

SPE-201282-MS

Segmentation of X-Ray Images of Rocks Using Deep Learning

Ibrahim Ar Rushood, Saudi Aramco D&WO; Naif Alqahtani and Ying Da Wang, The University of New South Wales; Mehdi Shabaninejad, The Australian National University; Ryan Armstrong and Peyman Mostaghimi, The University of New South Wales

Copyright 2020, Society of Petroleum Engineers

This paper was prepared for presentation at the SPE Annual Technical Conference & Exhibition originally scheduled to be held in Denver, Colorado, USA, 5 – 7 October 2020. Due to COVID-19 the physical event was postponed until 26 – 29 October 2020 and was changed to a virtual event. The official proceedings were published online on 21 October 2020.

This paper was selected for presentation by an SPE program committee following review of information contained in an abstract submitted by the author(s). Contents of the paper have not been reviewed by the Society of Petroleum Engineers and are subject to correction by the author(s). The material does not necessarily reflect any position of the Society of Petroleum Engineers, its officers, or members. Electronic reproduction, distribution, or storage of any part of this paper without the written consent of the Society of Petroleum Engineers is prohibited. Permission to reproduce in print is restricted to an abstract of not more than 300 words; illustrations may not be copied. The abstract must contain conspicuous acknowledgment of SPE copyright.

Abstract

Segmentation of X-ray images of rocks is an important step in digital rock technology. Current segmentation methods often suffer from operator bias and can be time-consuming. To overcome these limitations, an automated image segmentation model is created for pore scale images through the use of convolutional neural networks. A dataset of micro-Computed Tomography (CT) images of sandstones is considered. To create the ground truth data for training, Scanning Electron Microscopy 2D and micro-CT 3D images are used to obtain accurate segmentation masks. Three models are trained with the available data: One with limited images, one with the full dataset, and one with augmented data. The data augmentation is achieved by increasing sample size through image partitioning. The data set for each model is divided into training, validation, and testing with a 60/30/10 split, respectively. The U-Net architecture, designed to work with limited training data, is used to develop the models. Further validation of the models is performed on a different dataset unseen by the models. The Minkowski functionals and permeability of the volume generated by the segmented images are computed and compared with the ground truth segmentation. Against the unseen dataset, the models scored a dice coefficient of 0.9479, 0.9518, and 0.9599, respectively. We discuss the potential for improvement by data augmentation and fine-tuning. One limitation of the models is with the deficiency of variety in training data as both SEM and micro-CT is required to obtain the segmentation masks. We also discuss the performance of the model on unseen samples and show the potential improving efficiency in the digital rock technology. The models provide a quick and accurate segmentation for images of sandstones without the influence of operator bias, and the method shows promise for further development and improvement.

Introduction

The use of pore-scale imaging and digital rocks has seen a rapid rise in the last twenty years. Pore-scale images are, as the name implies, three dimensional scans of physical rock samples resolving microstructures at the pore level. Applications for this technology are important for porous media characterization. Such applications include the calculation of rock properties and simulation of flow (Blunt et al., 2013).

The main method through which digital rocks are obtained is a three-dimensional reconstruction of two-dimensional X-ray projections taken of a sample from multiple angles. The images are in a grey-scale format. They require image processing as well by filtering and segmentation. Filtering is the removal of artifacts (noise, beam hardening, etc.) affecting the quality of the image, and is an already established topic in the medical field. Segmentation is assigning each pixel to a phase in the image. Therefore, it is a critical and defining step for any quantitative analysis to be performed on the sample. There has been a variety of segmentation methods reported in the literature. However, there is no standardized method, and a sample segmented by two individuals may look different and achieve different results (Baveye et al., 2010; Tuller et al., 2013).

For pore scale images, segmentation is assigning each pixel/voxel of an image to a phase in the rock (pore space, solid grains (general or to a specific minerals), fluids if present) based on the properties of the pixel compromising the image (e.g. color, intensity, texture, etc.) (Kohler, 1981, Sridevi and Mala, 2012). For pore-scale imaging, segmentation is critical to define the features of sample. Segmentation of two features of interest is the simplest type of segmentation. This is commonly known as binarization (Tuller et al., 2013), and the two features may be referred to as the foreground and the background (Baveye et al., 2010). For pore-scale imaging, the two features of interest are the void and the matrix of the imaged sample. Multi-phase (e.g. voids, matrix, and saturating fluid) segmentation is possible, however, it is generally more complex (Tuller et al., 2013).

Micro-CT images of rocks have one greyscale channel unlike coloured images. the intensity of each pixel (or voxel for three dimensional images) is a definite number. This number lies between 0 to 255 in the 8-bit grayscale range, or 0 to 65535 in the 16-bit grayscale range (Tuller et al., 2013). Through the use of the greyscale intensity of each voxel, binarization is easily achieved in simple cases by choosing a threshold at a certain greyscale intensity value to divide the foreground and the background phases (Frucci and di Baja, 2008). Thresholding can be done globally (i.e. a threshold that applies on the image as a whole) or locally (through the use of a variable threshold across multiple regions in the image) (Tuller et al., 2013, Iassonov et al., 2009).

Segmentation has a crucial impact on any subsequent analysis done on the resulting image. In an experiment done by Baveye et al. (2010), three test images were made available for pre-processing and segmentation. Experts were invited to pre-process and segment the images and were encouraged to use the standard procedures they use for their research. Participating experts were not informed of the intended use of the resulting segmented images. Automated segmentation methods were applied by the authors of the paper as well. The resulting images were vastly different between both the experts' segmented images and the auto-segmented images showing operator bias was prevalent as well. These differences in a sample segmentation propagate forward to any subsequent analysis done on the sample (e.g. flow simulation and pore network modelling).

On the other hand, deep learning is a subset of machine learning methods that perform feature learning through multiple layers of non-linear functions, with each layer amplifying the features of importance. Convolutional neural networks (CNNs) are a class of neural networks specialized in pattern recognition and image-based problems. CNNs has proven useful in object detection (Cai et al., 2016), speech recognition (Sainath et al., 2013), and various applications of semantic segmentation (Long et al., 2015, Pereira et al., 2016) among many other applications. CNNs have been used in the digital rocks field including the development of super resolution pore scale images (Wang, 2018, Wang et al., 2019a, Wang et al., 2019b), image segmentation (Karimpouli and Tahmasebi, 2019, Wang et al., 2020), and estimation of sample's physical properties (Alqahtani et al., 2020).

This paper tests the image segmentation performance of the U-Net (a convolutional neural network) using high resolution pore-scale images as a means to overcome the problem of operator bias in image segmentation. The segmentation of the models is further tested by comparing several physical and

morphological properties of 3D volumes segmented using our model compared to the ground truth segmentation.

Dataset and Methodology

Training Dataset

The dataset used for training the segmentation models consists of 2D slices of a 3D image of Mt. Simon sandstone reservoir (Figure 1). A 3 mm diameter and 5 mm length core of Mt. Simon reservoir miniplug was imaged with micro-CT scanning at a resolution of 1.8 μm . The Mt. Simon sample is a naturally challenging sample, due to the irregular, rough, and slightly under-resolved clay minerals that coat the quartz grains, creating a complex pore space with many small chambers and tight closures which present a challenge to segmentation, whether it be by traditional methods or by CNNs.

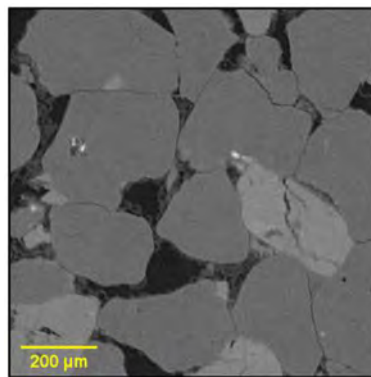


Figure 1—Micro-CT slice of Mt. Simon sandstone reservoir.

The same core was scanned and imaged with Quantitative Evaluation of Minerals by Scanning Electron Microscopy (QEMSCAN) (Wang et al., 2020). QEMSCAN retrieves high resolution digital images through Scanning Electron Microscopy (SEM). At the same time, it obtains energy dispersive x-ray (EDX) spectra and backscattered-electron image (BEI) information and uses that information to identify the mineralogy of each pixel (Ayling et al., 2012). QEMSCAN returns 2D slices of the core with accurate multi-mineral segmentation, which is then registered into the 3D micro-CT image. Despite the availability of the multi-mineral segmentation, the main interest of this paper is the development of a binary segmentation of solid and pore phases. Therefore, all the multi-mineral voxels are considered as solid voxels.

The dataset consists of 2,200 slices of size 1000×1000 pixels. Three different models are created with this dataset. One model is trained on the full dataset. This model will be referred to as Model 2200. Another model is trained with data augmentation through an increased sample size. The increased sample size is achieved by sample duplication to a total of 4,400 slices and image partitioning to four quarters per slice, with a quarter size of 700×700 pixels with an overlap of 300 pixels between each quarter for a total of 17,600 slices. This model will be referred to as Model 17600. The third model is trained with only 600 slices from the middle of the sample, with each slice is $900 \text{ pixels} \times 900 \text{ pixels}$. Its purpose is to measure the network's performance when given a limited dataset. This model will be referred to as Model 600. For each model, the dataset is divided into training, validation, and testing set with 60%, 30%, and 10% split.

Testing Dataset

Besides testing each model on 10% of the Mt. Simon training dataset allocated for testing, each model is further tested on a separate sample previously unseen by the model. High resolution SEM image (with 6100×6100 pixels) of the North sea sandstone sample are resgistrred to the micro CT image (Figure 2). Both of the micro-CT image and the SEM image will be used separately to test the models. The micro-CT

images of the sample are segmented using Otsu's method (Otsu, 1979). The SEM images are segmented by the watershed-based method using Avizu software (Niu et al., 2020), and these two segmentations will be considered as the ground truth. To test the models, 2,000 random 700×700 pixels slices of the micro-CT image are segmented using each model, and the model's segmentation is compared against the ground truth. The SEM image will be used similarly.

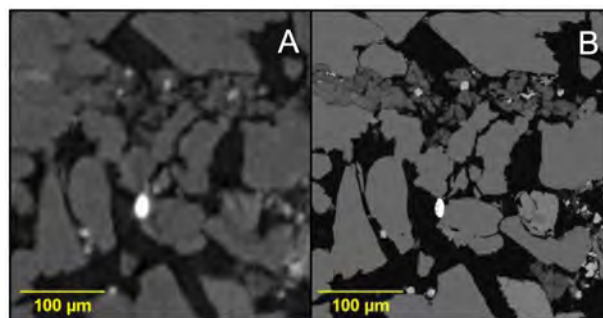


Figure 2—Micro-CT image (A) and SEM image (B) of North Sea sandstone reservoir.

Network Architecture

The model was developed using U-Net architecture. U-Net is a CNN developed specifically for biomedical image segmentation. Prior to U-net development, CNNs were achieving excellent results in image classification where the result is a single label (e.g. use of CNNs to identify the animal in an image). On the other hand, segmentation requires a classification label for each pixel. Another problem with CNNs at the time is the quantity of training data required for the model to achieve the required results, which is a luxury not necessarily available for the biomedical field (Ronneberger et al., 2015). The same CNNs limitations that prompted the development of U-Net apply for digital rock images, which make the U-Net architecture a suitable candidate for digital rock images segmentation.

U-Net architecture (Figure 3) is not much different than the standard CNNs with encoding and decoding sides. The major difference lies in concatenating the feature map from an encoding layer to its equivalent decoding layer on the decoding side. This addition enables the U-Net architecture to achieve greater pixel-wise accuracy (Wang et al., 2020) and overall better results (Isola et al., 2018). Referring to Figure 3, the encoding side of U-Net network starts with a double 3×3 convolution, with each convolution followed by a rectified linear unit (ReLU) activation function. The double convolution is followed by 2×2 max pooling with a 2-step stride. The decoding side starts with a 2×2 upsampling followed by the same double convolution sequence in the encoding side. The decoding side ends with a final convolution to create the output (Ronneberger et al., 2015). The implementation of U-Net in this paper adds batch normalization between the convolution and ReLU. However, it has to be pointed out that all the model training was done with a batch size of one. Therefore, in essence, this paper implementation of U-Net is no different from the original U-Net described by Ronneberger et al (2015).

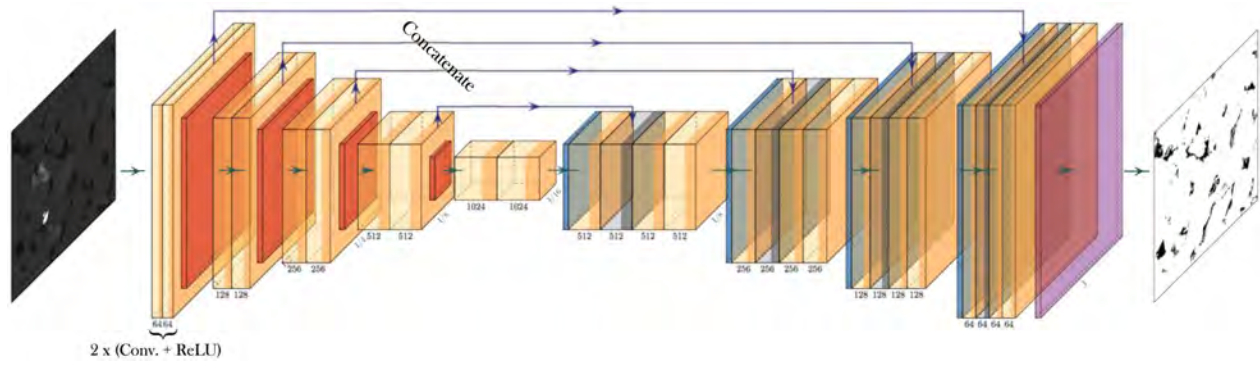


Figure 3—Representation of the U-Net architecture. Double convolution and ReLU followed by max pooling occur for every block on the encoding side. The decoding side starts with upsampling followed by the same double convolution and ReLU for every block.

Loss Function and Optimizer

The model is trained with binary cross-entropy loss as the loss function. The loss function measures the performance of the network and the weights it currently applies to the inputs. The objective of the network is to minimize the loss function (Nielsen, 2015). Assuming y_n is the target output and \hat{y}_n is the predicted output, the binary cross entropy loss (L_{BCE}) measures the performance as follows:

$$L_{BCE} = y_n \log(\hat{y}_n) + (1 - y_n) \log(1 - \hat{y}_n) \quad (1)$$

Root Mean Square Propagation (RMSprop) is the optimization method used for training the model. RMSprop is an adaptive learning rate method that accelerates the loss minimization. RMSprop updates the learning rate by dividing the learning rate of a weight by the moving average of the squared gradients of that weight (Hinton et al., 2012). This is done as follows:

$$E[g^2]_{wt} = \beta E[g^2]_{wt-1} + (1 - \beta) g_{wt}^2 \quad (2)$$

$$w_{t+1} = w_t - \frac{\eta}{\sqrt{E[g^2]_{wt}}} g_{wt} \quad (3)$$

Where w_t is the weight used at time t , g_{wt} is the gradient of weight at time t , $E[\]$ denotes the moving average of the value within, β is a constant between 0 and 1 and is usually assumed to be 0.9, and η is the initial learning rate.

Model Evaluation

Model Segmentation Accuracy. The accuracy of the model is evaluated using two metrics: dice coefficient metric (DCM) and intersection over union (IoU) metric (also known as Jaccard index). Both metrics are used to measure the similarity between two sets of data. Both have a maximum value of one (indicating the two samples are the same) and a minimum value of zero (indicating the two samples are completely different). Consider two possibly intersecting areas A and B. DCM is the ratio of twice the intersecting area of A and B over the sum of areas A and B (Zijdenbos et al., 1994) while IoU is the ratio of the intersecting area of A and B over the union of the areas A and B (Rezatofighi et al., 2019). The metrics are calculated as follows:

$$DCM = \frac{2 \times (A \cap B)}{A + B} \quad (4)$$

$$IoU = \frac{A \cap B}{A \cup B} \quad (5)$$

The two metrics are positively correlated. However, IoU penalizes the mistakes considerably more than DCM considering the inverse correlation between the intersecting area and the union area. Therefore, IoU is biased towards the worst performance while DCM is more of an indication of the average performance.

Measurement of Physical Properties. To evaluate the models' segmentation consistency and the validity of utilizing models trained with 2D to create 3D volumes, it is important to compute the physical properties of volumetric images. Further evaluation is performed on the 3D properties of the models' segmentation. A volume is obtained through segmenting 600 slices of size 600×600 pixels from the middle of the training dataset by each model, then the resulting slices are stacked to create the required volumes. Four properties are measured for each volume: porosity, surface area, permeability in all three directions, and Euler characteristic (EC). The porosity of each volume is the count of voxels defined as pore space over the total voxel count and is more indicative of models' segmentation accuracy rather than consistency. The same point can be made about the surface area. The permeability is measured utilizing the method detailed by Chung et al. (2019).

EC (also known as Euler number and Euler-Poincaré characteristic) is a topological property of an object, it is defined as (Toriwaki and Yonekura, 2002):

$$EC = b_0 - b_1 \quad (6)$$

Where b_0 is the count of individual disconnected bodies, and b_1 is the count of holes (loops) within these disconnected bodies.

In relation to pore scale images, EC is an indication of connectivity of the pores, where negative EC indicates higher connectivity (Vogel, 1997). Measurements are made based on the functions developed by Legland (2020) and using a 6-neighborhood configuration. It has to be noted that EC was measured considering the volumes as a closed system with the edges closed. Therefore, while the measurements may not reflect the true EC value for the sample, they are sufficient for relative comparison between the ground truth and the models' segmentation.

Device Specification and Training Time

All three models are trained on an Nvidia GTX 1080Ti GPU with 11 Gigabytes of GDDR5X memory installed on a PC with an Intel Xeon CPU E5-2623 at 3.00 GHz processor and 128 Gigabytes RAM. One training epoch takes one hour and 36 minutes for Model 17600, 24 minutes for Model 2200, and less than seven minutes for Model 600. Image prediction for any model is performed in less than 1 second for all models. This information is provided for benchmarking purposes.

Results and Discussion

Model Segmentation Accuracy

Figure 4 shows the average loss per epoch for each model during the training run. All three models are run for ten epochs. The loss for all three models converges within five epochs with the subsequent five epochs showing no noticeable major improvement in the loss function. However, a continuous and slight decrease in loss occurs during the last five epochs. In general, and as expected, Model 17600 achieves the best results. The average DCM and IoU scores of the Mt. Simon testing dataset segmented by the model are 0.9967 and 0.9934 respectively (Figure 4). However, considering the amount of data and the time taken to train Model 17600, the model shows modest returns. One training epoch for Model 17600 takes 1 hour and 36 minutes, while it takes 24 minutes for Model 2200. Model 2200 returns an excellent segmentation with DCM and IoU scores that are close to the scores achieved by Model 17600. Figure 5 shows a random slice from the Mt. Simon sandstone image, its ground truth segmentation, and its segmentation by the three models. Both of Model 17600 and Model 2200 managed to match the details of the region highlighted in Figure 5 (among others), with the two segmentations being nearly identical with the ground truth and each other. The trained network can identify the slightest changes in greyscale values. For example, the clay materials are segmented as solid in the ground truth segmentation of our image and the trained networks can easily distinguish clay from pore space edges (Figure 5). Model 600, despite the limited number and

variety of samples, returns a segmentation with comparable accuracy. In any case, the models are expected to perform well against its training data, even when a testing set is set aside before the training runs.

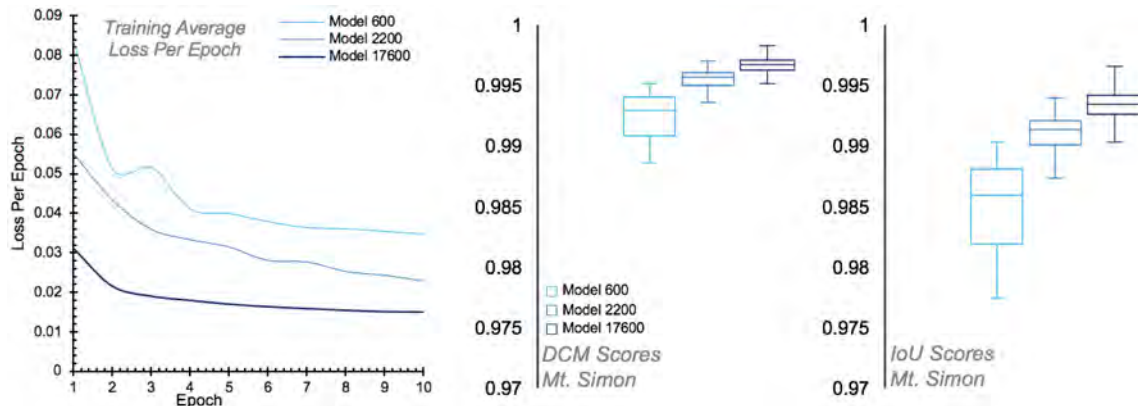


Figure 4—Average loss per epoch for the training data, and DCM and IoU scores for model segmented Mt. Simon test dataset.

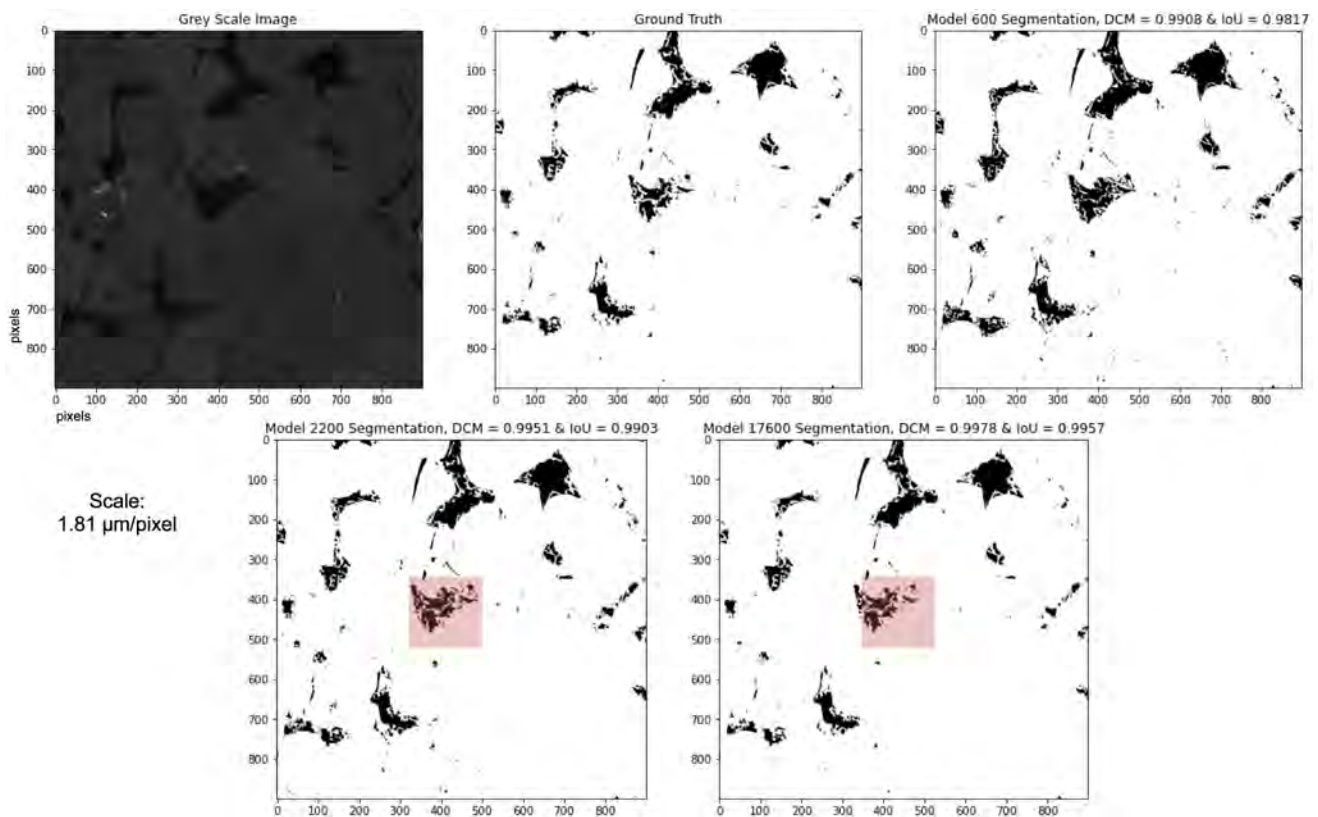


Figure 5—One slice from Mt. Simon sandstone reservoir dataset, its segmentation ground truth, and its segmentation by the three models. All three models return excellent segmentation results, with Model 2200 and Model 17600 results looking visually identical. Highlighted region contains fine details captured by both Model 2200 and Model 17600.

On the other hand, applying the three models on the unseen North Sea sandstone sample returns segmentation with more variable, yet acceptable, accuracies. The North Sea sandstone dataset is of a rock with the same lithology as the training dataset and has similar features. The pixels have varied grey-scale values and higher contrast, which is also similar to the training set. However, the micro-CT images have a lower resolution and blurred edges compared to the training dataset or its equivalent SEM images. Distorted regions due to lack of resolution have some effects on the models' segmentation accuracy as illustrated in

Figure 6. The average DCM scores for Model 17600, Model 2200, and Model 600 are 0.9599, 0.9518, and 0.9479 respectively, and IoU scores for the same models are 0.9229, 0.9079, 0.9010.

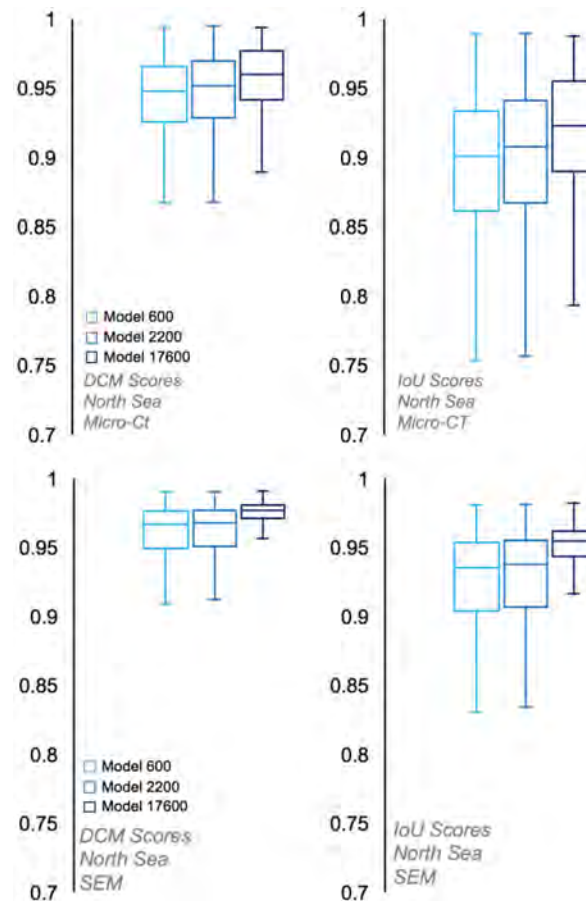


Figure 6—DCM and IoU scores for the models' segmented North Sea micro-CT and SEM images.

The effect becomes clearer when checking the three models' segmentation of a random North Sea micro-CT slice (Figure 7). The models' segmentation, while accurate to a degree, is sensitive to changes in greyscale and blur assigning more solid pixels to microporous regions compared to Otsu Thresholding (as highlighted in red in Figure 7). Also, all three models, and especially Model 17600, are identifying the noise as solid pixels (highlighted in blue in Figure 7). This probably can be mitigated by including diverse images with different qualities and noise levels in the training dataset of the models with their accurate segmentation masks such as the one we use for training our dataset.

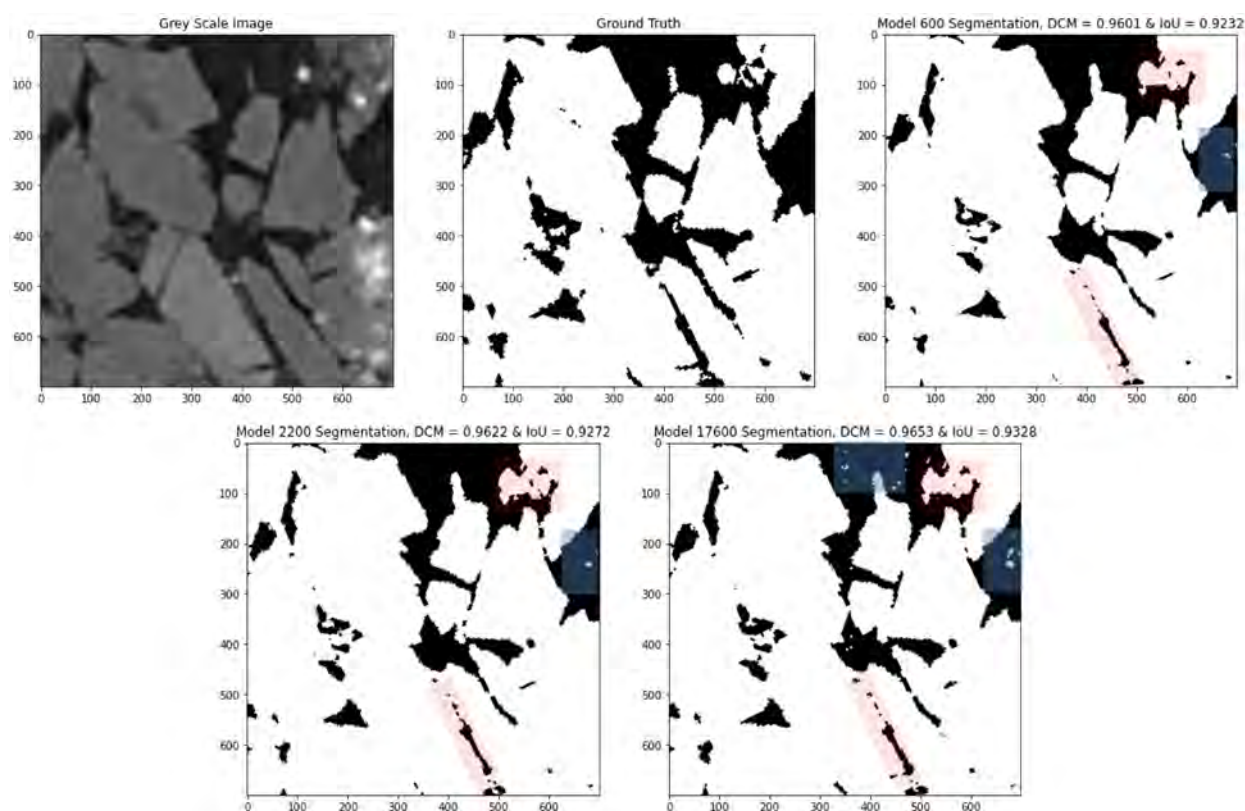


Figure 7—A slice from North Sea sandstone micro-CT images, its segmentation ground truth, and its segmentation by the three models. Highlighted in red are features missed. Highlighted in blue are non-existent features introduced.

Figure 8 shows an example slice of North Sea SEM image with its ground truth segmentation (segmented by Avizo using the watershed threshold method) and the results of the three models. The average DCM scores for Model 17600, Model 2200, and Model 600 are 0.9768, 0.9679, and 0.9667 respectively, and IoU scores for the same models are 0.9546, 0.9378, 0.9356 (Figure 6). Despite the image high resolution, watershed thresholding missed the edges and connectivity. All three models return a detailed segmentation with fine features missed by using watershed method segmentation, several of which are highlighted in Figure 8. This brings a question to the validity of using some of the available data segmentation methods as ground truth, especially when operator bias might have been affecting the segmentation.

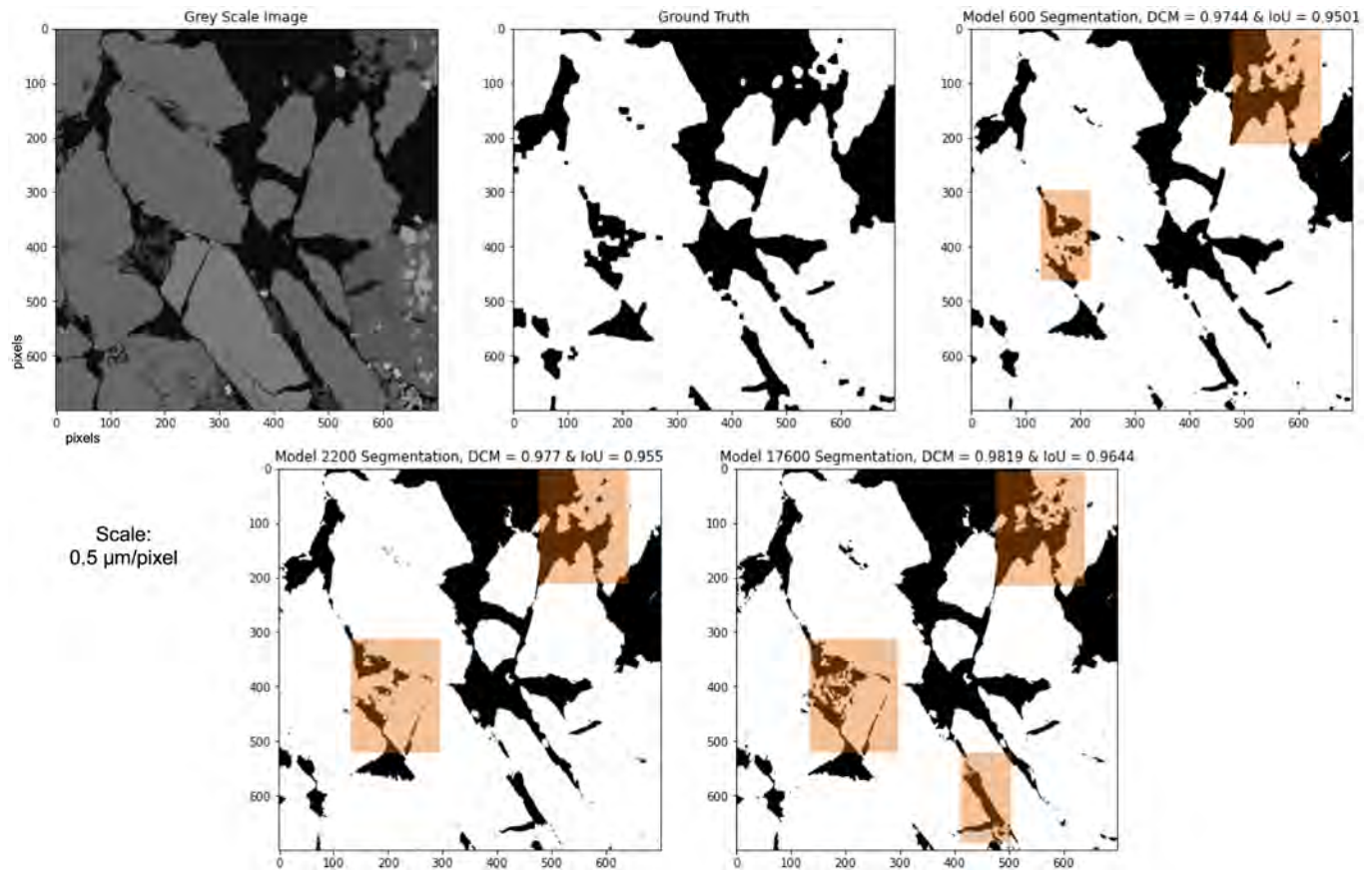


Figure 8—A slice from North Sea sandstone SEM images, its segmentation ground truth, and its segmentation by the three models. Highlighted are identified features not present in the ground truth.

Measurement of Physical Properties

It is vital to assess the accuracy of segmentation based on some of the physical properties of the 3D segmentation masks (stacking 2D slices). Table 1 shows the measured 3D properties of three generated volumes and the ground truth. In general, the models' performance corresponds to the data quantity used for training, with Model 17600 matching the ground truth in most measured properties. The porosity, as mentioned earlier, is a measurement of the models' segmentation accuracy, and the results follow the DCM and IoU scores. Both of Model 17600 and Model 2200 porosities are approximately equal to the ground truth (0.10315, 0.09874, and 0.10306 respectively). Model 600 volume low porosity shows the model's inability to capture the micro-porous areas of the sample. The same rationale is observed for the surface area as the models trained with less data tend to assign more solid pixels to microporous regions. EC shows a positive number for both Model 2200 and Model 600, suggesting a lack of connectivity between pores and the tendency to create many unconnected grains (cavities) in the image. The permeability measurement for both models supports that conclusion as well. Model 2200 volume retains some permeability with 38.3 mD, 31.5 mD, and 48.2 mD in each direction respectively. However, it is still far from the permeability of the ground truth (63.8 mD, 38.5 mD, 163 mD for the same directions). Model 600 volume lost almost all of the permeability of the sample, with a maximum of 10 mD in one direction, and nearly zero permeability for the other two directions. Model 17600 permeability matches the ground truth across all directions, suggesting the connectivity of the images is preserved, but with the existence of some small suspended grains in the pore space.

Table 1—Model segmented volumes' measured properties.

Property	Ground Truth	Model		
		17600	2200	600
Porosity	0.10306	0.10315	0.09874	0.06655
Surface Area, μm^2	6.8942×10^6	6.9540×10^6	6.9969×10^6	7.4125×10^6
Permeability (X-axis), mD	63.8	60	38.3	0
Permeability (Y-axis), mD	38.5	39	31.5	8.40×10^{-4}
Permeability (Z-axis), mD	163	159	48.2	10
EC (Raw)	-3239	-1675	11017	1108
EC (Isolated Pores Removed)	-11945	-12685	-925	-13546
EC (Isolated Pores and Grains Removed)	-12091	-15617	-14503	-20816

EC is measured in 3 stages of processing to identify sources of error associated with isolated pores, isolated grains, and finally the topology of the resulting hydraulically connected, clean pore space. EC measurement shows that overall, the Model 17600 generated volume possesses the best connectivity accuracy. Nevertheless, Model 17600 raw EC is half of the measured EC for the ground truth segmentation (-1675 compared to -3239), indicating an excess in isolated pores. Model 600's volume low connectivity follows its low permeability. On the other hand, Model 2200 shows extremely low connectivity with its EC being ten times higher than Model 600's EC (11017 compared to 1108). Such high EC value for Model 2200 can be either an indication of excessive presence of isolated fake pores (false negatives), the closure of pore loops due to suspended grains, or both. Processing the volumes by removing either or both of the isolated pores and grains then measuring the EC value of the processed volumes would help to determine the reason for the non-conforming EC value. Removing the isolated non-hydraulic pores will remove all fake pores and should mean the EC values would conform with the EC value of the ground truth. If the EC values do not move closer to match the EC value of the ground truth, then the closure of pore loops is the reason for non-conforming EC values and, likewise, the EC values should match. Table 1 shows the EC values of the processed volumes after the removal of isolated pores, then the removal of the isolated grains. The EC values of Model 17600 further show that Model 17600 is generally accurate. Its EC values follow the trend of the EC values of the ground truth and closely match it throughout (-12685 then -15617 for Model 17600, and -11945 then -12091 for the ground truth). For Model 2200, the EC of the model matches the EC of the ground truth after the removal of both isolated pores and grains (-925 then -14503), indicating the presence of both is the reason for the model's initial high EC value and the failure of the model to make accurate predictions around the edges where the clay surfaces are geometrically complex. Model 600 EC values should be ignored despite matching the EC of the ground truth after the removal of the isolated pores (-13546 vs. -11945), especially considering the inaccuracy of the other measured properties of the model's generated volume.

The performance of the models shows potential for the use of such models in the automation of image segmentation, and as a result, bypassing the operator bias. At the moment, the three models presented in this paper have limitations. One such issue that is easily solved is insufficient training. Ten epochs are an unusually short training duration. All three models would benefit from extra training (ideally to 100s of epochs), and the slow decline of the average loss function per epoch in Figure 4 is an indication. One other limitation is limited data available for training, both in quantity and variety. Model 17600 performance compared to the other two models is evidence of the effect of data quantity. It has to be mentioned that the existence of an extensive diverse dataset as is the case in this paper is rare. The variety of the sample provided for training impacts the developed model capability. As stated earlier, the three models were trained with one sample of sandstone with high greyscale intensity variance and high contrast. Therefore, the models

perform well for any image that matches these qualities. The North Sea sandstone features happened to have similar features. Consider the Savonnières limestone in Figure 9a. The best performing model of this paper does not produce a sensible segmentation for the image. However, with adjusted contrast (Figure 9b), the same sample when segmented with Model 17600 produces a rudimentary segmentation. The issue of sample variety can be mitigated when only one set of data is available by augmenting the original data with a modified version of it, whether that be through contrast adjustment, rotation, or mirroring.

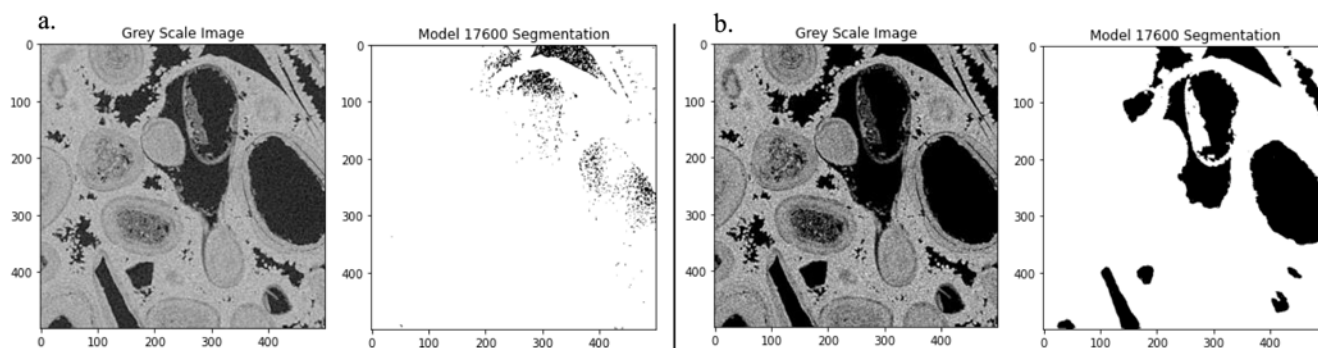


Figure 9—a. Savonnières limestone sample segmented with Model 17600. b. Savonnières limestone sample with adjusted contrast and segmented with Model 17600

Conclusion

Image segmentation is a crucial step in assessing the physical properties of porous media using their X-ray images. The image segmentation involves processing steps and the predefined parameters set by the operator. Therefore, it is susceptible to operator and methodology error and bias. In this paper, three deep learning models are used to overcome these limitations by training with high resolution data comprising of SEM-registered micro-CT images, which ensure accurate segmentation of the ground truth data. U-net CNN architecture is applied to develop surrogate models for pore-scale image segmentation. The high potential of these models in automating routine digital rock processing and analysis and eliminating user bias have been shown. The best performing model achieves high average scores of 0.9599 and 0.9229 in DCM and IoU metric scores respectively when tested against unseen micro-CT sandstone images. However, the performance can be enhanced through data augmentation and increasing the size of the dataset. The development of robust models for production purposes still faces some limitations such as the accuracy of ground truth segmentation masks and the diversity of input images in terms of rock type, imaging quality and noise levels.

References

- ALQAHTANI, N., ALZUBAIDI, F., ARMSTRONG, R. T., SWIETOJANSKI, P. & MOSTAGHIMI, P. 2020. Machine learning for predicting properties of porous media from 2d X-ray images. *Journal of Petroleum Science and Engineering*, 184.
- AYLING, B., ROSE, P., PETTY, S., ZEMACH, E. & DRAKOS, P. QEMSCAN[®] (Quantitative Evaluation of Minerals by Scanning Electron Microscopy): capability and application to fracture characterization in geothermal systems. Thirty-Seventh Workshop on Geothermal Reservoir Engineering, 2012 Stanford University, Stanford, California.
- BAVEYE, P. C., LABA, M., OTTEN, W., BOUCKAERT, L., DELLO STERPAIO, P., GOSWAMI, R. R., GRINEV, D., HOUSTON, A., HU, Y., LIU, J., MOONEY, S., PAJOR, R., SLEUTEL, S., TARQUIS, A., WANG, W., WEI, Q. & SEZGIN, M. 2010. Observer-dependent variability of the thresholding step in the quantitative analysis of soil images and X-ray microtomography data. *Geoderma*, **157**, 51-63.
- BLUNT, M. J., BIJELJIC, B., DONG, H., GHARBI, O., IGLAUER, S., MOSTAGHIMI, P., PALUSZNY, A. & PENTLAND, C. 2013. Pore-scale imaging and modelling. *Advances in Water Resources*, **51**, 197-216.
- CAI, Z., FAN, Q., FERIS, R. S. & VASCONCELOS, N. 2016. A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection. European conference on computer vision.

- CHUNG, T., WANG, Y. D., ARMSTRONG, R. T. & MOSTAGHIMI, P. 2019. Approximating Permeability of Microcomputed-Tomography Images Using Elliptic Flow Equations. *SPE Journal*, 24.
- FRUCCI, M. & DI BAJA, G. S. 2008. From Segmentation to Binarization of Gray-Level Images. *Journal of Pattern Recognition Research*, 3, 1-13.
- HINTON, G., SRIVASTAVA, N. & SWERSKY, K. 2012. Neural Networks for Machine Learning Lecture 6a Overview of Mini-Batch Gradient Descent. Cited
- IASSONOV, P., GEBRENEGUS, T. & TULLER, M. 2009. Segmentation of X-ray computed tomography images of porous materials: A crucial step for characterization and quantitative analysis of pore structures. *Water Resources Research*, 45.
- ISOLA, P., ZHU, J.-Y., ZHOU, T. & EFROS, A. A. 2018. Image-to-Image Translation with Conditional Adversarial Networks.
- KARIMPOULI, S. & TAHMASEBI, P. 2019. Segmentation of digital rock images using deep convolutional autoencoder networks. *Computers & Geosciences*, 126, 142-150.
- KOHLER, R. 1981. A Segmentation System Based on Thresholding. *Computer Graphics and Image Processing*, 15.
- LEGLAND, D. 2020. Geometric measures in 2D/3D images. MATLAB Central File Exchange.
- LONG, J., SHELHAMER, E. & DARRELL, T. Fully Convolutional Networks for Semantic Segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition 2015.
- NIELSEN, M. 2015. Neural Networks and Deep Learning. San Francisco, CA, USA: Determination press.
- 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 Resources Research*, 56.
- OTSU, N. 1979. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems*, 9, 62-66.
- PEREIRA, S., PINTO, A., ALVES, V. & SILVA, C. A. 2016. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE Trans Med Imaging*, 35, 1240-1251.
- REZATOFIGHI, H., TSOI, N., GWAK, J. Y., SADEGHIAN, A., REID, I. & SAVARESE, S. 2019. Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression.
- RONNEBERGER, O., FISCHER, P. & BROX, T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Available: <https://arxiv.org/abs/1505.04597>.
- SAINATH, T. N., MOHAMED, A.-R., KINGSBURY, B. & RAMABHADRAN, B. 2013. Deep Convolutional Neural Networks for LVCSR. IEEE international conference on acoustics, speech and signal processing.
- SRIDEVI, M. & MALA, C. 2012. A Survey on Monochrome Image Segmentation Methods. *Procedia Technology*, 6, 548-555.
- TORIWAKI, J. & YONEKURA, T. 2002. Euler Number and Connectivity Indexes of a Three Dimensional Digital Picture. *Forma*, 17, 183-209.
- TULLER, M., KULKARNI, R., FINK, W., ANDERSON, S. H. & HOPMANS, J. W. 2013. Segmentation of X-Ray CT Data of Porous Materials: A Review of Global and Locally Adaptive Algorithms. Soil-Water-Root Processes: Advances in Tomography and Imaging.
- VOGEL, H. J. 1997. Morphological Determination of Pore Connectivity as A Function of Pore Size Using Serial Sections. *European Journal of Soil Science*, 48, 365-377.
- WANG, Y. 2018. Reconstruction of High-Resolution Pore Structures from Micro-Tomography Images. Doctor of Philosophy, The University of New South Wales.
- WANG, Y., TENG, Q., HE, X., FENG, J. & ZHANG, T. 2019a. CT-image of rock samples super resolution using 3D convolutional neural network. *Computers & Geosciences*, 133.
- WANG, Y. D., ARMSTRONG, R. T. & MOSTAGHIMI, P. 2019b. Enhancing Resolution of Digital Rock Images with Super Resolution Convolutional Neural Networks. *Journal of Petroleum Science and Engineering*, 182.
- WANG, Y. D., SHABANINEJAD, M., ARMSTRONG, R. T. & MOSTAGHIMI, P. 2020. Physical Accuracy of Deep Neural Networks For 2d And 3d Multi-Mineral Segmentation Of Rock Micro-CT Images.
- ZIJDENBOS, A. P., DAWANT, B. M., MARGOLIN, R. A. & PALMER, A. C. 1994. Morphometric Analysis of White Matter Lesions in MR Images: Method and Validation. *IEEE Transactions on Medical Imaging*, 13.