construction and calculation of edge values.
- Graph component analysis.
- Markov Random Field segmentation.

3.1. Field representation

A field F (see Figure 1, left) is represented as a triplet $F = (B, E, L)$. B is $H \times W$ array with BSE measurements (an image of width W and height H). E and L are matrices

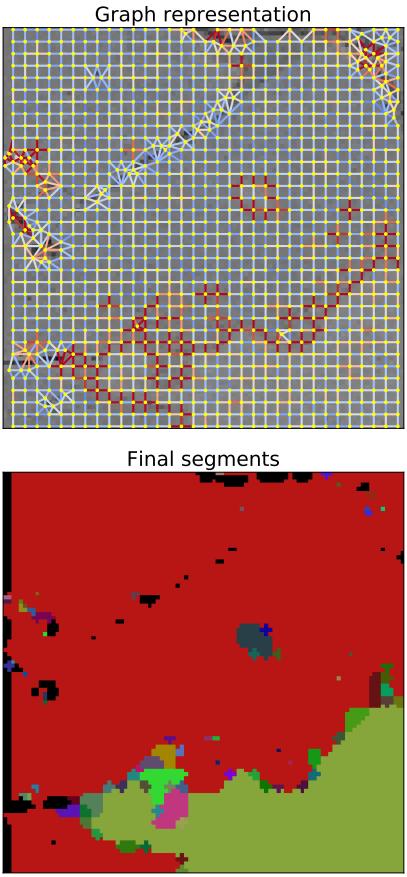


Figure 2: The proposed method constructs a graph representation of the field (edge color encodes δ_z , see Section 3.6) and decomposes the graph to obtain compact segments. Details in the text.

containing N rows with EDS measurements

$$E = \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix} \quad L = \begin{pmatrix} l_1 \\ \vdots \\ l_N \end{pmatrix}$$

where E is $N \times D$ matrix with N spectra with dimensionality D , and L is a $N \times 2$ matrix with locations of EDS measurements in the BSE array. The spectrum s_i corresponds to location l_i in the BSE array.

The proposed method assumes full coverage by B , and sparse coverage by EDS measurements, whose locations are determined at scan-time by a separate mechanism internal to the electron microscope. Although the proposed method poses no restrictions on the locations of the sparse EDS measurements, they are best placed at regular intervals on discrete materials, and ideally do not miss small separate grains.

3.2. Mineral dataset

For training the embedding model we use a dataset of 4,507 distinct materials relevant for the application. Each material is characterized by its ideal spectrum. In case of polymorphs (materials with identical chemistry, but different crystal structure e.g. graphite and diamond), which are

indistinguishable by EDS, only one of the polymorphs was used in the training dataset. During the training, we generate spectra from the ideal profiles by random sampling with a random number of events to simulate data from a microscope. These spectra are close to what is captured by the microscope, but it is not perfect since the ideal profiles do not include recurring artefacts of pulse processing electronics like sum peaks or escape peaks, and other unmodeled effects (Ritchie, 2009). These differences influence the overall performance of the proposed method on the real data especially in situations where multiple materials are mixed (e.g. on boundaries between two materials) since we train on pure spectra with no “deformations”.

3.3. Spectral decomposition with deep learning

The key part of the algorithm is to determine if two spectra capture the same material, which are composed from the same elements in the same ratio. While this can be solved with relatively low error by, e.g., mean absolute difference or a two-sample K-S test on high count spectral data, in presence of noise due to low number of X-ray events per pixel the task becomes very difficult. In our solution, we process the spectra with a deep learning model to provide a compact descriptor which can differentiate between spectra with different compositions. The descriptor is obtained by embedding to a low dimensional space – latent space Z as shown in Figure 3. Transforming EDS measurements into this latent space serves as a preprocessing step to separate image-space segmentation from material composition and characteristic line differentiation, but this model is made on resampled ideal spectra, which cannot contain all noise and variations of real measurements. By training the mapping to differentiate thousands of randomly selected materials, the network learns to generalize even unseen materials. Therefore, during segmentation, which is not trained for specific materials, the presented approach is not limited to previously seen minerals and solid solutions. This is demonstrated in Figure 9, where chemically-distinct domains with different Nb/Ta ratio in grains of columbite-tantalite solid solution are properly segmented despite the fact, that no spectrum corresponding to such chemical composition was in the training dataset.

3.4. Neural network design

The convolutional neural network (CNN, see Figure 4) comprises three parts — *normalization*, *feature extraction* and *embedding*. The input of the network is the vector of the spectrum s with dimensionality D followed by a normalization with mean μ and standard deviation σ calculated from ground truth spectra. We use the 200 eV to 15,560 eV spectral band with 10 eV bins (therefore the input to the network are spectra with $D = 1,536$). There are two reasons: first, coefficients outside this band are mostly uninformative and second, reducing the number of coefficients reduces the number of network parameters and this positively affects the execution speed of the network. Nothing prevents using the whole spectrum, though. Feature extraction contains 3 blocks with a sequence of convolution, batch normalization and ReLU activation (Nair and Hinton, 2010) followed by

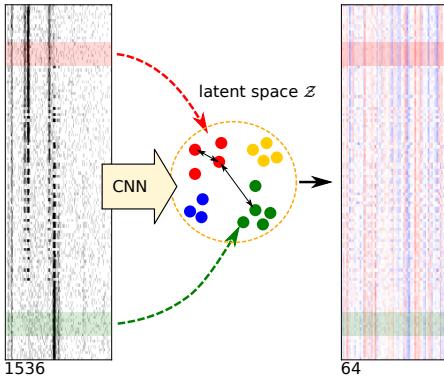


Figure 3: The original spectra are projected by CNN to a latent space \mathcal{Z} with 64 dimensions, where similar materials are mapped to similar locations

average pooling. Embedding processes the features with a sequence of dense layers and normalizes the output to unit length, producing a latent vector z .

Although almost any network design can be used here, we chose a simple network inspired by the VGG architecture (Simonyan and Zisserman, 2014) mainly for performance reasons. We changed the input shape, number of layers and blocks, and use *average pooling* instead of *max pooling* for downsampling since it better utilizes input data and has the ability to smooth the noise (inherent to spectral data). We did not observe any severe effect on the final results regarding the choice of downsampling strategy in preliminary experiments. Inference speed of this network is sufficient. On a single field, usually tens of thousands of measurements must be processed (in extreme cases it can be hundreds of thousands). This network can produce its output in a matter of seconds on low-end hardware which is still acceptable.

3.5. Embedding to low dimensional space

We train a neural network model M which produces a unit vector $z = M(s)$ for a spectrum s . The main purpose of this network is to make a compact, low dimensional representation of spectra. We reduce the input dimensionality from 3 000 to 64. Setting the dimensionality too low would lead to inefficient embedding which would not capture the differences in analyzed materials, setting it too high would increase memory and time requirements. We fixed the dimensionality to 64 in our experiments as a tradeoff between embedding accuracy and its size. We use the spectral information as an input and material class (i.e. its integer ID) as an output and optimize the network with ADAM (Kingma and Ba, 2014) optimizer with Semi-hard Triplet loss (Schroff et al., 2015).

The main property of the latent space \mathcal{Z} is that the difference $\|M(s_1) - M(s_2)\|_2$ is close to zero if s_1 and s_2 are from the same material and it is larger if the materials are different, whatever the materials are. By multiplying matrices M and E , individual EDS measurements s are transformed into the spectrum decomposition z .

3.6. Graph construction and labeling

From the known locations of EDS measurements in the field, we build a planar graph $G = (V, E)$ by Voronoi analysis (Bowyer, 1981). Vertices $V = (l_i, b_i, z_i)$ are locations of the measurements, each assigned with its location l_i , BSE value b_i and spectrum decomposition z_i obtained by the neural network. Edges are formed between spatially close measurements. The edges between measurements with distance larger than the $2\times$ sampling distance are not considered. Each edge $E_k = (i, j)$ is assigned with values δ_k^b and δ_k^z capturing difference in BSE and latent space representation of the spectrum.

$$\delta_k^b = |b_i - b_j|/2^{16} \quad (1)$$

$$\delta_k^z = \|z_i - z_j\|_2 \quad (2)$$

Where the value of δ^b is in the $(0, 1)$ interval and $\delta^z >= 0$. We remove the edges with $\delta_k^b > \tau_b$ or $\delta_k^z > \tau_z$ which disconnects measurements with high difference in BSE or spectrum. The parameters τ_d and τ_z can be tuned by the user to obtain different segmentations of the input data. The components of the resulting graph then constitute sets of measurements with the same material. As a result, each vertex is assigned a value $c_i \in N$ which is the label of the component it belongs to. These labels are the starting point for the Markov Field which generates the final pixel-wise label map.

3.7. Pixel labelling with dense MRF

Markov Random Field segmentation (Li, 1994) is a flexible optimisation method to create unlabeled segments given neighbouring pixel similarities and a sparse marker initialization. The field of pixel-wise BSE measurements is set on a fully connected grid, and we initialize the segmentation map with known labels from EDS graph analysis. By using *a priori* probability of sampling different measurements from the same distribution (class), and the probability of adjacency of different classes, the entire system becomes a joint distribution whose *a posteriori* distribution can be maximised by computing conjugate gradients for a given segmentation estimate. The distribution is iteratively optimized until convergence, where classes correspond to individual class probabilities. The maximum class at each pixel is taken to be its class, producing full segmentation.

4. Evaluation

Qualitative and quantitative evaluation is performed on selected representative fields. For these, two scans are performed — a fast evaluation scan and a detailed ground-truth scan (Section 4.1).

The evaluation scans were processed using the standard industrial solution — modified multispectral watershed-based segmentation and the proposed method (GDLS). Parameters of the methods were set by an expert who ensured that all relevant information is present in the resulting segmentation maps.

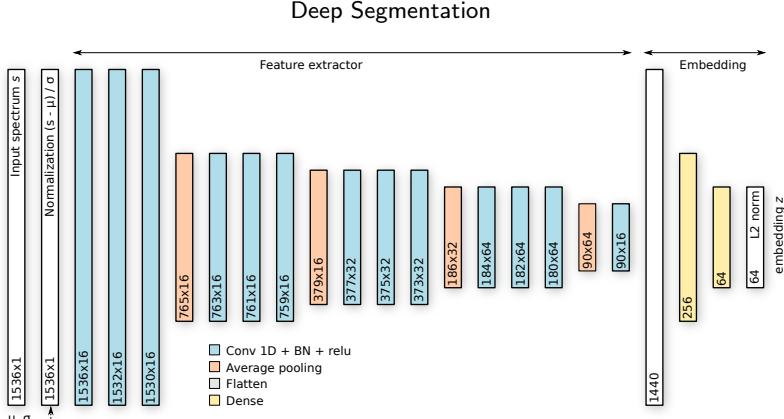


Figure 4: The neural network for mapping spectral information to latent space \mathcal{Z} . It processes the input s with normalization, feature extraction and embedding and produces the vector z .

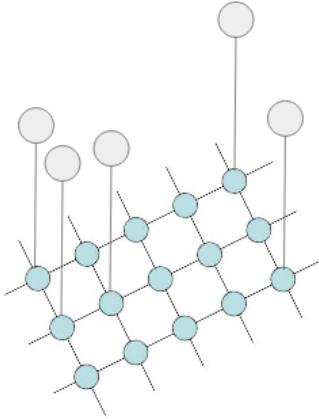


Figure 5: Dense per-pixel measurements (blue), sparse markers (gray), and relationships in a Markov Random Field

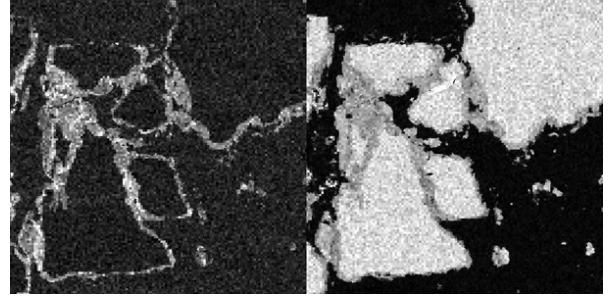


Figure 7: Element distribution maps from evaluated dot-mapping data. Left: Al-K line intensity, Right: Fe-K line intensity. Note the high level of noise and the lack of some details (e.g. the primary growth zoning).

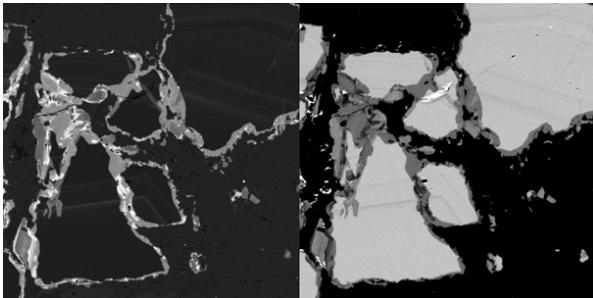


Figure 6: Low-noise, high-resolution element distribution maps from ground-truth data. Left: Al-K line intensity, Right: Fe-K line intensity. Note primary growth zoning and younger replacement zones around the edges.

We evaluated the methods using standard metrics (Section 4.2)

4.1. Data acquisition

The performance of both algorithms was evaluated on a set of 4 “Ground Truth” fields representing different types of geologically relevant samples (different kinds of ores and rocks). Since, to the best of our knowledge, there does not exist any industry standard or community-recognized bench-

mark, which could be used to evaluate the correctness of mineral identification and quantification from BSE and EDS data, we had to create the ground truth internally. The ground truth data were always acquired in High resolution mapping mode (Hrstka et al., 2018) with 10,000 – 100,000 EDS events per pixel to reduce statistical noise. The definitions of individual phases (sets of ranges for normalized EDS peak intensities and/or BSE intensity for each phase) were prepared in the TIMA SW based on inspection BSE image, elemental maps, and spectra of manually selected areas of the high-count, noise-free data. The set of individual phase definitions comprise the so-called classification scheme (Hrstka et al., 2018) The High-count noise-free data were subsequently classified pixel-by-pixel using the classification scheme. The resulting phase image represented ideal segmentation (Ground Truth). At the same time, standard data with 1,000 X-ray events per pixel in high-resolution mapping were acquired from the same area and were used to compare the performance of both algorithms. For 12 fields, 1,000 X-ray events per pixel data were also acquired using the Dot mapping mode (Hrstka et al., 2018) with a 3 times coarser EDS grid compared to BSE. The following microscope settings were used: accelerating voltage 25 kV, beam current 10 nA, working distance 15 mm. The acceleration voltage of 25 kV was chosen in order to properly excite all relevant X-ray lines including e.g. Mo-K, Cu-K, Fe-K to be able to properly iden-

Field	Ground Truth				Evaluation			
	BSE [μm]	EDS [μm]	X-ray events	Time [s]	BSE [μm]	EDS [μm]	X-ray events	Time [s]
A	3	3	20,000	8,577	3	9	1,000	55
B	3	3	20,000	7,233	3	9	1,000	49
C	1	1	20,000	20,269	2	6	1,000	42
D	10	10	100,000	2,074	10	10	1,000	26

Table 1

Overview of acquisition parameters and acquisition time for individual fields evaluated in table 2, where μm distances denote measurement spacing for each configuration. The X-ray events field specify number of X-ray events that were recorded in each measurement point and the Time field denotes the total acquisition time per field.

tify the minerals from their EDS spectra. Measurements performed on the four fields used are described in table 1, and the difference in accuracy of EDS measurements can be seen in figures 6 and 7.

4.2. Evaluation metrics

We use standard metrics for superpixel segmentation evaluation (Stutz et al., 2018). We modify the measures to take "invalid" pixels into account. Ground truth pixels can be invalid because even an expert on high-count data could not decide to which segment the pixels belong, or they were *a priori* marked as background by the device. The pixels in the segmented image can be invalid for two reasons — 1) the instrument marked them as background or 2) the method could not mark them. We do not want these pixels to participate in the evaluation. Let $S = \{S_j\}_1^K$ and $G = \{G_i\}$ be partitions of the same image where S represents the segmentation and G is ground truth segmentation.

Boundary precision (P) and recall (R) - measures adherence of the method to the edges and reflects false positive and false negative edges. Let $TP(G, S)$, $FN(G, S)$ and $FP(G, S)$ be a number of *true positive*, *false positive* and *false negative* edges. The P and R are defined as

$$R(G, S) = \frac{TP(G, S)}{TP(G, S) + FN(G, S)} \quad (3)$$

$$P(G, S) = \frac{TP(G, S)}{TP(G, S) + FP(G, S)} \quad (4)$$

In other words, the value of Precision reflects the ratio of real edges in the detected edges, and the value of Recall reflects the ratio of the real edges that were detected. Higher values means better adherence to edges, and consequently better segmentation.

Undersegmentation error (UE) - Measures leakage of detected segments with respect to G . It is defined as

$$UE(G, S) = \frac{1}{N} \sum_{G_i} \sum_{S_j \cap G_i \neq \emptyset} \min \{|S_j \cap G_i|, |S_j - G_i|\} \quad (5)$$

Achievable segmentation accuracy (ASA) - Measures the ratio of area that falls to the segment with highest in-

tersection.

$$ASA(G, S) = \frac{1}{N} \sum_{S_j} \max \{|S_j \cap G_i|\} \quad (6)$$

4.3. Quantitative Evaluation

Evaluation results of the 4 fields are summarized in Table 2. Figure 8 show visualizations of segmentation for fields C and D along with metrics for different parameters settings.

The new method achieves better *average segmentation accuracy* and *undersegmentation* which means that the segments better correspond to the structures in the fields. This is also supported by the values of *precision* and *recall* metrics (as visible in Figure 8). Watershed-based segmentation can achieve similar results for the price of oversegmentation. GDLS generates lower number of segments which are more compact. The exception is the field D which was scanned in the high resolution mapping mode, and contains a large number of single-pixel segments which is impossible to achieve with watershed algorithm since it makes 1 pixel borders around segments. Despite the higher number of segments on field D, GDLS better captures homogeneous areas in the field.

4.4. Visual Evaluation

Field C (see Figure 8) consists of grains of andradite garnet ($\text{Ca}_3\text{Fe}_2(\text{SiO}_4)_3$ — red color) containing chemically distinct zones enriched in grossularite component ($\text{Ca}_3\text{Al}_2(\text{SiO}_4)_3$ — dark red color), mostly along their rims. According to EDS quantitative analysis, the composition difference of the dark red and red zones is about 3–4 wt. % of Al and Fe. The garnet grains are enclosed in wollastonite — yellow color). From the High-count Ground Truth data (Figure 6), it is seen that the grossularite-enriched zones (darker red) are younger, and they cross-cut the primary oscillatory growth zoning of the andradite crystals.

The watershed and GDLS were compared on a dot mapping dataset acquired with lower resolution and lower number of X-ray events, summarized in Table 1, therefore, lots of features and small details, especially on the elemental maps (e.g. the brightest areas on the Al distribution map), are not visible due to lower resolution and noise (compare maps in figures 6 and 7). The criterion for setting the parameters was proper recognition of the grossularite-rich zones (light

Deep Segmentation

Field	Watershed					GDLS				
	ASA	UE	P	R	K	ASA	UE	P	R	K
A	0.97	0.059	0.61	0.97	4,080	0.98	0.047	0.81	0.98	3,427
B	0.89	0.22	0.8	0.95	12,036	0.89	0.22	0.85	0.93	8,247
C	0.93	0.13	0.3	0.99	21,822	0.93	0.15	0.69	0.83	5,102
D	0.89	0.22	0.44	1.0	2,829	0.94	0.12	0.52	1.0	5,093

Table 2

Metrics on the four testing fields evaluated for segmentation with parameters set by an expert. On average, the proposed method gives better results with lower number of segments.

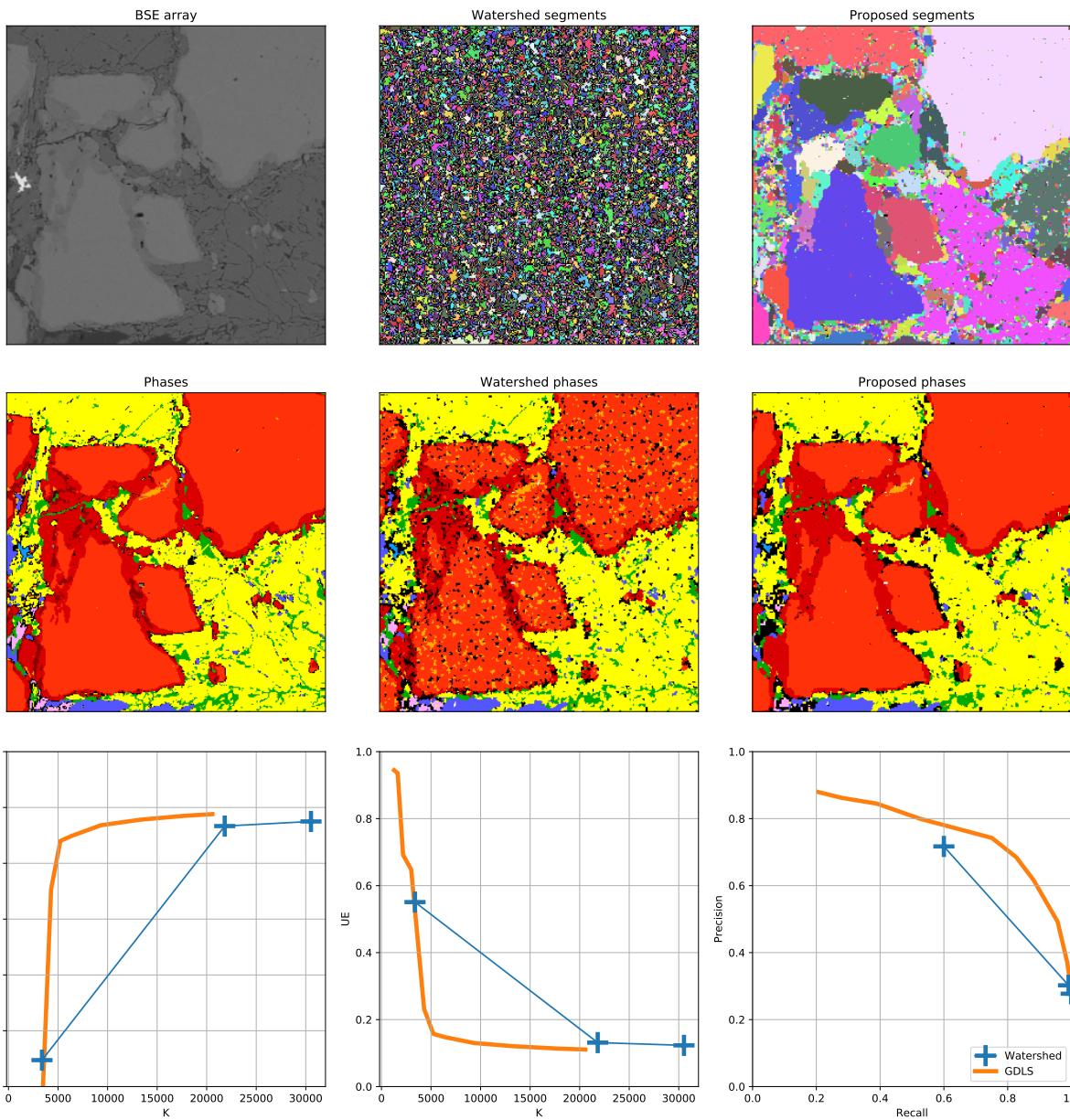
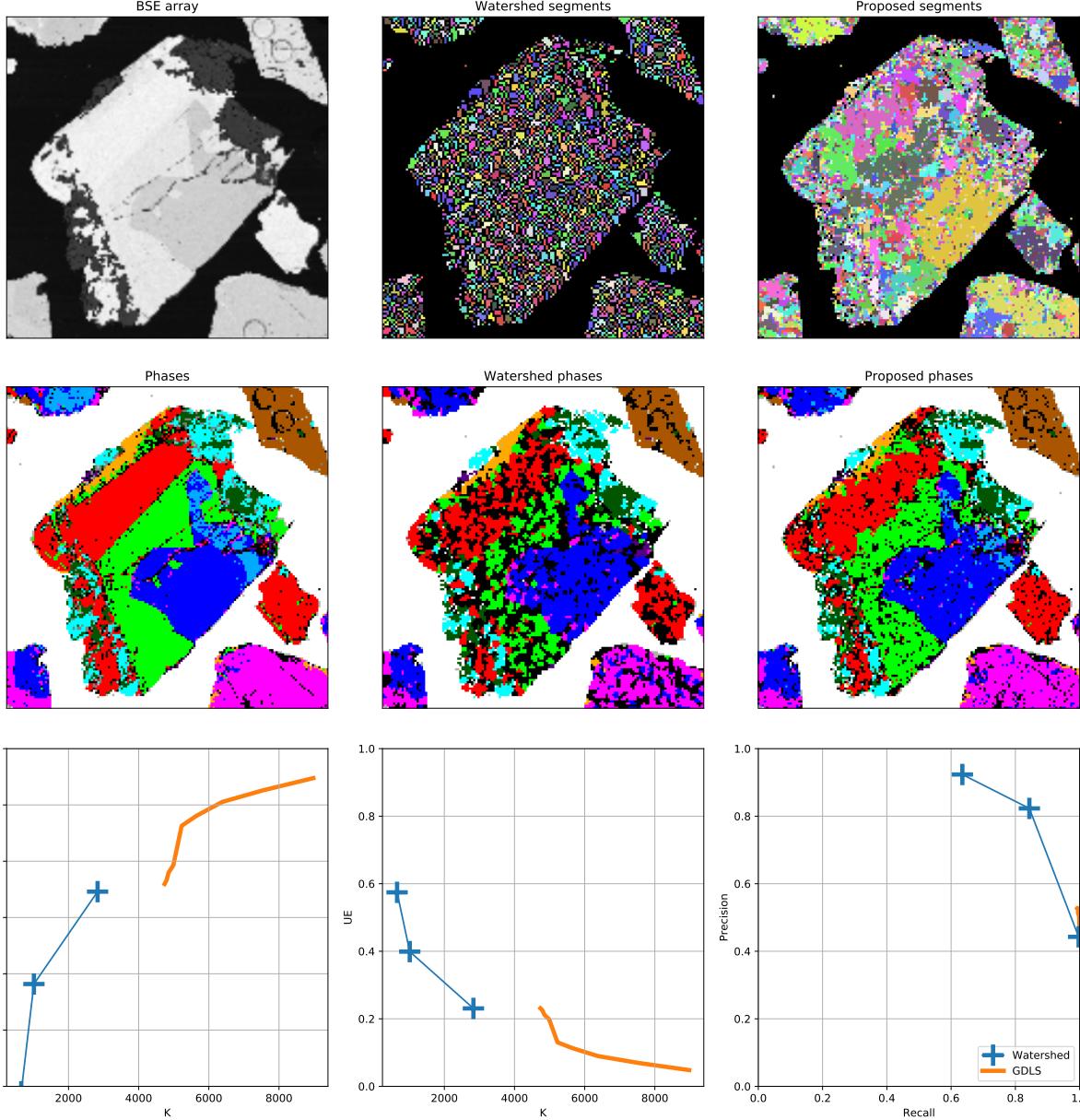


Figure 8: Example of segmentation and evaluated metrics of field C. For each field, first two rows (from left to right) show the BSE array of the field and ground truth phases, segments and phases obtained from Watershed-based segmentation, and segments and phases obtained by GDLS. The last row show the values of metrics for the different settings of watershed (blue) and GDLS (orange). The field contains andradite garnet (red) with Al-enriched, Fe-depleted (grossularite substitution) rims (dark red). The GDLS is able to distinguish these rims without significant oversegmentation and produces much clearer final phase map with less artifacts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Deep Segmentation



44
45
46
47
48
49
50
Figure 9: Example of segmentation and evaluated metrics of field D. Subfigures are divided as in Figure 8. The field contains
51 minerals of columbite-(Mn)-tantalite-(Mn) solid solution with distinct domains with different Nb/Ta ratio (red, green, blue and pink domains). Although GDLS produces more segments than watershed in this case, the final GDLS phase map contains less
52 artifacts because most of the image area consist of comparatively larger segments.

53
54
55
56
57
58
59
on the Al and darker on the Fe elemental map, respectively).
For watershed algorithm, separation of phases 18 had to be
chosen, resulting in serious oversegmentation (figure 8).

60
61
62
63
64
65
In the case of GDLS, the grossularite-rich zones were
recognized properly without any need for decreasing the thresh-
old so much, so the segments are large, and their boundaries
follow real phase boundaries. Unlike watershed, GDLS pro-
duces a much clearer phase map after classification, see Fig-
ure 8.

The field D consists of grains of Nb-Ta oxides (columbite-
(Mn) to tantalite-(Mn)) containing zones with different Nb/Ta
ratio (red, green, blue and pink colours on Figure 9). In this

case, both the Ground truth and the evaluation data were ac-
quired with dense EDS array (see table 1). The criterion for
parameter settings was proper recognition of red and green
zone, which differ by about 4 wt. % of Ta substituting for Nb.
For watershed algorithm, the threshold needs to be lowered
until it produces random small segments without any relation
to phase boundaries (Figure 9) resulting in lot of artifacts in
the final phase map. Compared to the dot mapping measure-
ments, the GDLS tend to produce more segments, especially
on phase boundaries and surface inhomogeneities. Despite
that, the GDLS segments most of the image with much larger
segment size and produces a phase map with significantly

Deep Segmentation

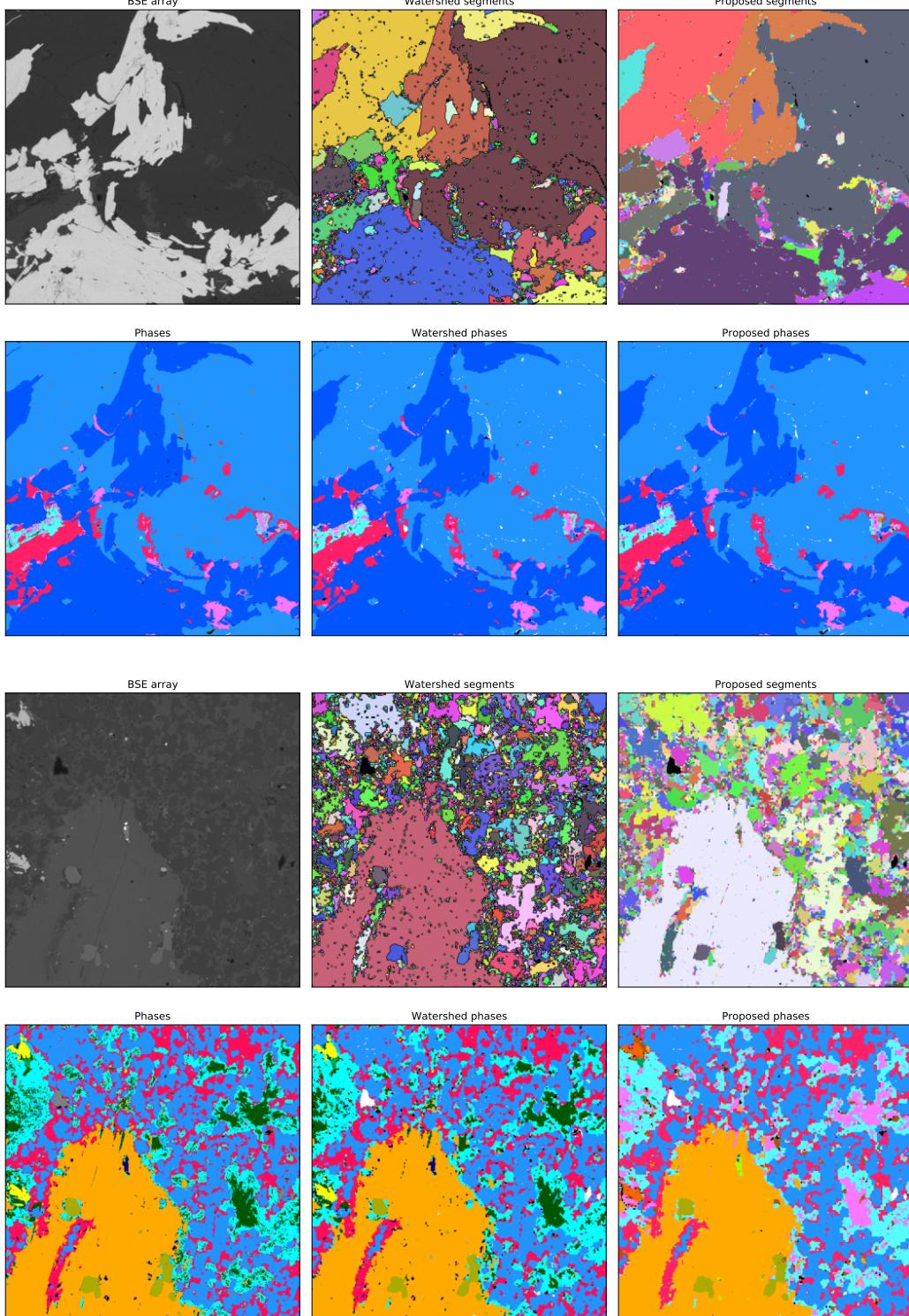


Figure 10: Comparison of the new method to Watershed of fields A and B. The evaluation metrics can be found in Table 2.

less artifacts (Figure 9). Therefore, the GDLS surpasses the capabilities of watershed also on the dense EDS data array in practical evaluation.

Figs 10 and 11 show fields A and B, respectively, where the GDLS performs comparably with watershed-based seg-

mentation from the visual point of view.

4.5. Discussion

As can be gathered from the review of existing methods (Section 2), all existing commercial solutions rely heavily on

BSE images to separate mineral grains. EDS data are used to classify phases pixel by pixel. In such cases, it is not possible to reliably distinguish mineral phases/individual zones in zoned mineral grains, which have similar but distinct composition. Depending on the actual composition differences, one needs to distinguish, it is necessary to acquire 1–2 orders of magnitude more X-ray events from each pixel (table 1) to correctly label all pixels and reduce the occurrence of artifacts (misclassified or unclassified pixels, which are classified incorrectly or not classified at all, as they do not fulfill the classification rules for any class due to statistical noise arising due to low number of X-ray events in the spectrum). But the 1–2 orders of magnitude more X-ray events per pixel means 1–2 orders slower data acquisition. Unlike the other solutions, the Mineralogic mining uses results of quantitative analysis of the EDS spectra (calculated wt. % of individual elements) and not the normalized peak intensities to classify the individual spectra (Keulen et al., 2020). However, this approach inherently suffers from the same limitations, because if the ratios of the peak intensities on individual pixels representing the same material are different due to statistical noise, inevitably, the calculated chemical composition from that spectra must be different as well. Applying any fine classification rules on such data to correctly label chemically similar areas and subsequently extract their morphological characteristics would not work.

TIMA is able to overcome the issues described above by producing cumulative spectra of segments and applying the classification rules on the cumulative spectra but not on every pixel, effectively eliminating statistical noise. On high-count cumulative spectra, it is possible to distinguish fine differences in chemical composition without misclassification artifacts. On the other hand, for this approach to work, it is essential that the segment boundaries follow real boundaries of the chemically coherent areas and that distinct phases are not merged together. In such cases, the cumulative spectrum does not represent any of the merged areas and gives the wrong result. For segmentation to avoid merging real phases together, one can lower the threshold (both on watershed and GDLS), which typically leads to oversegmentation, and therefore fewer X-ray events in the cumulative spectra and the same problems described above.

It is seen from images and graphs in Figure 8, that up to a certain point, both methods produce comparable results, but once the threshold is lowered, GDLS will maintain large segments much longer, achieving much better overall correct classification and a clearer phase map.

The overall performance of the watershed-base segmentation and GDLS was evaluated on a desktop PC (hardware: Intel® Core™ i7-4790 @ 3.60GHz, 16 GB RAM and NVIDIA GeForce GTX 1660 SUPER, software: Windows 10 64-bit, Python 3.7.1 and TensorFlow 2.4.1). The duration of analysis for the test cases A, B, C and D (see table 1) are summarized in the table 3. The timing of GDLS was measured with GPU support enabled and disabled. In the current state, GDLS is about one order of magnitude slower than the watershed-based segmentation (WSS). The optimum implementation

of the inference phase was not the goal of this study, there is a room for optimization in future.

One of the still unresolved weak points of both the state-of-the-art and the GDLS algorithms is the behavior on the phase boundaries, where the spectra are mixed due to the overlapping EDS excitation volumes. Since these spectra does not correspond to any material, they are recognized as separate segment and after classification, they are left as unclassified since they do not fulfill classification rules for any material. Such cases need to be specifically approached in later steps e.g. by adjusting the classification scheme or performing some kind of automated classification based on similarity with neighboring segments during data post-processing.

5. Conclusion

The goal of the study we present in this paper was to develop novel graph-based segmentation of BSE and EDS hyperspectral images using deep learning. We were able to go beyond the state of the art in terms of the quality of mineralogy analysis with the same input data compared to previous approaches. We have also demonstrated that the proposed approach is readily acceptable for mineralogy experts and improves the quality of their outputs.

We evaluated the output both numerically and visually with convincing results. The new method yields superior results especially when natural solid solutions (e.g. Nb-Ta oxides, garnets etc.) containing domains with similar chemistry are present in the sample, a situation that is very common. In such cases, the new method provides larger average segment size while still correctly recognizing the boundaries of chemically similar domains. Classification of the cumulative spectra from larger segments reduces the probability of erroneous classification resulting from low X-ray events in the spectrum, which manifests in fewer artifacts in the final phase map. The new method enables better chemical/mineralogical resolution of similar phases while retaining the same speed of analysis.

Future work will focus on improvements of spectra simulation for NN training, improvements of the NN performing the mapping to energy regions, element weight concentrations or stoichiometry, with custom weights for individual elements and noise suppression. Denoising the BSE image, element map reconstruction from sparsely sampled EDS data, and better handling of mixed spectra on phase boundaries would also be beneficial.

6. Computer Code Availability

The algorithm described in this work has been tested using Python and the Tensorflow framework. Visualizations are prepared with Matlab. The code necessary to replicate these results is freely available at

<https://github.com/RomanJuranek/graph-segmentation>

7. Acknowledgements

to be completed after review

Field	Field size	EDS points	WSS CPU only	GDLS	
				CPU only	GPU support
A	500 × 500	31763	1.1	23.0	12.2
B	500 × 500	29917	1.1	23.4	13.2
C	400 × 400	21717	0.7	14.2	6.8
D	150 × 150	13772	0.2	7.5	2.7

Table 3

Comparison of timing of the segmentation methods on the testing fields A, B, C and D. *Field size* = size of field (width × height), in pixels; *EDS points* = number of EDS measurement points; *WSS* = duration of the watershed-based segmentation, in seconds, *GDLS* = duration of GDLS, in seconds.

References

- Ayling, B., Rose, P., Zemach, E., Drakos, P., Petty, S., 2012. QEM-SCAN (Quantitative Evaluation of Minerals by Scanning Electron Microscopy): capability and application to fracture characterization in geothermal systems, in: AGU Fall Meeting Abstracts, pp. 1158–.
- Bowyer, A., 1981. Computing dirichlet tessellations. *The computer journal* 24, 162–166.
- Breiter, K., Ďurišová, J., Hrstka, T., Korbelová, Z., Hložková Vaňková, M., Vašinová Galiová, M., Kanický, V., Rambousek, P., Kněsl, I., Dobeš, P., Dosbaba, M., 2017. Assessment of magmatic vs. metasomatic processes in rare-metal granites: A case study of the Cínovec/Zinnwald Sn–W–Li deposit, Central Europe. *Lithos* 292–293, 198–217. doi:10.1016/j.lithos.2017.08.015.
- Di Zenzo, S., 1986. A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing* 33, 116 – 125. doi:10.1016/0734-189X(86)90223-9.
- Drewniok, C., 1994. Multi-spectral edge detection. some experiments on data from Landsat-TM. *International Journal of Remote Sensing* 15, 3743–3765.
- Fandrich, R., Gu, Y., Burrows, D., Moeller, K., 2007. Modern SEM-based mineral liberation analysis. *International Journal of Mineral Processing* 84, 310–320.
- Gloy, G., 2016. High speed differentiation of mineral phases of similar BSE intensity with AMICS. https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/X-rayDiffraction_ElementalAnalysis/Microanalysis_EBSD/LabReports/App_eds_16_amics_mineral_phases_lores.pdf. Accessed: 2020-07-02.
- Goldstein, J., Newbury, D., Michael, J., Ritchie, N., Scott, J., Joy, D., 2018. *Scanning Electron Microscopy and X-ray Microanalysis*. doi:10.1007/978-1-4939-6676-9.
- Gottlieb, P., Wilkie, G., Sutherland, D., Ho-Tun, E., Suthers, S., Perera, K., Jenkins, B., Spencer, S., Butcher, A., Rayner, J., 2000. Using quantitative electron microscopy for process mineralogy applications. *Jom* 52, 24–25.
- Grant, G., Reid, A.F., 1981. An efficient algorithm for boundary tracing and feature extraction. *Computer Graphics and Image Processing* 17, 225–237. doi:10.1016/0146-664X(81)90003-4.
- Gray, R., Neuhoff, D., 1998. Quantization. *IEEE Transactions on Information Theory* 44, 2325–2383. doi:10.1109/18.720541.
- Haluzová, E., Ackerman, L., Pašava, J., Jonášová, v., Svojtka, M., Hrstka, T., Veselovský, F., 2015. Geochronology and characteristics of Ni-Cu-(PGE) mineralization at Rožany, Lusatian granitoid complex, Czech Republic. *Journal of Geosciences* 60, 219–236.
- Hrstka, T., Gottlieb, P., Skála, R., Breiter, K., Motl, D., 2018. Automated mineralogy and petrology-applications of TESCAN Integrated Mineral Analyzer (TIMA). *Journal of Geosciences* 63, 47–63.
- Jasiewicz, J., Stepinski, T., Niesterowicz, J., 2018. Multi-scale segmentation algorithm for pattern-based partitioning of large categorical rasters. *Computers Geosciences* 118, 122 – 130. URL: <http://www.sciencedirect.com/science/article/pii/S0098300417305204>, doi:<https://doi.org/10.1016/j.cageo.2018.06.003>.
- Karimpouli, S., Tahmasebi, P., 2019. Segmentation of digital rock images using deep convolutional autoencoder networks. *Computers Geosciences* 126, 142 – 150. URL: <http://www.sciencedirect.com/science/article/pii/S0098300418303911>, doi:<https://doi.org/10.1016/j.cageo.2019.02.003>.
- Karvelis, P.S., Tzallas, A.T., Fotiadis, D.I., Georgiou, I., 2008. A multichannel watershed-based segmentation method for multispectral chromosome classification. *IEEE Transactions on Medical Imaging* 27, 697–708.
- Keulen, N., Malkki, S.N., Graham, S., 2020. Automated quantitative mineralogy applied to metamorphic rocks. *Minerals* 10, 47.
- Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- Li, S.Z., 1994. Markov random field models in computer vision, in: European conference on computer vision, Springer. pp. 361–370.
- Miller, P.R., Zuiderwyk, M.A., Reid, A.F., 1983. QEM*SEM image analysis in the determination of modal assays, mineral associations and mineral liberation, in: XIV International Mineral Processing Congress, pp. 1–20.
- Motl, D., Filip, V., 2013. Method and apparatus for material analysis by a focused electron beam using characteristic X-rays and back-scattered electrons. US Patent App. 13/398,114.
- Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines, in: Proceedings of the 27th International Conference on International Conference on Machine Learning, Omnipress, Madison, WI, USA. p. 807–814.
- Reed, S.J.B., 2005. Electron microprobe analysis and scanning electron microscopy in geology. Cambridge university press.
- Reid, A., Gottlieb, P., Macdonald, K., Miller, P., 1985. QEM*SEM image analysis of ore minerals: volume fraction, liberation, and observational variances. *Liberation and Observational Variances. Applied Mineralogy*, 191–204.
- Ritchie, N.W., 2009. Spectrum simulation in DTSA-II. *Microscopy and Microanalysis* 15, 454–468. doi:10.1017/S1431927609990407.
- Sandmann, D., 2015. Method development in automated mineralogy. Ph.D. thesis. Technische Universität Bergakademie Freiberg, Deutschland.
- Sandmann, D., Gutzmer, J., 2015. Nature and distribution of PGE mineralisation in gabbroic rocks of the Lusatian Block, Saxony, Germany Art und verbreitung der PGE-Mineralisation in den gabbroiden Gesteinen des Lausitzer Blocks, Sachsen, Deutschland. *Zeitschrift der Deutschen Gesellschaft für Geowissenschaften* 166, 35–53.
- Schroff, F., Kalenichenko, D., Philbin, J., 2015. FaceNet: A unified embedding for face recognition and clustering, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE. pp. 815–823. doi:10.1109/cvpr.2015.7298682.
- Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. [arXiv:1409.1556](https://arxiv.org/abs/1409.1556).
- Stutz, D., Hermans, A., Leibe, B., 2018. Superpixels: An evaluation of the state-of-the-art. *Computer Vision and Image Understanding* 166, 1–27.
- Sutherland, D.N., Gottlieb, P., 1991. Application of automated quantitative mineralogy in mineral processing. *Minerals Engineering* 4, 753–762.
- Výravský, J., Hreus, S., Cempírek, J., Vašinová Galiová, M., 2019. Assessment of scandium deportment and recovery potential in Li-Sn-W greisen deposit Cínovec/Zinnwald using EPMA, LA-ICP-MS and TIMA automated mineralogy, in: Proceedings of the 15th SGA Biennial Meeting,

27-30 August 2019, Glasgow, Scotland, pp. 1638–1641.
Xu, C., Kynický, J., Song, W., Tao, R., Lü, Z., Li, Y., Yang, Y., Pohanka,
M., Vašinová Galiová, M., Zhang, L., et al., 2018. Cold deep subduction
recorded by remnants of a Paleoproterozoic carbonated slab. *Nature
communications* 9, 1–8.