

ORIGINAL PAPER

Open Access



# Direct mineral content prediction from drill core images via transfer learning

Romana Boiger<sup>1\*</sup> , Sergey V. Churakov<sup>1,2</sup>, Ignacio Ballester Llagaria<sup>1,3</sup>, Georg Kosakowski<sup>1</sup>, Raphael Wüst<sup>4,5</sup> and Nikolaos I. Prasianakis<sup>1</sup>

## Abstract

Deep subsurface exploration is important for mining, oil and gas industries, as well as in the assessment of geological units for the disposal of chemical or nuclear waste, or the viability of geothermal energy systems. Typically, detailed examinations of subsurface formations or units are performed on cuttings or core materials extracted during drilling campaigns, as well as on geophysical borehole data, which provide detailed information about the petrophysical properties of the rocks. Depending on the volume of rock samples and the analytical program, the laboratory analysis and diagnostics can be very time-consuming. This study investigates the potential of utilizing machine learning, specifically convolutional neural networks (CNN), to assess the lithology and mineral content solely from analysis of drill core images, aiming to support and expedite the subsurface geological exploration. The paper outlines a comprehensive methodology, encompassing data preprocessing, machine learning methods, and transfer learning techniques. The outcome reveals a remarkable 96.7% accuracy in the classification of drill core segments into distinct formation classes. Furthermore, a CNN model was trained for the evaluation of mineral content using a learning data set from multidimensional log analysis data (silicate, total clay, carbonate). When benchmarked against laboratory XRD measurements on samples from the cores, both the advanced multidimensional log analysis model and the neural network approach developed here provide equally good performance. This work demonstrates that deep learning and particularly transfer learning can support extracting petrophysical properties, including mineral content and formation classification, from drill core images, thus offering a road map for enhancing model performance and data set quality in image-based analysis of drill cores.

**Keywords** Convolutional neural networks, Transfer learning, Core analysis, Lithology, Mineral composition

## 1 Motivation and background

Geological exploration of the underground is important for the mining, mineral, oil, and gas industries. In Switzerland, geological exploration of the underground for the deep geological disposal of radioactive waste is currently in progress and includes remote sensing (seismic data analysis), geophysical surveys (log analysis) and drill core laboratory analysis. Although the combination of these techniques provides precise and reliable results, laboratory investigations are often labor-intensive, time-consuming, and costly when large sample numbers and volumes are present. Combining conventional field and laboratory analytical techniques with machine learning may help enhance data analysis and provide a deeper

Editorial handling: György Hetényi.

\*Correspondence:

Romana Boiger  
romana.boiger@psi.ch

<sup>1</sup> Laboratory for Waste Management, Paul Scherrer Institute, Forschungsstrasse 111, Villigen PSI 5232, Switzerland

<sup>2</sup> Institute of Geological Sciences, University of Bern, Baltzerstrasse 1+3, Bern 3012, Switzerland

<sup>3</sup> ETH Zürich, Rämistrasse 101, Zürich 8092, Switzerland

<sup>4</sup> Nagra, Hardstrasse 73, Wettingen 5430, Switzerland

<sup>5</sup> Earth and Environmental Science, James Cook University, Townsville 4811, Australia



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

insight into data (Jooshaki et al., 2021; Jung and Choi, 2021; Woodhead and Landry, 2021).

In this paper, we demonstrate the capacity of machine learning respectively, deep learning, particularly emphasizing transfer learning to extract selected petrophysical properties, like mineral content or sample formation from drill core images.

Machine learning has been successfully used in the past to classify the lithology based on drill core images: Many of these studies applied convolutional neural networks (CNN), a special type of neural network that has proved to be successful for handling images. Particularly pre-trained CNN architectures such as VGG16, DenseNet, ResNet, ResNeSt, or ResNext applied to drill core samples from Norway, South Australia, Gulf of Mexico & North Sea, China and Switzerland achieved good results for lithology classification with accuracies for the test data sets ranging from 60% up to 99.6% (Alguliyev et al., 2023; Alzubaidi et al., 2021; Falivene et al., 2022; Fu et al., 2022; Lauper et al., 2021). Furthermore, in this context, autoencoders (Park and Jeong, 2023; Solum et al., 2022) and vision transformer architectures (Koeshidayatullah et al., 2022) have also been applied for this task. These techniques were utilized for several thousand image slices from drill cores from Western Australia, Gulf of Mexico & North Sea and Russia. The testing accuracy varied from 70% up to 96.4%, and showed a strong dependency on the heterogeneity of the images and the number of classes that the samples were categorized.

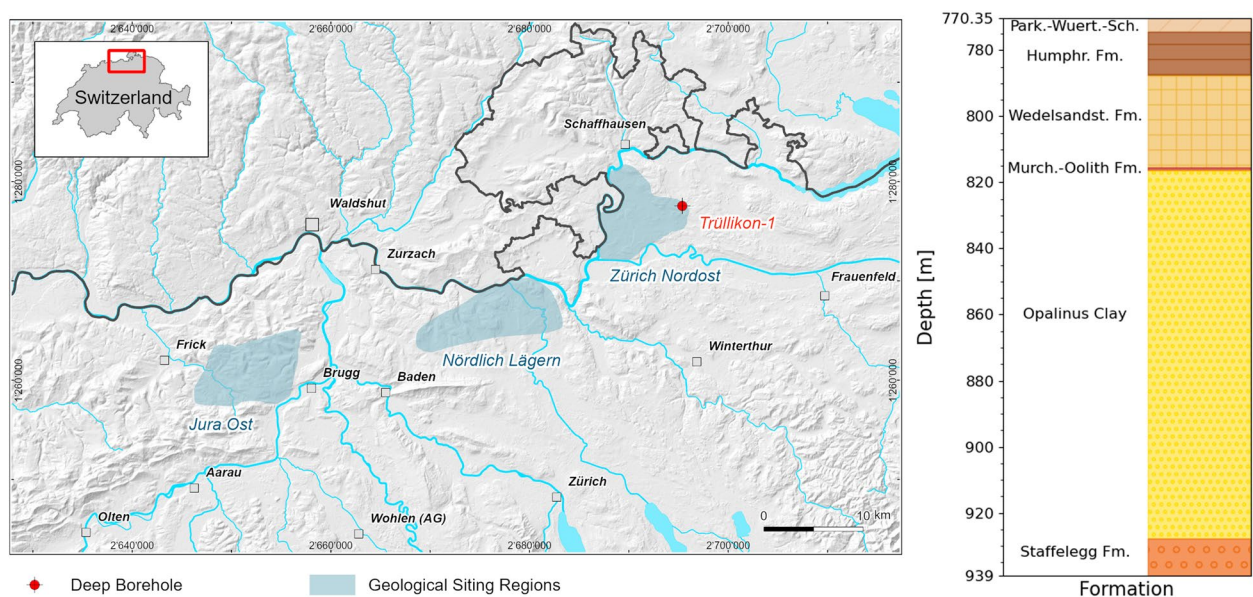
The analysis of the model performances reveals that the sample preparation and data preprocessing are paramount steps when applying machine learning methods. The necessary steps include the automatic detection of trays and cores (Baraboshkin et al., 2022; Gunther et al., 2021), the assessment of rock quality, the classification of intact and non-intact cores, as well as the recognition of empty tray areas and non-rock objects (Alzubaidi et al., 2022; Li et al., 2023). Depending on the application, for preprocessing the automated crack detection (Alzubaidi et al., 2022) can also be important. Moreover, the classification of lithology is done based on a narrow core interval and not all cores have the same length, thus the preprocessing step includes slicing the available data set into smaller images, ranging from 0.5x0.5cm up to 10x10cm, as used in the aforementioned works. Besides full drill core images, thin section images (Faria et al., 2022; Zhou et al., 2023), rock sample images (Shi et al., 2023), or rock debris (Xu et al., 2022) images were used to perform rock type classification. Some approaches to predict lithology from drill core images involve additional data to increase accuracy, like elemental information (Xu et al. 2021, 2022), petrophysical and geochemical data

(Houshmand et al. 2022), or pXRF measurements (Trott et al. 2022), and the use of machine learning models designed for data fusion.

Assessing the mineral content has so far only been done with different types of data, not drill core images alone. In Tuşa et al. (2020), mineral abundance predictions relied on a combination of hyperspectral short-wave infrared data and, for small areas, additional Scanning Electron Microscopy-based images with different machine learning approaches: Random Forest (RF), Support Vector Machine (SVM), and Neural Networks (NN). A model based on RF was developed in Barker et al. (2021) that estimated the mineral proportion from Long-wave infrared spectra. For labeling the training data, micro-X-ray fluorescence measurements were used. Spectral and geochemical data were utilized to train a CNN model that predicts the Cu concentration in Guerra Prado et al. (2023). In Krupnik and Khan (2020); Kupssinskü et al. (2022), hyperspectral data combined with different machine learning approaches like SVM, NN, and Spectral Angle Mapper were investigated for mineral mapping and porosity estimation. Also, spectral data, in terms of multi-sensor spectral imaging, together with SVM, were used to distinguish between six mineralogically meaningful classes, and the corresponding probability estimates of each class were derived in Lorenz et al. (2019). In the study discussed in Mishra et al. (2022), core plug samples were combined with continuous Kimeleon colorlith logs, which use information from the apparent matrix density, neutron porosity, and gamma-ray logs. K-means clustering was then used to classify the different rock types. From these continuous rock types, rock properties like permeability were calculated and up-scaled.

The goal of this work is to train neural networks that take only drill core images as input and can assess the lithological classes and the mineral content, respectively. We first perform a simplified lithological classification of drill core images from Northern Switzerland into distinct geological formations using pretrained CNNs designed for image classification. Inspired by the recent progress in automated image processing, we analyze the core images stemming from deep drilling exploration of the Mesozoic underground conducted by Nagra, the Swiss National Cooperative for the Disposal of Radioactive Waste (<https://nagra.ch>), in the context of the national program on site selection for Swiss deep geological repository for radioactive waste (Nagra, 2023).

The second step and principle objective of this research is to assess the performance of neural networks in predicting the mineral content (amount of carbonate, silicate and total clay) from solely the drill core images. Because none of those studies mentioned above considers the



**Fig. 1** Left: Cartographic representation illustrating the geographical region of northern Switzerland. The map specifically features the borehole site of Trüllikon 1–1 which is selected for the data analysis of this study. Background: ©Data:swisstopo and hillshading from NASA SRTM. Right: Abstracted lithological profile of the section Trüllikon 1–1 showing the interval between 770.35 m and 939 m depth. The blue bars denote the available data segments

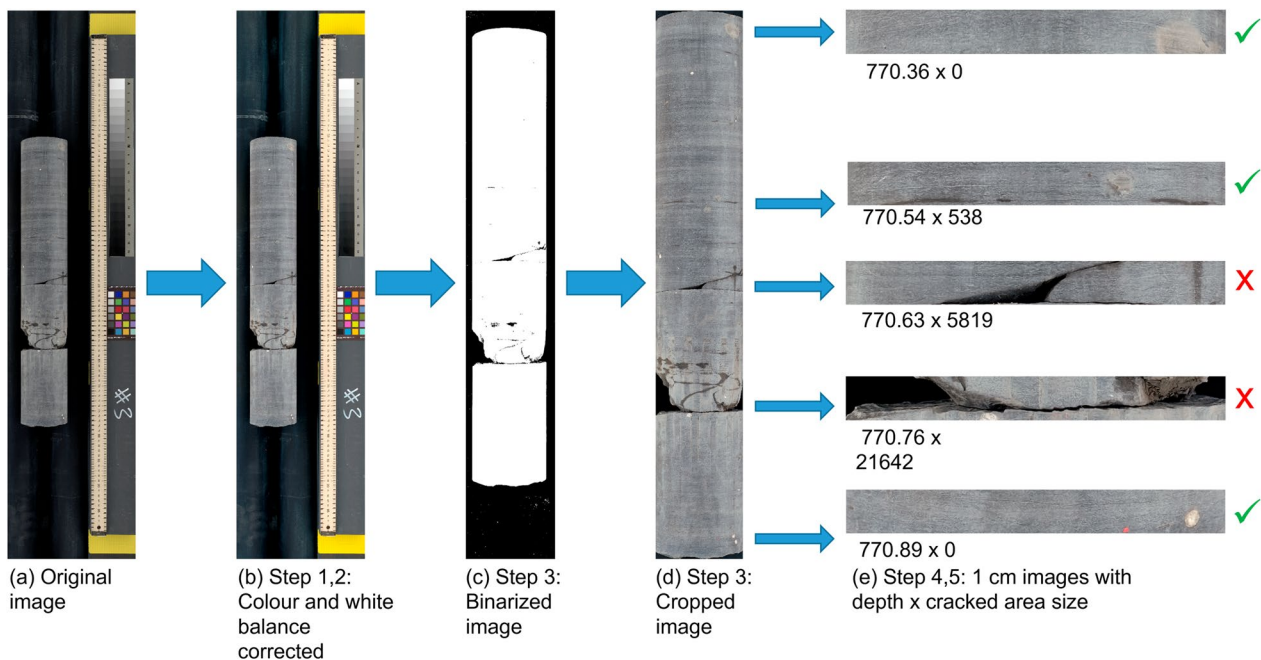
drill core images alone to predict the mineral content. For this purpose, parts of the classification model are used via transfer learning. Thereby, even with a relatively limited dataset (in our case 361 data points) the mineral content regression from drill core images is possible. For training the CNN, the images are labeled with mineral content data retrieved from a multidimensional log analysis of the petrophysical logs from one borehole. The so trained model is then applied to previously unseen drill core images of the same borehole and the predictions are compared to bulk XRD measurements of the mineral content.

The paper is structured in the following way: First, the available data set, consisting of the drill core images and additional measurements for labeling, is described in detail (Sect. 2). Section 3 introduces the methodology, covering data preprocessing, machine learning methods, and transfer learning. Results for the formation classification and mineral content regression are presented in Sect. 4 as well as details on the comparison of the regression model predictions to measured mineral contents. This is followed by final remarks and outlook for further research investigations in Sect. 5. This study not only advances geological analysis but also underscores the potential of machine learning to enhance subsurface exploration and characterization.

## 2 Data set

The samples and data used in this study are from the borehole Trüllikon 1–1 in the siting region of Zürich Nordost (left Fig. 1) in Switzerland, (<https://nagra.ch>). The borehole is part of Nagra’s national program on site selection for Swiss deep geological repository for radioactive waste, (Nagra, 2023). The total drilling depth was 1310 m. The section between 498–1029 m was cored with wireline coring. At the time of the study, only a fraction of the cored section was measured, validated, and released, comprising non-continuous segments of in total 55 m of core between 770.35 m and 939 m depth of sedimentary rocks. The available data segments are visualized by the blue bars next to the lithological profile shown in Fig. 1, right. Detailed information on all drilling procedures and data can be found at Nagra Arbeitsbericht NAB 20-09 (2021); Ammen and Palten (2021).

The extracted drill cores were washed, dried and photographed using a standardized procedure together with a ruler and reference cards (X-Rite ColorChecker Classic Mini color chart and BST14 gray scale chart) for samples referencing and colour calibration. For the high-resolution (10px/mm) core photographs, a DMT CoreScan system from avaluar GmbH with a CCD-Colour camera CSc3b1.26 was used. Note that only single core images, not the 360 degree photos were considered in this work. Further details on the photographing conditions can be found in Kaehr and Gysi (2021). The photographs were



**Fig. 2** Preprocessing pipeline, consisting of five distinct steps, from the original image (a) to the final 1 cm image segments (e). The 1 cm segments with the red cross next to the image are excluded from the analysis, since the cracked area size is bigger than 5000 pixels

stored along with the corresponding drilling depth. An example of such photographs is shown in Fig. 2a.

Along with the photographs, data from the petrophysical Multiminerall Log Analysis (MultiMin) (Marnat and Becker, 2021) was used for the labeling in the regression task. The MultiMin method is a stochastic workflow for log analysis based on the assumption that every borehole log measurement is determined by the mineral and fluid content of the rock that surrounds the borehole. If the linear or non-linear relations between measured properties and rock composition are known, it is possible to calculate the theoretical log response for a given rock composition or use all available log information to estimate the rock properties of interest. By comparing measured and predicted log values, it is possible to assess the quality of the estimated properties. For details see Chapter 3.3 in Marnat and Becker (2020). Petrophysical logs available for the analysis included caliper log, gamma ray log, spectral gamma ray log for U, Th and K, neutron hydrogen index, gamma-gamma density log, element spectroscopy, electrical resistivity, and a sonic log. In addition, multi-sensor core logger data (bulk density, compressional (P) wave velocity, spectral gamma ray curves for K, Th and U, and X-ray fluorescence elemental analysis for Fe, Si, Ca, Al, Ti and S) were available for Opalinus Clay and its confining units. These logs and lab measurements of mineralogy (XRD), porosity and density were used for the MultiMin modeling of porosity and

mineral composition; details on that can be found in the Nagra working reports (Marnat and Becker, 2021, 2020).

The MultiMin model was utilized to calculate mineral composition approximately every 15 cm in terms of clay minerals (kaolinite, illite, smectite, and chlorites), other (alumino) silicates (quartz, potassic feldspars, plagioclase - this group of minerals will be referred to as silicates for the rest of this work), carbonates (calcite, siderite, dolomite, ankerite), iron oxide, evaporates (anhydrite), and organic carbon (kerogen). For the sake of simplicity and robustness of the CNN model, only the total amount of total clay, carbonate, and silicate was considered for the regression.

The core samples are grouped to six distinct formations: 1. Parkinsoni-Württembergica-Schichten (738.97–774.55 m), 2. Humphriesoolith Formation (774.55–787.50 m), 3. Wedelsandstein Formation (787.50–815.51 m), 4. Murchisonae-Oolith Formation (815.51–816.42 m), 5. Opalinus Clay (816.42–927.91 m), 6. Staffelegg Formation (927.91–971.68 m); with the numbers in brackets giving the measured core depth. A visual representation of the abstracted lithology is provided in Fig. 1, right.

For validating the trained neural network model, bulk XRD measurements were used. Within the 55 m of core, a total of 23 bulk XRD measurements were available. The exact positions within the core are detailed in Table 3, Appendix 3, and are visually accessible in the

core photograph overview in the Appendix 4 (highlighted by blue rectangles).

### 3 Methodology

#### 3.1 Image data preprocessing

To facilitate robust and automated analysis of drill core photographs, a standardization process for colour balance and image dimensions was performed, and artifacts, such as cracks or missing fragments were labeled. Accordingly, a five-step preprocessing routine was introduced for all images. First, the photographs have been corrected for color (Step 1) and white balance (Step 2). Then, an automated algorithm for core segmentation from background (Step 3), core splitting into equally sized segments (Step 4), and crack detection (Step 5) was applied.

The preprocessing was implemented using the Python Imaging Libraries PILLOW and scikit-learn, (Clark 2015; Pedregosa et al. 2011). The five distinct steps of the automatic preprocessing pipeline are described in detail below and are as follows:

1. The first step of the color correction involved creating a color profile (ICC profile) using the ColorChecker from X-rite library. The ICC profile was then subsequently applied to the images. This step ensured that the color representation across the data set was consistent throughout the images.

2. The second step implemented the correction of the white balance using the white colour patch (A) as a ground truth for true white. By comparing the color values of the white patch with the rest of the image, the pixel values of the image were normalized relative to the maximum value found on the white patch. Figure 2a shows one example from the data set. The color and white balance corrected image is depicted in Fig. 2b.

3. The third step consisted of subtracting the background and extracting the domain representing the drill core from the photograph (Fig. 2c, d). The segmentation algorithm first used the length and depth of the core piece, recorded in the name of the photograph, to approximately locate and cut the core domain from the background (Fig. 2c). This was possible, since the core was always centered in the middle of the image. Further, Otsu's method (Otsu 1979) was used to separate the core from the background based on intensity distribution. Morphological operations were then applied to refine the segmentation. Next, continuous regions that belong to the core were identified and used to create a bounding box that accurately encompasses the core. The image was subsequently cropped using this bounding box, focusing the analysis exclusively on the core region for further processing and evaluation (Fig. 2d).

4. In the fourth step of the preprocessing pipeline, the extracted core images were divided into equally sized 1 cm depth segments. The segmentation algorithm was based on the core length obtained from the image filename, resulting in images with a width of around 100 pixels and a height of 850 pixels corresponding to a 1 cm segment of the drill core with a certain depth. (Fig. 2e). If a segment overlapped two continuous core images, they were merged to form a cohesive segment. This division resulted in 5496 core segments from all available photos, out of which 361 could be used for creating the regression model to predict the mineral content from the 1 cm drill core images. The small number of images for the regression task is due to the number of corresponding data from the MultiMin analysis. The remaining core segments (i.e. 5135) were used for the classification task.

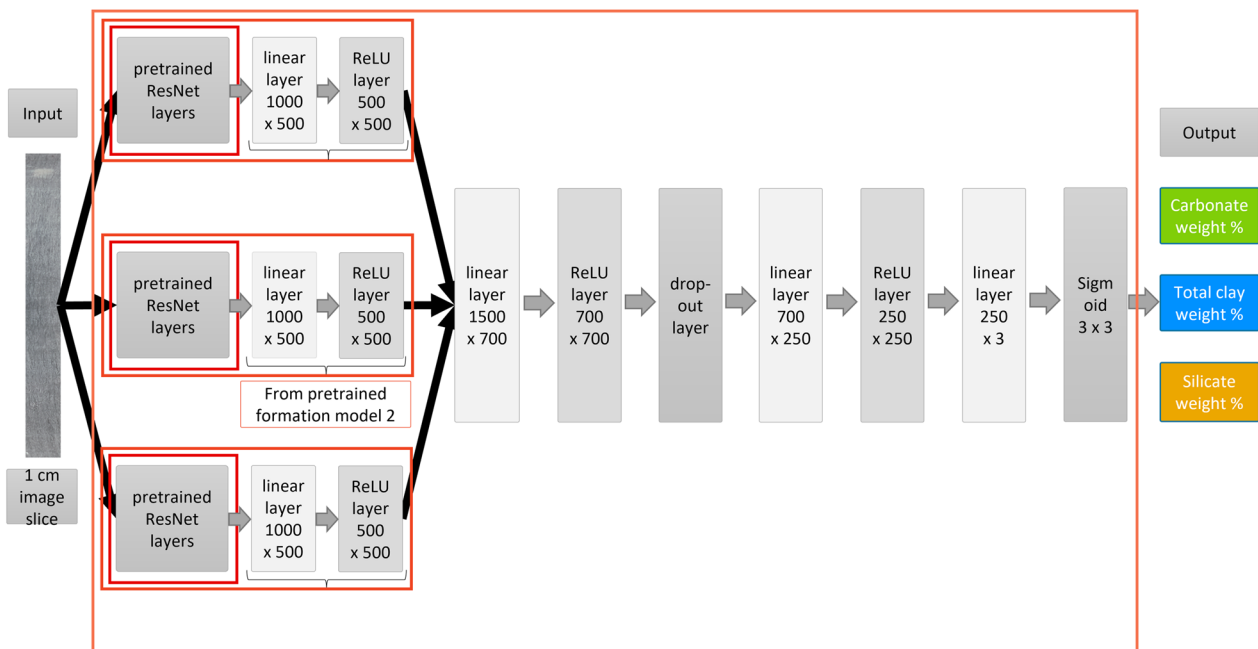
5. In the final, fifth, step, cracked segments were detected by binarizing the image again with Otsu's method. The images containing a crack were not excluded from the beginning, but the size of the cracked area per segment was stored in the filename. The segment was saved as a TIFF file with the naming convention: "DxC.tif" where D was the depth and C was the cracked area size in pixels (Fig. 2e). These images provided the input for the machine learning models. Images, where the cracked area size exceeded a certain threshold of, e.g., 5000 pixels, were removed from the data set for data quality improvements and thus better model performance.

#### 3.2 Neural networks and transfer learning

Neural networks have emerged as powerful tools for modeling a functional relationship between generic input and output. In particular, convolutional neural networks (CNN) have been proven specifically useful in image analysis, as they can autonomously learn feature engineering through the use of filters. Here, the basic principles of CNN usage are described, without providing intricate details, those can be found in Geron (2019).

Two neural network architectures were applied in this work. The first NN architecture was designed and trained to take the 1 cm drill core images as an input and predict the formation class as an output (i.e. classification task). The second NN architecture was intended for regression, and it was set up to take the 1 cm drill core images as an input, and estimate the content of clay, carbonate and silicate minerals in the core as output.

For that, the concept of transfer learning was used rather than creating and training a neural network architecture from scratch. In this process, a model trained for a similar but different task served as a backbone architecture and additional layers were placed on top of the pre-trained layers, see Fig. 3. Training such a neural network



**Fig. 3** Architecture of the neural network for the mineral content regression with three formation models as a backbone architecture

can be done in two ways, either all weights and biases are learned, or the ones of the pretrained model are fixed and only the new ones are updated.

In this work, pretrained deep convolutional neural networks (VGG16, ResNet18, ResNet34, ResNet50, ResNet101, ResNet152 (He et al., 2016) were used to generate appropriate models. The residual neural networks (ResNet) are particularly powerful because compared to very deep convolutional neural networks (like VGG) they overcome the vanishing gradient problem by using shortcuts (skip-connections). The classification and regression models using the different pretrained convolutional neural networks were implemented using pytorch (Paszke et al. 2019), a machine learning library for python and C++. The pretrained models implemented in pytorch, were originally trained on the ImageNet data set, storing their weights and biases. The ImageNet data set comprises over 14 million hand-annotated images, spanning thousands of categories, such as vehicles, animals, persons, fruits, and geological formations. For our models, pretrained models on this data set are responsible for extracting general features from images. Therefore the weights and biases of these pretrained backbone architectures were kept fixed during the training process of our specific models. While the last layers added on top of the pretrained models were adapted for the specific tasks. This is the common state-of-the-art procedure in image classification applications.

For training neural networks, typically the loss function is minimized, which computes the difference between the actual and predicted output targets. Different optimization algorithms are available. In this work, the Adam optimization algorithm was used together with a cross-entropy loss for the classification and a mean squared error loss for the regression. The training is additionally controlled by an early stopping procedure. This entails continuing training as long as the loss is decreasing and the performance on the validation data set is increasing. If this is not the case for a few epochs, then the training is stopped. The metrics used for evaluating the model are the accuracy for the formation classification and the (root) mean squared error and coefficient of determination  $R^2$  for the mineral content regression. To see the generalization performance of the neural network and to not introduce a bias through transfer learning, the data set was divided as follows: 5135 image segments were used solely for classification, and another 361 (i.e. 6.6%) solely for regression. For both tasks, the data were subdivided into three sets, the training and validation data set used for training, and the test data set used solely for testing after the training had been completed. More specifically for the classification task with a crack threshold below 5000 pixels, 4658 images were used for training (Parkinsoni-Württembergica-Schichten: 330 images, Humphriesiolith Formation: 911 images, Wedelsandstein Formation: 250 images, Murchisonae-Oolith Formation: 43 images, Opalinus Clay: 2498 images,

Stafflegg Formation: 626 images), 114 images for validation (19 images per class) and 113 images for testing (Parkinsoni-Württembergica-Schichten: 18 images and 19 images in all other classes). For the regression task, 276 images were used for training, 34 images for validation and 34 images for testing in the case of a crack threshold of 5000; for the threshold of 1000, it was 254, 32 and 31 images for training, validation and testing. Among the images for the regression task, in all images there were mixtures of clay, carbonate and silicate and not purely one element. It is important to mention that the images used for the regression task are independent from the images used for the classification task. So no images were used twice in the whole process. Also all the training, validation and testing images were independent from each other. All the images amongst many others can be found in Kaehr and Gysi (2021), in this document the cores are extracted from the background.

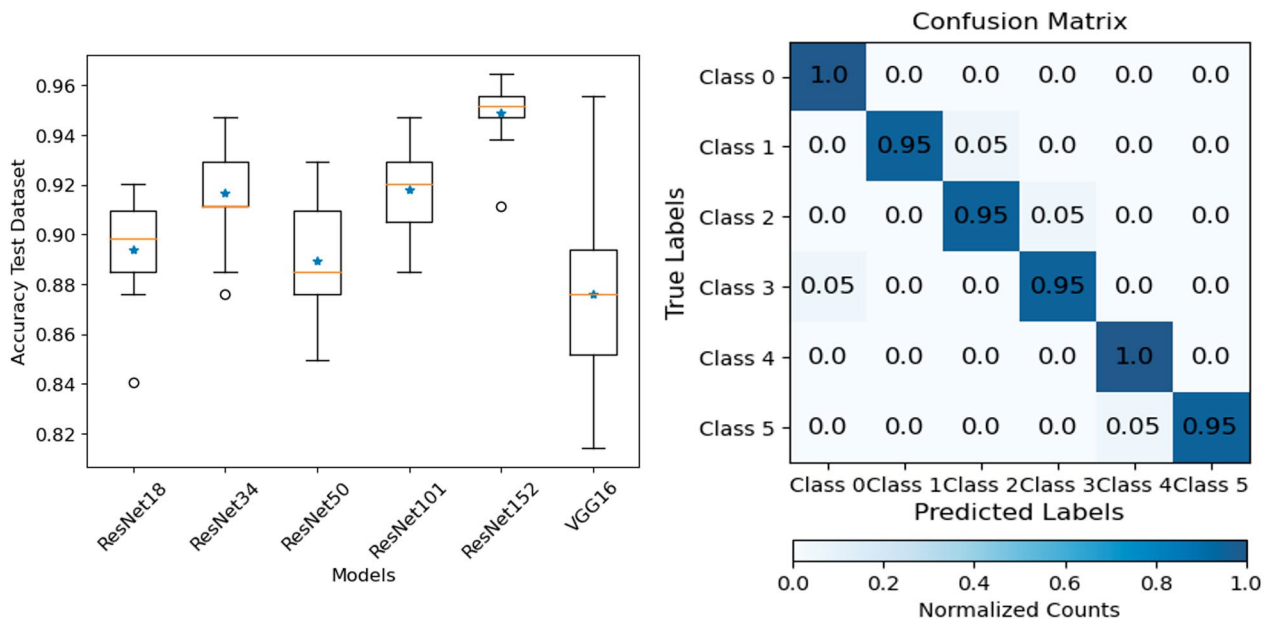
## 4 Results and discussion

### 4.1 Formation classification

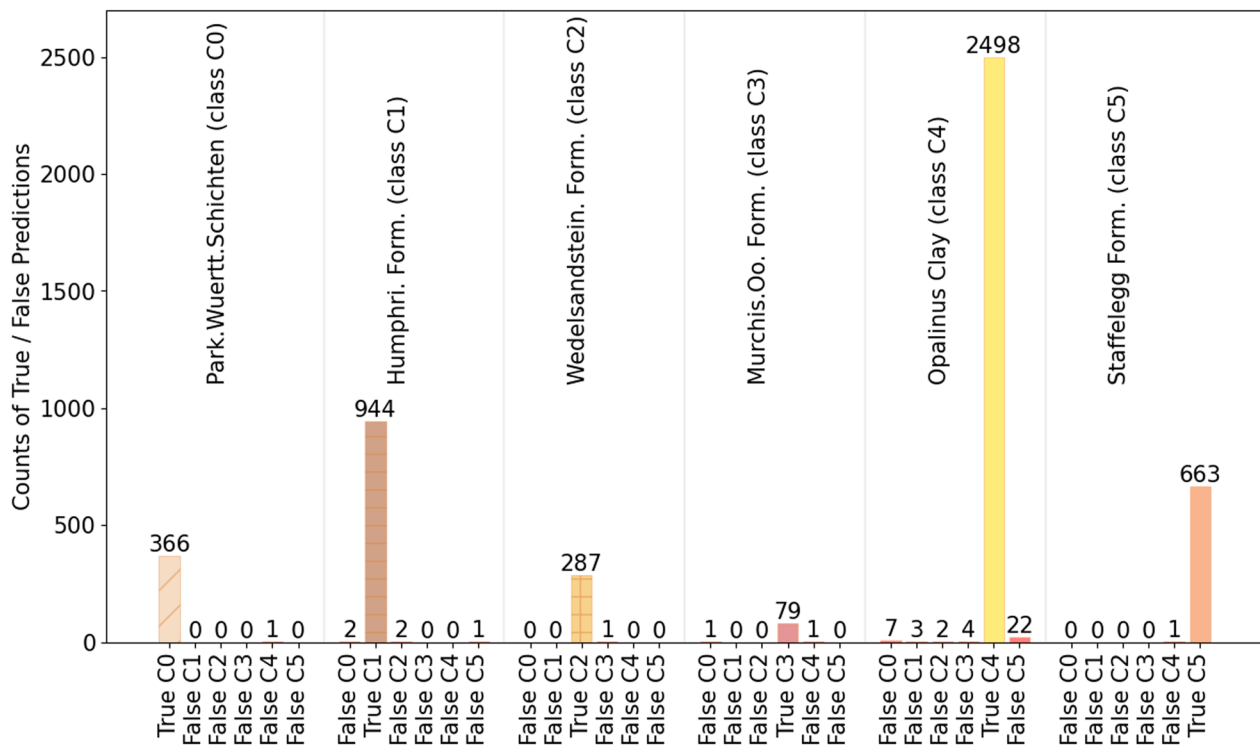
Six different pretrained backbone architectures (ResNet18, ResNet34, ResNet50, ResNet101, ResNet152, VGG16) were benchmarked for the formation classification task. The ResNet architecture requires same size and normalized input images. To use the image segments as input for the ResNet, the segments were loaded into a range between 0 and 1, normalized by using the mean and standard deviation, and resized to 850x100 pixels.

The output of the ResNet was then a 1000-dimensional vector. During training, the loaded weights and biases of the pretrained network were fixed. The 1000-dimensional vector returned by the backbone architecture served as input for the classification part. For the formation classification, the pretrained neural network was complemented with a linear layer of input 1000 and output 500 dimensions (1000x500), a ReLU layer (500x500) and again a linear layer (500x6). The last layer had 6 output dimensions due to the chosen encoding. Since the formation classes were categorical, each class needed to be encoded to a number; this was done with one-hot encoding, where the first class was represented as (1,0,0,0,0,0), the second as (0,1,0,0,0,0), and so forth until the last class (0,0,0,0,0,1). The implemented and trained model gave in the end the number of the predicted class, as an integer among {0, 1, 2, 3, 4, 5}.

After the training phase, the performance of the models was evaluated exclusively with the test data set in terms of accuracy. The graphical representation of the results, depicted as box plots derived from 14 different seeds for training initialization, is presented in Fig. 4a. In these plots, the blue star denotes the mean, the horizontal orange line represents the median, while the box encapsulates the lower and upper quartiles. The “whiskers,” short horizontal lines, denote the values for the first (lower) and third (upper) quartiles  $\pm 1.5$  times the interquartile range, which is the length between the first and third quartiles. Outliers are marked as individual circles.



**Fig. 4** Left: Formation classification results for the test data set for different backbone architectures with 14 different seeds each. Right: Confusion matrix for the formation classification model based on ResNet152 with artefact threshold  $T=5000$  pixels for the test data set



**Fig. 5** Counts of true and false predicted classes for each formation. In the x-axis, the classes are numbered; C0 corresponds to Park.Wuertt. Schichten, C1 to Humphr. Formation, ..., C5 to Staffelegg Formation. For a better understanding of the figure, consider e.g. the Park. Wuertt. Schichten, here the first 6 x-labels are relevant, the number of counts, 366 for True C0 means that 366 images were correctly classified as Park. Wuertt.Schichten, so class C0; no image was classified as class C1, C2, C3, or C5 and only one image that should have been a Park. Wuertt.Schicht was classified as class C4 (Opalinus Clay), so this was misclassified

As can be seen in Fig. 4a, the highest performing models are the ones with ResNet152 as a backbone architecture with the best one having an accuracy of 96.7%. For the overall best model, the confusion matrix is depicted in Fig. 4b. This matrix illustrates the comparison between the predicted and true formation classes for the test data set. Considering, for example the true class 2 (i.e. Wedelsandstein Formation) in row 3, one can see in the matrix that 95% of the images were classified correctly as class 2, whereas 5% were wrongly classified by the model as class 3. The plot reveals that misclassified images primarily correspond to neighboring classes. To investigate this further, the analysis was extended to the whole data set. Notably, out of 4885 images, only 48 were falsely classified (1%), with the majority of these misclassifications occurring between adjacent classes. Figure 5 illustrates the number of true and false predicted image slices for each formation. Among the 48 falsely predicted image slices, 8 images between 891.03 to 891.18m depth were predicted to fall into the Staffelegg formation instead of the Opalinus Clay. This misclassification is attributed to the high lithological similarity between the Opalinus Clay and the Staffelegg clay-rich rocks.

Similarly, 14 images were misclassified at the boundary between these two formations which again is attributed to little lithological difference between the samples at the boundary. Therefore, it is important to acknowledge that for classification performance of the CNN on core samples, stark differences between rock compositions facilitate differentiation, while similar or weak differences may lead to false classifications and further details would need to be considered.

The accuracy achieved at this stage is considered satisfactory and no further refinement of the models was pursued.

## 4.2 Mineral content regression

### 4.2.1 Comparison of neural network model architectures

Initially, the same architecture as for the formation classification, but with a linear layer of output dimension 3 was tested for the mineral content regression. The performance of this model was very poor with an average  $R^2$  value of  $-0.23$  over 15 different seeds. Hence, a different strategy was developed for the mineral content regression. Since the formation classification model performed very well, the trained formation classification neural



**Table 1** Mean and maximum R2 values for the test data set for the three different model types

	R2 test mean	R2 test max
T5000_1m	0.560	0.641
T5000_3m	0.617	0.691
T1000_3m	0.580	0.673

network, except the last layer, was used as the backbone architecture for the regression model. A linear layer (500x250), a ReLU layer (250x250), followed by a dropout layer and finally a linear layer (250x3) were added on top to solve the regression task (see Fig. 10 in the Appendix 2). With such an architecture, the classification of the images according to the rock type is performed first and helps to perform mineral composition analysis. The same strategy was also tested not only with one but with the three formation classification models showing the best performance, which were concatenated, and the following layers were added on top, see Fig. 3: A linear layer (1500x700), a ReLU layer (700x700), a dropout layer, a linear layer (700x250), a ReLU layer (250x250), a linear layer (250x3), and a sigmoid layer (3x3). Additionally, models were trained using a reduced crack detection/deletion threshold of 1000 pixels, as opposed to the previous threshold of 5000 pixels, for the model employing the ensemble of three formation classification models.

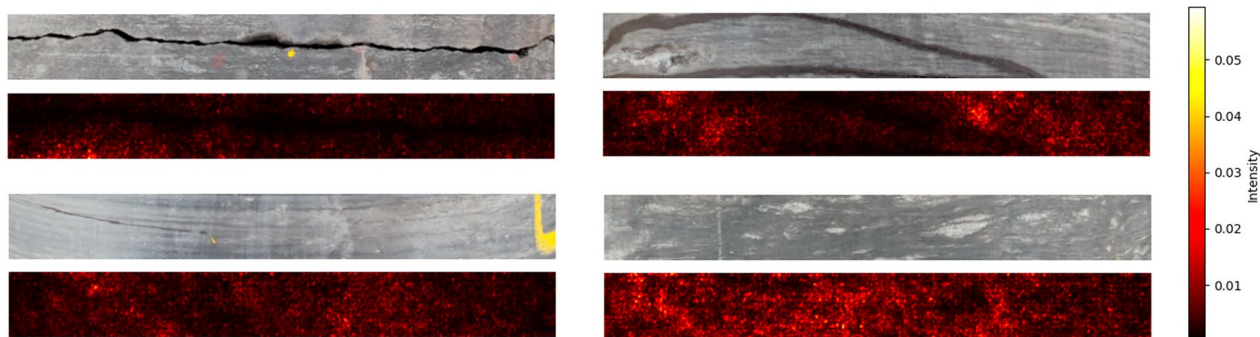
In summary, three model types were trained and tested for mineral content regression. These were: 1) the best formation classification model as backbone architecture with crack detection/deletion threshold of 5000 pixels (T5000-1 m); 2) the best trio of formation classification models as backbone architecture with a threshold of 5000 (T5000-3 m); 3) the best trio of formation classification models as backbone architecture with a threshold of 1000 (T1000-3 m).

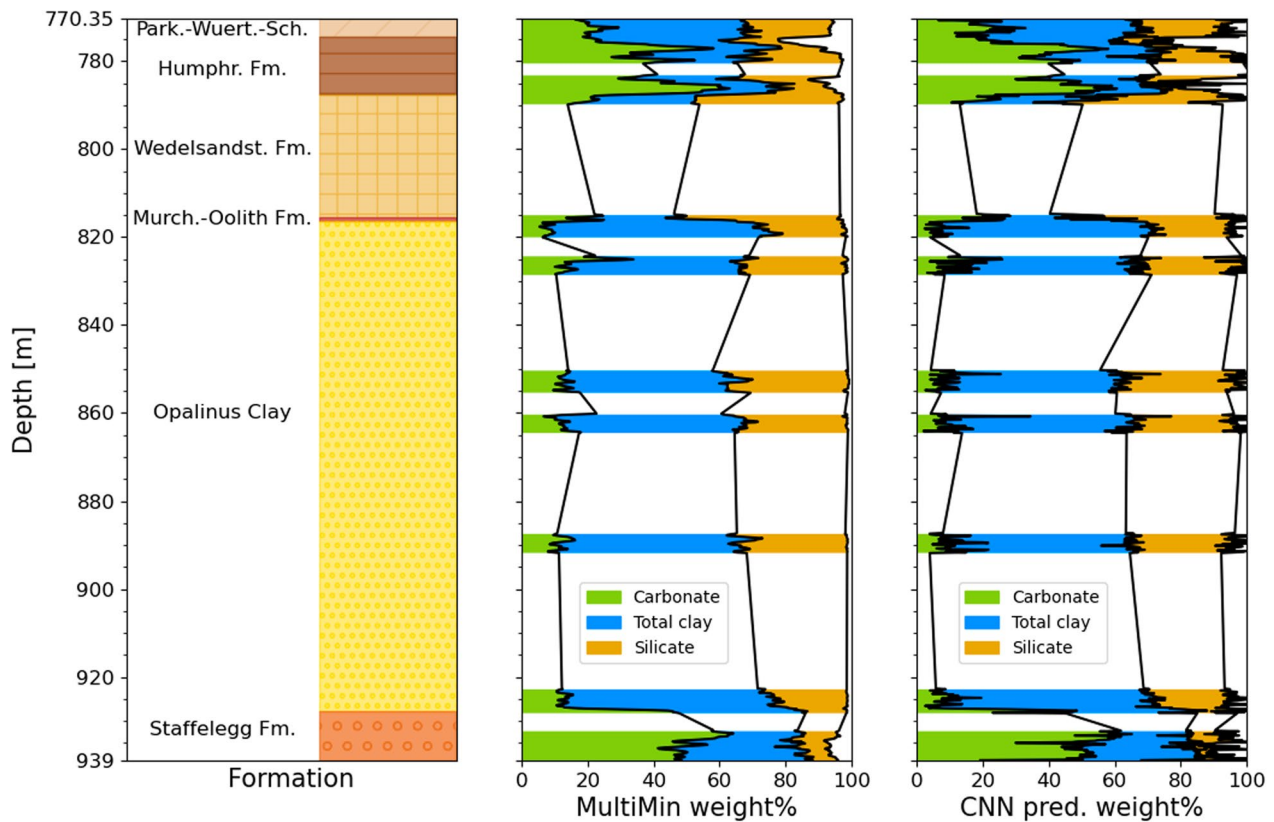
Each model type was trained 18 times with different seeds. The outcomes in terms of the coefficient of determination (R2) for the test data set are reported in Table 1 and are visualized in Fig. 11 in the Appendix 2. Data interpretations need to be aware that data used for training are derived from a model with inherent uncertainty.

The data in Table 1 suggest a subtle better performance in the regression models that utilize the ensemble of three formation models as a backbone architecture. The R2 values for the models with a larger crack deletion area of 5000 pixels are higher, but further investigations are needed to determine the significance of detected cracks on the mineral content regression modeling. Therefore, saliency maps were considered, to better understand the importance of crack detection and further need to dismiss these images. A saliency map for CNNs highlights the most relevant regions or features within an input image, helping to identify areas of significance for the particular task. The assigned dark colors refer to the “unimportant” domains, whereas the lighter red domains are the most “significant” ones for the neural network. From these maps, as depicted in Fig. 6, one can see that cracks or color marks appear very dark in the saliency maps, meaning that those features are not so relevant for solving the regression task. Hence, the size of the crack deletion area was not considered further within this work.

#### 4.2.2 Performance of best mineral content regression model

The best model according to the R2 values on the test data set was the one with three formation models as a backbone architecture and a threshold of 5000 pixels and a seed of 1000. The comparison of the multiMin model data and the CNN predicted mineral content is depicted in Fig. 7. For this best mineral content regression model, the absolute and relative errors as well as the R2 values for each individual mineral within the test data set were computed: The model exhibits a moderate level of

**Fig. 6** Saliency maps of 4 different 1 cm image slices, upper images show the true image, lower images the corresponding saliency maps



**Fig. 7** The first column shows the various formations along the depth, the second shows the predictions of mineral content with MultiMin model, whereas the third column shows the CNN predicted mineral content in this rock. Note that images were only available for the coloured areas

**Table 2** Comparison between the MultiMin log data and the CNN model prediction in terms of absolute (abs.err.) and relative (rel.err.) error of all three minerals for each formation. nr gives the number of data points available per formation. nr gives the number of data points available per formation

Formation	nr	Carbonate abs. err.	Total clay abs. err.	Silicate abs.err.	Carbonate rel. err. (%)	Total clay rel. err. (%)	Silicate rel.err. (%)
Park	26	0.03	0.025	0.025	16.42	5.21	10.87
Humphr	69	0.037	0.037	0.037	9.62	13.37	19.86
Wedel	22	0.033	0.035	0.022	15.67	11.24	4.89
Murch	5	0.033	0.038	0.067	21.06	11.17	16.00
Opa	173	0.051	0.036	0.027	41.14	6.73	10.17
Staffel	49	0.045	0.036	0.034	8.96	10.75	36.62

accuracy in predicting carbonate mineral content, with an absolute error of 0.059 and a relative error of 39.1%. While the model's R2 value of 0.609 indicates a reasonable fit to the data, further refinement may be necessary to enhance the accuracy of the model. In contrast, the model performed much better in predicting the silicate mineral content, with an absolute error of 0.038 and a relative error of 18.7%. The higher R2 value of 0.707

underscores the model's ability to explain the variance in silicate content. The best results with the trained models were obtained from the prediction of the total clay content, with an absolute error of 0.046 and a lower relative error of 10.7%. The R2 value of 0.811 refers to a robust correlation that emphasizes the model's capability to capture the variability in total clay content effectively (see also Fig. 7). The high prediction accuracy for the total

clay content may be due to several reasons including better spectral correlations with the clay phases.

In Appendix 3 Fig. 12, another direct comparison, of the MultiMin log data on the x-axis versus the predictions on the y-axis is given. The overall trend looks quite assuring and is supported by the metrics on the test data set, although there are some differences when considering all the details. Nevertheless, the model is already well suited for distinguishing between low and high mineral content for each species.

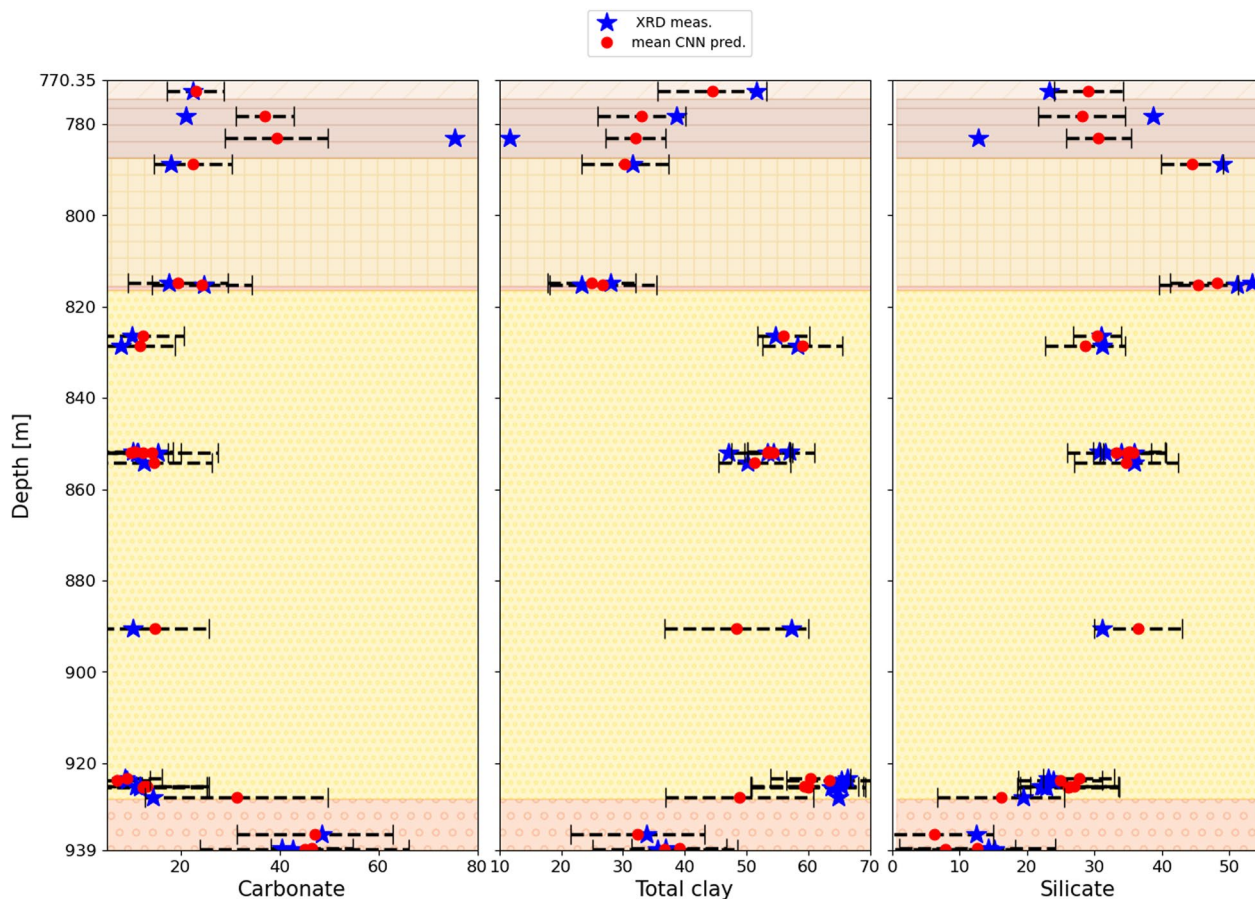
#### 4.2.3 Model performance across formations

To investigate the model performance further, the absolute and relative errors for the mineral content predictions for each formation were investigated: The results are depicted in Table 2, where nr gives the number of considered data points per formation. From the table, one can see that the performance varies a lot among the different formations. Summarizing, the lowest error values were achieved as follows: carbonates content was predicted best for the Humphriesiolith and

the Staffelegg Formation with a relative error below 9.7%. Clay minerals content predictions showed a relative error below 6.8% for the Parkinsoni-Württembergica-Schichten and the Opalinus Clay, while silicates predictions had a relative error below 5.1% for the Wedelsandstein Formation. Throughout the analysis, across all formations a systematic better prediction emerged for the mineral phases, which were prevalent in the selected formations. One reason for the differences among the various formations could be, that the number of data points for training the model was not the same for each formation class. In addition to improving the results, assuming that many more data points may be available, one model for each formation could be trained.

#### 4.3 Comparison with bulk XRD measurements

The training of the CNN for the mineralogical analysis was performed using data provided by MultiMin log model, which provides an indirect prediction of mineral content based on drill-logs data. The only measured data available for the true evaluation of the model's performance are the actual bulk XRD measurements performed on the cores.



**Fig. 8** Mean (red dots) and  $2\sigma$  of the CNN predicted mineral content from images at XRD depth  $\pm 5$  cm and the according bulk XRD measurements (blue stars)

As explained in Marnat and Becker (2021, 2020) the core data of porosity, mineralogy and rock density were used for calibration of log measurements in the MultiMin workflow, although this does not mean that the MultiMin model data exactly reproduce the measured core data. The difference between modelled and measured data allows us to evaluate the performance of the model. Therefore, the analysis of the performance of CNN predictions and MultiMin log model were both performed against true XRD core measurements.

In total 23 mineralogical data points obtained by XRD measurements on samples from core intervals used in image analysis were available for comparison with the model predictions. The comparison of core and log mineral compositions needs to consider small differences and uncertainties. For example, the rock samples taken for the analysis represent a core fragment with a volume of 1–10 cm<sup>3</sup>. These samples are not necessarily located on the surface of the core. Quite in contrary, the samples are taken from the core center to avoid contamination or any alteration processes. Thus, the prediction based on the visual information from the surface can not be more accurate than the typical variation of mineralogical content within a domain of 10×10×10 cm<sup>3</sup>, which was taken for the laboratory analysis. Accordingly, 11 image segments were taken for each data point as input for the model and the mean and standard deviation,  $\sigma$ , of the predicted mineral content were computed. The full table of results can be found in the Appendix 3, Table 3. In Fig. 8 the mean (red dot) and the 95% confidence interval, i.e. the  $2\sigma$  region, of predictions are depicted for each mineral, together with the real measurements (blue stars). Almost all the measurements lie within the  $2\sigma$  region. The only significant deviation is observed in one sample from Humphriesiolith. The photographic interpretation shows an unusually high carbonate content and is part of a thin layer of elevated carbonate content present within the heterogeneous Humphriesiolith unit.

To have a comparison to the MultiMin log data, also for the predicted model data the Spearman correlation coefficient ( $cc$ ) was calculated for each mineral, taking the prediction at the MultiMin log data depth:

Carbonate:  $cc_{MultiMin} = 0.80$  ( $p_{val} = 4.21E - 06$ ),  
 $cc_{CNN} = 0.90$  ( $p_{val} = 6.90E - 09$ ),

Clay:  $cc_{MultiMin} = 0.93$  ( $p_{val} = 1.42E - 10$ ),  
 $cc_{CNN} = 0.90$  ( $p_{val} = 6.43e - 09$ ),

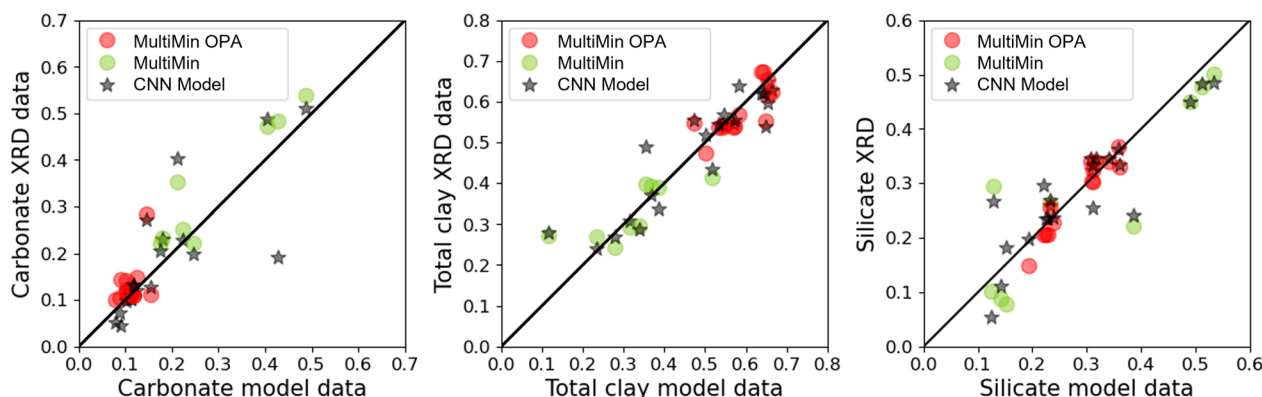
Silicate:  $cc_{MultiMin} = 0.84$  ( $p_{val} = 5.51E - 07$ ),  
 $cc_{CNN} = 0.81$  ( $p_{val} = 3.31E - 06$ ).

All values, including the relative and absolute errors; can be found in the Appendix 3, Table 4. This metric was chosen, since this was the metric used in Marnat and Becker (2021) to compare the MultiMin log data to the XRD measurements. The neural network model predictions are in good agreement with the MultiMin log data and show similar accuracy as illustrated in Fig. 9.

In summary, the mineral content regression model presents a promising tool for mineral estimation from drill core images for the selected formations within this core, although further refinement of the model to increase the accuracy would be desired. In the future, an expansion of the dataset and inclusion of different lithological core sections will be considered.

## 5 Conclusion and outlook

This study successfully established and tested an automatic workflow for preprocessing and analysis of drill core images based on machine learning methods. A pretrained ResNet architecture was trained as part of the workflow to classify the 1 cm drill core segments into 6 different formation classes (for which data and images were available). The classification model achieved a prediction accuracy of 96.7% for unseen images. Big parts of this formation classification model were subsequently used as a backbone architecture to establish a NN model to predict the mineral content (silicate, total clay and carbonate) from only drill core images.



**Fig. 9** Bulk XRD measurements (only representing carbonate, total clay, and silicate compositions) plotted versus MultiMin log data (red and green dots) and overlain with CNN predicted mineral content at MultiMin log data depth (black stars). The red dots correspond to the Opalinus clay, whereas the green dots to all the other formations

The CNN mineral content regression predictions were compared to the XRD measurements and showed comparable good correlations as the MultiMin log data. The critical technique used for both models was transfer learning, which involved the usage of models that were trained on related tasks and hence, the re-use of the obtained knowledge. It has to be emphasized that due to the limited number of experimentally measured mineralogical data used in this study (23), the CNN model was trained using a model based on the MultiMin log data set (361) derived from drill log data. The data analysis shows that both models, CNN and MultiMin log, demonstrate the same prediction performance when benchmarked against the measured mineralogical lab (XRD) data. In contrast to other models in the literature, the CNN model constructed in this study relies on images only and can be used for the image-based interpolation of mineralogical data down to 1 cm resolution.

Although both presented CNN models (formation classification, mineral content regression) show already good accuracy, there are several options to possibly further improve the performance, which lie outside the scope of the current paper: First, the hyperparameters, so the number of layers and neurons, as well as the activation functions were chosen to be fixed. A thorough search for optimal hyperparameters could drastically increase the performance of the model. Second, another choice of the backbone architecture (here ResNet was used) could improve the model and is planned for future work (like e.g. transformer networks or self-attention mechanisms).

In general, the transfer learning strategy from pretrained models seems to perform better than starting training from scratch. The ideal case would be to have a model that was trained on a huge data set of drill core images from different places all over the world to extract already important features from the images. The model fine-tuning could then be done for the specific applications and data sets.

The number of data points is essential for the performance of a machine learning model. In this study, the formation classification model is based on more than 5000 images, whereas the mineral content regression model was based on only 361 images. Increasing the amount of data would definitely improve the model and make it more robust and capable of generalization.

The data set in this study could benefit from further augmentation by incorporating images from other boreholes of the same formations of complementary geological units. This expansion would not only increase the data set's size but also enhance the model's ability to generalize across diverse and possibly challenging geological formations and thus rock compositions. Another cost-efficient option to augment the data set would be to create a synthetic data set with geostatistical methods.

The performance of the models might also be influenced by the choice of the segmentation size of the images, in this work we used 1 cm image segments. Decreasing the width or height of the segments would be a way of enlarging the data set, although each segment would have less information. Increasing the width would have the opposite effect, the data set available would be reduced, but each segment would contain more information. Optimizing this aspect can be crucial for achieving a balance between data quantity and quality.

To enhance the data quality, the radial distortion of the images, due to the position of the camera in relation to the drill core, should be considered. In future work, the implementation of a transformation function for correcting the radial distortion is planned.

In summary, this study provides a solid foundation for the regression of mineral content from drill core images using deep learning techniques. The outlined future research directions provide a road map for improving model performance, data set robustness, and overall predictive accuracy. Ultimately image-based regression methodologies will have the potential to support the field of geological analysis and drilling technologies.

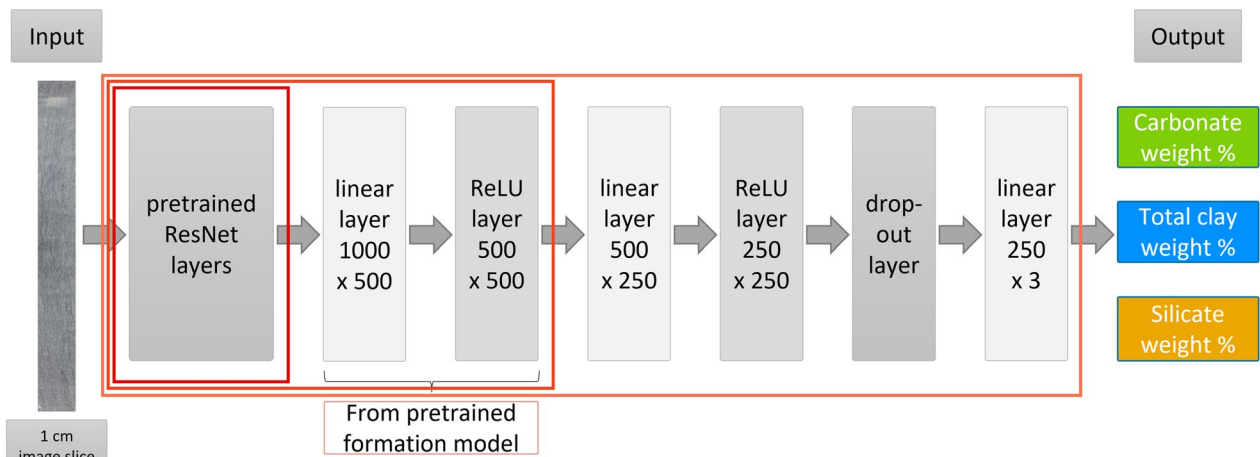
## Appendices

### 1. List of abbreviations (alphabetical order)

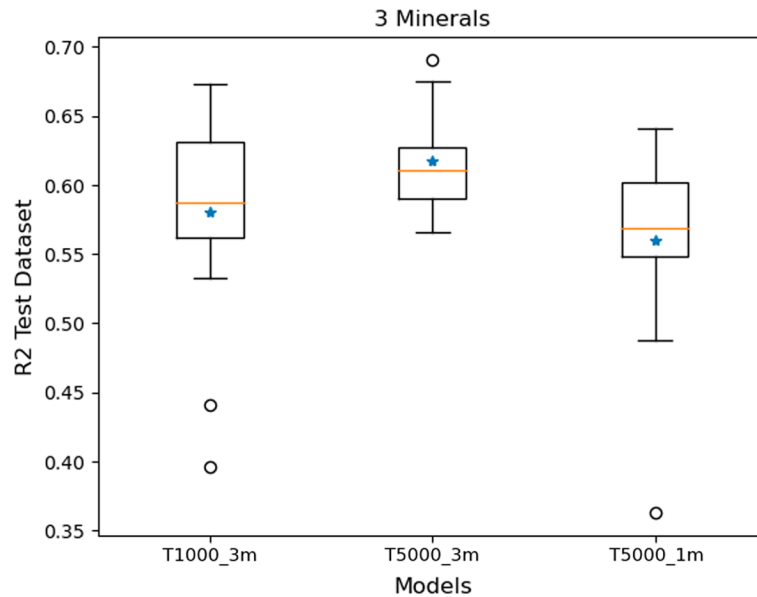
- cc: Correlation coefficient
- chemical elements: U: Uranium; Th: Thorium; K: Potassium; Fe: Iron; Si: Silicon; Ca: Calcium; Al: Aluminium; Ti: Titanium; S: Sulfur;
- CNN: Convolutional neural network
- ICC: International Color Consortium (colour profile)
- NN: Neural network
- $p_{val}$ : P-value (probability)
- pXRF: Portable X-ray Fluorescence analysis
- R<sup>2</sup>: Coefficient of determination
- ReLU: Rectified Linear Unit
- ResNet: Residual Neural Network
- RF: Random Forest
- SVM: Support Vector Machine
- TIFF: Tagged Image File Format
- VGG: Visual Geometry Group (a deep CNN architecture)
- XRD: X-ray Diffraction

### 2. Additional figures for mineral content regression

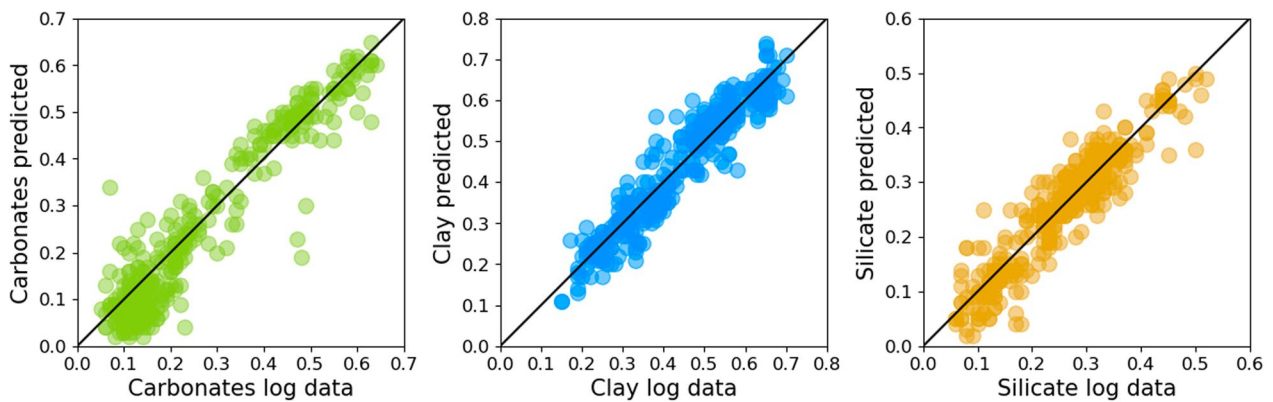
See Figs. 10, 11 and 12.



**Fig. 10** Architecture of the neural network for the mineral content regression with one formation model as a backbone architecture



**Fig. 11** Boxplot for comparing the three different model types for the mineral content regression in terms of the R2 values of the test data set



**Fig. 12** MultiMin log data vs CNN model predicted data for each mineral separately, the x and y axis show the according weight %

**Table 3** Comparison of model predicted data and bulk XRD measured data. The mean depth is given in meter [m]. The columns Sil., Carb., Clay show the measured mineral content. The columns Sil. pred.mean, Carb.pred.mean, Clay pred. mean and Sil.pred.std., Carb. pred.std., Clay pred.std. denote the mean and standard deviation of the predictions with the CNN model for images of the according measurement depth  $\pm 5$ cm. The mean depth is given in meter [m]. The columns Sil., Carb., Clay show the measured mineral content. The columns Sil. pred.mean, Carb.pred.mean, Clay pred.mean and Sil.pred.std., Carb.pred.std., Clay pred.std. denote the mean and standard deviation of the predictions with the CNN model for images of the according measurement depth  $\pm 5$ cm

	Mean depth [m]	Sil.	Sil. pred mean	Sil. pred std	Carb.	Carb. pred mean	Carb. pred std	Clay	Clay pred mean	Clay pred std	Lith-ology
0	772.73	23.3	29.16	2.58	22.4	23.01	2.88	51.66	44.43	4.41	Park. Würt. Sch.
1	778.32	38.7	28.14	3.25	21.1	36.99	2.93	38.71	32.98	3.54	Humphr. Form.
2	783.13	12.8	30.63	2.38	75.34	39.36	5.2	11.62	32.02	2.44	Humphr. Form.
3	788.76	49	44.52	2.3	17.98	22.41	3.95	31.6	30.33	3.54	Wedels. Form.
4	814.82	53.4	48.2	3.47	17.52	19.4	5.05	27.93	24.97	3.56	Wedels. Form.
5	815.28	51.2	45.46	2.93	24.68	24.33	5.04	23.32	26.74	4.32	Wedels. Form.
6	826.39	31	30.44	1.77	10.17	12.28	4.19	54.68	56	2.09	Opalinus Clay
7	828.51	31.17	28.67	2.97	7.87	11.68	3.53	58.24	59.01	3.22	Opalinus Clay
8	851.84	30.66	35.15	2.7	10.41	10.66	3.87	57	53.53	1.69	Opalinus Clay
9	851.88	31.65	35.73	2.4	11.27	9.87	3.75	54.37	53.77	1.82	Opalinus Clay
10	851.95	33.99	34.93	1.77	11.04	14.19	6.62	53.4	53.36	1.87	Opalinus Clay
11	852.05	35.97	33.3	3.66	15.43	12.26	3.92	47.16	54.28	3.39	Opalinus Clay
12	854.18	35.89	34.72	3.84	12.49	14.49	5.94	50.11	51.23	2.92	Opalinus Clay
13	890.39	31.2	36.48	3.26	10.25	14.73	5.46	57.31	48.36	5.84	Opalinus Clay
14	923.34	23.2	27.72	2.64	8.67	9.2	3.5	66.37	60.36	3.25	Opalinus Clay
15	923.81	23.89	24.99	3.11	9.03	7.19	3.33	65.41	63.46	3.53	Opalinus Clay
16	925.11	22.11	27.02	3.28	10.96	13.19	6.28	65.41	59.32	4.37	Opalinus Clay
17	925.16	22.39	26.36	3.69	11.7	12.46	6.43	64.35	59.84	4.55	Opalinus Clay
18	925.17	22.86	26.12	3.79	11.73	12.41	6.43	63.71	59.99	4.6	Opalinus Clay
19	927.41	19.38	16.12	4.71	14.4	31.27	9.23	64.86	48.86	5.96	Opalinus Clay
20	935.59	12.52	6.24	4.37	48.63	47.13	7.84	33.77	32.39	5.43	Staffel. Form.
21	938.53	14.2	12.6	5.8	40.43	46.45	4.15	36.85	39.13	3.85	Staffel. Form.
22	938.9	15.2	7.93	5.14	42.77	45.05	10.55	35.6	36.77	5.85	Staffel. Form.

**Table 4** Absolute (Abs.err.), relative (Rel.err.) error, coefficient of determination (R2) and Spearman correlation coefficient (cc) of the three minerals and the according p-values (p-val), computed for the 23 XRD measurement data points compared to the MultiMin Log data and the CNN model predictions. The metrics are computed for the predicted values at the corresponding MultiMin Log data depth. The metrics are computed for the predicted values at the corresponding MultiMin Log data depth

	XRD MultiMin Carbonate	XRD MultiMin Total clay	XRD MultiMin Silicate	XRD CNN Carbonate	XRD CNN Total clay	XRD CNN Silicate
Abs.err.	5.36	3.91	3.81	5.69	4.26	3.87
Rel.err.	25.29%	12.84%	17.88%	24.08%	13.59%	17.38%
R2	0.68	0.88	0.75	0.61	0.84	0.78
cc	0.80	0.93	0.84	0.90	0.90	0.81
p-val	4.21E-06	1.42E-10	5.51E-07	6.90E-09	6.43E-09	3.31E-06

### 3. Additional tables for comparison between model and bulk XRD measurements

See Tables 3 and 4.

### 4. Overview of core photographs

The following section contains the core images available and used in this study, already preprocessed until step 3, so the original image was color and white balance

corrected and cropped from the background. For each photo, the specific depth range is written above the photo together with the corresponding lithology class. The blue rectangles indicate where XRD laboratory measurements were performed. For model validation purposes, the slices at the depth given for the XRD measurements  $\pm 5$  cm are considered (blue rectangles).

**Drill core photographs, extracted from the background**

770.35m – 770.92m, Parkinsoni-Württembergica Schichten



770.92m-771.75m, Parkinsoni-Württembergica Schichten



771.75m-772.58m, Parkinsoni-Württembergica Schichten



772.58m-773.06m, Parkinsoni-Württembergica Schichten



773.06m-773.55m, Parkinsoni-Württembergica Schichten



773.55m-773.73m, Parkinsoni-Württembergica Schichten



773.73m-773.80m, Parkinsoni-Württembergica Schichten



773.80m-773.93m, Parkinsoni-Württembergica Schichten



773.93m-774.39m, Parkinsoni-Württembergica Schichten



774.39m-774.55m, Parkinsoni-Württembergica Schichten 774.55m-774.83m Humphriesoolith Formation



774.83m-775.23m, Humphriesoolith Formation



775.23m-775.63m, Humphriesoolith Formation





775.63m-776.36m, Humphriesoolith Formation



776.36m-776.92m, Humphriesoolith Formation



776.92m-777.80m, Humphriesoolith Formation



777.80m-778.22m, Humphriesoolith Formation



778.22m-778.54m, Humphriesoolith Formation



778.54m-779.44m, Humphriesoolith Formation



779.44m-779.93m, Humphriesoolith Formation



779.44m-779.93m, Humphriesoolith Formation



779.93m-780.63m, Humphriesoolith Formation



782.94m-783.76m, Humphriesoolith Formation



783.76m-784.60m, Humphriesoolith Formation



784.60m-785.20m, Humphriesoolith Formation



785.20m-785.94m, Humphriesoolith Formation



785.94m-786.74m, Humphriesoolith Formation



786.74m-787.50m, Humphriesoolith Formation; 787.50m-787.63m Wedelsandstein Formation



787.63m-788.43m, Wedelsandstein Formation



788.43m-788.99m, Wedelsandstein Formation



788.99m-789.92m, Wedelsandstein Formation



814.70m-815.51m, Wedelsandstein Formation; 815.51m-815.38m, Murchisonae-Oolith Formation



815.38m-815.68m, Murchisonae-Oolith Formation



815.68m-816.02m, Murchisonae-Oolith Formation



816.02m-816.42m, Murchisonae-Oolith Formation



816.42m-817.32m, Opalinus Clay



817.32m-818.04m, Opalinus Clay



818.04m-818.63m, Opalinus Clay



818.63m-819.38m, Opalinus Clay



819.38m-819.41m, Opalinus Clay



819.41m-819.56m, Opalinus Clay



819.56m-819.75m, Opalinus Clay



819.75m-820.15m, Opalinus Clay



824.04m-824.42m, Opalinus Clay



824.42m-825.40m, Opalinus Clay



825.40m-825.57m, Opalinus Clay



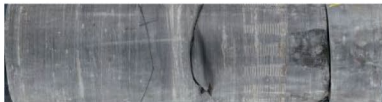
825.57m-825.66m, Opalinus Clay



825.66m-826.48m, Opalinus Clay



826.48m-826.82m, Opalinus Clay



826.82m-827.02m, Opalinus Clay



827.02m-827.28m, Opalinus Clay



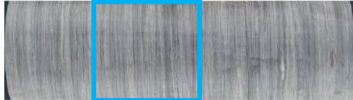
827.28m-827.75m, Opalinus Clay



827.75m-828.38m, Opalinus Clay



828.38m-828.70m, Opalinus Clay



850.04m-850.95m, Opalinus Clay



850.95m-851.70m, Opalinus Clay



851.70m-852.20m, Opalinus Clay



852.20m-852.30m, Opalinus Clay



852.30m-853.13m, Opalinus Clay



853.13m-853.36m, Opalinus Clay



853.36m-853.60m, Opalinus Clay



853.60m-853.75m, Opalinus Clay



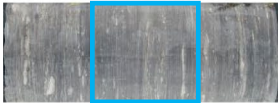
853.75m-853.90m, Opalinus Clay



853.90m-854.05m, Opalinus Clay



854.05m-854.30m, Opalinus Clay



854.30m-854.84m, Opalinus Clay



854.84m-855.47m, Opalinus Clay



860.03m-860.85m, Opalinus Clay



860.85m-861.22m, Opalinus Clay



861.22m-861.54m, Opalinus Clay



861.54m-862.08m, Opalinus Clay



862.08m-862.39m, Opalinus Clay



862.39m-862.88m, Opalinus Clay



862.88m-862.95m, Opalinus Clay



862.95m-863.56m, Opalinus Clay



863.56m-864.00m, Opalinus Clay



864.00m-864.62m, Opalinus Clay



887.15m-887.83m, Opalinus Clay



887.73m-888.13m, Opalinus Clay



888.13m-889.09m, Opalinus Clay



889.09m-889.97m, Opalinus Clay



889.97m-890.50m, Opalinus Clay



890.50m-890.85m, Opalinus Clay



890.85m-891.03m, Opalinus Clay



891.03m-891.89m, Opalinus Clay

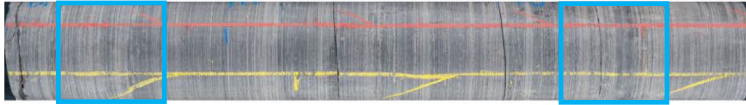


**Drill core photographs, extracted from the background**

922.49m-923.24m, Opalinus Clay



923.24m-923.94m, Opalinus Clay



923.94m-924.08m, Opalinus Clay



924.08m-924.96m, Opalinus Clay



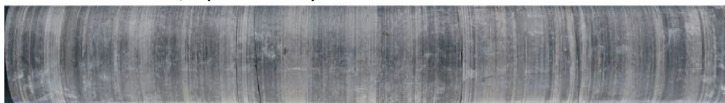
924.96m-925.41m, Opalinus Clay



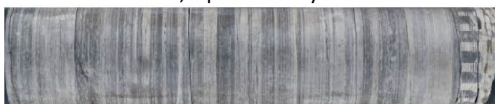
925.41m-926.03m, Opalinus Clay



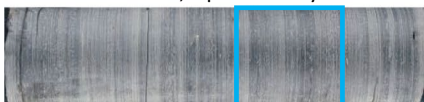
926.03m-926.70m, Opalinus Clay



926.70m-927.15m, Opalinus Clay



927.15m-927.53m, Opalinus Clay



927.53m-927.91m Opalinus Clay; 927.91m-928.31m Staffelegg Formation



932.00m-932.96m, Staffelegg Formation



932.96m-933.20m, Staffelegg Formation



933.20m-933.29m, Staffelegg Formation



933.29m-934.16m, Staffelegg Formation



934.16m-935.03m, Staffelegg Formation



935.03m-935.31m, Staffelegg Formation



935.31m-935.94m, Staffelegg Formation



935.94m-936.18m, Staffelegg Formation



936.18m-937.17m, Staffelegg Formation



937.17m-938.13m, Staffelegg Formation



938.13m-939.01m, Staffelegg Formation



### Acknowledgements

We sincerely thank Marc Pollefeys (ETHZ) for his insights and discussions regarding the implementation of neural networks in computer vision. We appreciate the internship of Geeta Goyal, that gave a starting point for this work. We are thankful for thoughtful discussions with Maximilian Mandl (Nagra) beforehand on the usage of machine learning for drill cores. Thanks also go to the Database Management Group at Nagra which provided maps and core images. We appreciate a lot the enlightening conversations with Thomas Gimmi (PSI) on the available datasets.

### Author contributions

RB, SC and NP conceptualized the work. RB, ILB and NP designed the methodology. RB and ILB developed the software framework and validated the models. RW provided the data and accompanying information. GK contributed to the data analysis and the data and model interpretation. RB drafted the first version of the manuscript. SC, GK, ILB, RW and NP substantially contributed to the manuscript. All authors read and approved the final manuscript.

### Funding

EURAD-DONUT WP received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 847593.

### Availability of data and materials

The data that support the findings of this study are available from Nagra but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Nagra. The code framework is available from the authors upon reasonable request.

### Declarations

#### Ethics approval and consent to participate

Not applicable

#### Consent for publication

All authors have given their consent for publication.



### Competing interests

The authors declare no conflict of interest.

Received: 22 November 2023 Accepted: 11 March 2024

Published online: 07 May 2024

### References

- Alguliyev, R., Imamverdiyev, Y., & Sukhostat, L. (2023). Automatic facies detection based on oilfield core images. *Petroleum Science and Technology*, 41(17), 1641–1664. <https://doi.org/10.1080/10916466.2022.2094954>.
- Alzubaidi, F., Mostaghimi, P., Si, G., Swietojanski, P., & Armstrong, R. T. (2022). Automated rock quality designation using convolutional neural networks. *Rock Mechanics and Rock Engineering*, 55(6), 3719–3734. <https://doi.org/10.1007/s00603-022-02805-y>.
- Alzubaidi, F., Mostaghimi, P., Swietojanski, P., Clark, S. R., & Armstrong, R. T. (2021). Automated lithology classification from drill core images using convolutional neural networks. *Journal of Petroleum Science and Engineering*, 197, 107933. <https://doi.org/10.1016/j.petrol.2020.107933>
- Ammen, M., & Palten, P.-J.: TBO Trüllikon-1-1: Data Report, Dossier I, Drilling, Wettingen, Switzerland (2021). Retrieved from [https://nagra.ch/wp-content/uploads/2022/08/e\\_nab20-009-Dossier-I.pdf](https://nagra.ch/wp-content/uploads/2022/08/e_nab20-009-Dossier-I.pdf)
- Baraboshkin, E. E., Demidov, A. E., Orlov, D. M., & Koroteev, D. A. (2022). Core box image recognition and its improvement with a new augmentation technique. *Computers & Geosciences*, 162, 105099. <https://doi.org/10.1016/j.cageo.2022.105099>
- Barker, R. D., Barker, S. L. L., Cracknell, M. J., Stock, E. D., & Holmes, G. (2021). Quantitative mineral mapping of drill core surfaces II: Long-wave infrared mineral characterization using pxrf and machine learning. *Economic Geology*, 116(4), 821–836. <https://doi.org/10.5382/econgeo.4804>.
- Clark, A. (2015). Pillow (PIL Fork) Documentation. readthedocs. Retrieved from <https://buildmedia.readthedocs.org/media/pdf/pillow/latest/pillow.pdf>
- Falivene, O., Aughter, N. C., Lima, R., Kleipool, L., Solum, J. G., Zarian, P., Clark, R. W., & Espejo, I. (2022). Lithofacies identification in cores using deep learning segmentation and the role of geoscientists: Turbidite deposits (Gulf of Mexico and North Sea). *AAPG Bulletin*, 106(7), 1357–1372. <https://doi.org/10.1306/03112221015>.
- Faria, E. L., Coelho, J. M., Matos, T. F., Santos, B. C. C., Trevizan, W. A., Gonzalez, J. L., Bom, C. R., Albuquerque, M. P., & Albuquerque, M. P. (2022). Lithology identification in carbonate thin section images of the Brazilian pre-salt reservoirs by the computational vision and deep learning. *Computational Geosciences*, 26(6), 1537–1547. <https://doi.org/10.1007/s10596-022-10168-0>.
- Fu, D., Su, C., Wang, W., & Yuan, R. (2022). Deep learning based lithology classification of drill core images. *PLOS ONE*, 17(7), 1–25. <https://doi.org/10.1371/journal.pone.0270826>
- Geron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, Inc., Retrieved from <https://raw.githubusercontent.com/data-science-projects-and-resources/Data-Science-EBooks/main/Machine>
- Guerra Prado, E. M., Souza Filho, C. R., & Muico Carranza, E. J. (2023). Ore-grade estimation from hyperspectral data using convolutional neural networks: A case study at the olympic dam iron oxide copper-gold deposit, Australia. *Economic Geology*. <https://doi.org/10.5382/econgeo.5023>.
- Gunther, C., Jansson, N., Liwicki, M., & Simistira-Liwicki, F. (2021). Towards a machine learning framework for drill core analysis. In: 2021 Swedish Artificial Intelligence Society Workshop (SAIS), pp. 1–6. IEEE, Sweden. <https://doi.org/10.1109/SAIS53221.2021.9484025>.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778 <https://doi.org/10.1109/CVPR.2016.90>
- Houshmand, N., Goodfellow, S., Esmaili, K., & Ordóñez Calderón, J. C. (2022). Rock type classification based on petrophysical, geochemical, and core imaging data using machine and deep learning techniques. *Applied Computing and Geosciences*, 16, 100104. <https://doi.org/10.1016/j.acags.2022.100104>.
- Jooshaki, M., Nad, A., & Michaux, S. (2021). A systematic review on the application of machine learning in exploiting mineralogical data in mining and mineral industry. *Minerals*, 11(8), 816. <https://doi.org/10.3390/min11080816>.
- Jung, D., & Choi, Y. (2021). Systematic review of machine learning applications in mining: Exploration, exploitation, and reclamation. *Minerals*, 11(2), 148. <https://doi.org/10.3390/min11020148>.
- Kaehr, D., & Gysi, M. (2021). TBO Trüllikon-1-1: Data Report, Dossier II, Core Photography, Wettingen, Switzerland. Retrieved from [https://nagra.ch/wp-content/uploads/2022/07/e\\_nab20-00920Dossier20II.pdf](https://nagra.ch/wp-content/uploads/2022/07/e_nab20-00920Dossier20II.pdf)
- Koeshidayatullah, A., Al-Azani, S., Baraboshkin, E. E., & Alfarraj, M. (2022). FaciesViT: Vision transformer for an improved core lithofacies prediction. *Frontiers in Earth Science*, 10, 992442. <https://doi.org/10.3389/feart.2022.992442>.
- Krupnik, D., & Khan, S. D. (2020). High-resolution hyperspectral mineral mapping: Case studies in the Edwards Limestone, Texas, USA and Sulfide-Rich Quartz Veins from the Ladakh Batholith, Northern Pakistan. *Minerals*, 10(11), 967. <https://doi.org/10.3390/min10110967>.
- Kupssinskü, L. S., Guimarães, T. T., Cardoso, M.d.B., Bachi, L., Zanotta, D., Souza, I., Falcão, A.X., Velloso, R.Q., Cazarin, C.L., Veronez, M.R., & Gonzaga, L. (2022). Hyperspectral data as a proxy for porosity estimation of carbonate rocks. *Australian Journal of Earth Sciences*, 69(6), 861–875. <https://doi.org/10.1080/08120099.2022.2046636>
- Lauper, B., Zimmerli, G. N., Jaeggi, D., Deplazes, G., Wohlwend, S., Rempfer, J., & Foubert, A. (2021). Quantification of lithological heterogeneity within opalinus clay: Toward a uniform subfacies classification scheme using a novel automated core image recognition tool. *Frontiers in Earth Science*. <https://doi.org/10.3389/feart.2021.645596>
- Li, J. X., Tsang, M., Zhong, R., Esterle, J., Pirona, C., Rajabi, M., & Chen, Z. (2023). Automatic coal mine roof rating calculation using machine learning. *International Journal of Coal Geology*, 274, 104292. <https://doi.org/10.1016/j.coal.2023.104292>.
- Lorenz, S., Seidel, P., Ghamisi, P., Zimmermann, R., Tusa, L., Khodadadzadeh, M., Contreras Acosta, I. C., & Gloaguen, R. (2019). Multi-sensor spectral imaging of geological samples: A data fusion approach using spatio-spectral feature extraction. *Sensors*. <https://doi.org/10.3390/s19122787>
- Marnat, S., & Becker, J.K. (2020). Petrophysical log analysis of deep and shallow boreholes: Methodology report, Wettingen, Switzerland. Retrieved from <https://nagra.ch/downloads/arbeitsbericht-nab-20-30/>
- Marnat, S., & Becker, J.K. (2021). TBO Trüllikon-1-1: Data Report, Dossier X, Petrophysical Log Analysis, Wettingen, Switzerland. Retrieved from [https://nagra.ch/wp-content/uploads/2022/08/e\\_nab20-009-Dossier-X.pdf](https://nagra.ch/wp-content/uploads/2022/08/e_nab20-009-Dossier-X.pdf)
- Mishra, A., Jyoti, A., & Haese, R. R. (2022). Irida: A machine learning based code for the automated derivation of site-specific rock type logs and their properties using Kimeleon colourlith image logs. *Applied Computing and Geosciences*, 16, 100102. <https://doi.org/10.1016/j.acags.2022.100102>
- Nagra. Swiss National Cooperative for the Disposal of Radioactive Waste. Tiefbohrung Trüllikon-1. <https://nagra.ch/wissensforum/tiefbohrung-truelikon-1/>. Accessed 23 Aug 2023
- Nagra Arbeitsbericht NAB 20-09, (2021). TBO Trüllikon-1-1: Data Report, Wettingen, Switzerland x Retrieved from <https://nagra.ch/en/downloads/arbeitsbericht-nab-20-09-2/>
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66. <https://doi.org/10.1109/TSMC.1979.4310076>
- Park, J., & Jeong, J. (2023). Assessment of the effectiveness of a convolutional autoencoder for digital image-based automated core logging. *Geoenergy Science and Engineering*, 227, 211802. <https://doi.org/10.1016/j.geoen.2023.211802>.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., & Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. In: Wallach, H., Larochelle, H., Beygelzimer, A., Alché-Buc, F., Fox, E., Garnett, R. (eds.) Advances in Neural Information Processing Systems 32. Curran Associates, Inc., Retrieved from <http://papers.nips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011).

- Scikit-learn: Machine Learning in Python. <https://dl.acm.org/doi/10.5555/1953048.2078195>
- Nagra. Tiefbohrung Trüllikon-1. <https://nagra.ch/wissensforum/tiefbohrung-truellikon-1/>. Accessed 28 Aug 2023.
- Shi, H., Ma, W., Xu, Z., & Lin, P. (2023). A novel integrated strategy of easy pruning, parameter searching, and re-parameterization for lightweight intelligent lithology identification. *Expert Systems with Applications*, 231, 120657. <https://doi.org/10.1016/j.eswa.2023.120657>.
- Solum, J. G., Auchter, N. C., Falivene, O., Cilona, A., Kleipool, L., & Zarian, P. (2022). Accelerating core characterization and interpretation through deep learning with an application to legacy data sets. *Interpretation*, 10(3), 71–83. <https://doi.org/10.1190/INT-2021-0189.1>.
- Trott, M., Leybourne, M., Hall, L., & Layton-Matthews, D. (2022). Random forest rock type classification with integration of geochemical and photographic data. *Applied Computing and Geosciences*, 15, 100090. <https://doi.org/10.1016/j.acags.2022.100090>.
- Tuşa, L., Khodadadzadeh, M., Contreras, C., Rafiezadeh Shahi, K., Fuchs, M., Gloaguen, R., & Gutzmer, J. (2020). Drill-core mineral abundance estimation using hyperspectral and high-resolution mineralogical data. *Remote Sensing*, 12(7), 1218. <https://doi.org/10.3390/rs12071218>.
- Woodhead, J., & Landry, M. (2021). Harnessing the power of artificial intelligence and machine learning in mineral exploration-opportunities and cautionary notes. *SEG Discovery*, 127, 19–31. <https://doi.org/10.5382/Geo-and-Mining-13>.
- Xu, P., Gan, C., Wang, L., & Cao, W. (2022). A Multi-feature Extraction-based Image Identification Method for Rock Debris in The Drilling Process. In: 2022 China Automation Congress (CAC), pp. 6666–6671. IEEE, Xiamen, China. <https://doi.org/10.1109/CAC57257.2022.10054959>.
- Xu, Z., Shi, H., Lin, P., & Liu, T. (2021). Integrated lithology identification based on images and elemental data from rocks. *Journal of Petroleum Science and Engineering*, 205, 108853. <https://doi.org/10.1016/j.petrol.2021.108853>.
- Xu, Z., Shi, H., Lin, P., & Ma, W. (2022). Intelligent on-site lithology identification based on deep learning of rock images and elemental data. *IEEE Geoscience and Remote Sensing Letters*, 19, 1–5. <https://doi.org/10.1109/LGRS.2022.3179623>.
- Zhou, Z., Yuan, H., & Cai, X. (2023). Rock thin section image identification based on convolutional neural networks of adaptive and second-order pooling methods. *Mathematics*, 11(5), 1245. <https://doi.org/10.3390/math11051245>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.