

# Segmentation of shale SEM images using machine learning

David G. Tang\*, Kyle T. Spikes  
The University of Texas at Austin

## Summary

The segmentation procedure in a digital rock physics workflow is often very challenging and time consuming. Here, we present an alternative method for quickly segmenting digital rock physics images that utilizes machine learning. Elemental SEM images of a core sample serve as inputs into a neural network. The network then outputs the probability of a pixel belonging to a certain class. Segmentation is implemented by choosing the class with the highest probability. This process allows for the uncertainty to be quantified in mineral phase identification. After training the algorithm, the rest of the image and subsequent images can be quickly segmented. We demonstrate the segmentation process on a shale sample with six different phases.

## Introduction

The microstructure of a rock controls its physical properties (Arns et al., 2002). Laboratory measurements on core often serve as the source of this microstructural information. However, accurately making these measurements can be difficult (Dvorkin et al., 2008). These measurements are crucial because they are used to calibrate rock physics models. One solution to this problem is to use digital rock physics (DRP). The DRP process entails imaging a core sample and then performing numerical simulations on the digitized rock. These non-destructive experiments can provide insight as to how the pore-scale features affect the macroscopic properties (Andr  et al., 2013). Various studies have demonstrated the viability of DRP on various rock properties including elasticity (Andr  et al., 2013 and Makarynska et al., 2008), permeability (Dvorkin et al., 2011), and electrical conductivity (Sun et al., 2013).

The process of identifying mineral phases in the image, or segmentation, is often the most difficult, yet important step in digital rock physics. Misidentifying phases can lead to erroneous results. The most widely used segmentation processes use either thresholding methods, statistical methods, region-growing and contour-evolution methods (Fredrich et al., 2014), or some combination of the aforementioned methods. Currently, studies using DRP have been performed on rocks with relatively few phases such as sandstone (e.g., Madonna et al., 2012). In a more complex rock such as shale, many phases (e.g., feldspar and quartz), can look visually similar to one another, making it difficult to discern among different minerals. Thresholding in this case would be difficult because the amplitude histograms

would contain significant overlap among different phases. Region-growing methods require the user to manually select and identify seed regions on which to grow. Additionally, traditional segmentation methods do not provide an easily quantifiable measure of uncertainty.

We instead propose that if SEM elemental images are collected during the imaging process, then these images can serve as inputs into a machine learning algorithm that segments the image. The machine learning approach is relatively fast to implement. Once trained, other images within the same sample can be automatically segmented. Machine learning methods also provide probabilities of minerals belonging to a certain class, which allows for an estimate of uncertainty. We demonstrate this process on a shale sample.

## Theory and Methodology

The first step in DRP is to obtain a high-resolution image of a sample. Many different imaging techniques can be used, including computed tomography (CT) and scanning electron microscopy (SEM). For shale samples, the higher resolving power of an SEM is necessary to capture nanometer scale features such as pores and organic material (Rodriguez et al., 2014). An SEM image results from focusing a beam of high energy electrons across each point in the sample. The interactions between the electron beam and the sample surface produces secondary and back-scattered electrons, which are used to image the structure of the sample. These interactions also produce X-rays, which when paired with an energy-dispersive X-ray spectroscopy (EDS) sensor, can be used to determine the elements present in a sample. The elemental composition is useful in identifying the minerals present in the sample.

Figure 1a-1g shows examples of the elemental SEM images collected for a shale sample. We applied a bilateral filter to these images to denoise the image while preserving the edges (Tomasi and Manduchi, 1998). Figure 1h shows an RGB composite image that was created by summing together the elemental images from silicon, calcium, and sodium. Minerals that would be difficult to distinguish and segment before (i.e., in a grayscale CT image) now separate out quite well. These images are used to identify six different mineral phases (calcite, feldspar, quartz, total organic content, clay, and pore space). For example, the large amount of carbon present in the carbon elemental image indicated that the mineral classification should be organic content. Similarly, the presence of aluminum, sodium, and silicon helped discriminate feldspar from quartz. Clearly, as

## Segmentation and machine learning

the RGB composite demonstrates, there exists some relationship among various elements and mineral type (i.e., chemical formula). However, the exact relationship between pixel amplitudes between different elemental images is unknown. Fortunately, supervised machine learning is a well suited approach to determine this relationship, allowing us to segment the image.

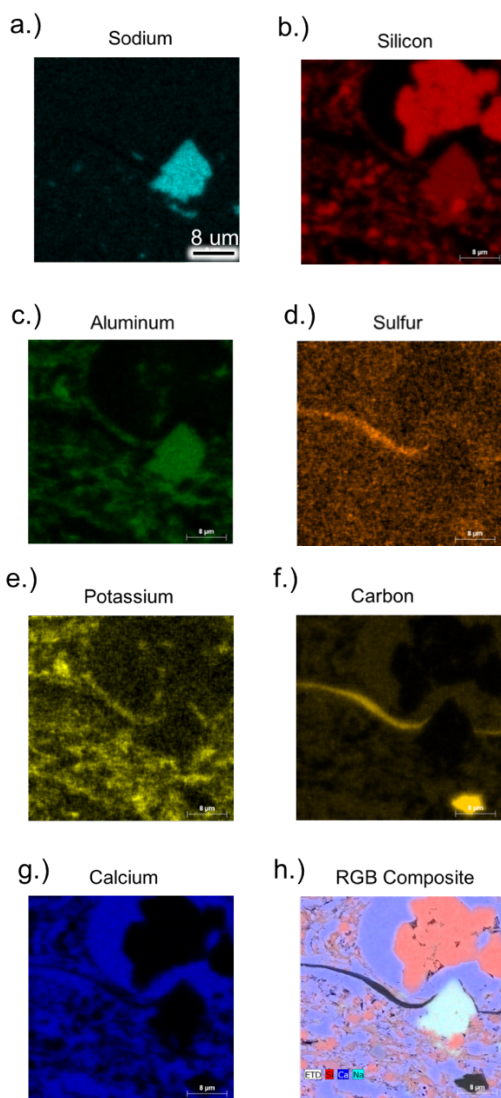


Figure 1: (a-g) Elemental SEM images of sodium, silicon, aluminum, sulfur, potassium, carbon, and calcium. (h) RGB composite image created by summing up elemental images. Six different mineral phases are identified from this image including: calcite (purple), feldspar (light blue), quartz (pink), total organic content (black), and pore space (black).

In supervised learning, the goal is to learn a model that predicts the output (e.g., mineral phase) from the input features (e.g., various elemental image amplitudes). The first step is to train the algorithm on a dataset where the labels of the output are known. In this study, we tested the effectiveness of a neural network on the segmentation process. Neural networks are modeled after the brain by building a network of interconnected neurons (Figure 2). The first layer of the neural network constitutes the input layer, into which the training features are fed. We then add hidden layers to form relationships between the input and output layers. In the hidden layer, the value of a node,  $z_j$ , is the weighted linear summation of the outputs,  $y_i$ , of the nodes that are connected to it from the previous layer (Equation 1). A non-linear activation function (e.g., hyperbolic tangent, ReLU) is then applied to  $z_j$  (Equation 2), allowing for the neural network to capture non-linear relationships (Schmidhuber, 2015). These outputs,  $y_j$ , serve as the inputs into the next layer. In the final layer, the softmax function is used to output the probabilities of the input belonging to each class. We then compare the output of the neural network to the known answers, and adjust the weights accordingly using the backpropagation algorithm (Rumelhart et al., 1986).

$$z_j = \sum_i w_{ij} y_i \quad (1)$$

$$y_j = f(z_j) \quad (2)$$

Where:

$z_j$  is the total input to node  $j$

$y_i$  is the output of the unit  $i$  connected to node  $j$

$w_{ij}$  is the weighting term

$y_j$  is the output from node  $j$  into the next layer

$f$  is a non-linear activation function

To create the training dataset, we manually selected pixels where the mineral phase could be easily identified. Each input node to the neural network corresponded to an elemental image intensity at the same pixel location across each image (Figure 2). The output layer value would be “1” for the correct mineral phase, and “0” for all other phases. A randomly chosen portion of this training data (15%) was set aside to serve as a test set. Each image was also normalized to have a mean of 0 and a standard deviation of 1. This avoided having any one feature from dominating the others. After training, the learned weights were then applied to the test set to evaluate the performance of the chosen parameters (e.g., number of nodes in each layer, number of hidden layers, regularization parameters). A grid search over a range of parameters was used to decide on the final model. The F1 scoring metric (Equation 3) was used to compare the various models. The F1 score is a weighted average between precision and recall. Precision is defined as the number of true positives versus the number of total positives (true and false), whereas recall is defined as the probability of detection (true positives versus false positives and false negatives). For perfect recall and precision, the F1 score is

## Segmentation and machine learning

equal to 1. The implementation for the training process and model selection made use of an open source Python package called scikit-learn (Pedregosa et al., 2011).

$$F1 = 2 * \frac{p*r}{p+r} \quad (3)$$

Where:

$p$  is the precision

$r$  is the recall

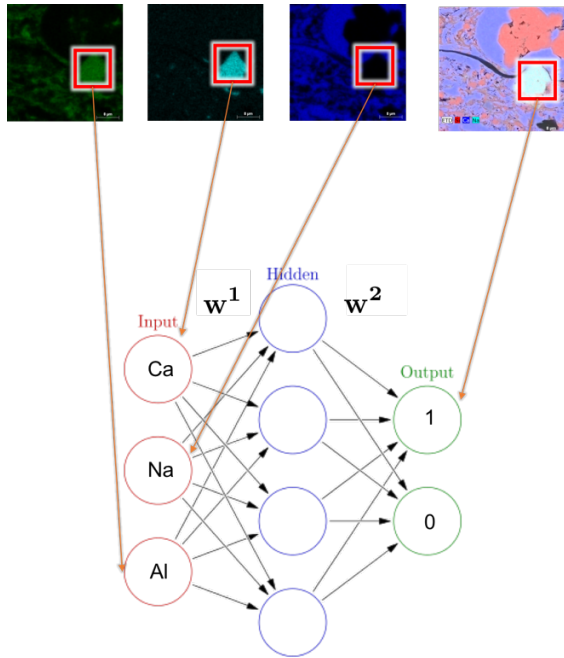


Figure 2: Example architecture of a neural network consisting of an input layer, a hidden layer, and an output layer. The input layer consists of elemental image pixel intensity values at a given location, and the output layer contains the probabilities of a given classification.  $w^1$  and  $w^2$  represent the weights that are learned from the training dataset through an optimization process. In this example, the same pixel is chosen across all elemental images for the input layer. The output layer should predict “1” in the node for calcite, and “0” for all other minerals.

## Results

It was difficult to distinguish between pore space and clay, so for the purposes of training the image, they were grouped into a single category. Figure 3 shows the confusion matrix for the optimal choice of parameters. A confusion matrix compares the prediction from the neural network against the true, manually picked label. The values along the diagonal of the matrix represent true positives. Off diagonal values represent false positives and false negatives. Overall, the neural network does very well on calcite, feldspar, quartz, and total organic content (TOC). When the true label is calcite (first row in the confusion matrix), the neural network

predicted calcite correctly every time. When the true label is clay/pore, the neural network incorrectly labeled 10 out of 78 pixels as quartz.

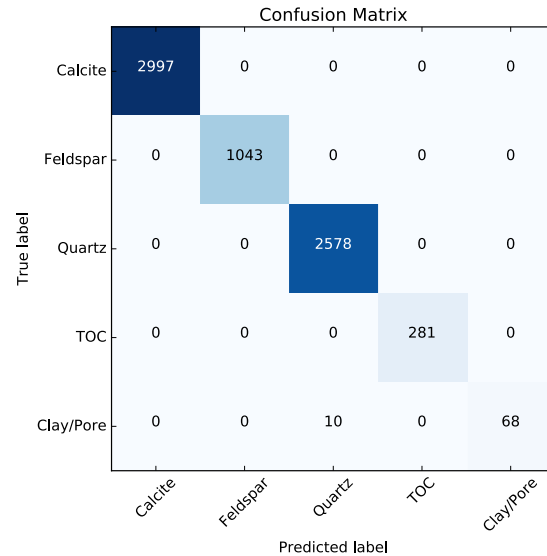


Figure 3: A confusion matrix. The Y axis is the true mineral phase of a pixel, whereas the X axis the phase the neural network predicts. The numbers along the diagonal of the matrix represent the number of pixels predicted correctly, whereas the off diagonal numbers represent false positives and false negatives. If there is a value in the second column of the first row, this would indicate that a pixel for calcite was predicted incorrectly as feldspar. If there is a value in the last row of the first column, this would indicate that a clay/pore pixel was mislabeled as calcite.

The output layer of the neural network contains the probabilities of a pixel belonging to a certain class (Figure 4). This provides a measure of uncertainty in the classification process. Red indicates high probability of belonging to a certain class, whereas blue indicates a low probability. A “transition zone” can be seen along grain boundaries, where the probabilities decrease very quickly from one phase to another. This is likely due to the limits in fully resolving the grain to grain contacts during the imaging process. We segment the image by choosing the class with the highest probability.

Figure 5 shows the overall results of the segmentation process. Overall, the neural network captures the major features of the image reasonably well. For example, the large quartz grain in the upper right corner has been correctly identified. In addition, the segmentation process also captures the streak of organic content and large feldspar grain. Figure 6 shows the weights learned from the training process applied to a different image from the same sample.

## Segmentation and machine learning

The results from Figure 6 demonstrate that the weights learned from the training data in Figure 5 can be applied to images elsewhere in the same sample.

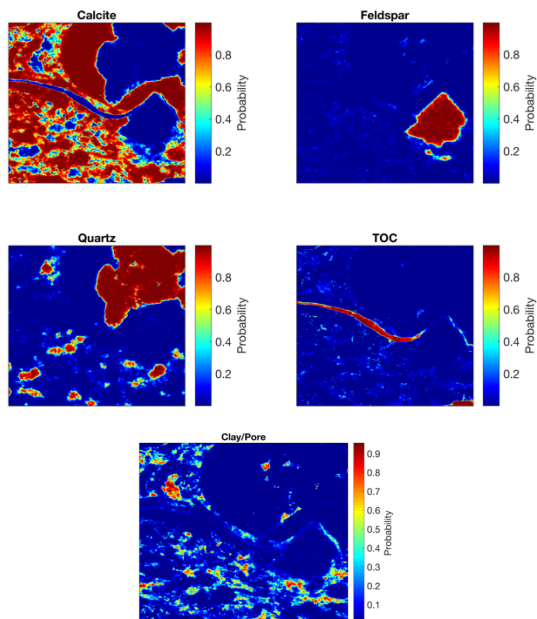


Figure 4: Probability of each class at a given pixel location. Red values indicate high probability of belonging to a certain class, whereas blue areas indicate low probability. Most mineral phases had greater than 90% probability when predicted, except for the clay/pore class.

## Conclusions

Machine learning is a promising method for segmenting SEM images with EDS elemental images. We can train the neural network algorithm with a couple of manually labeled examples. Once the initial training process is complete, it is relatively quick to automatically segment the rest of the image. We can also apply the learned weights to segment other images within the same sample. This allows for more images on which to run the DRP numerical simulations, allowing for a more representative solution for the physical properties of a sample. The probability map for additional images can indicate the presence of additional minerals that have yet to be identified.

This process, however, is still limited by the resolution of the image, along with the quality of the image itself (e.g., presence of streaks or other imaging artifacts). Like other segmentation methods, image processing techniques will have to be used if such features are present. The main challenge we encountered was properly identifying the minerals used in training the algorithm. In the future, additional effort must be put in to properly differentiating

between clay and pore. Future work will also test the effectiveness of other machine learning algorithms and additional image processing techniques.

## Acknowledgements

The authors would like to thank the sponsors of the EDGER Forum for their support.

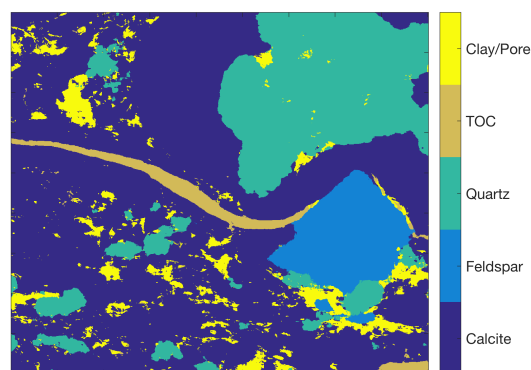


Figure 5: Segmentation result of the SEM image using a neural network. Colors in the image correspond with various mineral types. The major features of the image have been correctly captured, including the feldspar crystal embedded within quartz, and the streak of organic material.

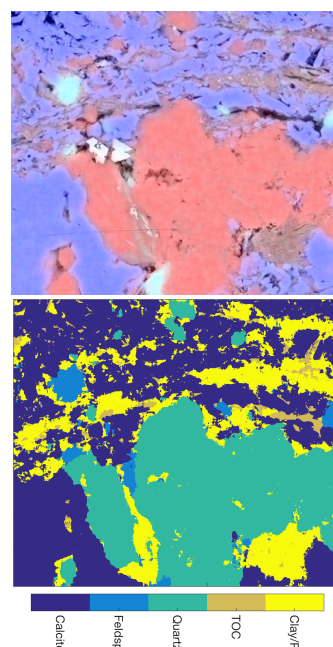


Figure 6: (Top) RGB composite from a different SEM image within the same core sample. (Bottom) Segmentation result after applying the weights learned from Figure 5.

## EDITED REFERENCES

Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2017 SEG Technical Program Expanded Abstracts have been copyedited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

## REFERENCES

- Andra, H., N. Combaret, J. Dvorkin, E. Glatt, J. Han, M. Kabel, Y. Keehm, F. Krzikalla, M. Lee, C. Madonna, M. Marsh, T. Mukerji, E. Saenger, R. Sain, N. Saxena, S. Ricker, A. Wiegmann, and X. Zhan, 2013, Digital rock physics benchmarks — Part I: Imaging and segmentation: *Computers and Geoscience*, **50**, 24–32, <http://doi.org/10.1016/j.cageo.2012.09.005>.
- Arns, C. H., M. A. Knackstedt, W. V. Pinczewski, and E. J. Garboczi, 2002, Computation of linear elastic properties from microtomographic images: Methodology and agreement between theory and experiment: *Geophysics*, **67**, 1396, <http://doi.org/10.1190/1.1512785>.
- Dvorkin, J., M. Armbruster, C. Baldwin, Q. Fang, N. Derzhi, C. Gomez, B. Nur, A. Nur, and Y. Mu, 2008, The future of rock physics: Computational methods vs. lab testing: *First Break*, **26**, 64–68.
- Dvorkin, J., N. Derzhi, E. Diaz, and Q. Fang, 2011, Relevance of computational rock physics: *Geophysics*, **76**, no. 5, E141–E153, <http://doi.org/10.1190/geo2010-0352.1>.
- Fredrich, J. T., D. L. Lakshtanov, N. M. Lane, E. B. Liu, C. S. Natarajan, D. M. Ni, and J. J. Toms, 2014, Digital rocks: Developing an emerging technology through to a proven capability deployed in the business: *Society of Petroleum Engineers*, <http://doi.org/10.2118/170752-MS>.
- Madonna, C., B. S. G. Almqvist, and E. H. Saenger, 2012, Digital rock physics: numerical prediction of pressure-dependent ultrasonic velocities using micro-CT imaging: *Geophysical Journal International*, **189**, 1475–1482, <http://doi.org/10.1111/j.1365-246X.2012.05437.x>.
- Makarynska, D., B. Gurevich, R. Ciz, C. H. Arns, and M. A. Knackstedt, 2008, Finite element modelling of the effective elastic properties of partially saturated rocks: *Computers and Geoscience*, **34**, 647–657, <http://doi.org/10.1016/j.cageo.2007.06.009>.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, 2011, Scikit-learn: Machine learning in Python: *Journal of Machine Learning Research*, **12**, 2825–2830.
- Rodriguez, R., D. Crandall, X. Song, C. Verba, and D. Soeder, 2014, Imaging techniques for analyzing shale pores and minerals: NETL-TRS-6-2014, NETL Technical Report Series; U. S. Department of Energy, National Energy Technology Laboratory.
- Rumelhart, D. E., G. E. Hinton, and R. J. Williams, 1986, Learning representations by back-propagating errors: *Nature*, **323**, 533–536; <http://doi.org/10.1038/323533a0>.
- Schmidhuber, J., 2015, Deep learning in neural networks: An overview: *Neural Networks*, **61**, 85–117, <http://doi.org/10.1016/j.neunet.2014.09.003>.
- Sun, J., J. Zhao, X. Liu, H. Chen, L. Jiang, and J. Zhang, 2014, Pore- scale analysis of electrical properties in thinly bedded rock using digital rock physics: *Journal of Geophysics and Engineering*, **11**, <http://doi.org/10.1088/1742-2132/11/5/055008>.
- Tomasi, C., and R. Manduchi, 1998, Bilateral filtering for gray and color images: Sixth International Conference on Computer Vision, Bombay, 839–846, <http://doi.org/10.1109/ICCV.1998.710815>.