

URTeC: 2846

## Machine-Learning-Assisted Segmentation of FIB-SEM Images with Artifacts for Improved Pore Space Characterization of Tight Reservoir Rocks

Andrey Kazak (Center for Hydrocarbon Recovery, Skoltech); Kirill Simonov\* (Schlumberger Moscow Research, Skoltech); Victor Kulikov (PicsArt Inc., Skoltech)

Copyright 2020, Unconventional Resources Technology Conference (URTeC) DOI 10.15530/urtec-2020-2846

This paper was prepared for presentation at the Unconventional Resources Technology Conference held in Austin, Texas, USA, 20-22 July 2020.

The URTeC Technical Program Committee accepted this presentation based on information contained in an abstract submitted by the author(s). The contents of this paper have not been reviewed by URTeC, and URTeC does not warrant the accuracy, reliability, or timeliness of any information herein. All information is the responsibility of and is subject to corrections by the author(s). Any person or entity that relies on any information obtained from this paper does so at their own risk. The information herein does not necessarily reflect any position of URTeC. Any reproduction, distribution, or storage of any part of this paper by anyone other than the author without the written consent of URTeC is prohibited.

### Abstract

The focused-ion-beam scanning electron microscopy (FIB-SEM) technology allows imaging of nano-porous tight reservoir rock samples in 3D at a resolution up to 3 nm/voxel. However, the quality and efficient segmentation of FIB-SEM images is still a complicated and challenging task. Correct porosity determination from FIB-SEM images requires fast and robust segmentation. Typically, a trained operator spends days/week for subjective and semi-manual labeling of a single FIB-SEM dataset. The presence of FIB-SEM artifacts, such as pore backs, requires the development of a new methodology for efficient image segmentation. We developed a robust approach for an automated highly-efficient multimodal segmentation of FIB-SEM datasets using machine-learning (ML) based methods.

A representative collection of rocks samples was formed based on the petrophysical interpretation of well logs for a complex tight gas reservoir rock of the Berezov formation (West Siberia, Russia). The core samples passed through a multiscale imaging workflow for pore space structure upscaling from nanometer to log scale. FIB-SEM imaging resolved the finest scale using FEI Versa 3D DualBeam analytical system. Image segmentation utilized an architecture based on a convolutional neural network (CNN) in the DeepUnet configuration. The implementation utilized the Pytorch framework in a Linux environment. Computation exploited high-performance computing system based on Intel(R) Core (TM) i7-6700 CPUs and NVIDIA GTX 1080i TitanBlack GPUs.

The target data included three 3D FIB-SEM datasets with a physical size of around  $20 \times 15 \times 25 \mu\text{m}$  with a voxel size of 5 nm. A professional geologist manually segmented (labeled) a fraction of slices. We split the labeled slices into training (TD) and validation data (VD). We then augmented training data to increase TD size. The developed CNN delivered promising results. The model performed automatic segmentation with the following average quality indicators based on VD: accuracy of 96.66%, the precision of 91.67%, recall of 67.57%, and F1 score of 77.00%. We achieved a significant boost in segmentation speed of 14.5 Mpx/min. compared to 0.18–1.45 Mpx/min. for manual labeling, yielding at least 10 times of efficiency increase.

The presented research work improves the quality of quantitative petrophysical characterization of complex reservoir rocks using digital rock imaging. The development allows the multiphase segmentation of 3D FIB-SEM data complicated with artifacts. It delivers correct and precise pore space segmentation resulting in insignificant turn-around time saving, as well as increased quality of porosity data. Although image segmentation using CNNs is a mainstream in the modern ML world, it is an emerging novel approach for reservoir characterization tasks.

## Introduction

To date, the worldwide petroleum industry considers low-permeable (tight) rocks, such as organic-rich mudstones (shale formations) as reservoirs containing considerable resources and reserves of hydrocarbons (oil, gas, and condensate) (EIA-2017). Such unconventional reservoirs differ from conventional porous media by the contribution of nanoscale pores (Nelson-2009) to the total effective porosity and their heterogeneity in size, shape, connectivity (Kuila-2013).

Hydrocarbon resource assessments followed by reserves estimation require not only the application of routine (API-1998), but also special (Luffel-and-Guidry-1992, Durand-et-al.-2019, Nikolaev-and-Kazak-2019, Kazak-and-Kazak-2019) petrophysical core analysis methods. However, the determination of effective properties on the core scale turned out to be not practical enough while being applied to well log data interpretation, RCAL techniques, classic fluid-flow modeling. A viable workaround is the application of novel techniques involving direct imaging of the rock's void space structure. One of the most developed approaches in this area is digital rock physics.

Digital Rock Physics (DRP), also referred to as Digital Core Analysis (DCA), Digital Rock (DR), is a technology permitting to run the numerical simulation and extracting petrophysical parameters from a digital twin of real rock samples. Throughout the two last decades, the need for understanding pore-scale features and processes, and macroscopic petrophysical properties calculation contributed to the development of the DRP technology (Andra-et-al.-2013b). The key paradigm of DRP is “image and compute.” The approach involves imaging and digitizing the pore space and mineral matrix of natural rock and then numerically simulating various physical processes in this digital object to obtain macroscopic rock properties, including absolute permeability (Sun-et-al.-2017), electrical conductivity (Liu, Sun, and Wang-2009), and elastic moduli (Andra-et-al.-2013b). Globally, the DRP workflow consists of three-step flow (Dvorkin-et-al.-2011): 1) imaging of rock plugs allowing to resolve pore-scale features; 2) processing the raw images to eliminate the noise and segment voids space and mineral – organic matter – fluid matrix phases; 3) numerical simulation of effective properties to obtain cross-property relations.

Within the DRP workflow, the focused-ion-beam scanning electron microscopy FIB-SEM 3D microscopy became a widely applicable, although the invasive, high-resolution method for 3D-imaging of a sample's microstructure. The operation of milling and imaging by FIB runs ion particle and solid body interactions (Giannuzzi-and-Stevie-2005). Modern laboratory setups, so-called Dual Beam tools, include an ion column integrated with an electron column in a single device. When applied to reservoir rocks holding a complex void space structure, FIB-SEM delivers digital rock models with voxel size down to  $2.5 \times 2.5 \times 5$  nm (Goral, Deo, and Andrew-2018, Goral-et-al.-2019).

The essential step within the DRP workflow is the building of 2D and 3D digital rock models (DRM) that, in turn, heavily rely on image segmentation. Here we consider segmentation as a process of partitioning a digital image into multiple segments and assigning a label to every pixel in an image such that pixels with the same label share specific characteristics (Frucci, Sanniti, and Sanniti di Baja-2008). Adequate characterization of the pore geometry leads to correct spatial void space structure characterization, as well as the results of DRP simulations. For this reason, segmentation is the crucial step of a quality DRM.

The quality DRMs based on FIB-SEM images require the application of dedicated segmentation techniques that are different from those for other imaging methods. Mostly, it is associated with image artifacts, which are intrinsic for FIB-SEM images like the “poreback” artifact, FIB curtaining effect, and “sample charging” artifact (Kelly-et-al.-2016). Our literature survey shows a shortage of automated segmentation methods developed and successfully implemented in practice.

The presented research work aims at improving the quality of quantitative petrophysical characterization of complex reservoir rocks using digital rock imaging. The task of the work includes the evaluation of modern machine-learning (ML) methods for an objective and automated segmentation of FIB-SEM reservoir rock images complicated with artifacts and the development of a viable model with higher performance in terms of segmentation quality and computing time than that of currently used manual and semi-automated segmentation techniques.

Quality FIB-SEM image segmentation is a complicated and challenging task. FIB-SEM artifacts, textural contrast, and the nature of imaging show the necessity for the development of a new methodology for image segmentation. The conducted literature review indicates that ML is a promising approach to solve this kind and segmentation problem. Therefore, we dedicated this work to test the feasibility of ML-based segmentation of complex FIB-SEM 3D datasets.

## State of the Art

Here we provide a concise summary of the most appropriate image segmentation techniques from the domain of general image processing and analysis. Although the described items are mostly general-purpose, their description provides a good starting point for the current state of the problem, as well as acts as a reference for more advanced segmentation methods.

Recently, segmentation methods based on artificial neural networks (NN) are becoming accessible and more applicable as computational power increases rapidly. The working principle of NN is more or less the same between all the types — to provide NN with learning similar to the cognitive process of the human brain. NN is composed of a large number of connected layers that have their weight.

Generally, NN-based image segmentation methods form two main classes: supervised and unsupervised. Supervised methods require manually segmented data for training, so the quality of training directly depends on the quality of input data. Unsupervised methods are semi- or fully automatic. User interaction might be required at some stage to assess or improve performance, but mainly the result should be more or less human independent. However, the thesis does not imply describing NN mechanics in detail. Introductory and advanced-level information one can find in ([Zurada 1992](#), [Gurney 1997](#)).

### Segmentation of FIB-SEM Images: Features and Issues

FIB-SEM imaging requires the application of dedicated segmentation techniques that are different from that for other imaging methods (data acquisition types). Mostly, it is associated with image artifacts, which are intrinsic for FIB-SEM images. Below we provide a brief overview of primary FIB-SEM image artifacts, but not pretending to describe all the details and techniques available for artifact removal (mitigation). ([Taillon et al. 2018](#), [Reuteler 2017](#)) provide a good starting point to detailed descriptions on the topic.

Image artifacts lead to difficulties in image analysis. Concerning geoscience application, mostly, we have a porous rock as a target sample presenting a complicated heterogeneous aggregate with different mineral and element compositions that consequently leads to a complex inner geometry of sample phases. This heterogeneity influences the milling as various materials have different milling rates, and thus, milled surfaces can be uneven throughout the sample. Moreover, possible pore destruction and generation due to ion beam milling are still in question.

The most common FIB-SEM image artifacts are the “poreback” artifact, FIB curtaining effect, and “sample charging” artifact ([Taillon et al. 2018](#), [Kelly et al. 2016](#)).

The “poreback,” “pore shine,” or “shine through” artifact is a topography-induced contrast artifact leading to misinterpretation of the pore-adjacent region. FIB-SEM slicing brings void space into the image. However, the depth of the poreback can be low, and electrons are not trapped to illustrate “black” pores that we are supposed to see in this region, and alternately show this region as bright. Otherwise, the redeposited material while milling can fill up the pore during imaging. In this case, the grey intensity value represents the matrix behind the void space instead of the void itself. Thus, these regions obtain the characteristics of the image body that further lead to misclassification and wrong calculation of image properties. During manual segmentation, a human interpreter can distinguish pore borders. However, this is a non-trivial challenge for automated segmentation as pixel intensity criteria are not enough for proper interpretation.

Another artifact caused by sample inner structure dissimilarity is the “curtaining” effect. One appears due to ion channeling while milling and takes shape as a vertical striation. Any inhomogeneity can lead to such an artifact, but it is widespread when there is an open pore structure.

Besides, FIB-SEM image can hold the so-called local charging artifact. Charging appears during imaging nonconductive or almost nonconductive materials. The electrons accumulate net positive charge in regions due to the absence of the conduit path. Objects of this type commonly appear as bright spots, warpings, and nonphysical fissures.

One of the critical issues in shale imaging is the correct interpretation of pores and kerogen. Relatively smaller nanoscale pores, which can be inside of kerogen, correspond to darker regions. At the same time, relatively large microscale pores illustrate a 3D structure depicted in the 2D surface with different shades of grey that simultaneously be complicated with pore-backs. Thus, a differentiation of pores based on greyscale values using global thresholding does not deliver adequate results.

There are several relatively new FIB-SEM image morphological algorithms developed for porous materials. Although not yet tested for reservoir rock images, they may potentially be applied to deal with the artifacts mentioned above.

FIB-SEM imaging technique and porous media presence imply a vision of structure on the 2D image before it appears in a section. For this reason, it is challenging even for the human eye to distinguish the prolongation of the structure in the Z-direction. Typically, an experienced interpreter does segmentation in such complicated cases. The (Salzer-et-al.-2012) presents a new segmentation approach of the FIB-SEM image of porous materials. Analysis of the FIB-SEM image stack identifies the last occurrence of a void-space structure according to the dramatic decrease of grey values. The corresponding fall marks the first appearance of the target object in the Z-direction. Afterward, the algorithm estimates the feasible threshold for the last appearance. On the last step, the algorithm propagates the local threshold back to all the previous slices. The behavior explains the method's name — Local Threshold Backpropagation (LTB). The threshold estimation and consequent backpropagation work as designed if the last appearance of the void structure is inside the stack. Thus, the method fails for structures penetrating an image stack in Z-direction. The technique compares intensity values in the Z-direction and detects the appearance and disappearance of structures based on this comparison rather than on absolute intensity values. The method produces excellent results when it is up to discovering structures in the Z-direction. Still, additional image filtering is required. In particular, dilation of the foreground adds small structure bounds and results in the expansion of the foreground into the background (Salzer-et-al.-2012). The LTB has an advantage to account of the nature of FIB-SEM images concerning the visibility of the structure located on different layers. However, although the global threshold removes foreground edges partially, it may lead to the full destruction of small-sized foreground objects (Salzer-et-al.-2012).

Another modern segmentation algorithm for FIB-SEM images with porous media is Valley Detection, somewhat extension of the LTB approach (Salzer-et-al.-2014). LTB algorithm relies on the assumption that the grey intensity slightly fluctuates within the observed structure. However, not all datasets can meet this condition. Geometry and various material properties can lead to one and influence the sharp variations of grey intensity. For instance, oblique and rough surfaces have higher grey intensity, while roundish smoothed surfaces tend to have lower grey intensity value. Besides, porebacks influence thin structures to a great extent, and the latter tend to have lower grey intensity value. Valley Detection attempts to mark the first and last occurrences of the target structures by choosing local minima and maxima in the scope of the structure. The corresponding intensity “the valley” along Z-direction marks the phase separating two distinct structures. The method considers the existence of local minima at the last slice holding structure.

Another FIB-SEM image segmentation algorithm is Morphological Segmentation (Morph) (Prill-et-al.-2013). The algorithm is sophisticated, multistage, and has five tuning parameters. The workflow includes contrast-enhancement, detection of morphological features in the Z-profiles of the images indicating matrix, void space, or artifacts followed by constrained watershed transformation. The authors report an overall error rate of 12.3% on the synthetic 3D FIB-SEM dataset.

The traditional image segmentation methods might induce significant errors (Andra-et-al.-2013a). Generally, raw datasets include physical and nonphysical artifacts and noise, leading to poor quality of segmentation results using conventional techniques. While reviewing digital images, the human visual perception system automatically smoothes noise and recognizes specific patterns (features), but ignores artifacts. However, the process is difficult to formalize and implement in a computational workflow. Naturally, there was one common way to handle this problem — laborious and time-consuming manual segmentation. However, over time, datasets become larger, digital images get complex and multidimensional (multimodal), while manual approaches consume a lot of human and time resources.

### Machine Learning Approach

The last 20 years show the gradual development of Machine Learning (ML) approach in data science. The previous decade revealed a technological boom, and ML methods started to occupy the market. One of the main factors of ML rise is the extensive growth of computational power. Mostly, it addresses advanced parallel computing using an array of general-purpose (GP) graphics performance units (GPU), with a representative example of the NVIDIA CUDA programming model.

The efficiency of similar algorithms applying to microscopy image segmentation implies the ability to generate random classifiers that work in high dimensional space in comparison to those handling only 2D images output from the microscope's detector. So-called multimodal images result from data integration of different imaging and computational sources. The data sources include image acquisition using different detectors, voltages, energies, angles. The additional dimensions may hold the original images after additional filtering (e.g., gradient smoothing,

textural and compositional transformations) to extract different local and non-local features. The extracted features form a single vector to characterize every pixel on local, non-local, or textural data.

Although these technologies enriched many domains where data science shows robust performance, for example, medical diagnosis and target advertisement, the analysis of digital rock images (CT, optical, or electron microscopy) within geoscience is still under development. A small number of recent publications related to ML-based FIB-SEM image segmentation ([Andrew·2018](#), [Tracey·et·al·2019](#)) confirm this statement.

In geoscience, there are many rock types to analyze, and each rock type has its representative elementary volume (REV). Examples span from high porous and high permeable conventional terrigenous and carbonate reservoir rocks to unconventional shales. Understanding the latter requires multiscale imaging. [Andrew·\(2018\)](#) addresses the quantitative benchmarking of ML-based classifiers against traditional processing of image workflow for different rock samples from petroleum geoscience. Assessment of the effect from different segmentation techniques involved three 3D (data that are not obtained by direct imaging) datasets synthetically created based on real rock sample images, including the FIB-SEM dataset of the Vaca Muerta shale and micro-CT of a Bentheimer sandstone. The paper compares Otsu thresholding, seeded watershed region growing, and ML-based segmentation. The comparison highlighted several features. Firstly, the misclassification rate for conventional techniques could be high even at low noise. Secondly, the ML-based segmentation used textural information forming a vector that allowed differentiating in classification. This technique kept the misclassification rate below 15% within a 30% noise range.

[Tracey·et·al·\(2019\)](#) discusses integrating a random forest classifier with the deep-learning residual neural network U-net (ResUnet) to segment pore-backs from a FIB-SEM image stack for a membrane sample with 80% porosity. The training included 5–20 images for each case, resulting in a total of four trials. The results indicate that the training sample distribution has a significant impact on prediction quality. The augmented ResUnet delivers a segmentation accuracy approaching 90% for 3D image stacks of polymer membranes if the training set is more than 20 images evenly distributed throughout the dataset.

In summary, quality FIB-SEM image segmentation is a complicated and challenging task. FIB-SEM artifacts, textural contrast, and the nature of imaging show the necessity for the development of a new methodology for image segmentation. The conducted literature review indicates that ML is a promising approach to solve this kind and segmentation problem. Therefore, this work studies the feasibility of ML-based segmentation of complex FIB-SEM 3D datasets.

## Materials

The presented study utilizes data from an R&D project dedicated to DRP characterization of one most promising tight gas formation located in West Siberia of the Russian Federation — Lower Berezov formation. [Kazak·et·al·\(2017\)](#) provide a comprehensive description of the geology, rock sample collection, and supplementary data on the samples. For the sake of brevity, we outline only relevant information.

The reservoir rock is mostly represented by dark-grey gaze(opoka)-like clays, siliceous mudstones with rare interbeds of argillaceous siltstones, and fine-grained sandstones. Hydrocarbon accumulations of the Berezov formation are associated with siliceous remains of radiolarians, shells, or frustules of diatoms, spicules of sponges, etc., dwelled in shallow sea areas.

From the petrophysical perspective, the samples had porosity 18–26% with an average value of 22% and gas permeability 0.2–60 mD with an average value of 11 mD. High porosity and relatively low permeability were initially associated with the presence of a large number of micro-scale voids having interconnecting nano-channels.

The whole core pieces were split into several representative standard plugs Ø30×30–90 mm. After core-plug petrophysical tests, the plugs, in turn, were separated by representative mini-plugs Ø3×10 mm. Then each mini-plug was cut into two halves: the main and archive. We imaged a flat cut surface of the main half, identified and marked the target regions, and performed FIB-SEM imaging. The acquisition of 3D FIB-SEM models involved the application of the FEI Versa 3D DualBeam analytical system.

## Target Datasets

In this work, we handled 3 three target FIB-SEM datasets ([Table 1](#)). The exact locations of the FIB-SEM models correspond to rock-specific features, shed light on the void space nanostructure, and explain the transport and storage capacity of the studied reservoir rock. For example, one of the FIB-SEM models included a contact fracture between



the clay mineral grain and the surrounding siliceous matrix (Figure 1). Another FIB-SEM model captures the void space of a stylolite seam (Figure 2).

Table 1. Specifications of the target FIB-SEM datasets.

ID	SEM Mode / Detector	Bit Depth	Image Dimension X (px)	Image Dimension Y (px)	Pixel Size (nm)	Number of Slices	Slice Thickness (nm)	Image Physical Dimensions H×W×D (μm)
3.1.1	SE	8	3,072	2,048	5.6	100	25	17.3×11.5×15
7.1.1	SE	8	4,096	3,536	4.5	553	25	18.3×15.8×15
7.1.2	SE	8	4,096	3,536	4.5	115	25	18.3×15.8×15

It is worth noting that the resulting voxel size had anisotropy due to the difference in spatial resolution in XY- and Z-direction, and this affected the connectivity of both voids and matrix in the resultant 3D digital rock models. This effect is an intrinsic feature of the FIB-SEM technology, but hardware tuning and image processing steps may mitigate it effectively.

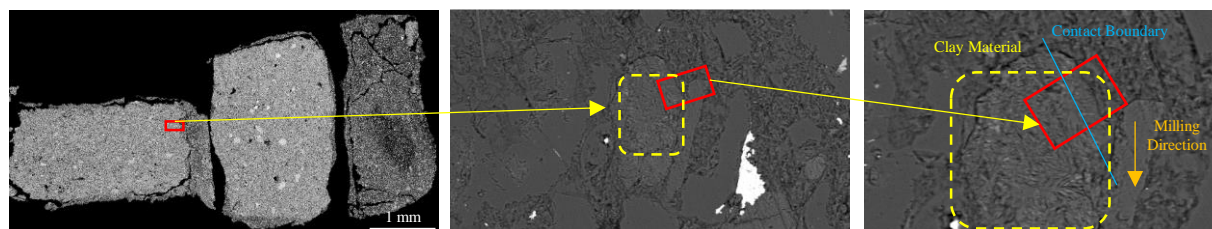


Figure 1. The location of Dataset 3.1.1 capturing the contact zone between rock matrix and clay material.

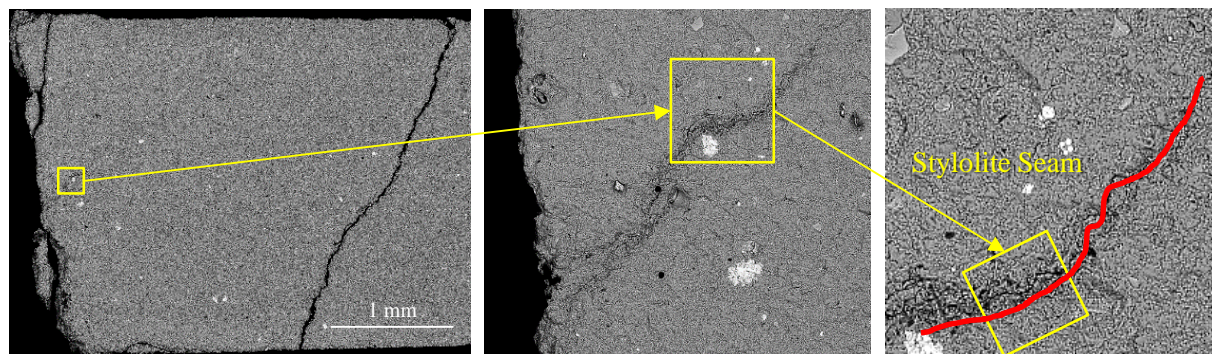


Figure 2. The location of Dataset 7.1.1 capturing the microstructure of a stylolite seam.

### Labeled Dataset

We manually segmented (labeled) the filtered slices using the Wand (tracing) tool of the Fiji 1.52h software package. Due to image processing features, some of the resulting labels had dimensions smaller than those of target-dataset slices by 30–40%. For example, if the target-dataset slices had dimensions of 4,096×3,536 px, and the corresponding TD slices on average had dimensions of 3,700×1,600 px. To ensure spatial data consistency, we registered (in terms of coordinates) each label to the corresponding image. We then selected the slices for labeling in such a way to the main specific void-space features and traced them throughout the whole dataset (Figure 3).

Before labeling of image slices from Dataset 7.1.1, we preprocessed them using FEI PerGeos software. The image processing workflow included fast-Fourier-transform filtration to remove periodical artifacts in the XY-slice plane; individual slice alignment; image stack crop to the volume intersecting all the slices; non-local means smoothing and unsharp masking filtering to remove high-frequency noise and increase edge contrast.

Eventually, we obtained a total of 21 manually labeled XY slices, including 19 slices from Dataset 7.1.1, a single slice from Dataset 7.1.2, and a single slice Dataset 3.1.1.

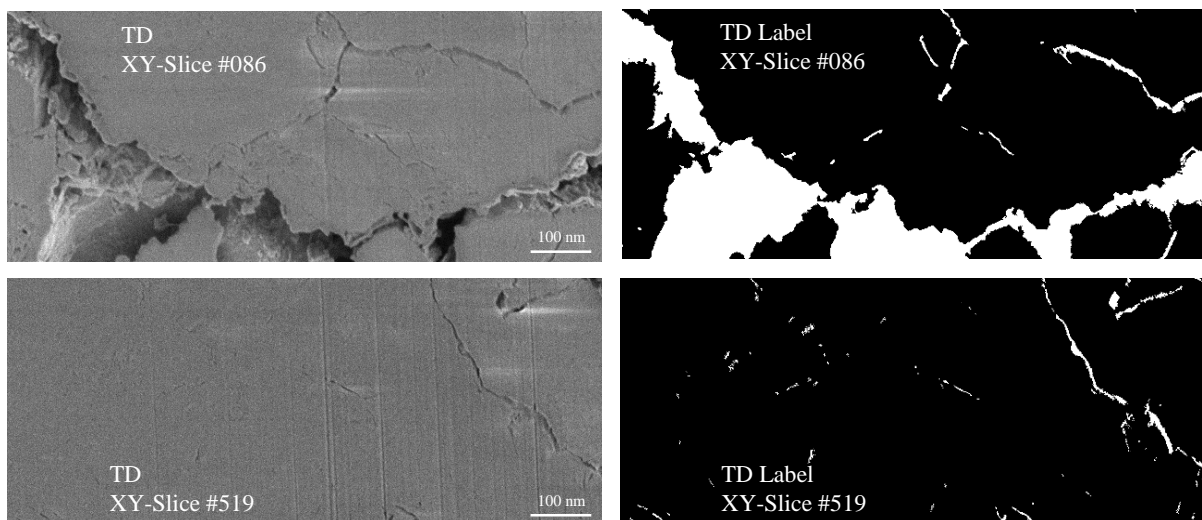


Figure 3. Example of original slices from the target dataset 7.1.1 (left) and the corresponding results of manual segmentation (right).

### Training Datasets

The training data included a set of slice pairs. Each pair consisted of an XY slice from the target dataset and the corresponding label.

For CNN training, we used the target XY slice images without additional processing for two reasons. Firstly, we know that in the general case, any data processing reduces the amount of valuable information (signal) from the input dataset and, in the best case, keeps the amount at its original level. Secondly, we realize that during training, a deep CNN captures poorly-understood and weakly-defined features of the training data, including that part of the image that we consider noise. Therefore, we decided to train using the raw FIB-SEM slices, rather than filtering them before feeding to the CNN.

The base TD included a total of 10 labeled slices. However, testing of a single CNN configuration required expansion of the base TD with a few extra labeled slices and areas (patches) from the same target Dataset 7.1.1 (Figure 4).

The patches correspond to regions of the original slices outside of the base TD. Thus, the patch dimensions 2,000–3,000×1,000–1,300 px that are smaller than those of the base TD slices.

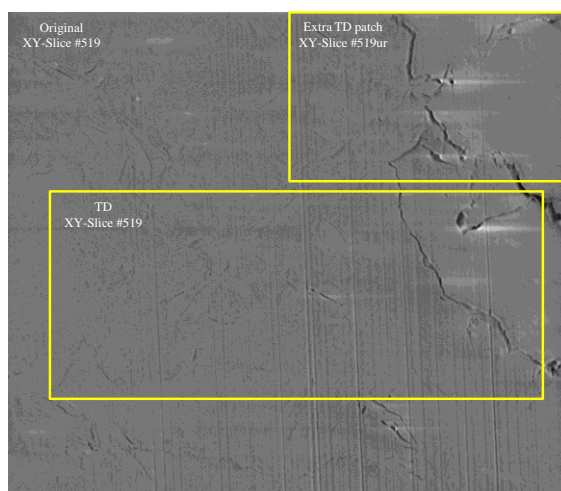


Figure 4. Location of the extra labeled patch #519ur of the original XY slice #519 for TD expansion.

### Validation Datasets

Validation data (VD) included a subset of labeled slices and provided for the evaluation of CNN's performance. CNN training stage did not use the VD images, that is, CNN did not know about their existence. If this essential condition fails, then a CNN performance assessment becomes complicated, since the model has learned to derive from the TD or has just captured these observations.

The base VD included a total of 4 labeled slices Dataset 7.1.1.

### Methods

#### ML-based Approach

Semantic segmentation targets to provide image information on a pixel level. The goal of semantic segmentation is a classifying of image pixels that belong to the same object class. Currently, there are plenty of NN types and architectures. To attack the problem of segmentation of FIB-SEM rock images, we chose CNN called DeepUnet since its design can handle the segmentation of high-resolution images.

DeepUnet (Ronneberger, Fischer, and Brox 2015) is an end-to-end fully convolutional network with two other kinds of short connections called U- and connections with an hour-glass architecture, allowing to extract image features at different scales. The central concept of DeepUnet is a couple the layers in the contractive track with the extensive track. High-resolution features obtained during the contracting track are integrated into the upsampled output. Thus, a consecutive convolutional layer can learn to agglomerate more precise outputs, adopting derived features. Additionally, to decrease a loss while extracting high-level features, DeepUnet develops the contractive track along with an extensive track by proposing the DownBlock, UpBlock, and Plus connection.

Both DownBlock and UpBlock (Figure 5) complement the features before and after passing through the convolution layers. Such a structure yields a deeper and computationally efficient convolution neural architecture (Li et al. 2018). The DownBlock consists of two convolutional layers, sequentially linked to the Batch Normalization layer and concatenated with the ReLu layer. Both convolutional layers apply a  $3 \times 3$  convolution kernel, a  $1 \times 1$  step size. The input of the block is 64 - dimension features. The second convolution layer output has the same feature size as the input to the DownBlock. The Plus layer sums the data after the convolution.

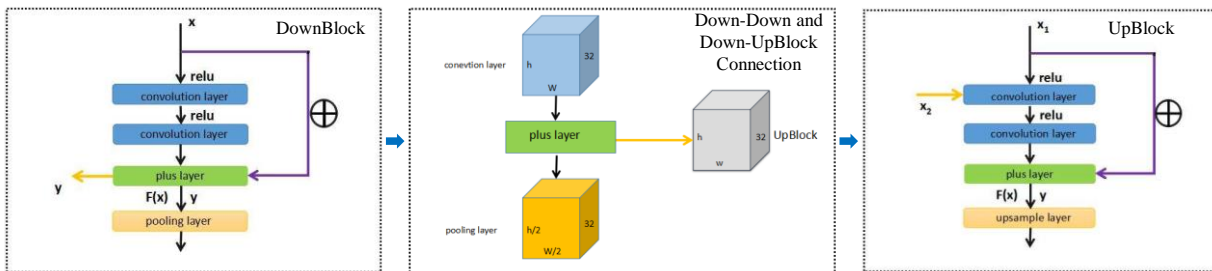


Figure 5. The internal structure and interconnection of both DownBlock and UpBlock (Li et al. 2018).

Both input greyscale images illustrating reservoir rock with void space and matrix materials, and the output images are binary segmentation maps. CNN is not a fully connected network type and uses only a  $1 \times 1$  convolution layer for dimension reduction. Finally, the Softmax activation layer is applied to transform the tensor values into maximum arguments and treat them as a two-class problem.

If the Plus layer produces  $y$  and  $x$  is the input of the DownBlock, the Plus layer passes  $x$  and the second convolution output into (1), and the filtering outputs keep the optimal value  $y$  further into the max-pooling layer. In (1),  $W_1$  means the first convolution, and  $W_2$  means the second convolution, while  $\sigma$  describes the ReLU function.

$$y = W_{2\sigma}(W_1x) + x \quad (1)$$

The max-pooling layer in the DownBlock has a kernel size of  $2 \times 2$  and a step size of  $2 \times 2$ . At this stage,  $y$  proceeds to max-pooling layer, and integration of feature maps to the correspondent UpBlock takes place. The DeepUnet maintains a fancy symmetric structure. The UpBlock is filled with feature maps to provide more precise upsampling. The UpBlock mirrors the main components of the DownBlock as it is similarly composed of two convolutional layers



and layers. However, it is clear from the name; an upsampling layer replaced the max-pooling one. Integrated feature map  $x$  is the input for the convolution layer in (2).

$$x = [\delta, x_1, x_2] \quad (2)$$

In (2),  $x_1$  is the feature map from the previous UpBlock, and  $x_2$  is that from the DownBlock through U-connection and is the upsampling operator. Owing to the DeepUnet structure, an integrated feature map is passed to the UpBlock before max pooling so that convolutional layers process it and update the data to obtain more accurate results during upsampling. To keep the original image dimensions, the resolution of DownBlocks' output should be the same as that of the UpBlocks' input.

### Preprocessing of Training Data

The limited number of TD slices made data augmentation a necessary step. We applied a row of image transforms, which included vertical and horizontal flip, rotation, blurring, histogram normalization, gray level expansion. A total number of TD images results from the number of input samples  $\times$  the number of image transforms. The base number of TD images is 10, for which the augmentation increased the TD size to 150 images.

### Computational Resources

The computations utilized a modern high-performance computing (HPC) cluster (Table 2). The software stack included Python 2.7, CUDA 8.0, Pytorch 1.1. For coding, we used the Jupyter Notebook.

Table 2. Technical specifications of the utilized HPC environment.

<b>CPU</b>	Intel(R) Core (TM) i7-6700 CPU @ 3.40GHz
<b>GPU</b>	NVIDIA GTX 1080i TitanBlack
<b>RAM</b>	32 GB
<b>OS</b>	Ubuntu 16.04

### Tuning Parameters

Since we picked DeepUnet as the central CNN architecture, we focused on tuning its parameters rather than testing other CNN architectures. The latter, however, forms the scope of future research endeavors. There are several parameters, which we studied to evaluate the CNN performance and conduct sensitivity analysis (Table 3).

Table 3. A list and description of the DeepUnet tuning parameters.

Group	Tuning Parameter		Description
Loss Function	Cross-Entropy		Cross-entropy loss or log loss measures the performance of a classification configuration whose output is a probability value between 0 and 1 ( <a href="#">Murphy 2012</a> ).
	Focal Loss	$\gamma = 1$	DeepUnet with Focal loss function smoothly adjusting the rate at which easy examples are down-weighted ( <a href="#">Lin et al. 2020</a> ).
		$\gamma = 2$	DeepUnet with Focal loss function sharply adjusting the rate at which easy examples are down-weighted ( <a href="#">Lin et al. 2020</a> ).
Optimization Algorithm	Adam		Adaptive Moment Estimation (Adam) ( <a href="#">Kingma and Ba 2015</a> ).
	SGD		Stochastic gradient descent (SGD) is an iterative method for optimizing an objective function with suitable smoothness properties ( <a href="#">Zhang, Choromanska, and LeCun 2015</a> ).
	AMSGrad		AMSGrad optimizer is an alternative from Adam using lower learning rates and consolidating the instinct of slowly decaying the effect of past gradients on the learning rate ( <a href="#">Tran Thi Phuong 2019</a> ).
TD and VD order	K-Folding		Cross-validation delivers a more accurate estimate of configuration prediction performance ( <a href="#">Seni and Elder 2010</a> ).
Number of TD image pairs	Expansion		Expands TD by adding extra slices.

### CNN Configurations

We analyzed the impact of each tuning parameter on the CNN output and combined them in 8 configurations ([Table 4](#)). The arrangements have different settings of the loss function, optimizer algorithm, and variation of TD and VD. We, however, realize that the number of possible configurations exceeds 8 but chose this number due to limited resources.

Table 4. The working CNN configurations.

ID	Configuration	Description
1	CE	DeepUnet with the Cross-Entropy loss function and AMSGrad optimizer.
2	Fl_y1	DeepUnet with the Focal loss function, $\gamma=1$ , and AMSGrad optimizer.
3	Adam	DeepUnet with the Focal loss function, $\gamma=2$ , and Adam optimizer.
4	SGD	DeepUnet with the Focal loss function and SGD optimizer.
5	K1	DeepUnet with the Focal loss function, AMSGrad optimizer, and VD expanded with slices 001, 086, 277, 553.
6	K2	DeepUnet with Focal loss function and AMSGrad optimizer and VD expanded with slices 035, 439, 477, 519.
7	K3	DeepUnet with Focal loss function and AMSGrad optimizer and VD expanded with slices 101, 439, 519, 553.
8	Retrained	DeepUnet with Focal loss function, AMSGrad optimizer, and TD expanded with slices 001ll, 402, 443, 519ur, 553ur (001ll, 519ur, 553ur are parts of 519 and 553 slices, but not overlapping with the original TD).

### Performance Metrics

We evaluated the CNN performance by calculating standard metrics via a confusion matrix (Powers and Ailab 2011). A confusion matrix is a performance measurement giving combinations of predicted and actual values. The elements of the confusion matrix enable the calculation of derivative parameters such as accuracy, recall, precision, and F1 score. Accuracy is the ratio of correctly predicted examples by the total examples (3). The recall is the ratio of correct positive predictions to the total positive examples (4), while precision is the ratio of correct positive predictions to the total predicted positives (5). F1 score is a measure to compare both recall and precision (6).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$F1 \text{ Score} = 2 \frac{Recall \cdot Precision}{Recall + Precision} \quad (6)$$

### Results

At the first step, we tested configurations 1–4. To avoid GPU overload and memory shortage, we split the TD slices into 448×448 px tiles in CE configuration. As a result, we have got low segmentation quality due to patch stitching back after segmentation.

Afterward, we studied the results by FL\_y1, where the Focal Loss suppresses class imbalance, and  $\gamma$  smoothly adjusts the rate at which easy examples are down-weighted. Setting  $\gamma > 0$  reduces the relative loss for well-classified examples (probability > 0.5), putting more focus on hard, misclassified cases. We used  $\gamma = 1$  value along with the AMSGrad optimizer.

For the Adam configuration, we replaced AMSGrad with Adam optimizer and Focal Loss with  $\gamma = 2$ . This model showed consistent results (Figure 8), and we assumed that  $\gamma = 2$  is the best setting for Focal Loss.

To test the error rate convergence while selecting the optimizer, we switched to SGD. Even though the training stage took as twice as much time as Adam did — 5 h of training for 36,000 epochs and 10 h of training for the same number of epochs, the SGD provided reliable results (Figure 8). However, our visual inspection revealed mistakes in segmentation.

Having compared FL\_y1, Adam, and SGD configurations, we did not notice significant advantages of SGD, while the computation time doubled. We then decided to stop switching both optimizer and loss function and arrived at AMSGrad and Focal Loss with  $\gamma = 2$ .

We analyzed the segmentation performance of configurations 1–4 and revealed several slices (035, 101, 116, 339) with FN values higher than those for others (Figure 6). The low FN values mean underestimation of the pore space. FP values fall in a range of 0–3%, suggesting a minimal misclassification of the rock matrix. We consider that the correct classification of the rock matrix is more important than an underestimation of pore space. Thus, the configurations show a positive trend.

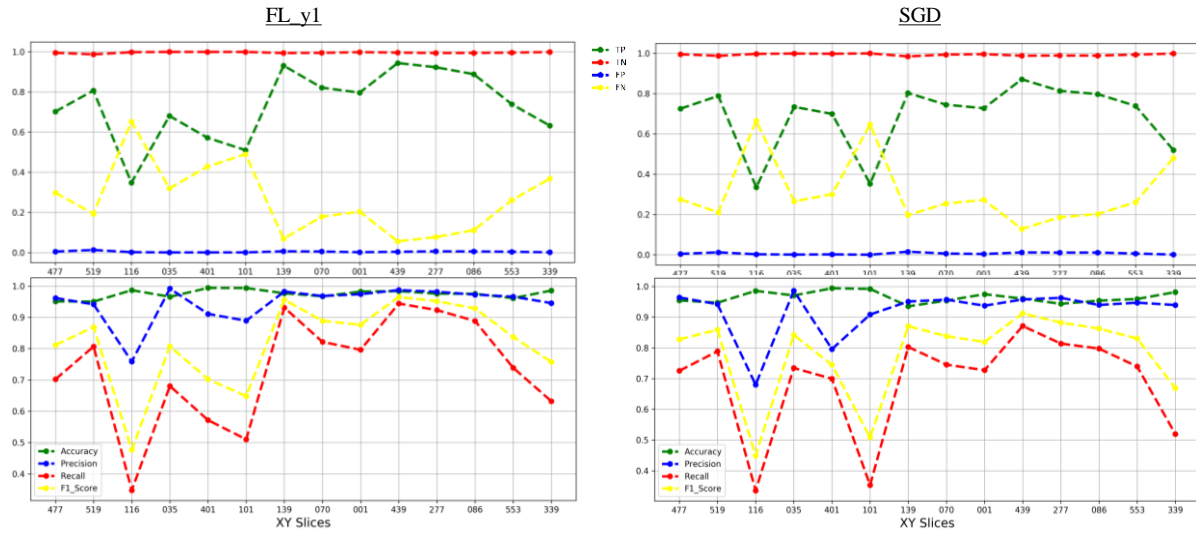


Figure 6. Performance analysis of 0FL\_y1 and SGD configurations. The top plots show to the components the confusion matrix, while the bottom plots present the derivative parameters.

An analysis of the segmentation performance of configurations 1–4 motivated the study of TD and VD composition. We used K-folding to create two more configurations identical to the host ones but with different TD and VD patches. The technique is a common approach to check model stability and its independence on of choice for slices within a single TD. For this purpose, we created configurations K1, K2, K3 with VD including slices 001, 086, 277, 553; slices 035, 439, 477, 519; and slices 101, 439, 519, 553 correspondingly.

K1 configuration showed excellent results, while K2 and K3 showed instable performance (Figure 7). Thus, we decided to expand TD rather than to exchange TD and VD.

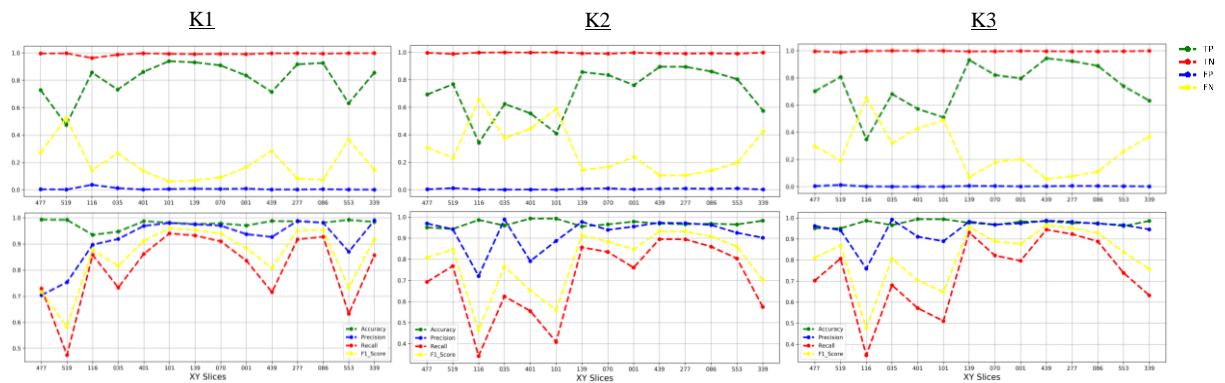


Figure 7. Performance analysis of configurations K1, K2, K3. The top plots show the components of the confusion matrix, while the bottom plots present the derivative parameters.

All the studied configurations show similar trends (Figure 7). We noticed that few slices had difficult-to-segment features, leading to a segmentation quality lower than that for most of TD and VD slices. The slices 035, 101, 116, 339, and 401 turned out to be the most problematic, while segmentation of all the rest yielded to lower error rates with all the configurations.

Moreover, the slices highly correlate with each other. Thus we have tested an idea to expand TD with extra labeled slices and areas (patches). The TD expansion included 5 items: slices 402 and 443, and patches 001ll, 519ur, 553ur, where ll stands for lower left and ur — upper right. The additional data items reside close to the problematic slices. Since K1 performed best, we used it as the basis for the Retrained configuration. Retrained originated from K1 by additional training on the expanded TD for 12,000 epochs. However, this effort has not delivered further growth in segmentation quality. Table 5 summarizes the results of our study.



Table 5. Summary of average segmentation quality metrics for all the studied configurations.

Configuration Name	Epochs	Training Time	Learning Rate	Segmentation Time (sec)	Accuracy	Precision	Recall	F1 Score
CE	36,000	05:30:00	0.0010	45–50	97.45	92.94	74.88	82.15
Fl_y1	36,000	05:30:00	0.0010	45–50	97.14	93.55	70.50	79.61
Adam	36,000	05:30:00	0.0010	45–50	97.66	91.00	78.52	84.09
SGD	58,000	10:40:00	0.0010	45–50	96.46	91.13	67.63	76.89
K1	36,000	06:00:00	0.0001	45–50	96.66	91.67	67.57	77.00
K2	36,000	06:00:00	0.0001	45–50	97.44	88.84	77.57	82.43
K3	36,000	06:00:00	0.0001	45–50	91.43	83.45	74.56	78.13
Retrained	48,000	09:00:00	0.0001	45–50	96.24	77.31	85.64	80.19

[Figure 8](#) illustrates the segmentation performance for all the studied configurations. We picked the slice 001 as a representative one because it is the “hardest to interpret” for all configurations. Furthermore, we evaluated the performance in the lower right part where there is a massive whitish spot, and almost all the configurations misclassified it to be a pore space region.

Interestingly, although the Adam configuration delivered the best metrics ([Table 5](#)), a visual inspection of the segmentation results clearly shows that K1 performs much better ([Figure 8](#)). We have not expected the observation, and thus it is a promising direction for further research.

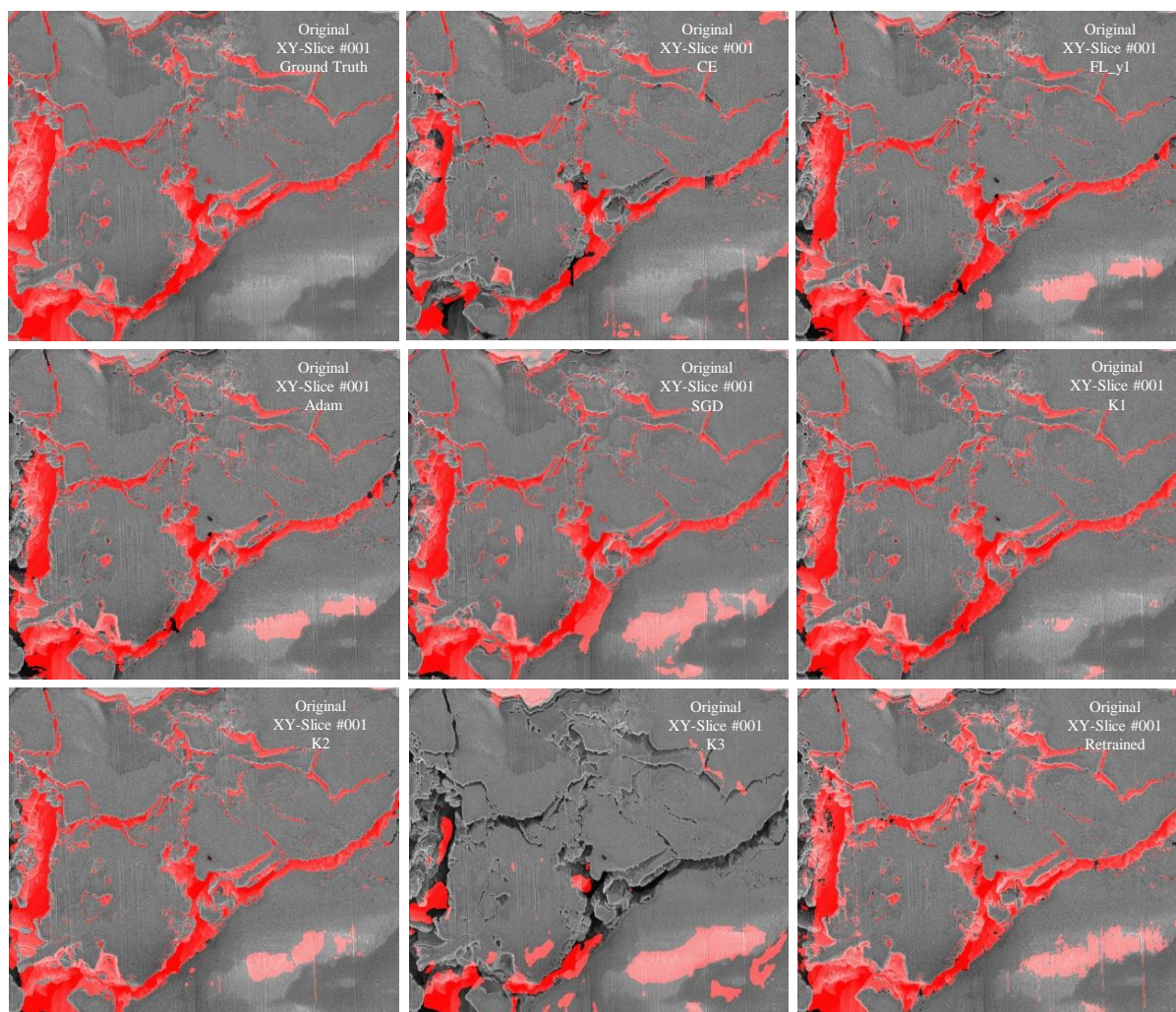


Figure 8. Visual comparison of segmentation results by all the studied configurations on the target Dataset 7.1.1, XY-Slice #001.

Besides, we analyzed the performance on Slice #001 Dataset 7.1.2, which was not a part of TD and is entirely new to all configurations. We observed fair performance results by almost all the studied configurations. However, both K2 and K3 showed massive misinterpretation, while the Retrained misinterpreted the image warping artifact as a pore (Figure 9).



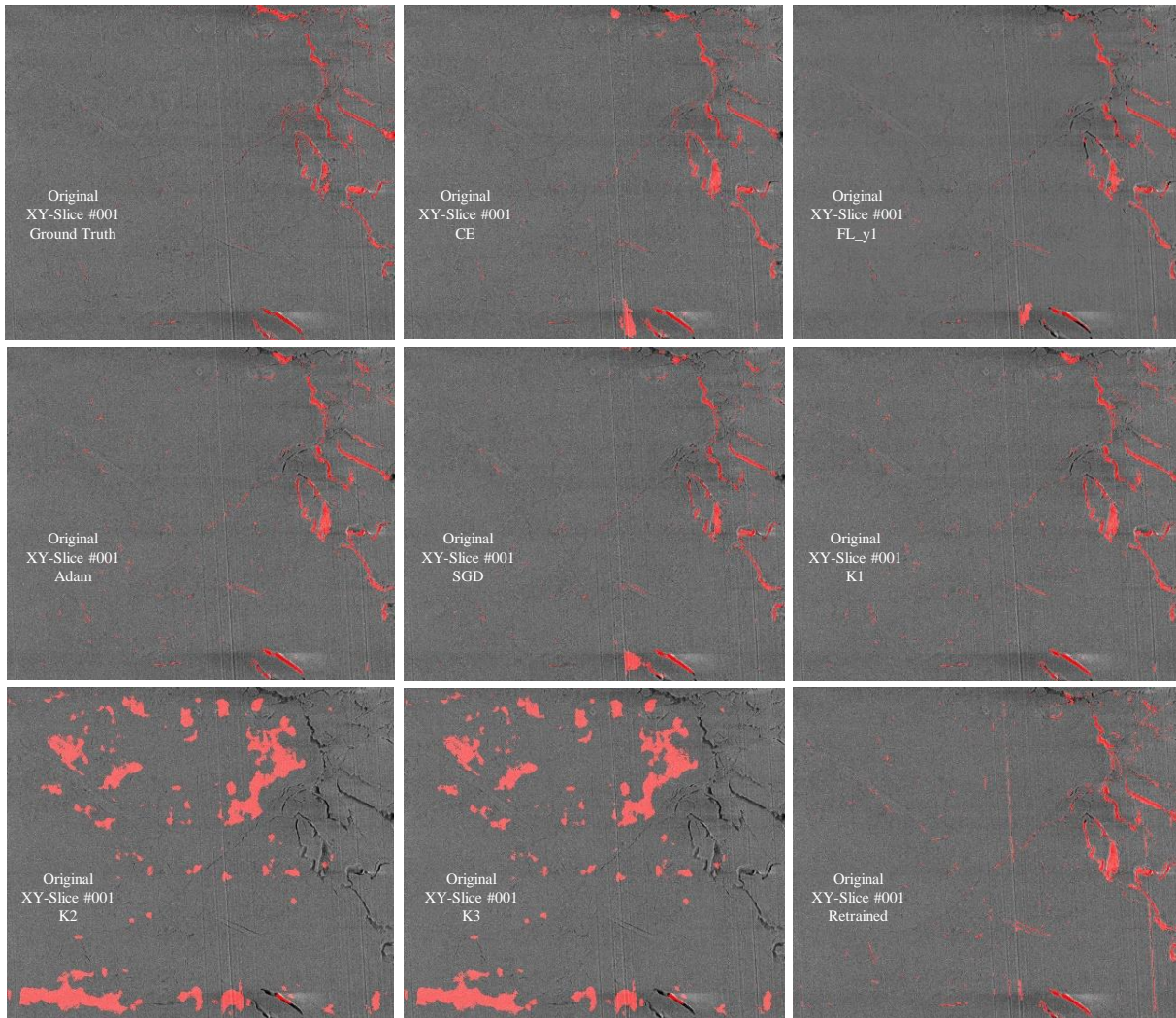


Figure 9. Visual comparison of segmentation results by all the studied configurations on the target Dataset 7.1.2, XY-Slice #001.

To provide the most stressful test of performance, we ran all configurations on an entirely new slice from another target dataset 3.1.1. [Figure 10](#) compares the application results of the so-called transfer learning technique. The idea is in overcoming the isolated learning paradigm and utilizing the knowledge acquired for one task to solve related ones ([Weiss, Khoshgoftaar, and Wang 2016](#)). All the configurations failed to accomplish the exercise. From this experience, we conclude that different FIB-SEM datasets (image stacks) obtained with different parameters and characterizing different rocks differ dramatically, disabling a straightforward application of transfer learning.

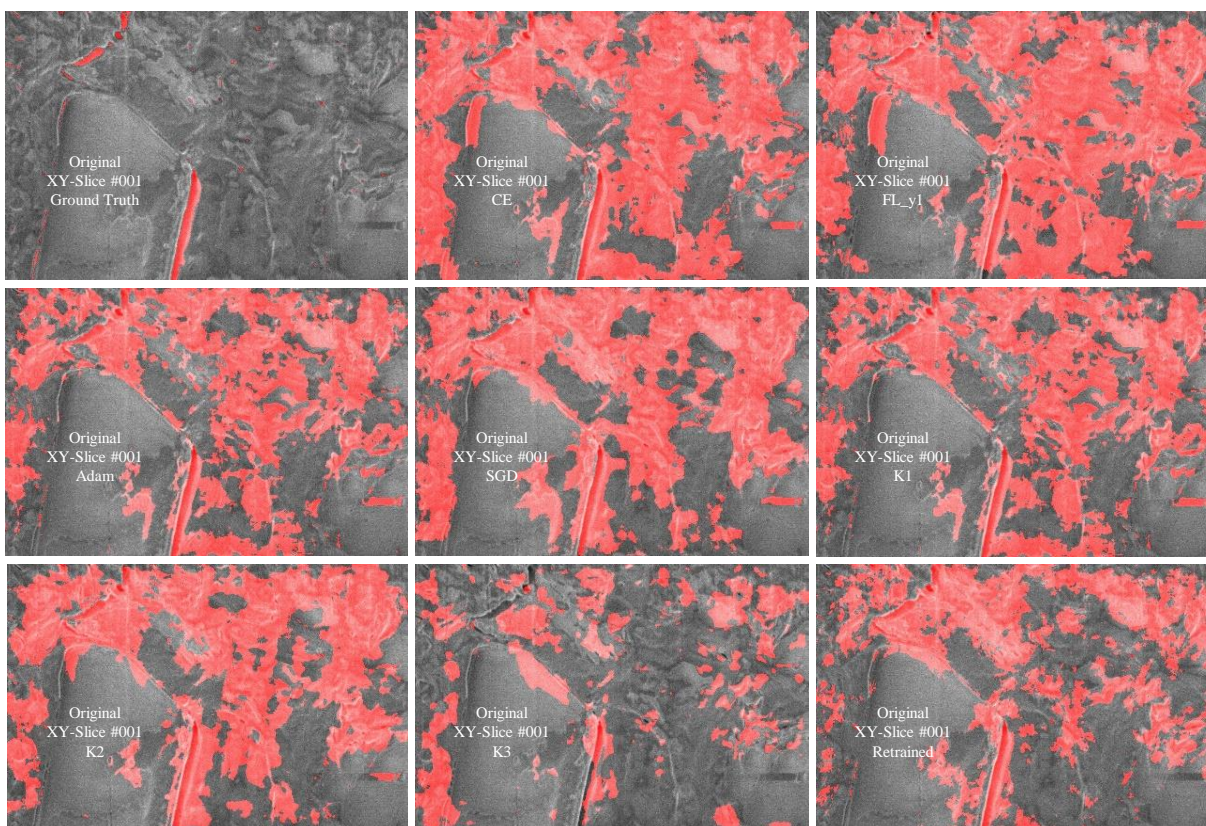


Figure 10. Visual comparison of segmentation results by all the studied configurations on the target Dataset 3.1.3, XY-Slice #001.

According to the integrated evaluation of performance metrics and results of visual inspection, K1 configuration delivers the best segmentation results. Adam configuration takes the second place, although Adam's formal metrics are higher than those of K1.

[Figure 11](#) compares the derivative performance indicators for K1, K2, K3 for arbitrary slices from all the target datasets and highlights sensitivity to the selection of target slices. For all three configurations, the accuracy gradually decreases from 100% for the TD and VD slices to nearly 80% for slices from Datasets 7.1.2 and 3.1.1 previously unseen by the CNN. At the same time, the results show a class imbalance towards a rock matrix of 70–90%.

K1 shows both precision and F1 score performance indicators above 80% for the slices from TD, VD, and Datasets 7.1.1 and 7.1.2. We consider these results pretty successful.

K2 demonstrated all the performance indicators of above 60% for the test slices from TD, VD, and 7.1.1. At the same time, K3 shows fair results only for TD and VD classification with a strong recommendation for further results validation by visual inspection.

The comparison shows that the CNN configurations strongly depend on the choice of TD for the correct segmentation of FIB-SEM images even for the same sets of tuning parameters.



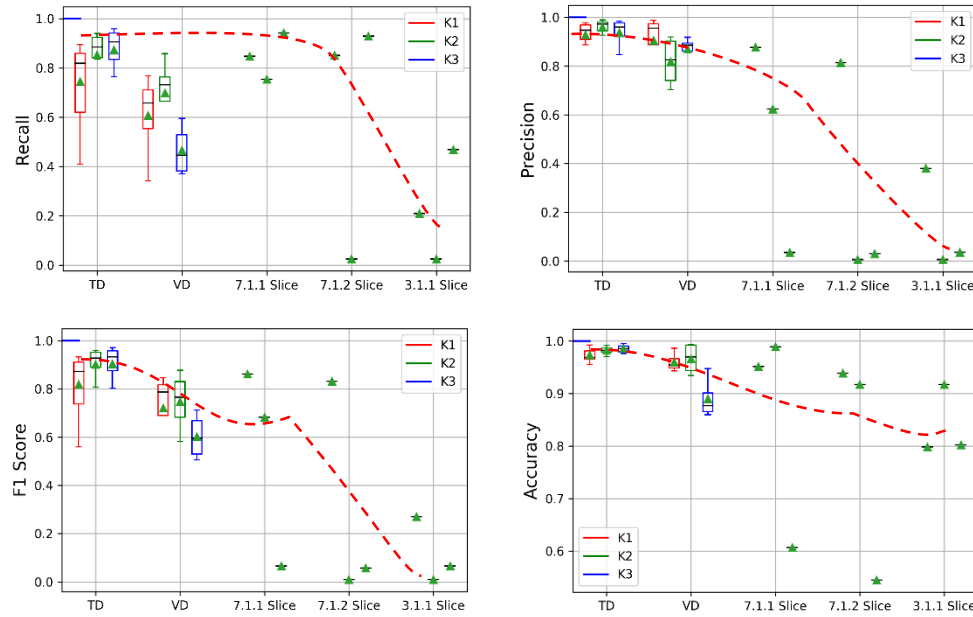


Figure 11. Comparison of segmentation quality metrics for the best CNN configurations for different VD selections, including TD, VD, arbitrary slices of Datasets 7.1.1, 7.1.2, and 3.1.1. The red dashed line shows the averaged results.

## Discussion

During image segmentation, we prefer to underestimate the pore space volume. An overestimation of porosity leads to an optimistic evaluation of both the storage (porosity) and transport (permeability) properties of the target reservoir rock.

Among all the tested configurations, we chose K1 because it has the best performance according to visual inspection. K1 correctly segmented the maximum number of regions, including complicated ones such as the bottom right corner of Figure 8. All other configurations segmented this region erroneously as the pore space.

With TD of only 14 slices, the configuration successfully segmented the slices of both Datasets 7.1.1 and 7.1.2 and showed a dramatic time reduction. We timed the manual segmentation of a single slice from Dataset 7.1.1 as 10–80 min., while the CNN segmentation took only 45–60 s. Our estimate of the time required for training and segmentation of Dataset 7.1.1 ranges within 13–14 h., while the manual segmentation can take 92–700 h. We consider this to be an excellent achievement.

However, segmentation of slices from Dataset 3.1.1 showed that even the best performing configuration currently does not provide satisfactory results, suggesting the transfer learning technique is likely not applicable in this case. Expansion of existing TD with slices from the unseen dataset may potentially mitigate the issue.

Another important feature would be the systematic underestimation of the pore space even if CNN performed adequately and successfully suppressed the FIB-SEM artifacts.

K-folding manipulation showed that the proper choice of TD is crucial and may lead to poor performance. Thus, the selection of TD should be carried out thoroughly using algorithms for image statistics evaluation to provide the most representative slices from the whole target dataset.

Although we did not vary the image dimensions too much, the CE configuration demonstrated that with a narrow field of view, the CNN does not capture enough features to segment image correctly.

In this study, we focused on testing the DeepUnet architecture without considering others. However, to date, there are many promising CNN candidates (Sultana, Sufian, and Dutta 2020) for a feasibility study in the scope of FIB-SEM image segmentation. We consider the selection of the optimal architectures and configurations as an important direction for further research on the topic.

## Conclusions

The correct determination of storage and transport properties from FIB-SEM images requires fast and robust segmentation. However, the quality and efficient segmentation of FIB-SEM images is still a complicated and challenging task. The presence of FIB-SEM artifacts, such as pore backs, requires the development of a new methodology for efficient image segmentation.

The conducted literature review demonstrated that the conventional segmentation techniques, including global thresholding and watershed thresholding, show poor performance for the segmentation of FIB-SEM images complicated with artifacts. Typically, a trained operator spends days/week for subjective and semi-manual labeling of a single FIB-SEM dataset.

We attacked the quality segmentation problem in the presence of imaging artifacts using DeepUnet CNN. The architecture has a unique design suitable for high-resolution images and features extraction from small local areas on images and training the weight of these features to classify objects.

The target datasets came from FIB-SEM imaging of a representative collection of tight-gas reservoir rock samples (Berezov Formation, West Siberia, Russia). The complex void space structure of the studied rock complicated the digital rock images with acquisition artifacts, including porebacks, curtaining, bright pore edges, and complicated 3D pores. For training, we used a subset of a limited number of cropped and manually labeled slices from the target dataset. We then augmented the training dataset by a set of image transforms, including vertical and horizontal flip, rotation, blurring, histogram normalization, gray level expansion. Eventually, we developed and implemented the configuration, which can automatically segment the target datasets.

The application of the DeepUnet to the target FIB dataset delivered promising results. Firstly, the CNN misclassification rate is 2%. Secondly, CNN pore space segmentation to two classes showed higher quality than that to a single class. Thirdly, CNN segmentation speed is 14.5 Mpx/min., which is more than 10 times higher than the average manual segmentation speed is 0.18–1.45 Mpx/min.

Directions for further development include testing of different architectures besides DeepUnet and intelligent choice of training data.

In summary, we developed a robust approach for an automated highly-efficient multimodal segmentation of FIB-SEM datasets using machine-learning (ML) based methods. The research results significantly improve the quality of quantitative petrophysical characterization of complex reservoir rocks using digital rock FIB-SEM imaging.

## Acknowledgments

*NOVATEK, Moscow, Russia:* for providing reservoir rock samples and fruitful discussions of the results.

*Prof. Mikhail Spasennykh, Prof. Dimitri Pissarenko, and Prof. Alexey Cheremisin* (Skolkovo Institute of Science and Technology, Moscow, Russia) for providing an opportunity to visit the UNSW and learn Digital Core Analysis techniques.

*Ignacio Arganda-Carreras* (Ikerbasque Research Fellow, UPV/EHU) for generous help in tuning Transform Virtual Stack Slices plugin of ImageJ for inverse registration of images.

## References

- Andra, H., Combaret, N., Dvorkin, J. et al. 2013a. Digital Rock Physics Benchmarks-Part I: Imaging and Segmentation. *Computers & Geosciences* **50**: 25-32. WOS:000313611100004. <http://dx.doi.org/10.1016/j.cageo.2012.09.005>.
- Andra, H., Combaret, N., Dvorkin, J. et al. 2013b. Digital Rock Physics Benchmarks-Part II: Computing Effective Properties. *Computers & Geosciences* **50**: 33-43. WOS:000313611100005. <http://dx.doi.org/10.1016/j.cageo.2012.09.008>.
- Andrew, M. 2018. A Quantified Study of Segmentation Techniques on Synthetic Geological XRM and FIB-SEM Images. *Computational Geosciences* **22** (6): 1503-1512. <http://dx.doi.org/10.1007/s10596-018-9768-y>.
- API. 1998. Recommended Practices for Core Analysis. Recommended Practice 40. Dallas, TX: American Petroleum Institute (API).

Durand, M., Nikitin, A., McMullen, A. et al. 2019. Crushed Rock Analysis Workflow Based on Advanced Fluid Characterization for Improved Interpretation of Core Data. SPWLA 60th Annual Logging Symposium, The Woodlands, Texas, USA, 2019/6/15/.

Dvorkin, J., Derzhi, N., Diaz, E. et al. 2011. Relevance of Computational Rock Physics. *Geophysics* **76** (5): E141-E153. WOS:000297946200009. <http://dx.doi.org/10.1190/Geo2010-0352.1>.

EIA. 2017. International Energy Outlook 2017 Overview (IEO2017). Washington, DC..

Frucchi, M., Sanniti, G. and Sanniti di Baja, G. 2008. From Segmentation to Binarization of Gray-Level Images. *Journal of Pattern Recognition Research* **3**. <http://dx.doi.org/10.13176/11.54>.

Giannuzzi, L. and Stevie, F. 2005. Introduction to Focused Ion Beams: Instrumentation, Theory, Techniques and Practice.

Goral, J., Deo, M. and Andrew, M. 2018. Pore Network Modeling of Marcellus Shale using Digital Rock Analysis with Machine Learning Image Segmentation. 7-10.

Goral, J., Walton, I., Andrew, M. et al. 2019. Pore System Characterization of Organic-Rich Shales Using Nanoscale-Resolution 3D Imaging. *Fuel* **258**. <http://dx.doi.org/10.1016/j.fuel.2019.116049>.

Gurney, K. 1997. An Introduction to Neural Networks. London: Taylor Francis, Inc..

Kazak, A., Chugunov, S., Nachev, V. et al. 2017. Integration of Large-Area SEM Imaging and Automated Mineralogy-Petrography Data for Justified Decision on Nano-Scale Pore-Space Characterization Sites, as a Part of Multiscale Digital Rock Modeling Workflow.

Kazak, E. S. and Kazak, A. V. 2019. A Novel Laboratory Method for Reliable Water Content Determination of Shale Reservoir Rocks. *Journal of Petroleum Science and Engineering* **183**: 106301. <http://dx.doi.org/10.1016/j.petrol.2019.106301>.

Kelly, S., El-Sobky, H., Torres-Verdín, C. et al. 2016. Assessing the Utility of FIB-SEM Images for Shale Digital Rock Physics. *Advances in Water Resources* **95**: 302-316. <http://dx.doi.org/https://doi.org/10.1016/j.advwatres.2015.06.010>.

Kingma, D. P. and Ba, J. 2015. Adam: A Method for Stochastic Optimization..

Kuila, U. 2013. "Measurement and Interpretation of Porosity and Pore-Size Distribution in Mudrocks. The Hole Story of Shales." Doctor of Philosophy (Petroleum Engineering), Colorado School of Mines.

Li, R., Liu, W., Yang, L. et al. 2018. DeepUNet: A Deep Fully Convolutional Network for Pixel-Level Sea-Land Segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **11** (11): 3954-3962. <http://dx.doi.org/10.1109/JSTARS.2018.2833382>.

Lin, T. Y., Goyal, P., Girshick, R. et al. 2020. Focal Loss for Dense Object Detection. *IEEE Trans Pattern Anal Mach Intell* **42** (2): 318-327. 30040631. <http://dx.doi.org/10.1109/TPAMI.2018.2858826>.

Liu, X., Sun, J. and Wang, H. 2009. Numerical Simulation of Rock Electrical Properties Based on Digital Cores. *Applied Geophysics* **6** (1): 1-7. <http://dx.doi.org/10.1007/s11770-009-0001-6>.

Luffel, D. L. and Guidry, F. K. 1992. New Core Analysis Methods for Measuring Reservoir Rock Properties of Devonian Shale. *Journal of Petroleum Technology* **44** (11): 1184-1190. <http://dx.doi.org/10.2118/20571-PA>.

Murphy, K. 2012. Machine Learning: A Probabilistic Perspective: MIT.

Nelson, P. H. 2009. Pore-Throat Sizes in Sandstones, Tight Sandstones, and Shales. *Geologic Note. AAPG Bulletin* **93** (3): 329-340.

Nikolaev, M. Y. and Kazak, A. V. 2019. Liquid Saturation Evaluation in Organic-Rich Unconventional Reservoirs. A Comprehensive Review. *Earth-Science Reviews* **194** (194C): 327-349. <http://dx.doi.org/10.1016/j.earscirev.2019.05.012>.

Powers, D. and Ailab. 2011. Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Markedness & Correlation. Vol. 2.

Prill, T., Schladitz, K., Jeulin, D. et al. 2013. Morphological segmentation of FIB-SEM data of highly porous media. *Journal of Microscopy* **250** (2): 77-87. <http://dx.doi.org/10.1111/jmi.12021>.

Reuteler, J. 2017. FIB Artifacts and Tricks to Overcome Them. .

Ronneberger, O., Fischer, P. and Brox, T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. .

Salzer, M., Spettl, A., Stenzel, O. et al. 2012. A Two-Stage Approach to the Segmentation of FIB-SEM Images of Highly Porous Materials. *Materials Characterization* **69**: 115-126. <http://dx.doi.org/https://doi.org/10.1016/j.matchar.2012.04.003>.

Salzer, M., Thiele, S., Zengerle, R. et al. 2014. On the Importance of FIB-SEM Specific Segmentation Algorithms for Porous Media. *Materials Characterization* **95**: 36-43. <http://dx.doi.org/10.1016/j.matchar.2014.05.014>.

Seni, G. and Elder, J. F. 2010. Ensemble Methods in Data Mining: Improving Accuracy Through Combining Predictions. *Synthesis Lectures on Data Mining and Knowledge Discovery* **2** (1): 1-126. <http://dx.doi.org/10.2200/S00240ED1V01Y200912DMK002>.

Sultana, F., Sufian, A. and Dutta, P. 2020. Evolution of Image Segmentation using Deep Convolutional Neural Network: A Survey. *Knowledge-Based Systems* **201-202**: 106062. <http://dx.doi.org/https://doi.org/10.1016/j.knosys.2020.106062>.

Sun, H., Yao, J., Cao, Y.-c. et al. 2017. Characterization of Gas Transport Behaviors in Shale Gas and Tight Gas Reservoirs by Digital Rock Analysis. *International Journal of Heat and Mass Transfer* **104**: 227-239. <http://dx.doi.org/10.1016/j.jheatmasstransfer.2016.07.083>.

Taillon, J. A., Pellegrinelli, C., Huang, Y.-L. et al. 2018. Improving Microstructural Quantification in FIB/SEM Nanotomography. *Ultramicroscopy* **184**: 24-38. <http://dx.doi.org/https://doi.org/10.1016/j.ultramic.2017.07.017>.

Tracey, J., Lin, S., Jankovic, J. et al. 2019. Iterative Machine Learning Method for Pore-Back Artifact Mitigation in High Porosity Membrane FIB-SEM Image Segmentation. *Microscopy and Microanalysis* **25**: 186-187. <http://dx.doi.org/10.1017/S1431927619001661>.

Tran Thi Phuong, L. T. P. 2019. On the Convergence Proof of AMSGrad and a New Version. *IEEE Access* **7** (1). <http://dx.doi.org/10.1109/ACCESS.2019.2916341>.

Weiss, K., Khoshgoftaar, T. M. and Wang, D. 2016. A Survey of Transfer Learning. *Journal of Big Data* **3** (1). <http://dx.doi.org/10.1186/s40537-016-0043-6>.

Zhang, S., Choromanska, A. and LeCun, Y. 2015. Deep Learning with Elastic Averaging SGD..

Zurada, J. 1992. Introduction to Artificial Neural Systems. St. Paul: West Publishing Co..