Adv Syst Sci Appl 2022; **04**:31–50

The Influence of Image Morphology on Neural Network-Based Segmentation Results

Efim V. Lavrukhin[1,2], Kirill M. Gerke [2]

[1] *Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow, Russia*
[2] *Schmidt Institute of Physics of the Earth of Russian Academy of Sciences, Moscow, Russia*

Abstract: In this article, we have provided initial results of evaluating the influence of training data morphology on convolutional neural networks segmentation quality. To do this, we selected four different porous media 3D images obtained with the help of X-ray microtomography (XCT) and segmented them using a local thresholding ”converging active contours” algorithm. With the help of synthetic tomography algorithm, we created synthetic XCT image based on these segmentations. As a result, we obtained true XCT – ground-truth pairs for considered porous media samples. Then with the help of stochastic reconstructions, we created a set of samples with varying degrees of morphological similarity to the original porous media samples. In order to check the effect of morphological properties on the quality of segmentation, we trained a U-Net model with ResNet50 encoder on pairs of synthetic XCT – ground-truth data, and assessed the segmentation quality of synthetic XCT images as obtained based on stochastically reconstructed images. Based on the metrics, we concluded that the quality of segmentation is more influenced by the morphological differences of the original porous media samples than by difference from the generated stochastic reconstructions. We discussed possible ways to improve the future experiments design in order to finally resolve the issue of training for XCT image segmentation with neural networks.

Keywords: porous medium, image processing, convolutional neural network, image segmentation, synthetic tomography, correlation functions

1. INTRODUCTION

It is not a secret that porous media structure, i.e., spatial distribution of forming materials (or phases), defines its major physical properties [2, 11, 58, 64]. In case structure information is known up to a relevant spatial resolution, a great number of such properties, including flow and transport characteristics, can be effectively obtained using so-called pore-scale modeling techniques that include direct voxel based [29,43,55], meshing-based [10] or computationally efficient pore-network modeling approaches [22,50,54]. Computed or/and upscaled from the first principles, flow properties are of primary importance to parameterize Darcian scale and to explain or design flow mechanics relevant for petroleum engineering [12, 30], hydrology [6, 27], food engineering [21] and numerous other applications. With recent progress in 3D imaging [19,68], stochastic reconstructions based on limited input data [3,17,35,39,63,71], or their combination [47,48], obtaining detailed structural information for a given porous media sample is not an issue. Moreover, recently proposed multiscale image fusion techniques [24] provide a solution for complex/hierarchical porous media samples such as complex rocks [19,25] and soils [40]. Regardless of the methodology employed, some input experimentally

---

[*] Corresponding author: e.lavrukhin@digital-core.ru

obtained structural information is necessary, and is mainly obtained by scanning electron microscopy (SEM), focused ion beam-SEM imaging, or X-ray computed tomography (XCT) [25]. This information comes in the form of grayscale images where pixels/voxels have the intensity proportional to some local physical property, e.g., for XCT it depends on local X-ray absorption [68]. Yet, as we previously defined structure as spatial distribution of constituting phases, the crucial step necessary to perform pore-scale simulations is the transition from local attenuation coefficients to local material(s) distribution. In the simplest case that is needed, for instance, to solve Stokes equation to simulate single phase flow and compute permeability [29, 43, 55], we need to distinguish two porous media components — pores and solids. Such a transition from local physical properties to phases (or materials) is called image segmentation, and, in case of two phases (e.g., pores and solids), segmentation is also referred to as binarization.

The biggest problem of any segmentation procedure is that it is never strictly rigorous and always contain some degree of uncertainty. Apart from numerous experimental and numerical problems associated with image acquisition, the main reason for this issue to appear is due to the limited imaging resolution – partial volume effects [19, 20, 28, 68], i.e., precise segmentation would be possible only for images with nearly infinite resolution. In other words, each pixel/voxel usually contain a mixture of different phases, but segmentation requires it to be affiliated with a relevant phase, ideally with the material which dominates the volume under study. Utilization of grayscale values to evaluate local abundance of a given material within each pixel/voxel is possible only for mono-component (consisting of a single material, which is rarely the case in relevant applications) or requires complex dual energy scanning with subsequent inverse modelling [70]. The primary target of the segmentation procedure in this context is to provide as accurate as possible approximation of the spatial distribution of phases under given image resolution conditions.

Numerous segmentation techniques are available for this task and can be roughly divided into four categories:

1. manual segmentation;
2. global thresholding;
3. local thresholding methods;
4. other approaches including machine learning and unsupervised learning algorithms.

Manual approach implies the choice of a single threshold grayscale value between two phases to be segmented by an operator and still a popular way to go.

Global segmentation approach is basically the same as manual one, but the threshold is chosen with the help of some automatic computational technique based on grayscale histogram (e.g., by minimizing intra-class variance in the method by Otsu [53]).

Local thresholding methods identify two thresholds for each two phases to be segmented out. These thresholds represent trustworthy boundaries which define, for example, 100% certainty that all voxels below the lower threshold are pores, while all voxels above the upper threshold are pores. The pixels/voxels in between these two bounds are treated and classified according to some algorithms: statistically kriged in the popular indicator kriging method [52], grown from seeds (100% certain areas) using region growth [31] or converging active contours [61]. These thresholds are chosen either manually or automatically [60]. Here we mentioned only some popular methods, more detailed list of techniques applied in porous media research can be found in the review of Schlüter et al. [59].

Machine learning methods fall into supervised and unsupervised bins, in the former case some input training information is provided to the algorithm, whilst in the latter case the algorithm makes all decisions by itself.

Despite the abundance of segmentation techniques and some auxiliary image processing methods such as filtering, none of the existing methodologies can pretend to provide universal and adequately accurate results. It is widely acknowledged that manual and global thresholding techniques are inferior to the local thresholding methods [34], which, in turn, suffer from necessity to choose confidence bound threshold values. Automatic choice of thresholds [59, 60] does not necessarily provide adequate results, as was observed in numerous studies [40].

An interesting study by Baveye et al. [7] concluded that human operator performs much better than any automatic thresholding algorithm. Whether it is true or not is still an open question, as the majority of segmentation techniques tested by Baveye et al. [7] are not necessarily suitable for porous media applications. The results of the unsupervised segmentation techniques are not robust in the sense of the number of phases [33] and accuracy [15, 16, 42].

Supervised segmentation methodologies are very limited in XCT porous media applications [38, 46, 65], but their recent usage in numerous other fields such as medical imaging [9], neuroscience [18], and satellite image segmentation for urban structures [4] and flood identification [51] motivated this study. Moreover, testing and verification of all existing methodologies in the area of porous media research is seriously hampered by the absence of the true data — precise segmentation results for different samples. Artificial (synthetic) XCT images produced by rescaling, noise and artifact addition (e.g., Wang et al. [67] or Schlüter et al. [59]) do not possess all necessary complexity and represent real XCT only to some extent.

At this point we can pose a relevant question — what the ideal segmentation methodology would look like? Any expert working with 3D XCT images or porous media would generally agree with the following definition: with the help of the ideal segmentation technique one should be able to take any XCT image and obtain accurate segmentation by using only a stack of images as input data, i.e., without any operator input or (fitting) parameters. From our above-mentioned analysis of current segmentation technology, we can confidently conclude that none of existing methods provide a solid performance according to that definition. Supervised machine learning algorithms seem to provide a basis for an ideal technique, but require ground-truth (precise segmentation results) data to train neural networks.

We believe that the holy grail of the ideal training data lies with synthetic XCT datasets. Such datasets can be created from any structure consisting of any number of materials or phases [8, 14, 69].

Now, we come to the problem of obtaining these original true structures or spatial phase distributions, especially considering that real porous media are infinite resolution structures, at least up to molecular scale. We argue that the original should be no more than an order of magnitude finer (better resolution) than the optimal resolution for the images we shall segment. For example, typical XCT images for sandstones has resolution of 2.5-5 μm and in some instances it is enough to see all details relevant for structure characterization and pore-scale modelling. For carbonates and soils this may be different due to a wide range of pore sizes, some of them well below 1 μm which is an approximate limit for resolution with current desktop scanners such as Bruker-SkyScan-1172. In short, if the images we want to segment will have the resolution up to 1 μm, the true phase distributions of 100 nm will suffice to create accurate projections assuming the receiver matrix (CCD camera in real scanners) to end up with 1 μm pixel size. Such ”magnified” images can be obtained easily using a number of techniques: discretization of a predefined grain packings for sandstones (e.g., [45]), stochastic constructions of a synthetic porous media [36, 41], multiscale stochastic reconstruction/image fusion [24, 62], or super-resolution stochastic technique [49]. The proxies for magnification can be obtained from real porous media images by segmenting them using current state-of-the-art techniques (online reference) enhanced by mineralogical composition overlay from 2D analysis [13, 32].

In the above analysis we arrive to a universal framework to create an ideal segmentation technology based on convolutional neural networks (CNN) or possibly other (better/emerging) type of machine learning methodology. The general scheme of this framework is shown in Fig. 1.1.

But in this case, some data on the spatial materials’ distribution in the porous media is required. Then the question arises about the size and diversity of the sample that will be

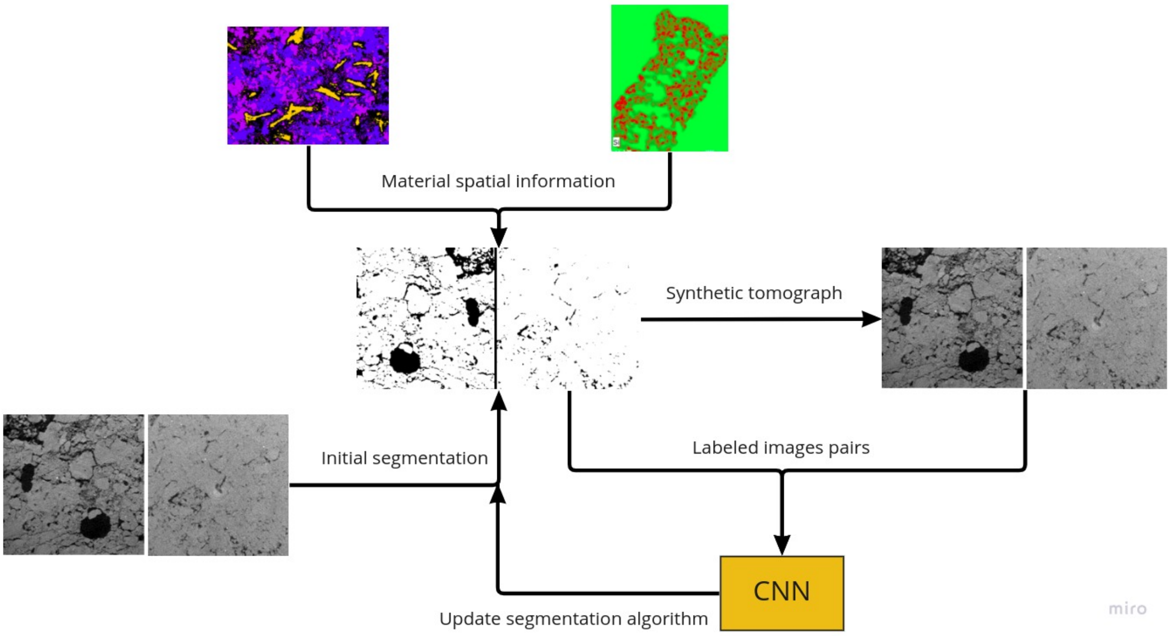


Fig. 1.1. Universal workflow to train machine learning based segmentation models.

used in training: is it necessary to use patterns of spatial distribution from different types of porous media, or how accurately is it necessary to reproduce the original (and unknown) morphological structure?

The objective of this paper is to understand how accurate the morphological structure of the samples used to construct a synthetic XCT should be – or in other words, how similar to the images we want to segment. We also want to elucidate how morphological structure errors of various sizes will affect the quality of the trained segmentation model. For this, we construct stochastic samples with different levels of likelihood relative to the original structures. For all these samples we create synthetic XCT images and train segmentation CNN model on original sample-synthetic XCT pairs. The segmentation quality of these samples analyzed depending on their likelihood, as well as on the morphological properties of the original structure.

2. METHODOLOGY

2.1. Convolutional Neural Networks (CNNs) for image segmentation

CNN for segmentation are tuned to segment entire input image into sub-regions of particular origin, e.g., in landscape images CNN can distinguish between grass, woods, sky, etc. Therefore, the task is to identify which object a given pixel of the image belongs to. Thus, in segmentation CNNs classifiers are applied to each pixel independently.

Furthermore, to reduce noise, segmentation CNNs use two procedures:

1. down-sampling (encoder), when original image is compressed in size;
2. up-sampling (decoder), when compressed image is restored to the size of the original image.

Then each pixel of the up-sampled image is a subject for it own classifier, which is independent from the classifiers of other pixels.

A general structure of the segmentation CNN‘s is presented in Fig. (2.2)

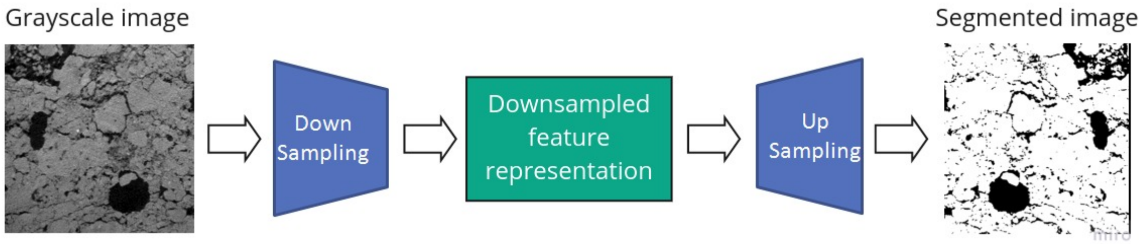


Fig. 2.2. General structure of the encoder-decoder CNNs architecture for segmentation.

The most prominent networks in image segmentations CNNs are U-net [56] and SegNet [5].

In this paper, we are solving a task of dissociating between porous and non-porous space within XCT images. Furthermore, there can be only either porous or non-porous space in the image, i.e., no background pixels. Therefore, we need to assign each pixel to be either of porous or non-porous type. Segmentation CNNs are natural solutions to this 2-class segmentation problem.

In this study, we used a specific version of the U-net architecture: a U-net with a Res-Net 50 [37] encoder that was pretrained on the ImageNet dataset [44]. We expect this combination provides a balance between the speed of the architecture and the quality of the segmentation result.

2.2. Porous media acquisition and specs

We gathered a library of XCT images containing 4 natural porous media samples of different genesis.

All 3D XCT images were obtained using SkyScan-Bruker 1172 desktop scanner with different resolutions, which depended on the size of the cylindrical samples subdrilled (in case of rocks) from core material.

All resolutions were carefully chosen to capture major pore sizes. Some amount of porosity could be below resolution and, thus, not represented on the images.

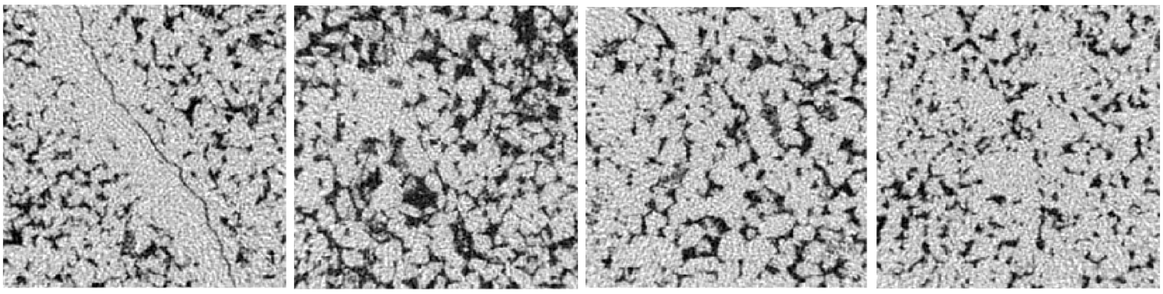
All these samples were chosen to be very different in pore structure. Moreover, one of the samples contains cracks and fissures, that provide additional complexity to the segmentation procedure.

After acquisition, we have cropped regions of interest with volumes of 1000^3 voxels from each full size scan. Subcropping is useful is useful for a number of reasons:

1. To remove all border effects due to cylindrical sample shape;
2. To make subvolume cubic, such shape is convenient for further processing, as it basically represents a 3D matrix of the same dimension;
3. To reduce input size, the size of 1000^3 voxels provides a good compromise between image representability and computational efforts needed for its processing.

Next, the original porous media sample was segmented into two phases using the conventional local thresholding approach – converging active contours method [61]. Obtained morphological binary structures were used as a basis for constructing synthetic XCT image. The pore space of these structures was filled with air, the solid phase was considered to be made of kaolinite ($Al_2Si_2O_5OH_4$).

While creating synthetic XCT (see details in the next subsection), we compressed the resulting images (as well as the corresponding markup images) in 5 times to 200^3 in order to obtain and study the partial volume effect – this made segmentation task on synthetic images closer to real XCT conditions.



Sample 1 Sample 2 Sample 3 Sample 4

Fig. 2.3. Examples of 2D slices through 3D synthetic XCT images.

The example 2D slices from all synthetic XCT images are shown in Fig. 2.3. This figure also identifies sample references used throughout this study: Sample 1, Sample 2, Sample 3, and Sample 4.

2.3. *Synthetic XCT algorithm*

The synthetic tomography algorithm consists of three stages:

1. Construction a sample model based on physical and morphological information;
2. Calculation of shadow projections;
3. Reconstruction of XCT image.

The first step is required to calculate absorption for each phase independently (2.1) using the numerical direct Radon transform:

$$A_m^l = \sigma_m \int_l \rho_m(l) dl, \tag{2.1}$$

where:

m - phase number;
l - direction to calculate absorption;
σ_m - attenuation coefficient of phase m material;
ρ_m - linear density of phase m material.

The total absorption (2.2) is calculated as the sum of the absorption of all materials (2.1):

$$A^l = \sum_m A_m^l. \tag{2.2}$$

Directional transmittance (2.3) is calculated using total absorption (2.1) according to:

$$T^l = e^{A^l}. \tag{2.3}$$

Further, to complete second step, the shadow projection for each detector pixel and each rotation angle of the sample (which are set by the direction l) can be calculated as the number of photons that reached the detector through the sample. The number of photons (2.4) is a random variable with a Poisson distribution, and with an average proportional to the number of photons N_0 released by source, and the transmittance in a given direction:

$$N^l \sim Pois(N_0 T^l). \tag{2.4}$$

Next step is to mathematically reconstruct the image from projections. This can be done using the numerical inverse Radon transform. Synthetic tomography was implemented

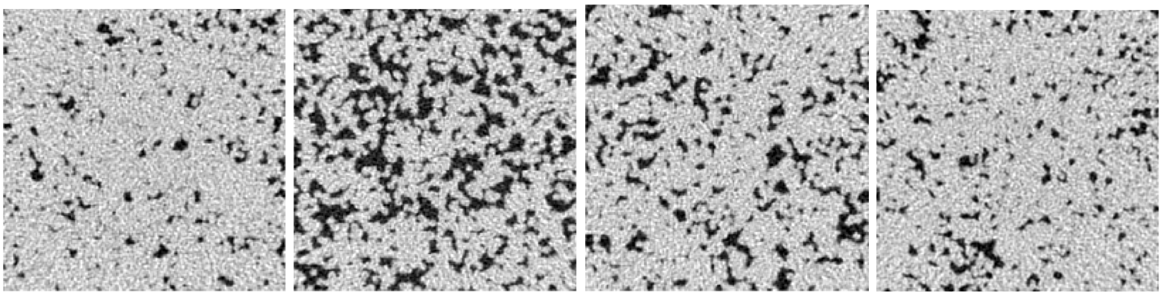

Sample 1 Sample 2 Sample 3 Sample 4

Fig. 2.4. Example 2D slices from obtained 3D stochastic reconstructions.

in python with the help of Astra toolbox [1], that includes forward and backward Radon transform.

At this point we have the ground-truth original phase distribution and its synthetic XCT 3D grayscale image. Moreover, during projections simulation and reconstruction procedures we can apply any software/hardware setups and parameters, including noise addition and specified scanning resolution.

2.4. Stochastic reconstruction algorithm

Stochastic reconstruction is an algorithm that iteratively performs permutations of voxels within an image, trying to match its statistical properties to the original (or training) image.

As input data, the stochastic reconstruction algorithm receives the calculated values of a set of two-point correlation functions (here: S_2, L_2, F_{ss}) in orthogonal and diagonal directions for a range of correlation lengths r. Each correlation function measures a probability of some event for a given correlation length r. Specifically, S_2 measures the probability that both ends of the r line segment lie within the pore phase, L_2 – the probability that the whole segment lies within the pore space, and F_{ss} – that both ends fall into a pore-solid interface. Comprehensive information about correlation functions can be found in [64].

The reconstruction algorithm starts with mono-disperse sphere packing with similar porosity and close S_2 correlation function. The porosity is known from $S_2(0)$ or $L_2(0)$ value. With each iteration, the algorithm swaps two voxels of an image to minimizes a special energy function (2.5), which is calculated from the difference between the current values of the image's correlation functions f and the specified correlation functions values $\hat{f}$ summed over all correlation lengths r:

$$E(I) = \sum_f \sum_r \big(f(r) - \hat{f}(r)\big)^2 \longrightarrow \min\,. \tag{2.5}$$

The energy function is minimized using the simulation annealing algorithm. The algorithm stops after minimizing the error close to zero, or after a large number of consecutive unsuccessful attempts. During annealing we saved the structure dynamics after some particular number of permutations attempts – after 1, 4, 16 and 64 $\times 10^4$ iterations. A more detailed description of this algorithm can be found in [23, 26]. Stochastic replicas were also used as input for synthetic XCT 3D images generation, see examples on Fig.2.4.

2.5. Image preparation for CNN

Segmentation CNN has a supervised learning architecture. This means that to train the network we need a pair of samples, i.e., the input examples and desired segmentation for each such image. Here we describe the way the images to be fed into the network have been prepared for the segmentation experiments.

We refer to each 3D volume of synthetic XCT as a $stack$. There are 200 images with resolution 200×200 pixels within each such stack. We consider each image of resolution 200×200 taken in parallel to the same plane as a single independent image, i.e., no 3D spatial information about image location in the stack is provided to CNN.

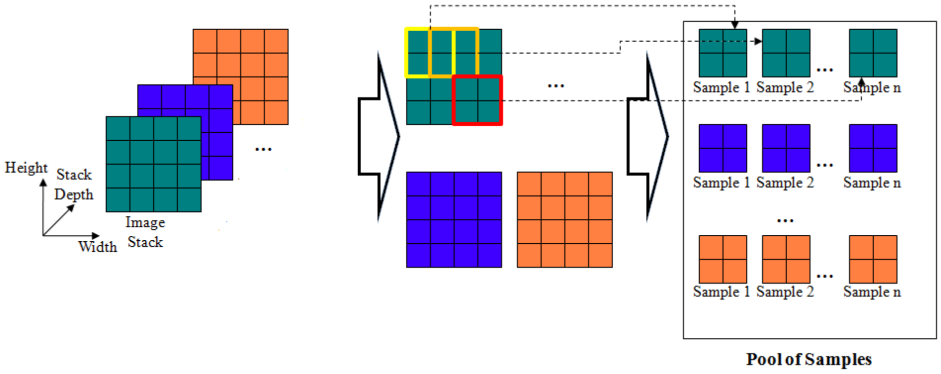


Fig. 2.5. Preparation of the network inputs: convert stack image into image samples.

Next, we collect a pool of $n = 32$ samples of 128×128 pixels from each image with minimal overlapping. This procedure is explained in detail on Fig. 2.5. The pixel intensity within each sample is standardized by scaling with mean and standard deviation from ImageNet (due to the use of an ImageNet-pretrained encoder). This pool of samples constitutes the input samples to be fed into the segmentation CNN.

2.6. Implementation

The stochastic reconstructions algorithm is implemented on the Julia programming language.

The synthetic tomography algorithm is implemented on the python programming language using the Astra framework for calculating linear projections and the Numpy and PyTorch libraries for matrix calculations parallelization.

All neural networks models were trained on the Python programming language using the PyTorch library.

We used a computer with Ubuntu 18.04 operation system, 128GB RAM and a NVIDIA Geforce rtx 2090 graphics card as a working station for computational experiments.

3. EXPERIMENTS

3.1. Model description

For the purpose of this study, we have selected the U-net segmentation CNN architecture [56], as shown in Fig.3.6. The U-net is a fully-convolutional encoder-decoder architecture with skip-connections, which are feedforward connections from encoder to decoder layers. U-net uses convolution and max-pooling layers for compression (down-sampling) in the encoder part. Transposed convolutional layers and concatenation with skip-connected encoder feature map for decompression (up-sampling) in the decoder part. Convolutional layers use only with 3x3 filters.

In this study we used a U-net architecture with a ImageNet-pretrained Res-Net 50 encoder.

3.2. Training process

As input to the network we use 128×128 from the pool of samples with corresponding binarized 128×128 desired output samples. These input samples are arranged into batches

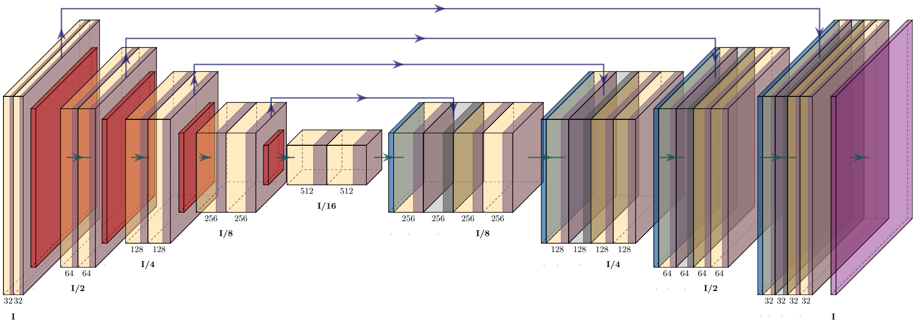


Fig. 3.6. U-net's architecture.

of size 32 for model training. Therefore, the output of the model constitutes tensor with pore and non-pore probabilities within each pixel.

We use combination of cross-entropy (3.6) and smoothed Intersection Jaccard Index (3.7) as loss function:

$$CE(x,y) = -\Big(y\log p(x) + (1-y)\log\big(1-p(x)\big)\Big), \tag{3.6}$$

$$Jaccard(x,y) = -\log\Big(\frac{p(x)y+\varepsilon}{p(x)+y-p(x)y+\varepsilon}\Big), \tag{3.7}$$

$$L(I,M) = \frac{1}{N}\sum_{i=1}^{N}\sum_{x,y\in I_i,M_i}\frac{1}{WH}\Big(CE(x,y)+\alpha Jaccard(x,y)\Big), \tag{3.8}$$

where:

N - total number of images;
I_i - i-th image to segment;
M_i - i-th true segmentation mask;
W - width of the output image;
H - height of the output image;
x, y - image/mask pixels with corresponding spatial positions;
$p(x)$ - neural network prediction for pixel x;
ε = 10^{-6};
α = 1.

We used Adam optimizer [57] with following parameters:

$lr = 10^{-2}$ - initial learning rate;
$wd = 10^{-2}$ - weight decay;
$\beta_1 = 0.95$ - coefficient used for computing running averages of gradient;
$\beta_2 = 0.999$ - coefficient used for computing running averages of gradient square;
$lr_d = 0.5$ - multiplicative learning rate decay on validation loss plateau;
$lr_{min} = 10^{-6}$ - minimal learning rate.

3.3. Experimental setup

After the preparation of binarized and synthetic XCT images as was described in the previous sections, we split each stack into 4–1–1 proportion – training, validation and testing datasets. We trained the model on the training samples (Fig. 3.7) and tuned the hyperparameters to

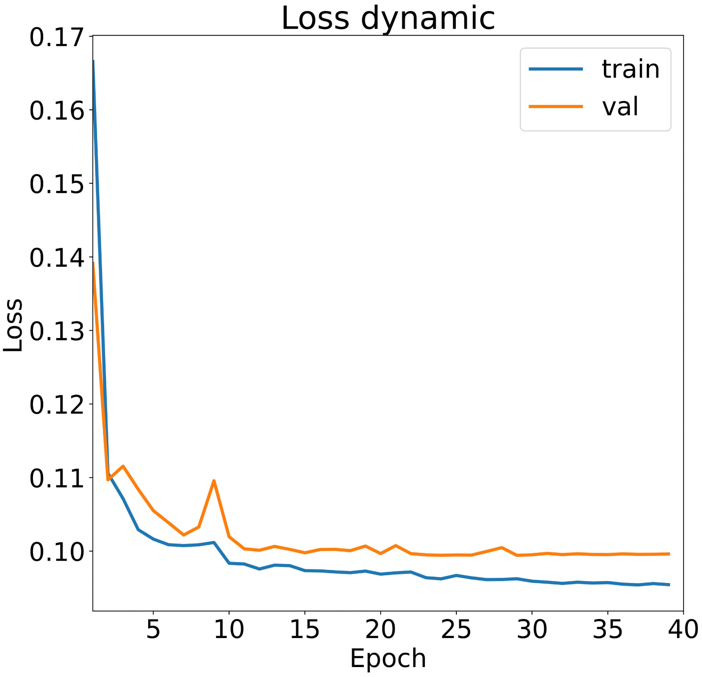


Fig. 3.7. U-Net training process.

minimize the loss function on the validation samples. As a result, we obtained the model trained on samples with real porous media morphological properties.

We used this model for segmentation of the images from the test dataset, as well as for the segmentation of four sequences of synthetic XCT images, that were created from stochastic reconstructions (Fig. 3.8).

Finally, we measured the segmentation quality metric by the initial morphological properties and by the number of annealing iterations (iterations numbers 1, 4, 16, 64) against the morphological properties original porous media ground-truth images.

3.4. Segmentation accuracy metrics

To measure model performance in our experiments we employed intersection over union, or IoU metric – **the Jaccard index**. It describes the ratio between the size of the intersection of true and predicted labels sets, and the size of the union of these sets. This metric is widely used to assess quality in segmentation studies. We calculate it for the positive class (i.e., pore phase) for greater sensitivity. In order to get the class labels from the probabilities that are returned by the model, we used the probability threshold of 0.5.

The IoU metric was measured for each synthetic XCT – ground-truth pair.

4. RESULTS

We obtained an average $IoU = 0.81$ for the pore phase, which is a relatively good result. It should also be noted that measurements are carried out against true ground-truth labels (which were used to produce synthetic XCT images). Most segmentation errors are located around the phase separation boundaries and can be explained by the influence of partial volume effect.

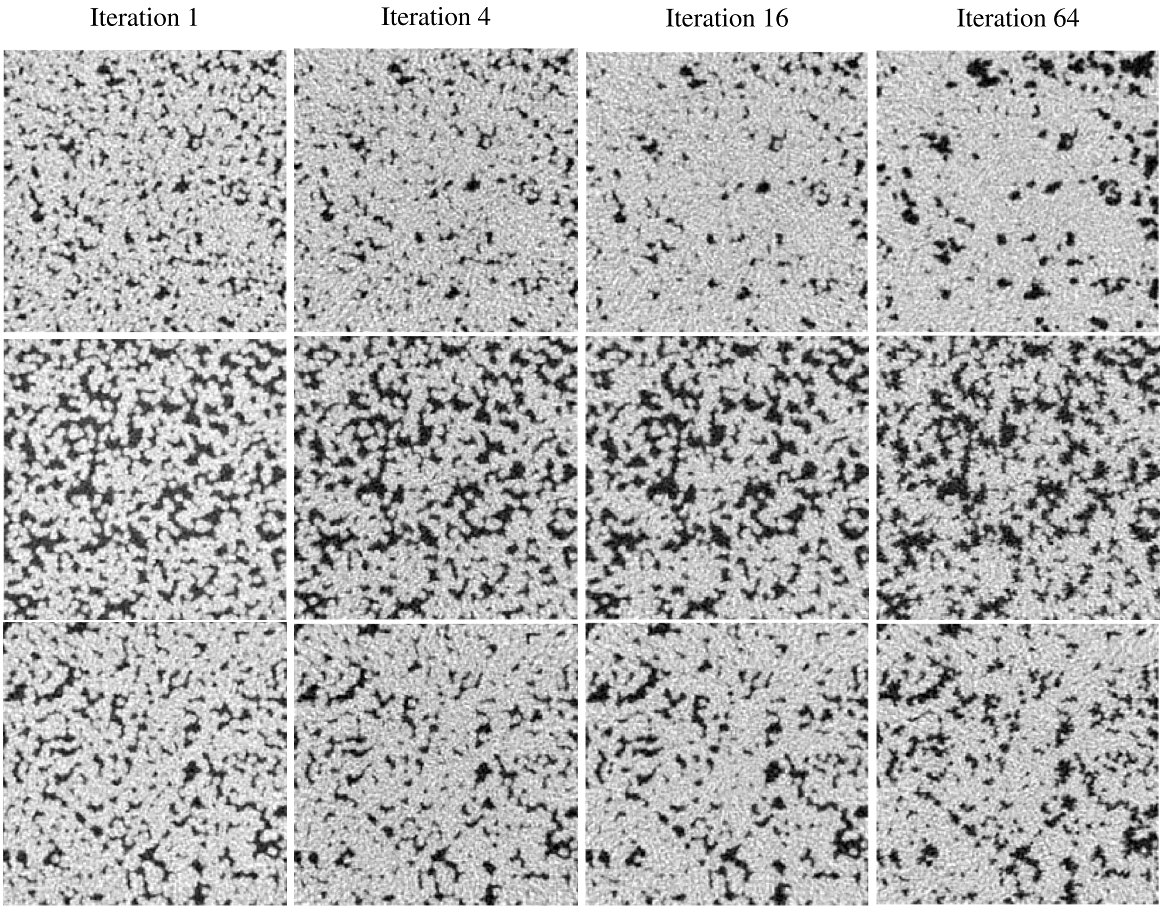


Fig. 3.8. Examples of 2D slices from 3D sequences of samples produced by stochastic reconstruction algorithm.

4.1. Visual segmentation results

In this section we show visual results of the CNN-based segmentation (Fig. 4.9, 4.10, 4.11, 4.12).

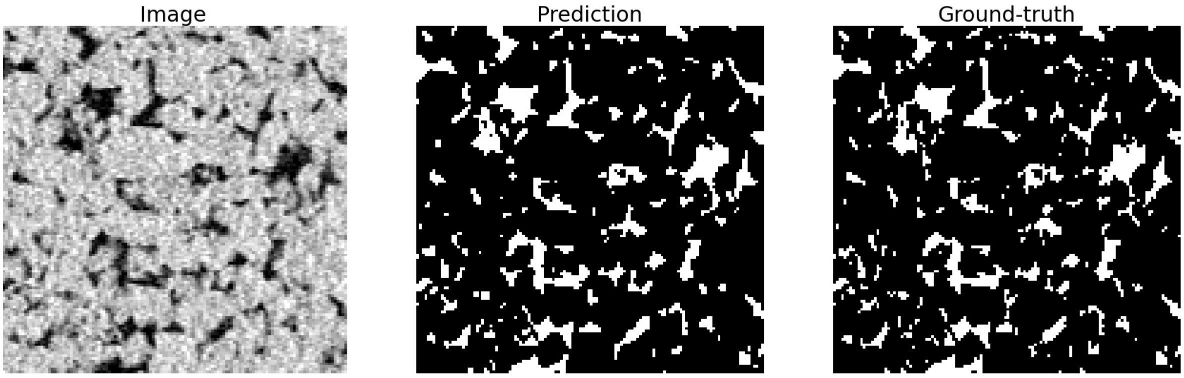


Fig. 4.9. Examples of 2D slices through 3D segmented images for Sample 1.

Also we can see (Fig. 4.13) the visualization of false-positive (shown in red) and false-negative (shown in blue) segmentation errors.

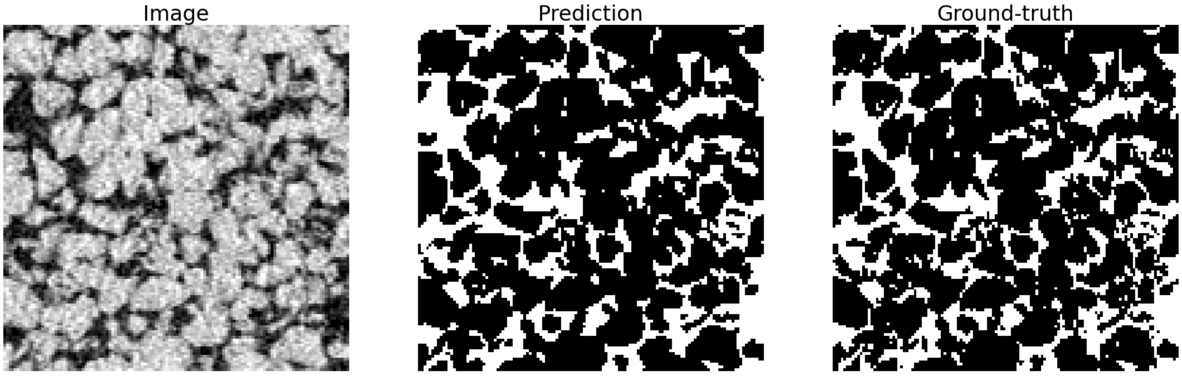


Fig. 4.10. Examples of 2D slices through 3D segmented images for Sample 2.

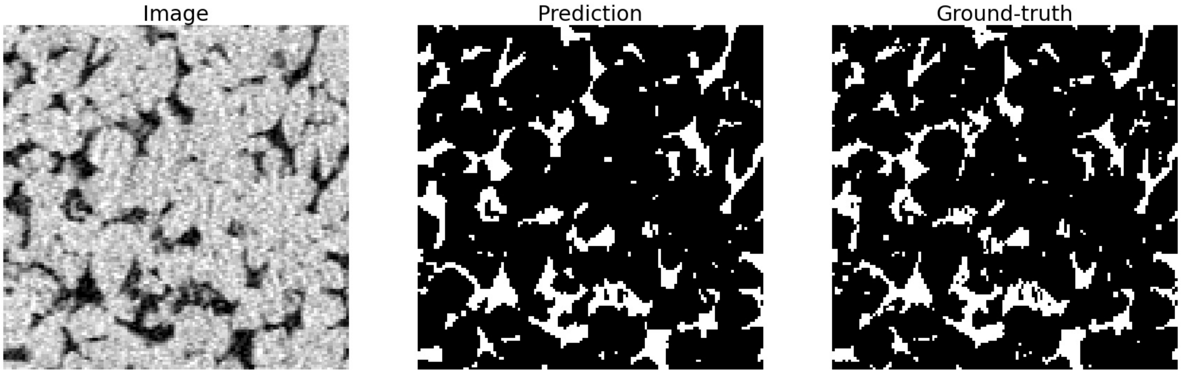


Fig. 4.11. Examples of 2D slices through 3D segmented images for Sample 3.

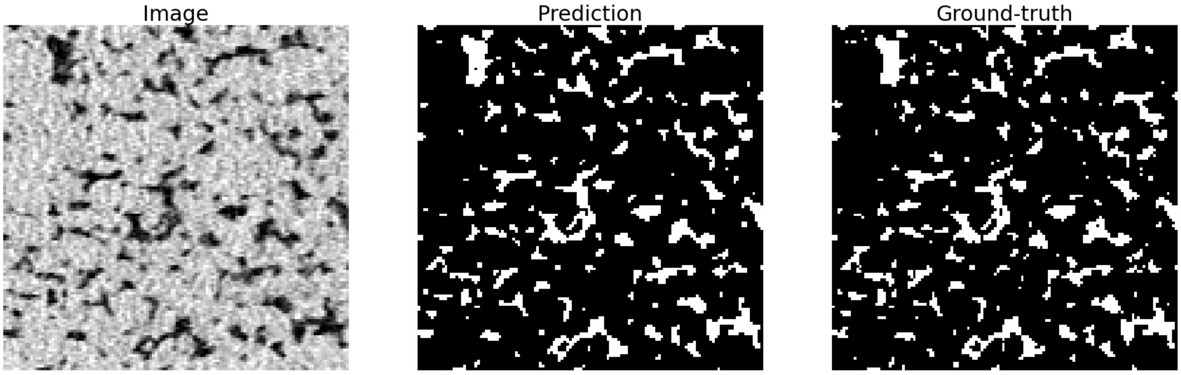


Fig. 4.12. Examples of 2D slices through 3D segmented images for Sample 4.

As we can see, our model is uncertain only about interface pixels, the amount of false-negative and false-positive errors is almost equal.

4.2. Quantitative segmentation results

To quantitatively describe the observed results, we built a bar charts to compare the following options:

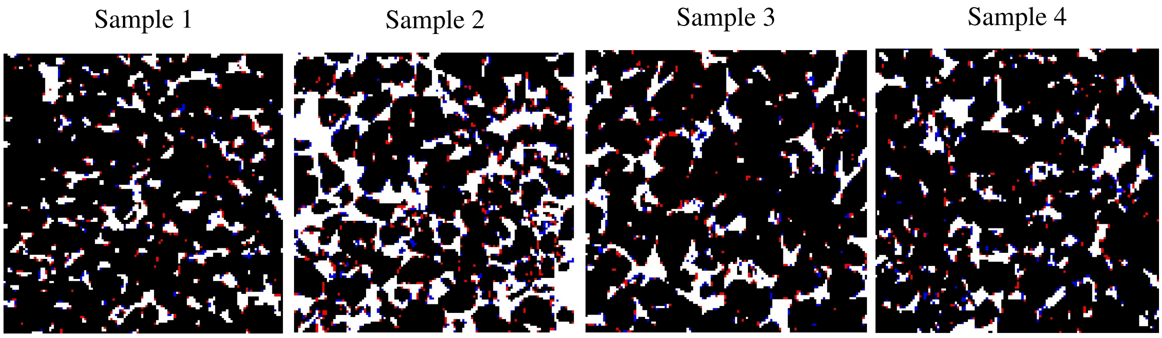


Fig. 4.13. Visualization of false-positive and false-negative segmentation errors.

1. Segmentation quality depending on the number of the annealing iterations for each sample separately (Fig. 4.14);
2. Segmentation quality of the original porous media stacks (Fig. 4.15);
3. Segmentation quality depending on the number of the annealing iterations in average over all stacks (Fig. 4.16).

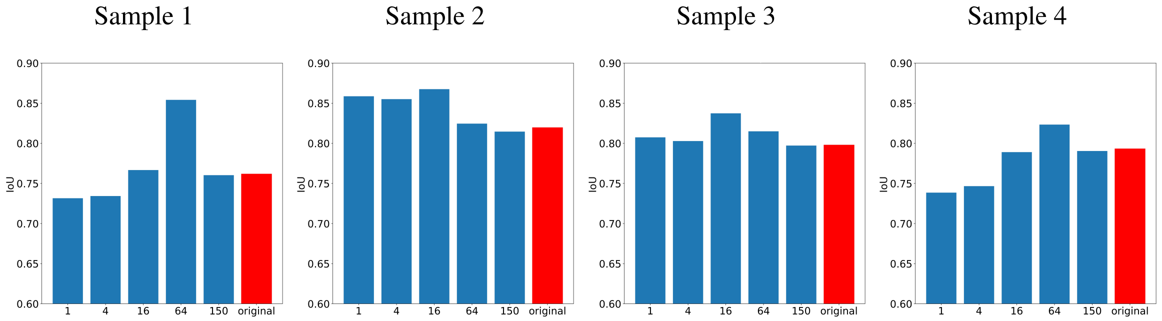


Fig. 4.14. Metrics for each stack.

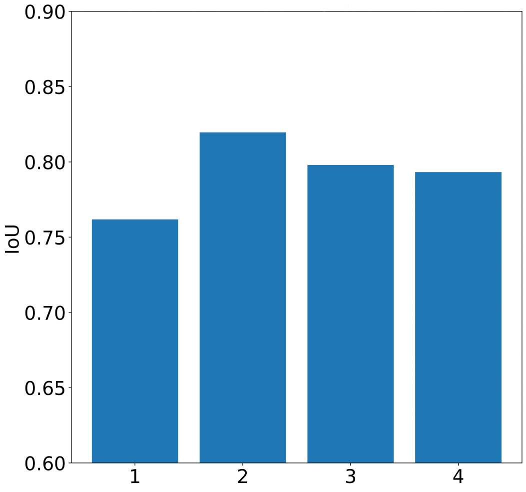


Fig. 4.15. Separate IoU metrics for original (real porous media) Samples.

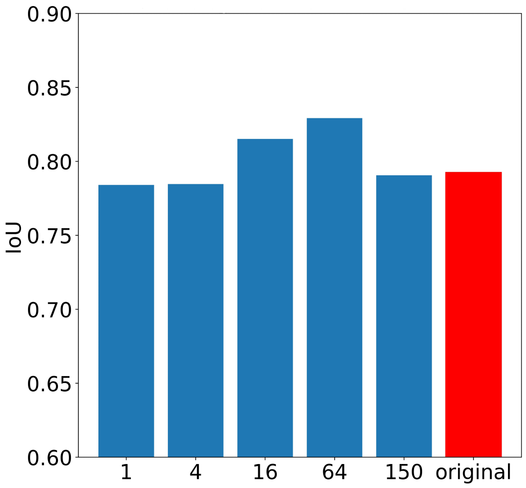


Fig. 4.16. Averaged IoU metrics over all Samples

5. DISCUSSION

First of all, we tried to understand why the segmentation quality for Sample 2 was much better than for the rest of samples. We believe that this is due to the fact that this sample differs from others in its morphological properties (in terms of correlation functions). Samples 1, 3, 4 are similar to each other in correlation functions L_2 (Sample 3 slightly differ from Samples 1, 4) and F_{SS} (all Samples are close in all directions), all Samples are pretty close to each other in S_2, the Sample 2 is not similar to the others in any correlation function, except S_2 (see Fig. 5.17).

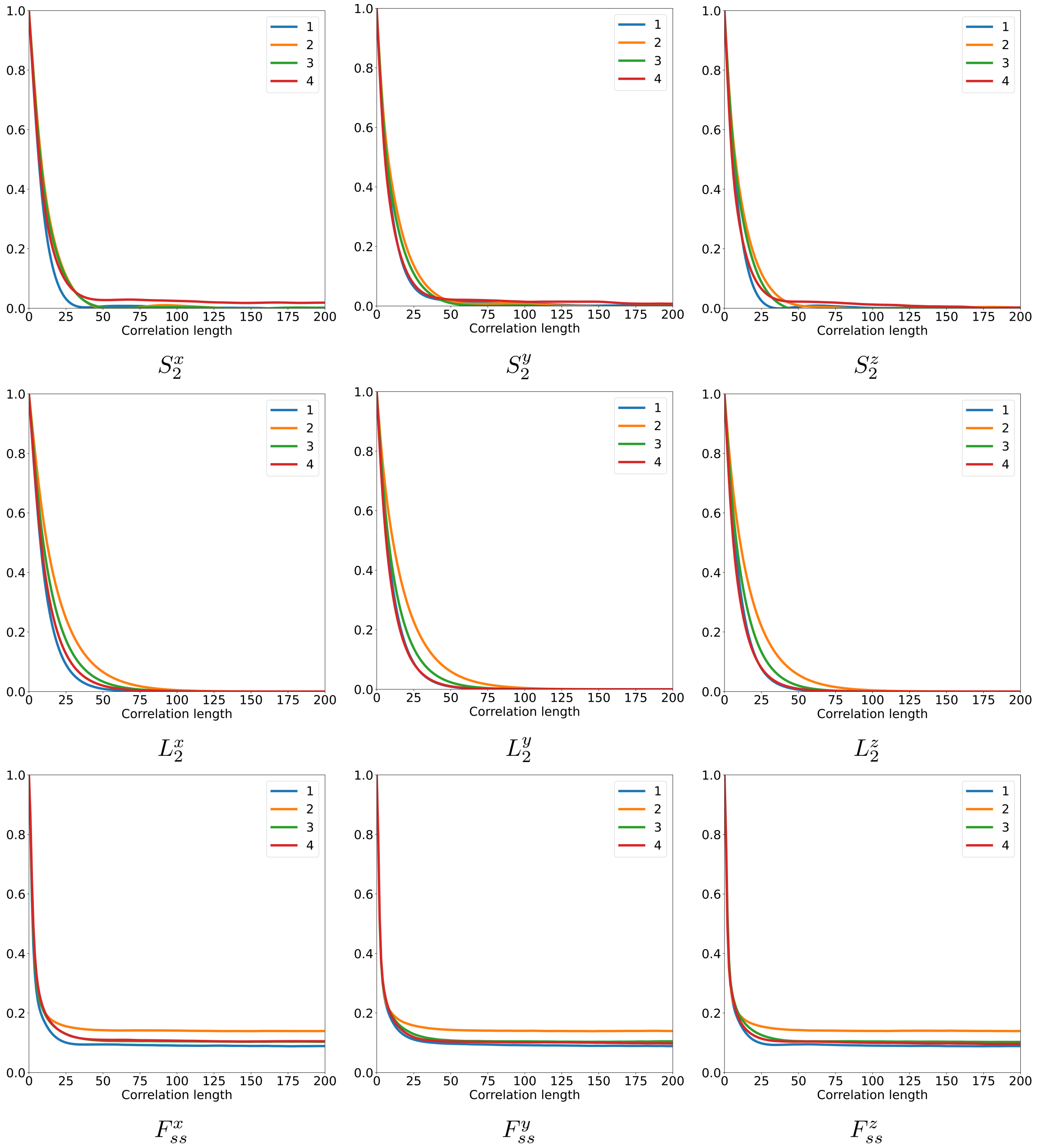
Next, we analyzed why metrics dynamics (changes in segmentation accuracy with each annealing iteration) for stochastic reconstructions is multidirectional for different samples. In particular Sample 2 is different to other samples. In this regard, we established that Sample 2 has the simplest structure (Fig. 2.3), and therefore showed the best segmentation metrics. Moreover, stochastic ”deconstruction” (structural dynamics in reversed direction on Fig.3.8) further simplifies its structure (Fig. 3.8) by increasing the IoU metric.

The presence of the peak in metrics at 16 and 64 iterations raised some questions. After reviewing the corresponding stochastic reconstructions (Fig. 3.8), we can conclude that they are visually much more similar to the original samples than the reconstructions at the final iterations. This means that the limited information content of the correlation functions does not fully convey the morphological properties of the original stacks.

The above-mentioned considerations allow to suspect that the deconstruction of the initial samples’ structure has less effect on the segmentation metrics than the initial difference in the morphological properties between replicas and original porous media images. To test this hypothesis, it is necessary to carry out a similar experiment with a large number of different original samples. Porosity of the sample can be of great importance on the results, therefore, it is necessary to neutralize its influence. This can be achieved by collecting a a library of original samples with the same initial porosity (stochastic reconstructions preserve the porosity of the original).

6. CONCLUSION

In this paper we presented the initial results of measuring the influence of samples morphology on CNN-based segmentation quality. To do this, we segmented four different



S_2^x S_2^y S_2^z

L_2^x L_2^y L_2^z

F_{ss}^x F_{ss}^y F_{ss}^z

Fig. 5.17. Plot of the used correlation functions set along all directions.

porous media samples using an unsupervised converging active contours algorithm. Next, we obtained the XCT – ground-truth pairs for these samples with the help of synthetic tomography algorithm. We acquired a set of samples with varying degrees of morphological properties similar to the original stacks with the help of stochastic reconstructions. We trained U-Net model with ResNet50 encoder on pairs for the original porous media samples. Next, we assessed the segmentation quality of synthetic XCT obtained from stochastically reconstructed samples, and checked the effect of morphological properties on the segmentation quality.

Based on the metrics, we concluded that the quality of segmentation is influenced more by the morphological differences between the original soil samples than by difference from the generated stochastic reconstructions. We discussed possible ways to improve the future model experiment design in order to finally uncover the relationship between morphology variations ans segmentation accuracies.

The results as presented here can help in simplifying the construction of a general segmentation pipelines. We showed that it is not necessary to segment the original images perfectly. To build a universal pipeline, it is necessary to continue this research direction with a much larger library of images with various morphological properties.

ACKNOWLEDGEMENTS

This work was supported by the Russian Foundation for Basic Research grant 20-31-90118\20 (The application of neural networks to solve the problem of optimal segmentation of tomographic images of geological rocks).

REFERENCES

1. Van Aarle, W. et al. (2015). The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography. *Ultramicroscopy*, **157**, 35–47.
2. Adler, P. (1992). *Porous media: geometry and transports*. Oxford, UK: Butterworth-Heinemann.
3. Adler, P.M., Jacquin, C.G. & Quiblier, J.A. (1990). Flow in simulated porous media. *Int. J. Multiph. Flow* **16**, 691–712. https://doi.org/10.1016/0301-9322(90)90025-E
4. Albert, A., Kaur, J., & Gonzales, M.C. (2017) Using Convolutional Networks and Satellite Imagery to Identify Patterns in Urban Environments at a Large Scale. *arXiv*: 1704.02965v5, [Online]. Available: https://arxiv.org/abs/1704.02965
5. Badrinarayanan V., Alex K.A., Cipolla R. (2016) SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *arXiv*: 1511.00561v3, [Online]. Available: https://arxiv.org/abs/1511.00561
6. Barbosa, L.A.P., Gerke, K.M. & Gerke, H.H. (2022). Modelling of soil mechanical stability and hydraulic permeability of the interface between coated biopore and matrix pore regions. *Geoderma*, **410**, 115673.
7. Baveye, P.C., Laba, M., Otten, W., Bouckaert, L., Sterpaio, P. D. et al. (2010). Observer-dependent variability of the thresholding step in the quantitative analysis of soil images and X-ray microtomography data. *Geoderma*, **157**(1–2), 51–63.
8. Berg, S., Saxena, N., Shaik, M. & Pradhan, C. (2018). Generation of ground truth images to validate micro-CT image-processing pipelines. *The Leading Edge*, **37**(6), 412–420.
9. Bernal, J., Kushibar, K., Asfaw, D.S., Valverde, S., Oliver, A. et al. (2017) Deep convolutional neural networks for brain image analysis on magnetic resonance imagesing: a review. *arXiv*: 1712.03747v3, [Online]. Available: https://arxiv.org/abs/1712.03747v3
10. Bilger, C., Aboukhedr, M., Vogiatzaki, K. & Cant, R. S. (2017). Evaluation of two-phase flow solvers using Level Set and Volume of Fluid methods. *Journal of Computational Physics*, **345**, 665–686.
11. Blunt, M. J. (2017). *Multiphase flow in permeable media: A pore-scale perspective*. Cambridge, UK: Cambridge University Press.
12. Blunt, M.J., Bijeljic, B., Dong, H., Gharbi, O., Iglauer, S. et al. (2013) Pore-scale imaging and modelling. *Advances in Water Resources*, **51**, 197–216. https://doi.org—10.1016/j.advwatres.2012.03.003.
13. De Boever, W., Derluyn, H., Van Loo, D., Van Hoorebeke, L. & Cnudde, V. (2015). Data-fusion of high resolution X-ray CT, SEM and EDS for 3D and pseudo-3D chemical and structural characterization of sandstone. *Micron*, **74**, 15–21.

14. Bultreys, T., Brabant, L., Boone, M., Cnudde, V., & Van Hoorebeke, L. (2013). Unsupervised segmentation evaluation measures for parameter optimization in indicator-Kriging. *Tomography of materials and structures : book of abstracts : posters*, pp. 147–150.
15. Chauhan, S., Rühaak, W., Anbergen, H., Kabdenov, A., Freise, M. et al. (2016). Phase segmentation of X-ray computer tomography rock images using machine learning techniques: an accuracy and performance study. *Solid Earth*, **7**(4), 1125–1139.
16. Chauhan, S., Rühaak, W., Khan, F., Enzmann, F., Mielke, P. et al. (2016). Processing of rock core microtomography images: Using seven different machine learning algorithms. *Computers & Geosciences*, **86**, 120–128.
17. Cherkasov, A., Ananev, A., Karsanina, M., Khlyupin, A. & Gerke, K. (2021) Adaptive phase-retrieval stochastic reconstruction with correlation functions: Three-dimensional images from two-dimensional cuts. *Physical Review E*, **104**(3), 035304.
18. Ciresan, D.C., Giusti, A., Gambardella, L.M., & Schmidhuber, J. (2012) Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images, *Proc. of the 25th International Conference on Neural Information Processing Systems*, Stateline, NV, 1–9.
19. Cnudde, V. & Boone, M.N. (2013) High-resolution X-ray computed tomography in geosciences: A review of the current technology and applications. *Earth-Science Reviews*, **123**, 1–17. https://doi.org/10.1016/j.earscirev.2013.04.003
20. Cnudde, V., Masschaele, B., Dierick, M., Vlassenbroeck, J., Van Hoorebeke, L. et al. (2006). Recent progress in X-ray CT as a geosciences tool. *Applied Geochemistry*, **21**(5), 826–832.
21. Derossi, A., Gerke, K.M., Karsanina, M.V., Nicolai, B., Verboven, B. et al. (2018). Mimicking 3D food microstructure using limited statistical information from 2D cross-sectional image. *Journal of Food Engineering*, **241**(5), 116–126.
22. Gerke, K.M. et al. (2020) Improving watershed-based pore-network extraction method using maximum inscribed ball pore-body positioning. *Advances in Water Resources*,**140**, 103576.
23. Gerke, K.M., Karsanina, M. V. (2015). Improving stochastic reconstructions by weighting correlation functions in an objective function. *Europhysics Lett.*, **111**, 56002. https://doi.org/10.1209/0295-5075/111/56002
24. Gerke, K.M., Karsanina, M. V. & Mallants, D. (2015). Universal Stochastic Multiscale Image Fusion: An Example Application for Shale Rock. *Sci. Rep.*,**5**, 15880. https://doi.org/10.1038/srep15880
25. Gerke K.M., Karsanina M.V., Sizonenko T.O., Miao X., Gafurova D.R. et al. (2017) Multi-scale image fusion of X-ray microtomography and SEM data to model flow and transport properties for complex rocks on pore-level. *SPE Russian Petroleum Technology Conference* Moscow, Russia. https://doi.org/10.2118/187874-MS.
26. Gerke, K.M., Karsanina, M. V., Vasilyev, R. V. & Mallants, D. (2014). Improving pattern reconstruction using directional correlation functions. *Europhysics Lett.*, **106**, 66002. https://doi.org/10.1209/0295-5075/106/66002
27. Gerke, K.M., Sidle, R.C. & Mallants, D. (2015). Preferential flow mechanisms identified from staining experiments in forested hillslopes. *Hydrol. Process.*,**29**, 4562–4578. https://doi.org/10.1002/hyp.10468
28. Gerke, K.M., Skvortsova, E.B. & Korost, D. V. (2012). Tomographic method of studying soil pore space: Current perspectives and results for some Russian soils. *Eurasian Soil Sci.*, **45**, 700–709. https://doi.org/10.1134/S1064229312070034
29. Gerke K.M., Vasilyev R.V., Khirevich S., Karsanina M.V., Collins D. et al. (2018) Finite-difference method Stokes solver (FDMSS) for 3D pore geometries: Software development, validation and case studies. *Computers & Geosciences*, **114**, 41–58. https://doi.org/10.1016/j.cageo.2018.01.005

30. Godinho, J.R., Gerke, K.M., Stack, A.G. & Lee, P.D. (2016) The dynamic nature of crystal growth in pores. *Scientific Reports*, **6**, 33086. https://doi.org/10.1038/srep33086
31. Hashemi, M. A., Khaddour, G., François, B., Massart, T. J. & Salager, S. (2014). A tomographic imagery segmentation methodology for three-phase geomaterials based on simultaneous region growing. *Acta Geotechnica*, **9**(5), 831–846. https://doi.org/10.1007/s11440-013-0289-5.
32. He Y., Pu C., Jing C., Gu X., Chen Q. et al. (2017). Reconstruction of a digital core containing clay minerals based on a clustering algorithm. *Physical Review E*, **96**(4), 043304.
33. Hu, D., Ronhovde, P. & Nussinov, Z. (2012). Replica inference approach to unsupervised multiscale image segmentation. *Physical Review E*, **85**(1), 016101.
34. 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**(9), 1–12.
35. Jiao, Y. & Chawla, N. (2014). Modeling and characterizing anisotropic inclusion orientation in heterogeneous material via directional cluster functions and stochastic microstructure reconstruction. *J. Appl. Phys.*, **115**, 093511. https://doi.org/10.1063/1.4867611
36. Jiao, Y., Stillinger, F.H. & Torquato, S. (2007). Modeling heterogeneous materials via two-point correlation functions: Basic principles. *Physical Review E*, **76**(3), 031110.
37. Kaiming, H., Xiangyu, Z., Shaoqing, R. & Jian, S. (2015) Deep Residual Learning for Image Recognition. *arXiv*: 1512.03385, [Online]. Available: https://arxiv.org/abs/1512.03385
38. Karimpouli, S. & Pejman, T. (2019) Segmentation of digital rock images using deep convolutional autoencoder networks. *Computers & Geosciences*, **126**, 142–150.
39. Karsanina, M.V. & Gerke, K.M. (2018) Hierarchical optimization: Fast and robust multiscale stochastic reconstructions with rescaled correlation functions. *Physical review letters*, **121**(26), 265501.
40. Karsanina, M.V., Gerke, K.M., Skvortsova, E.B., Ivanov, A.L. & Mallants, D. (2018) Enhancing image resolution of soils by stochastic multiscale image fusion. *Geoderma*, **314**, 138–145. https://doi.org/10.1016/j.geoderma.2017.10.055
41. Karsanina, M.V., Gerke, K.M., Vasilyev, R.V. & Korost, D.V. (2015). Using correlation functions to model material's structure with desired physical properties. *Math. Model. Comput. Simulations*, **27**, 50–63.
42. Khan, F., Enzmann, F. & Kersten, M. (2016). Multi-phase classification by a least-squares support vector machine approach in tomography images of geological samples. *Solid Earth*, **7**(2), 481–492
43. Khirevich, S., Ginzburg, I. & Tallarek, U. (2015). Coarse-and fine-grid numerical behavior of MRT/TRT lattice-Boltzmann schemes in regular and random sphere packings. *J. Comput. Phys.*, **281**, 708–742. https://doi.org/10.1016/j.jcp.2014.10.038
44. Krizhevsky, A., Sutskever I. & Hinton G.E. (2012) ImageNet classification with deep convolutional neural networks. *Proc. of the 25th International Conference on Neural Information Processing Systems*, Stateline, NV, 1097–1105.
45. Latief, F.D.E., Biswal, B., Fauzi, U. & Hilfer, R. (2010). Continuum reconstruction of the pore scale microstructure for Fontainebleau sandstone. *Physica A: Statistical Mechanics and its Applications*, **389**(8), 1607–1618.
46. Lavrukhin, E.V., Gerke, K.M., Romanenko, K.A., Abrosimov, K.N. & Karsanina, M.V. (2021) Assessing the fidelity of neural network-based segmentation of soil XCT images based on pore-scale modelling of saturated flow properties. *Soil and Tillage Research*, **209**, 104942.
47. Li, H., Chawla, N. & Jiao, Y. (2014). Reconstruction of heterogeneous materials via stochastic optimization of limited-angle X-ray tomographic projections. *Scripta Materialia*, **86**, 48–51.

48. Li, H., Chen, P. E. & Jiao, Y. (2018). Accurate Reconstruction of Porous Materials via Stochastic Fusion of Limited Bimodal Microstructural Data. *Transport in Porous Media*, **125**, 5–22.
49. Mariethoz, G., Renard, P. & Straubhaar, J. (2011) Extrapolating the fractal characteristics of an image using scale-invariant multiple-point statistics. *Mathematical Geosciences*, **43**(7), 783–797. https://doi.org/10.1007/s11004-011-9362-5
50. Miao X., Gerke, K.M. & Sizonenko, T.O., (2017). A new way to parameterize hydraulic conductances of pore elements: A step forward to create pore-networks without pore shape simplifications. *Adv. Water Resour.*, **105**, 162–172. https://doi.org/10.1016/j.advwatres.2017.04.021
51. Nogueira, K., Fadel, S.G., Dourado, I.C., de O. Werneck, R., Munoz, J.A.V. et al. (2017) Exploring ConvNet diversity for flooding identification. *arXiv*: 1711.03564v2, [Online]. Available: https://arxiv.org/abs/1711.03564
52. Oh, W. & Lindquist, B.W. (1999). Image thresholding by indicator kriging. *IEEE Trans. Pattern Anal. Mach. Intell.*, **21**, 590–602. https://doi.org/10.1109/34.777370
53. Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, **9**(1), 62–66.
54. Raeini, A.Q., Bijeljic, B. & Blunt, M. J. (2017). Generalized network modeling: Network extraction as a coarse-scale discretization of the void space of porous media. *Physical Review E*, **96**(1), 013312.
55. Raeini, A.Q., Blunt, M.J. & Bijeljic, B. (2012). Modelling two-phase flow in porous media at the pore scale using the volume-of-fluid method. *Comput. Phys.*, **231**, 5653–5668. https://doi.org/10.1016/j.jcp.2012.04.011
56. Ronneberger, O., Fischer P. & Brox T. (2015) U-net: convolutional networks for biomedical image segmentation. *arXiv*: 1505.04597v1, [Online]. Available: https://arxiv.org/abs/1505.04597v1
57. Ruder S. (2016) An overview of gradient descent optimization algorithms. *arXiv*: 1609.04747v2, [Online]. Available: https://arxiv.org/abs/1609.04747
58. Sahimi, M. (2011). Flow and transport in porous media and fractured rock: from classical methods to modern approaches. New York, NY: John Wiley & Sons.
59. Schlüter, S., Sheppard, A., Brown, K. & Wildenschild, D. (2014). Image processing of multiphase images obtained via X-ray microtomography: a review. *Water Resources Research*, **50**(4), 3615–3639.
60. Schlüter, S., Weller, U. & Vogel, H.J. (2010). Segmentation of X-ray microtomography images of soil using gradient masks. *Comput. Geosci.*, **36**, 1246–1251. https://doi.org/10.1016/j.cageo.2010.02.007
61. Sheppard, A.P., Sok, R.M. & Averdunk, H. (2004) Techniques for image enhancement and segmentation of tomographic images of porous materials. *Physica A: Statistical mechanics and its applications*, **339**(1), 145–151. https://doi.org/10.1016/j.physa.2004.03.057
62. Tahmasebi, P. (2018). Nanoscale and multiresolution models for shale samples. *Fuel*, **217**, 218–225.
63. Thovert, J.F. & Adler, P.M. (2011). Grain reconstruction of porous media: Application to a Bentheim sandstone. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, **83**, 056116. https://doi.org/10.1103/PhysRevE.83.056116
64. Torquato, S. (2013). Random heterogeneous materials: microstructure and macroscopic properties. New York, NY: Springer.
65. Varfolomeev, I., Yakimchuk, I. & Safonov, I (2019). An application of deep neural networks for segmentation of microtomographic images of rock samples. *Computers*, **8**(4), 72.
66. Vesely, M., Bultreys, T., Peksa, M., Lang, J., Cnudde, V. et al. (2015). Prediction and evaluation of time-dependent effective self-diffusivity of water and other effective transport properties associated with reconstructed solids. *Transport in Porous Media*,

110(1), 81–111. https://doi.org/10.1007/s11242-015-0557-y.

67. Wang, W., Kravchenko, A.N., Smucker, A.J.M. & Rivers, M.L. (2011). Comparison of image segmentation methods in simulated 2D and 3D microtomographic images of soil aggregates. *Geoderma*, **162**(3–4), 231–241.
68. Wildenschild, D. & Sheppard, A.P. (2013). X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems. *Advances in Water Resources*, **51**, 217–246. https://doi.org/10.1016/j.advwatres.2012.07.018
69. Yang, Y.S., Gureyev, T.E., Tulloh, A., Clennell, M.B. & Pervukhina, M. (2010). Feasibility of a data-constrained prediction of hydrocarbon reservoir sandstone microstructures. *Measurement Science and Technology*, **21**(4), 047001.
70. Yang, Y. S., Liu, K. Y., Mayo, S., Tulloh, A., Clennell, M. B. et al. (2013). A data-constrained modelling approach to sandstone microstructure characterisation. *Journal of Petroleum Science and Engineering*, **105**, 76–83.
71. Yeong, C. & Torquato, S. (1998). Reconstructing random media. II. Three-dimensional media from two-dimensional cuts. *Phys. Rev. E*, **58**, 224–233. https://doi.org/10.1103/PhysRevE.58.224