



MudrockNet: Semantic segmentation of mudrock SEM images through deep learning

Abhishek Bihani^{a,*}, Hugh Daigle^a, Javier E. Santos^a, Christopher Landry^b, Maša Prodanović^a, Kitty Milliken^c

^a Hildebrand Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, Austin, TX, USA

^b Center for Subsurface Energy and the Environment, University of Texas at Austin, Austin, TX, USA

^c Bureau of Economic Geology, University of Texas at Austin, Austin, TX, USA

ARTICLE INFO

Keywords:

Deep learning
Machine learning
Neural networks
Mudrocks
Digital rock physics
Image analysis
Image segmentation
Convolutional neural network

ABSTRACT

Segmentation and analysis of individual pores and grains of mudrocks from scanning electron microscope images is non-trivial because of imaging artifacts, variation in pixel grayscale values across images, and overlaps in grayscale values among different physical features such as silt grains, clay grains and pores, which make identifications difficult. Moreover, because grains and pores often have overlapping grayscale values, direct application of threshold-based segmentation techniques is not sufficient. Recent advances in the field of computer vision have made it easier and faster to segment images and identify multiple occurrences of such features in an image, provided that ground-truth data for training the algorithm are available. Here we propose a deep learning SEM image segmentation model, MudrockNet based on Google's DeepLab-v3+ architecture implemented with the TensorFlow library. The ground-truth data were obtained from an image-processing workflow applied to scanning electron microscope images of uncemented muds from the Kumano Basin offshore Japan at depths <1.1 km. The trained deep learning model obtained a pixel-accuracy > 90%, and predictions for the test data obtained a mean intersection over union (IoU) of 0.6663 for silt grains, 0.7797 for clay grains and 0.6751 for pores. We also compared our model with the random forest classifier using trainable Weka segmentation in ImageJ, and it was observed that MudrockNet gave better predictions for silt grains, clay grains and pores in most cases. The size, concentration, and spatial arrangement of the silt and clay grains can affect the petrophysical properties of a mudrock, and an automated method to accurately identify the different grains and pores in mudrocks can help improve reservoir and seal characterization for petroleum exploration and anthropogenic waste sequestration.

1. Introduction

The dominant sedimentary rocks on Earth, mudrocks are composed of silt- and clay-size particles, that typically include clay minerals, quartz, feldspar, and carbonate, that are smaller than 63 µm (Macquaker and Adams, 2003; Lazar et al., 2015b). These lithologies are important as capillary seals over hydrocarbon accumulations (Schowalter, 1979; Schlömer and Kroos, 1997), as caprocks for carbon capture and storage (Li et al., 2006), and as unconventional oil and gas reservoirs (e.g., Bustin et al., 2008). At higher concentration, silt grains provide the support framework for the mudrocks, and at lower concentration they shelter the larger pores from compaction (Oertel, 1983; Yang and Aplin, 2007; Desbois et al., 2009; Schneider et al., 2011; Pommer and Milliken,

2015). The size, concentration, and spatial arrangement of silt- and clay-size grains affect the petrophysical properties of mudrocks, thereby affecting fluid transport behavior (Potter et al., 2005; Lazar et al., 2015a; Bihani and Daigle, 2019). Hence, scanning electron microscope (SEM) images of mudrock samples can serve an effective tool for improved reservoir and caprock characterization.

A typical workflow for filtering and segmenting microstructural images of rocks consists of steps to identify the individual features, each requiring the input of the user, for example fixing the segmentation threshold pixel values for each image. However, with recent advances in machine learning, it is possible to reduce the number of steps and the degree of user intervention required for feature detection or image segmentation. Deep learning has also been applied by a number of

* Corresponding author.

E-mail address: abihani@utexas.edu (A. Bihani).

researchers to segment different rock image types for aiding petrophysical analysis, including sandstone SEM images (Karimpouli and Tahmasebi, 2019; Niu et al., 2020), microCT images (Marmo et al., 2005; Chauhan et al., 2016; Varfolomeev et al., 2019; Guntoro et al., 2019; Wang et al., 2019b; Anderson et al., 2020), cores (Ivchenko et al., 2018; Baraboshkin et al., 2018; Alzubaidi et al., 2021) or thin section images (Izadi et al., 2017; Dong et al., 2019; de Lima et al., 2020; Tang et al., 2020). Other researchers (Tang and Spikes, 2017; Andrew, 2018; Tian and Daigle, 2019; Wu et al., 2019; Knaup et al., 2019; Knaup, 2020; Z. Chen et al., 2020; Misra et al., 2020) have used machine learning techniques successfully for feature detection in high-resolution SEM images of shales or mudrocks.

Tian and Daigle (2019) used automated object detection algorithms on high-resolution backscattered electron images and energy-dispersive X-ray spectroscopy (EDX) images to characterize the location and size of microfractures and their preferential association with particular minerals in shales. However, the object detection algorithm is only suitable for recognizing features which have specific shapes and are limited to only a few occurrences in an image. If the algorithm is applied to an SEM image containing hundreds of pores and grains, the bounding boxes cannot capture all the features accurately. Tang and Spikes (2017) used high-resolution EDX images of shale samples as inputs in a neural network to detect calcite, feldspar, quartz, kerogen, and clay/pore, but were unable to distinguish between clays and pores. C. Li et al. (2021) used SEM-EDX images of Bakken shale to compare shallow machine learning and deep learning (U-Net) models and found that the random forest model outperformed the U-Net model predictions due to an imbalanced dataset. However, they found that the U-Net model was able to predict the isolated smaller particles better than the other machine learning models. Knaup et al. (2019) used a U-Net model to segment pores, organic matter, and mineral content from SEM images obtained from the Woodford shale and were able to achieve intersection over union (IoU) results over 90%. Their study used SEM-EDX color maps in addition to grayscale SEM images for segmentation. Since such methods exploit the differences in mineral colors in similarly sized grains for grain recognition, they cannot be applied for segmentation when only grayscale SEM images are available. Andrew (2018) used machine learning to segment grayscale shale data into five classes (porosity, organic phase, quartz, calcite, and pyrite) using differences in grayscale intensity and textural contrasts and investigated the effect of noise on model performance. He found that machine learning outperformed traditional techniques like Otsu and seeded watershed region growing at a wide range of noise levels. Wu et al. (2019) used feature extraction followed by the gradient boosting and random forest classifiers for segmentation of high-resolution grayscale shale SEM images. They found that the random forest classifier gave the best results for detecting the four types of features; pores/fractures, rock matrix (quartz, clay, calcite), pyrites and kerogen. However, they did not identify the different types of grains. Z. Chen et al. (2020) used a U-Net model for segmenting large mineral (silt size) grains, clay aggregates and organic matter from SEM images of Duvernay shale samples from Western Canada. They prepared 8000 random image slices from four labeled large image tiles and were able to achieve an IoU of 0.917 on the test dataset.

Thus, while researchers have used deep learning for segmentation of SEM images of shales and mudrocks, identification of individual pores and grains of specific types and sizes from grayscale SEM images which have a variation in texture and maturity remains an active area of research.

2. Background

In the field of computer vision, the task of assigning labels pixel by pixel in an image is termed semantic segmentation (L. C. Chen et al., 2018). Semantic segmentation requires identifying irregular object outlines in an image, and therefore, has stricter accuracy requirements

than conducting simple image-level classification or bounding-box based object detection in an image (Liu et al., 2019). We propose using it to detect pores as well as particular types of grains (e.g., silt) in the images, with sufficient amount of data from raw images and ground truth are provided for training. Moreover, whereas semantic segmentation can identify different object classes (pores and silt grains), it is possible to enhance the segmentation to instance-level segmentation, which entails detecting each instance of grain and pore in the image and assigning a unique identifier for each instance. This can also be done in a post-processing step, wherein all the unconnected grains and pores that have been detected are assigned a unique identifier to allow further analysis.

Convolutional neural networks (CNN) use small spatial filters that are reused throughout the domain which makes them effective for image segmentation. They have different types of layers whose parameters can be trained by data for performing particular tasks (Liu et al., 2019). Long et al. (2015) introduced fully convolutional neural networks to image segmentation by replacing fully connected layers (all neurons in a layer connected to all neurons in previous layer) with fully convolutional layers (neurons in a layer are connected to a small region of the previous layer). As shown in Fig. 1A, an input image passed through multiple convolutional layers of the encoder (black arrows), undergoes successive reductions in the height and width of the convolved image representations (feature maps). Thus, the encoder converts the input data into a feature map by convolving the original image with a smaller kernel (filter) to identify important features in the image which would not be easily identifiable to humans (Goodfellow et al., 2016). The first output feature map is then used as the input in the next convolution with another kernel and the process continues until reaching the bottleneck layer (blue dotted line), which has the highest compressed representation of the input data. Thereafter, a deconvolution network (decoder) can be used to up-sample the image through interpolation across successive layers to recover the feature map back to the original image size (Noh et al., 2015).

Ronneberger et al. (2015) built on this and modified it to create the U-net architecture, which worked with fewer training images but improved the results. It uses a symmetric encoder-decoder architecture with a large number of feature channels in the encoder part, which helps in preserving the information passed to the high-resolution layers, and the decoder layers are connected to the corresponding encoder layers to improve the resolution of predicted image. Since that time, further advances in semantic segmentation have occurred, including DeepLab, which is a state-of-the-art supervised segmentation model created and released as open source by Google (Chen et al., 2014). It was first released as DeepLab-v1 (Chen et al., 2014) and was followed by multiple improvements; DeepLab-v2 (Chen et al., 2017a), DeepLab-v3, (Chen et al., 2017b), DeepLab-v3+ (L. C. Chen et al., 2018).

MudrockNet is based on DeepLab-v3+, which also uses an encoding-decoding architecture to extract information using convolutional layers. However, whereas applying an encoder-decoder architecture can decrease computational times for neural network training and capture sharp object boundaries, the repeated down-sampling may miss rich semantic information when the resolution of training layers is smaller than the features to be identified (Chen et al., 2018). The architecture can also cause problems while recognizing objects at multiple scales. For example, if an SEM image of a mudrock has pores of multiple sizes, the smallest pores consisting of just few pixels may not be identified by models using conventional encoding-decoding architecture (Fig. 1A). The DeepLab-v3+ architecture overcomes these issues by combining two techniques, atrous convolutions and spatial pyramid pooling (atrous spatial pyramid pooling) (Fig. 1B). Applying parallel atrous (dilated) convolutions effectively increases the field of view, incorporating multi-scale context, because the atrous filters prevent the down-sampling operations in the last few layers by inserting holes between the filter weights (Chen et al., 2018). The spatial pyramid pooling uses versions of the same image at multiple scales which are later

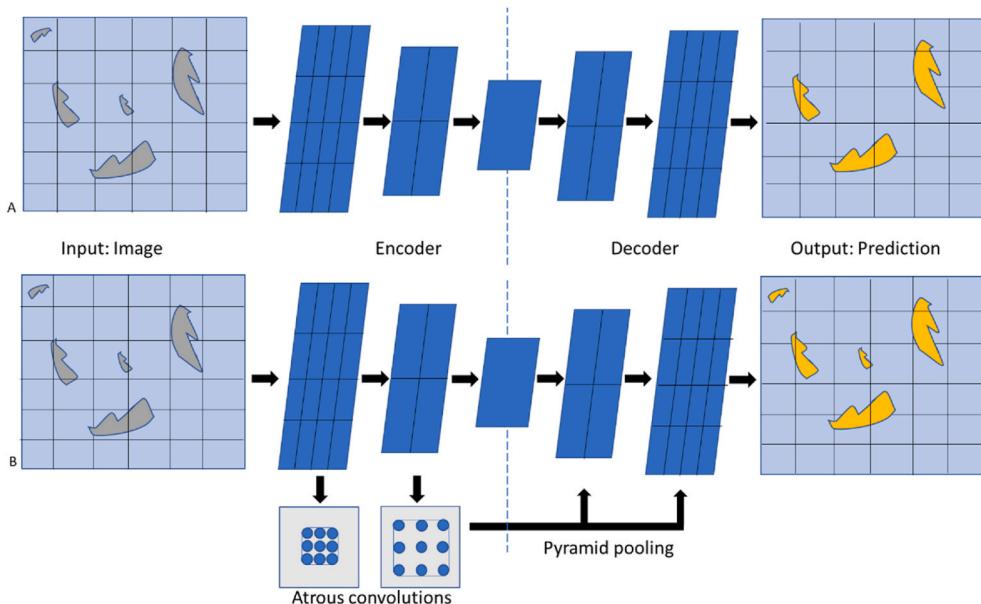


Fig. 1. Comparison of A) Normal encoder-decoder architectures, B) DeepLab-v3+ architecture with atrous convolutions and pyramid pooling. Modified after [Chen et al. \(2018\)](#). Convolutions between successive feature maps are shown by black arrows, and grid lines indicate the number of pixels in the particular image feature map. It can be observed that DeepLab-v3+ architecture is better at segmenting smaller sized features.

combined, and when used along with atrous convolutions and encoder-decoder modules can provide better results for semantic segmentation while retaining the spatial resolution.

Therefore, we trained the MudrockNet model to identify pores, silt-size and clay-size grains from the SEM images obtained at different depths with varying silt fractions. This model can be used for segmentation of other mudrock SEM images, and the individual pores and grains can be used for further analysis.

3. Materials and methods

Fig. 2 describes the workflow for semantic segmentation using MudrockNet and result comparison. The process consists of three parts: data pre-processing, model training and testing, and data post-processing.

The SEM images for the model were obtained from uncemented mud samples in the Kumano Basin, which is a large forearc basin in the

Nankai trough, offshore Japan ([Moore et al., 2013](#)). These core samples were acquired at Site C0002 during Integrated Ocean Drilling Program (IODP) expeditions 315 and 338 at depths <1.1 km below the sea floor ([Milliken et al., 2016](#)). Forty-nine SEM images from five core samples at different depths were used for the study. The sample surfaces for pore imaging were prepared by Ar-ion cross-section polishing and were coated in 6 nm iridium in a Leica EM ACE600 to minimize surface charging effects. The images were produced by secondary electron detection using the FEI Nova-NanoSEM 430 scanning electron microscope by a mixed signal of backscattered electron and secondary electron detection to reduce charging effects ([Nole et al., 2016](#)). All the images were scanned at a machine magnification of either 15,000 \times or 40,000 \times resulting in horizontal field widths (HFW), i.e. image widths of 20 μm or 7.5 μm respectively.

An image processing workflow was applied to SEM images following [Landry et al. \(2017\)](#) to obtain the ground truth data for the model. This is necessary since simple threshold-based segmentation techniques

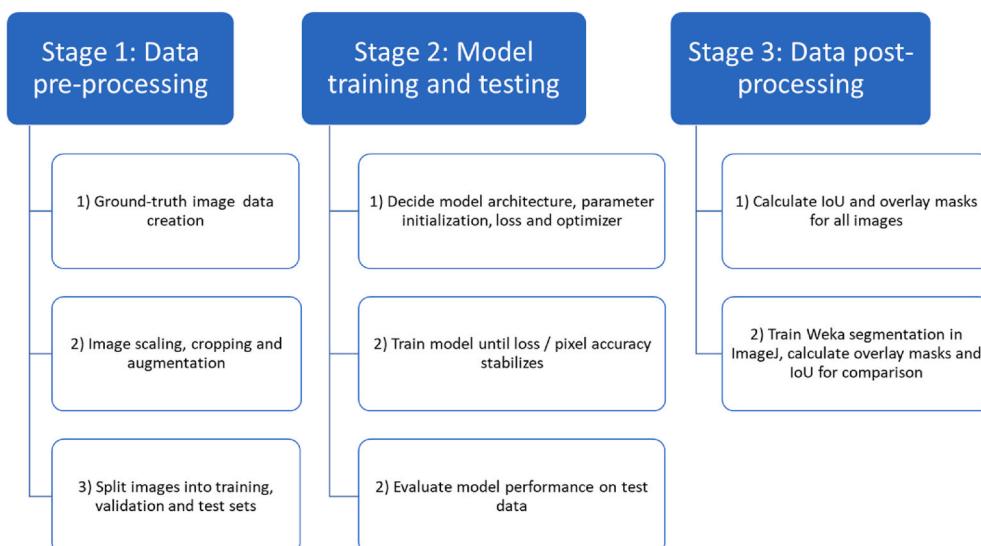


Fig. 2. Workflow for semantic segmentation and comparison of results.

without filtering cannot properly separate grains and pores as they can have overlapping grayscale values. For example, on applying simple segmentation to the image in Fig. 3A results in wrongly predicting darker grain pixels as pores (Fig. 3B). Therefore, the first step in the workflow was to apply a median filter to smooth the image to the targeted spatial scale of features to be segmented, followed by a combination of top-hat and bottom-hat filter to increase the local contrast. After smoothing and contrast enhancement, a grayscale threshold value is applied to separate the image into pores and grains without overlaps (Fig. 3C). The image segmentation was done at different spatial scales, to capture pores of all sizes and the final image was merged together, similar to pyramid pooling. The silt-size grains (equivalent circular diameter $>2 \mu\text{m}$) were identified by eroding and dilating the grain component multiple times as specified by the user, and the segmented images were reviewed and corrected if required. Segmented silt size grains (hereafter referred to as silt grains), clay size grains (hereafter referred to as clay grains) and pores are shown in Fig. 3D. This dataset of all original and segmented images (silt, clay, pores) used in this study using this conventional segmentation algorithm is available at Bihani et al. (2020).

To improve model training, all images are scaled to the same magnification of $15,000\times$. Thereafter, to maintain consistent image size and increase the total number of training images, we divided the SEM images into multiple equal parts, each equal to the height and width of the smallest scaled SEM image (400×343 pixels) as shown for an example image in Fig. 4A, B, C. Horizontal (Fig. 4D) and vertical (Fig. 4E) flipping of the images was carried out for data augmentation by

following the methods of Tian and Daigle (2018). Additionally, randomly selected images were subjected to color space transformations (Shorten and Khoshgoftaar, 2019), like modifying the image contrast (Fig. 4F) or modifying the image brightness (Fig. 4G), therefore, resulting in a total of 2239 images. The images (original and corresponding ground truth) were then split into training (~70%), validation (~15%), and test (~15%) datasets.

The MudrockNet model was trained using a NVIDIA GeForce GTX 1070 GPU with 8 GB memory for the three classes: silt, pore, and clay. Transfer learning was performed from a pre-trained ResNet-101 model (on ImageNet) and training was continued for 50 epochs until the loss became constant. We used the sparse softmax cross entropy loss, weight decay, and the momentum optimizer. The particular loss function is suitable for mutually exclusive classification of several classes (Wang et al., 2019a), and the inclusion of weight decay regularization helps prevent overfitting by penalizing large weight updates (Schmidhuber, 2015). The addition of momentum to the optimizer helps the stochastic gradient descent navigation during training by adding a fraction of the direction from the previous steps which reduces irrelevant oscillations (Sutskever et al., 2013). The training was evaluated by two metrics: pixel accuracy and IoU for pores, silt, and clay grains. IoU or the Jaccard index (Fig. 5) is defined as the area of intersection (common pixels) between the prediction and the ground truth divided by the area of their union (pixels present in both images) (Tian and Daigle, 2018). The pixel-accuracy is defined as the number of pixels labeled correctly (True Positive + True Negative) over the total number of pixels in the image (True Positive + True Negative + False Positive + False Negative). The

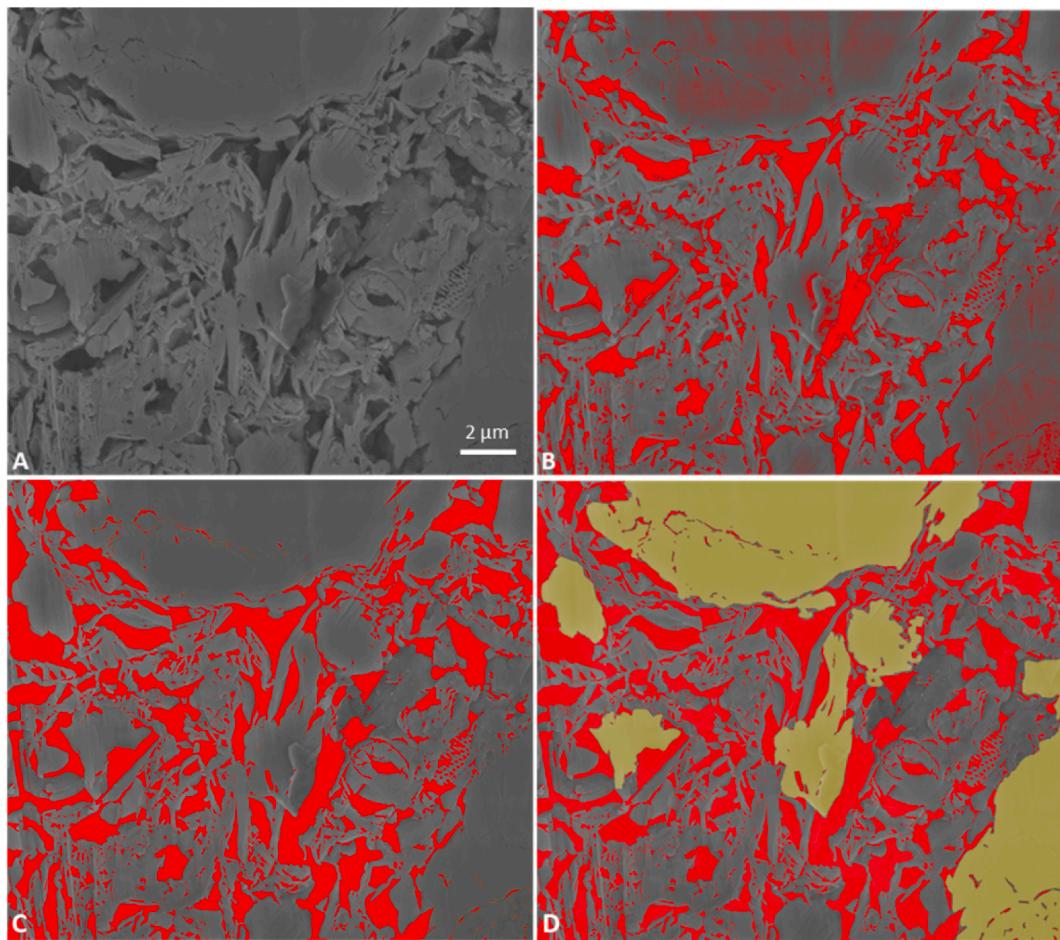


Fig. 3. SEM images: A) Original; B) Segmentation without filtering; C) Segmentation after filtering (predicted pores are shown in red), D) Segmentation with silt grains (yellow), clay grains (gray), and pores (red) separated. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

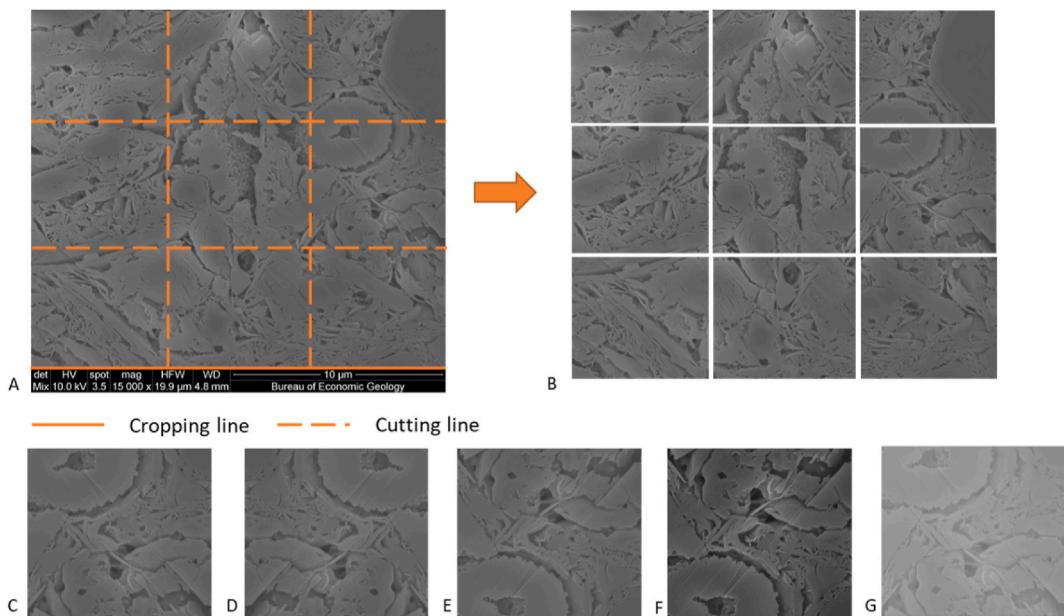


Fig. 4. A) Original image (magnification 15,000 \times), B) After cutting and cropping, C) Sample input image, D) Horizontally flipped image, E) Vertically flipped image, F) Contrast increased image, G) Brightness increased image.

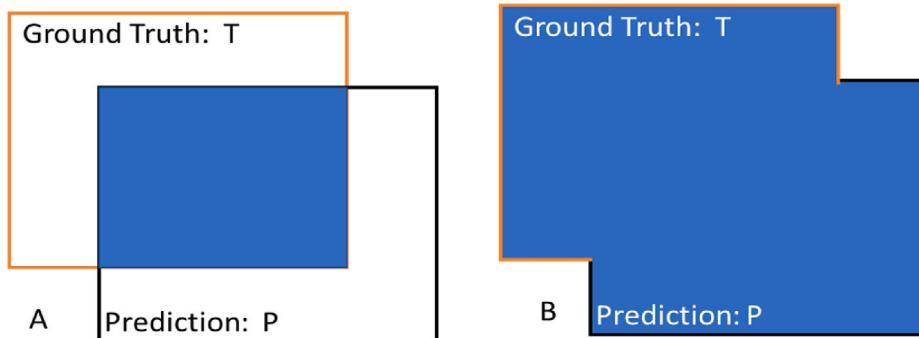


Fig. 5. A) Intersection area: ground truth and prediction, B) Union area: ground truth and prediction.

IoU is preferred over pixel accuracy since it is not affected by class imbalance (eg. in Fig. 3D, clay grain pixels > pore pixels > silt grain pixels) while still providing a metric per class (Rahman and Wang, 2016). Thereafter, the trained MudrockNet was tested on the withheld test dataset for an unbiased evaluation of its ability to predict the features in different cases. We found that using transfer learning gave better mean IoU results on the test data (silt grains: 0.6663, clay grains: 0.7797, pores: 0.6751), when compared with the results for a model trained from scratch only in our data (silt grains: 0.4671, clay grains: 0.7136, pores: 0.6592). The predicted segmentations are compared with ground-truth data by calculating the mean IoU for silt, clay and pores for each image, and typically using a low threshold, IoU values > 0.5 can be considered true positive (Y. Li et al., 2017). The overlay masks of the MudrockNet predictions and the ground truth (conventional segmentation algorithm) predictions over the raw image are also calculated for visual inspection.

The trainable Weka segmentation with a random forest classifier in ImageJ (Arganda-Carreras et al., 2017) was also used to perform image segmentation, to compare the trained model's performance with other machine-learning methods. Random forest is an ensemble tree-based learning algorithm that averages multiple decision trees from randomly selected subsets of training data to build class predictions (Liaw and Wiener, 2002). The algorithm was selected since it has been successfully applied for problems in rock SEM images (Maitre et al.,

2019; Wu et al., 2019), and being an ensemble method helps increase the accuracy while controlling overfitting. The Weka model was trained on 50 images from the training set with multiple samples for each class (silt, clay, pore). The training used edge-detection filters like the Sobel and Hessian filter along with a Fast Random Forest classifier, which is a multi-threaded version of the random forest classifier using 200 trees and 2 random features per node. The trained Weka classifier was then applied on the test data to compare predictions using overlay masks and IoU values.

4. Results and discussion

Fig. 6A shows the training and testing metrics of the MudrockNet model. The network training was stopped after 50 epochs, once the training and validation loss plateau (values 13.16 and 13.74 respectively), and the training and validation pixel-accuracy reached a plateau (values 0.9317 and 0.9107 respectively). The mean IoU values for the individual classes (silt, clay, and pores) for the training, validation and test datasets are shown in Table 1, and IoU frequency distributions for the three classes are seen in Fig. 6B, C, D.

In Table 1, the predicted pixel accuracy values are likely higher than the IoU values due to class imbalance, and so IoU values are preferred as a metric over pixel accuracy for further analysis. It can be seen from the comparison that the test and validation IoU values and are only slightly

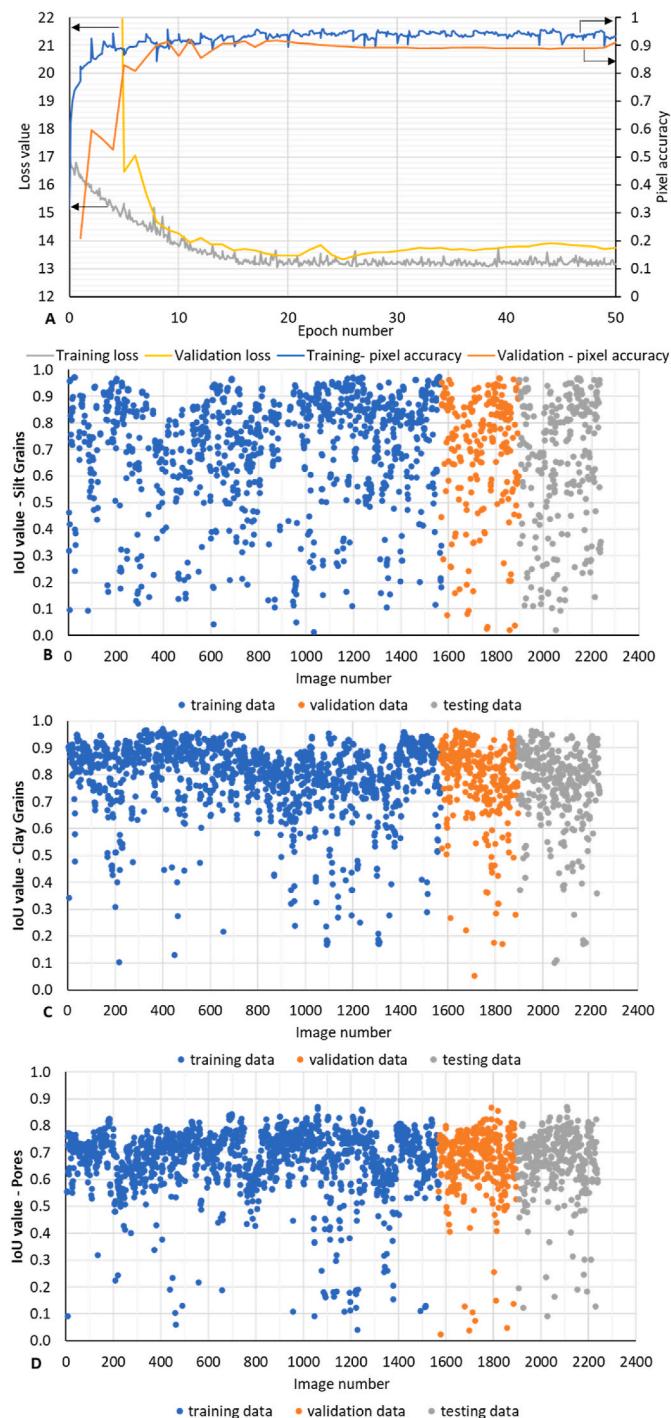


Fig. 6. Training and validation metrics for MudrockNet model: A) Training metrics (black arrows point to corresponding axes), B) Silt grain IoU values for individual images, C) Clay grain IoU values for individual images, D) Pore IoU values for individual images.

Table 1

Mean Intersection over union (IoU) and Pixel Accuracy values according to class.

Mean Values	Type	Training set	Validation set	Test set
Silt grains	IoU	0.7299	0.7070	0.6663
	Pixel Accuracy	0.9559	0.9513	0.9436
Clay grains	IoU	0.8084	0.7890	0.7797
	Pixel Accuracy	0.9311	0.9254	0.9195
Pores	IoU	0.6842	0.6727	0.6751
	Pixel Accuracy	0.9540	0.9528	0.9510

lower than the training set (low variance) and, thus, the model is able to generalize the mud segmentation problem. This indicates that the model can suitably identify the silt grains, clay grains and pores in mudrock SEM image sets that are of similar quality and size to those used here.

While a majority of the grain and pore IoU values in Fig. 6B, C and D are larger than 0.5 (low threshold for true positive), the silt and clay grain IoUs (Fig. 6B and C) have a larger amount of scatter than the pore IoUs (Fig. 6D), especially in the validation and test data. This may mean that MudrockNet is better trained to predict pores than the grains in new SEM images, which is possible due to various factors like the variation in silt grain size, greater number of pores than silt grains in the data, and due to the overlapping pixel values of silt grains and clay grains, which can lead to errors, especially for smaller silt grains. While data augmentation techniques like changing the contrast or brightness of images have been implemented to improve predictions for SEM images with too low or high brightness, the smaller number of such images in the dataset can also be a reason for reduced IoU values.

Fig. 7 shows the overlay mask of ground truth (conventional segmentation algorithm data) (A), MudrockNet model predictions (B), and trainable Weka model predictions (C), on five selected SEM images from the test set. The silt grains are in red, pores in green, clay in gray color, and the truth images show a scale bar for reference. Other comparisons are available online at https://github.com/abhishekdbihani/MudrockNet/tree/main/dataset/data/masks_test.

Table 2 shows the comparison of the IoU values for silt grains and pores by both the methods (MudrockNet and Weka model) for images shown in Fig. 7. In Fig. 7-1, we can see that the MudrockNet model predictions (1B) for both the silt size grains (red) and the pores of different sizes (green) match the ground truth image from conventional segmentation (1A) well, as confirmed from the IoU values. While the Weka model (1C) is also able to detect the silt grains, it is not able to adequately differentiate between silt and clay grains, as seen in the smaller red patches in the clay. In Fig. 7-2, MudrockNet (2A) is able to predict the entire silt grain (red in center) comparable to the ground truth image (2B). The Weka model (2C) is able to detect the larger structures, but overpredicts the silt grains and pores. Similar behavior is seen in Fig. 7-3, where the predictions from MudrockNet (3A) are very similar to the ground truth (3B), with high IoU values, while the Weka model's (3C) silt predictions underpredict the actual silt grain pixels and gives false positives. In Fig. 7-4, MudrockNet (4B) predicts silt grains and pores as seen in the ground truth image (4A), while the Weka model (4C) tends to predict silt pixels (red), in the clay grains and in the large central pore. Fig. 7-5 shows that the MudrockNet model (5B) gives low IoU values with overprediction of silt grains and underprediction or pores when compared to the ground truth (5A). This is possibly due to a combination of factors like low image brightness and low contrast between grains and pores which can lead to the neural network predicting clay grains and surrounding pores as silt grains. While the Weka model (5C) predictions have higher IoUs than the MudrockNet, it also overpredicts pores and underpredicts silt grains possibly due to the low brightness and contrast of the particular image.

Fig. 8 shows the comparison of ground truth (A), and MudrockNet model predictions (B) for a large test SEM image (magnification 15,000 \times) of 2048 \times 1767 pixels. It can be seen that despite the image being larger than the images used for training, MudrockNet can identify the pores (green), silt grains (red) and clay grains (gray) with a IoUs of 0.7819 (silt), 0.8365 (clay) and 0.6636 (pore). While there are some differences since MudrockNet may identify some smaller grains as silt size due to the larger size of the image, Fig. 8 shows the capability of the model to segment SEM images larger than it was trained on. Additionally, while the prediction time for the conventional segmentation algorithm (with default values) took 7.767 s, the MudrockNet model's prediction took only 0.218 s, which is faster by an order of magnitude.

Thus, from the overall trends observed in the example images in Figs. 7 and 8, and the corresponding IoU values in Table 2, we can infer that while the conventional filtering and segmentation algorithm used

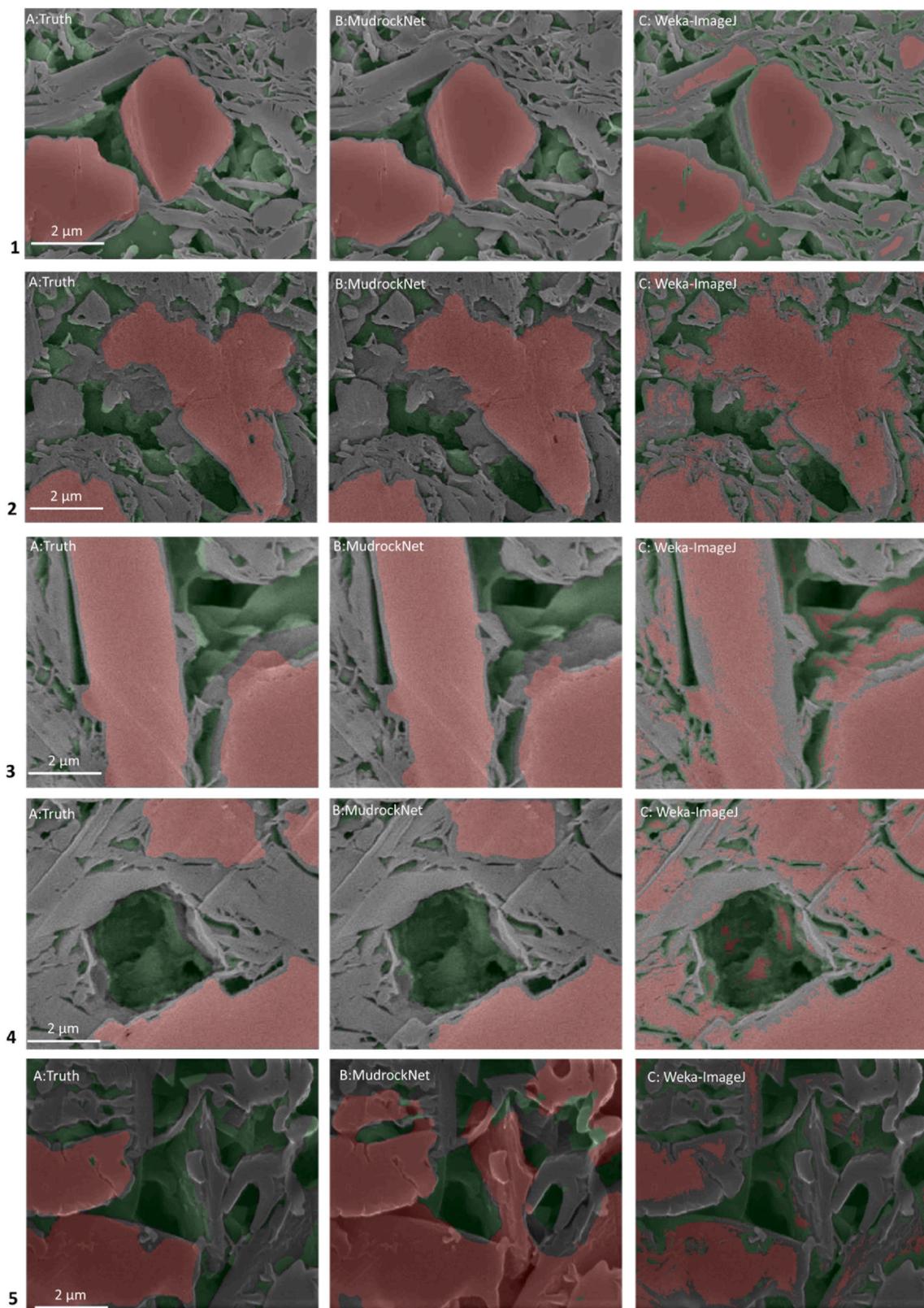


Fig. 7. Example SEM images of silt grains (red), pores (green), clay (gray)- A) ground truth (conventional segmentation algorithm) data, B) MudrockNet model predictions, C) trainable Weka model predictions. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 2

Silt and pore IoU value comparisons from MudrockNet and Weka model for images in Fig. 7.

IoU values	MudrockNet model			Weka model			
	Image	Silt grains	Clay grains	Pores	Silt grains	Clay grains	Pores
7-1	0.9117	0.8177	0.6138	0.6579	0.6138	0.5269	
7-2	0.8805	0.7658	0.6558	0.6374	0.5461	0.6747	
7-3	0.9329	0.7899	0.7647	0.6574	0.4768	0.6257	
7-4	0.8052	0.8673	0.8204	0.4693	0.4000	0.5772	
7-5	0.1957	0.3192	0.4356	0.5546	0.6897	0.6807	

for ground truth images was able to successfully isolate the individual pores and silt grains well, the deep learning approach (MudrockNet) was able to obtain similar results with less user intervention, fewer steps, and faster. This was possible by capturing the image features at multiple scales and combining them while conserving the spatial resolution. In contrast, the Weka model was unable to capture the silt grains and pores accurately, possibly due to limited training data and model limitations.

Some previous researchers have used SEM-EDX color maps in addition to grayscale SEM images for segmentation (Tang and Spikes, 2017; C. Li et al., 2021; Knaup et al., 2019, 2020) by using the differences in mineral colors in similarly sized grains for grain recognition, but such images may not always be available. Others (Wu et al., 2019; Misra et al., 2020) performed segmentation of pores/fractures, kerogen, and rock matrix in grayscale SEM images without identifying the different types of grains or their sizes. Andrew (2018) used machine learning to segment grayscale shale SEM images into five classes (pores, organics, pyrites, quartz, and calcite) but did not differentiate between silt and clay grains. While Z. Chen et al. (2020) segmented large mineral (silt) grains, clay aggregates and organic matter from grayscale shale SEM images, there seems to be less variation in image brightness and contrast in the images shown in their paper which makes it relatively easier to differentiate between organic matter and grains across different images in the dataset. There is also a distinguishable difference in the size and shape of the clay and the silt grains in the data despite having similar pixel values, since the mineral grains are considerably larger and well-rounded compared to the clay. Conversely, the core image data used for training the MudrockNet model is uncemented and less compacted since the samples are from depths <1.1 km below the sea floor (Milliken et al., 2016) with relatively fewer stark differences between grain sizes and shapes. Additionally, there is also a wider variety of brightness and contrast observed in our images and it is possible that the limited data availability despite data augmentation for images with low or high brightness/contrast, reduces the model performance and lowers

the predicted IoU. However, the wide variety of images in the dataset, with variations in image pixel intensities, porosity values, and grain shapes and sizes also allows the model to make better generalized predictions that are close to those obtained from expert user-guided segmentation (as seen in Fig. 7).

Since the images have varying pixel values for the same feature in different images due to the SEM image generation process, the prediction models need to be trained to capture this variation to prevent mislabeling. Moreover, the mean IoU observed for the silt grains was comparatively lower because of fewer training data compared to pores, and as the variable size silt grains in many cases have same pixel values as the clay grains. This can cause a difference in the predictions from the different methods, and they may have to be post-processed to remove grains below a specified size threshold. Finally, the uncertainty of the ground truth data due to the complex filtering and segmentation process despite manual corrections can also result in predictions of pores and silt grains by the MudrockNet model which may have been overlooked in the conventional method. We hypothesize that an active learning approach (like the one presented in Santos et al., 2020) would be helpful in querying the uncertainty space of these types of models.

5. Conclusion and future work

We propose a method for filtering and segmentation using deep learning to identify pore and grain features from grayscale mudrock SEM images. The MudrockNet model is trained based on Google's DeepLab-v3+ architecture from the TensorFlow library, and the predictions for the test data obtain a mean IoU of 0.6663 for silt grains, 0.7797 for clay grains and 0.6751 for pores. Comparisons with the trainable Weka segmentation model in ImageJ showed that MudrockNet gave superior predictions in most cases. Once trained, deep learning algorithms can provide a segmentation result way faster than a conventional segmentation algorithm (such as the one here used for ground truth). This is specifically important as the imaging resolution improves, and large area SEM images become easily available.

The ability to detect pores and grains without user input and multiple stages can make the segmentation process more streamlined and easier to use. An automated method to accurately identify silt grains, clay size particles, and pores at multiple scales across a wide variety of SEM images can improve the characterization of mudrocks and may even lead to a better understanding of the role of mudrocks as reservoirs or seals for petroleum exploration or sequestration of radioactive waste or carbon dioxide. If more SEM image datasets are available along with the labeled ground truth data, then the model accuracy and the ability to better detect pores and grains can be further improved.

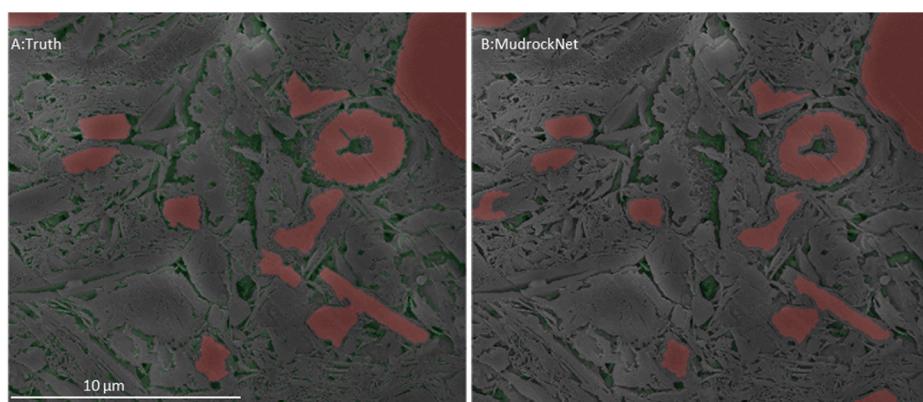


Fig. 8. Comparison of predictions for a large (magnification 15,000×) test SEM image with silt grains (red), pores (green), and clay (gray)- A) ground truth (conventional segmentation algorithm), B) MudrockNet model predictions. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

6. Data and code availability

The original mudrock SEM images used in this work are posted on Digital Rocks Portal, please see Milliken et al. (2016). The original images and the associated ground truth data is also available on Digital Rocks Portal (Bihani et al., 2020). Our trained MudrockNet model and the trainable Weka model for ImageJ are available in a GitHub repository at <https://github.com/abhishekdbihani/MudrockNet>.

CRediT authorship contribution statement

Abhishek Bihani: Writing – original draft, carried out the conventional segmentation for ground truth image labeling, trained the model and drafted the manuscript. **Hugh Daigle:** Supervision, Funding acquisition, Writing – original draft, supervised the work, acquired funding, and helped draft the manuscript. **Javier E. Santos:** Writing – original draft, helped in model training and evaluation and helped draft the manuscript. **Christopher Landry:** Writing – original draft, created the conventional segmentation code used for ground truth image labeling. **Maša Prodanović:** Writing – original draft, helped in implementing the conventional segmentation code and helped draft the manuscript. **Kitty Milliken:** Writing – original draft, acquired the mudrock samples and prepared the SEM images used for the project. All authors have read, reviewed, and approved the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the University of Texas at Austin for the support, and Equinor for funding this research through the Statoil (Equinor) Fellows program. We are grateful to the associate editor and three anonymous reviewers for their constructive feedback which made this manuscript stronger. Samples and data were provided by the Integrated Ocean Drilling Program (IODP). Funding for sample preparation and SEM imaging was provided by a post-expedition award (Milliken, P. I.) from the Consortium for Ocean Leadership. MudrockNet was implemented using the TensorFlow library in Python, and the implemented model and the architecture (DeepLab-v3+) has been modified from <http://github.com/rishizek/tensorflow-deeplab-v3-plus> made for the PASCAL VOC dataset. We would also like to thank Mr. Anurag Bihani for his assistance in creating MudrockNet.

References

- Alzubaidi, F., Mostaghimi, P., Swietojanski, P., Clark, S.R., Armstrong, R.T., 2021. Automated lithology classification from drill core images using convolutional neural networks. *J. Petrol. Sci. Eng.* 197, 107933.
- Anderson, T.I., Vega, B., Kovscek, A.R., 2020. Multimodal imaging and machine learning to enhance microscope images of shale. *Comput. Geosci.* 145, 104593.
- Andrew, M., 2018. A quantified study of segmentation techniques on synthetic geological XRM and FIB-SEM images. *Comput. Geosci.* 22 (6), 1503–1512.
- Arganda-Carreras, I., Kaynig, V., Rueden, C., Elceirí, K.W., Schindelin, J., Cardona, A., Sebastian Seung, H., 2017. Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics* 33 (15), 2424–2426.
- Baraboshkin, E.E., Ivchenko, A.V., Ismailova, L.S., Orlov, D.M., Baraboshkin, E.Y., Koroteev, D.A., 2018. Core photos lithological interpretation using neural networks. In: 20th International Sedimentological Congress.
- Bihani, A., Daigle, H., 2019. On the role of spatially correlated heterogeneity in determining mudrock sealing capacity for CO₂ sequestration. *Mar. Petrol. Geol.* 106 (106), 116–127. <https://doi.org/10.1016/j.marpetgeo.2019.04.038>.
- Bihani, A., Daigle, H., Prodanovic, M., Milliken, K., Santos, J.E., 2020. Mudrock images from Nankai trough [Data set]. Digital Rocks Portal. <https://doi.org/10.17612/BVXS-BC79>.
- Bustin, A.M.M., Bustin, R.M., Cui, X., 2008. Importance of fabric on the production of gas shales. In: Society of Petroleum Engineers (SPE), Unconventional Gas Conference, Keystone, Colorado, USA. SPE Paper 114167.
- Chauhan, S., Rühaak, W., Khan, F., Enzmann, F., Mielke, P., Kersten, M., Sass, I., 2016. Processing of rock core microtomography images: using seven different machine learning algorithms. *Comput. Geosci.* 86, 120–128.
- Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2014. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs arXiv preprint arXiv:1412.7062.
- Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2017a. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (4), 834–848.
- Chen, L.C., Papandreou, G., Schroff, F., Adam, H., 2017b. Rethinking Atrous Convolution for Semantic Image Segmentation arXiv preprint arXiv:1706.05587.
- Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 801–818.
- Chen, Z., Liu, X., Yang, J., Little, E., Zhou, Y., 2020. Deep learning-based method for SEM image segmentation in mineral characterization, an example from Duvernay Shale samples in Western Canada Sedimentary Basin. *Comput. Geosci.* 138, 104450.
- de Lima, R.P., Duarte, D., Nicholson, C., Slatt, R., Marfurt, K.J., 2020. Petrographic microfacies classification with deep convolutional neural networks. *Comput. Geosci.* 142, 104481.
- Desbois, G., Urai, J.L., Kukla, P.A., 2009. Morphology of the pore space in claystones—evidence from BIB/FIB ion beam sectioning and cryo-SEM observations. *eEarth Discuss.* 4 (1), 1–19.
- Dong, S., Zeng, L., Xu, C., Dowd, P., Gao, Z., Mao, Z., Wang, A., 2019. A novel method for extracting information on pores from cast thin-section images. *Comput. Geosci.* 130, 69–83. <https://doi.org/10.1016/j.cageo.2019.05.003>.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT press.
- Guntoro, P.I., Tiu, G., Ghorbani, Y., Lund, C., Rosenkranz, J., 2019. Application of machine learning techniques in mineral phase segmentation for X-ray microcomputed tomography (μ CT) data. *Miner. Eng.* 142, 105882. <https://doi.org/10.1016/j.mine.2019.105882>.
- Ivchenko, A.V., Baraboshkin, E.E., Ismailova, L.S., Orlov, D.M., Koroteev, D.A., Baraboshkin, E.Y., 2018. Core photo lithological interpretation based on computer analyses. In: IEEE Northwest Russia Conference on Mathematical Methods in Engineering and Technology, pp. 425–428.
- Izadi, H., Sadri, J., Bayati, M., 2017. An intelligent system for mineral identification in thin sections based on a cascade approach. *Comput. Geosci.* 99, 37–49. <https://doi.org/10.1016/j.cageo.2016.10.010>.
- Karimpouli, S., Tahmasebi, P., 2019. Segmentation of digital rock images using deep convolutional autoencoder networks. *Comput. Geosci.* 126, 142–150. <https://doi.org/10.1016/j.cageo.2019.02.003>.
- Knaup, A., Jernigen, J., Curtis, M., Sholeen, J., Borer, J.I., Sondergeld, C., Rai, C., 2019. Unconventional reservoir microstructural analysis using SEM and machine learning. In: Presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference. <https://doi.org/10.105530/urtec-2019-638>.
- Knaup, A., 2020. Deep Learning for Shale SEM Image Analysis. MS Thesis. University of Oklahoma.
- Landry, C.J., Prodanović, M., Reed, R., Eichhubl, P., Mohanty, K., 2017. Estimating mudrock oil-water relative permeability curves using digital rock physics. In: Unconventional Resources Technology Conference, Austin, Texas, 24–26 July 2017, pp. 2711–2731. <https://doi.org/10.15530/urtec-2017-2691701>.
- Lazar, O.R., Bohacs, K.M., Macquaker, J.H., Schieber, J., Demko, T.M., 2015a. Capturing key attributes of fine-grained sedimentary rocks in outcrops, cores, and thin sections: nomenclature and description guidelines. *J. Sediment. Res.* 85 (3), 230–246.
- Lazar, O.R., Bohacs, K.M., Schieber, J., Macquaker, J.H., Demko, T.M., 2015b. Mudstone Primer: Lithofacies Variations, Diagnostic Criteria, and Sedimentologic-Stratigraphic Implications at Lamina to Bedset Scales. SEPM (Society for Sedimentary Geology).
- Li, C., Wang, D., Kong, L., 2021. Application of machine learning techniques in mineral classification for scanning electron microscopy-energy dispersive X-ray spectroscopy (SEM-EDS) images. *J. Petrol. Sci. Eng.* 200, 108178.
- Li, Y., Qi, H., Dai, J., Ji, X., Wei, Y., 2017. Fully convolutional instance-aware semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2359–2367.
- Li, Z., Dong, M., Li, S., Huang, S., 2006. CO₂ sequestration in depleted oil and gas reservoirs-caprock characterization and storage capacity. *Energy Convers. Manag.* 47 (11–12), 1372–1382. <https://doi.org/10.1016/j.enconman.2005.08.023>.
- Liaw, A., Wiener, M., 2002. Classification and regression by random forest. *R. News* 2 (3), 18–22.
- Liu, X., Deng, Z., Yang, Y., 2019. Recent progress in semantic image segmentation. *Artif. Intell. Rev.* 52 (2), 1089–1106.
- Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440.
- Macquaker, J.H., Adams, A.E., 2003. Maximizing information from fine-grained sedimentary rocks: an inclusive nomenclature for mudstones. *J. Sediment. Res.* 73 (5), 735–744.
- Maitre, J., Bouchard, K., Bédard, L.P., 2019. Mineral grains recognition using computer vision and machine learning. *Comput. Geosci.* 130, 84–93.
- Marmo, R., Amadio, S., Tagliaferri, R., Ferreri, V., Longo, G., 2005. Textural identification of carbonate rocks by image processing and neural network: methodology proposal and examples. *Comput. Geosci.* 31, 649–659. <https://doi.org/10.1016/j.cageo.2004.11.016>.
- Milliken, K.L., Prodanovic, M., Nole, M., Daigle, H., 2016. Unconsolidated Muds from the Nankai Trough [Data set]. Digital Rocks Portal University of Texas at Austin. <https://doi.org/10.17612/P7F59W>.

- Misra, S., Ganguly, E., Wu, Y., 2020. Generalization of Machine Learning Assisted Segmentation of Scanning Electron Microscopy Images of Organic-rich Shales. Machine Learning for Subsurface Characterization. <https://doi.org/10.1016/B978-0-12-817736-5.00011-9>.
- Moore, G., Kanagawa, K., Strasser, M., Dugan, B., Maeda, L., Toczko, S., the Expedition 338 Scientists, 2013. NanTroSEIZE Stage 3: NanTroSEIZE Plate Boundary Deep Riser 2. IODP Preliminary Report, p. 338.
- Niu, Y., Mostaghimi, P., Shabaninejad, M., Swietojanski, P., Armstrong, R.T., 2020. Digital rock segmentation for petrophysical analysis with reduced user bias using convolutional neural networks. *Water Resour. Res.* 56, 1–11. <https://doi.org/10.1029/2019WR026597>.
- Noh, H., Hong, S., Han, B., 2015. Learning deconvolution network for semantic segmentation. *Proc. IEEE Int. Conf. Comput. Vis.* 1520–1528. <https://doi.org/10.1109/ICCV.2015.178>.
- Nole, M., Daigle, H., Milliken, K.L., Prodanović, M., 2016. A method for estimating microporosity of fine-grained sediments and sedimentary rocks via scanning electron microscope image analysis. *Sedimentology* 63 (6), 1507–1521.
- Oertel, G., 1983. The relationship of strain and preferred orientation of phyllosilicate grains in rocks—a review. *Tectonophysics* 100 (1–3), 413–447.
- Pommer, M., Milliken, K., 2015. Pore types and pore-size distributions across thermal maturity, Eagle Ford Formation, southern Texas. *AAPG Bull.* 99, 1713–1744. <https://doi.org/10.1306/03051514151>, 09.
- Potter, P.E., Maynard, J.B., Depetris, P.J., 2005. *Mud and Mudstones: Introduction and Overview*. Springer Science & Business Media.
- Rahman, M.A., Wang, Y., 2016, December. Optimizing intersection-over-union in deep neural networks for image segmentation. In: International Symposium on Visual Computing. Springer, Cham, pp. 234–244.
- Ronneberger, O., Fischer, P., Brox, T., 2015. October. U-net: convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, pp. 234–241.
- Santos, J.E., Mehana, M., Wu, H., Prodanovic, M., Kang, Q., Lubbers, N., Viswanathan, H., Pyrcz, M.J., 2020. Modeling nanoconfinement effects using active learning. *J. Phys. Chem. C* 124 (40), 22200–22211.
- Schmidhuber, J., 2015. Deep learning in neural networks: an overview. *Neural Network*. 61, 85–117.
- Schneider, J., Flemings, P.B., Day-Stirrat, R.J., Germaine, J.T., 2011. Insights into pore-scale controls on mudstone permeability through resedimentation experiments. *Geology* 39 (11), 1011–1014. <https://doi.org/10.1130/G32475.1>.
- Schlömer, S., Krooss, B.M., 1997. Experimental characterization of the hydrocarbon sealing efficiency of cap rocks. *Mar. Petrol. Geol.* 14 (5), 565–580.
- Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. *J. Big Data* 6 (1), 1–48.
- Schowalter, T., 1979. Mechanics of secondary hydrocarbon migration and entrapment. *AAPG (Am. Assoc. Pet. Geol.) Bull.* 63, 723–760. <https://doi.org/10.1306/2F9182CA-16CE-11D7-8645000102C1865D>, 0149.
- Sutskever, I., Martens, J., Dahl, G., Hinton, G., 2013. On the importance of initialization and momentum in deep learning. In: Proceedings of the 30th International Conference on Machine Learning, in PMLR, vol. 28, pp. 1139–1147, 3.
- Tang, D., Spikes, K., 2017. Segmentation of shale SEM images using machine learning. In: SEG Technical Program Expanded Abstracts 2017. Society of Exploration Geophysicists, pp. 3898–3902.
- Tang, D.G., Milliken, K.L., Spikes, K.T., 2020. Machine learning for point counting and segmentation of arenite in thin section. *Mar. Petrol. Geol.* 120, 104518.
- Tian, X., Daigle, H., 2018. Machine-learning-based object detection in images for reservoir characterization: a case study of fracture detection in shales. *Lead. Edge* 37 (6), 435–442. <https://doi.org/10.1190/le37060435.1>.
- Tian, X., Daigle, H., 2019. Preferential mineral-microfracture association in intact and deformed shales detected by machine learning object detection. *J. Nat. Gas Sci. Eng.* 63 (October 2018), 27–37. <https://doi.org/10.1016/j.jngse.2019.01.003>.
- Varfolomeev, I., Yakimchuk, I., Safonov, I., 2019. An application of deep neural networks for segmentation of microtomographic images of rock samples. *Computers* 8 (4).
- Wang, C., Sun, J., Xu, W., Chen, X., 2019a. Depth learning standard deviation loss function. *J. Phys. Conf.* 1176 (3), 032050). IOP Publishing.
- Wang, Y., Teng, Q., He, X., Feng, J., Zhang, T., 2019b. CT-image of rock samples super resolution using 3D convolutional neural network. *Comput. Geosci.* 133, 104314.
- Wu, Y., Misra, S., Sondergeld, C., Curtis, M., Jernigen, J., 2019. Machine learning for locating organic matter and pores in scanning electron microscopy images of organic-rich shales. *Fuel* 253 (May), 662–676. <https://doi.org/10.1016/j.fuel.2019.05.017>.
- Yang, Y., Aplin, A.C., 2007. Permeability and petrophysical properties of 30 natural mudstones. *J. Geophys. Res.: Solid Earth* 112 (3). <https://doi.org/10.1029/2005JB004243>.