Evaluation of Segmentation Methods Based on Classification Patterns for Micro-Tomography Applications in Rock Analysis

Gustavo J. Q. de Vasconcelos[*][†], Giovanna Antonieti[*][‡], Gustavo C. Libel[§], Paola R. R. Rosa[*],
Nathaly L. Archilha[*], Tiago J. Carvalho[‡], Helio Pedrini[†], Thiago V. Spina[*]

[*]Brazilian Synchrotron Light Laboratory, Brazilian Center for Research in Energy and Materials, Campinas, SP, Brazil
[†]Institute of Computing, University of Campinas, Campinas, SP, Brazil
[‡]Federal Institute of São Paulo – Campinas, Campinas, SP, Brazil
[§]Department of Informatics, Federal University of Technology – Paraná, Curitiba, PR, Brazil

***Abstract*—4D micro and nanotomography allows the study of time-resolved phenomena, such as understanding how water, oil, and gas interact within rock pore space to improve oil engineering. This category of experiment has only become possible due to the invention of powerful devices such as the MOGNO micro and nano-tomography beamline of Sirius, the new state-of-the-art Brazilian synchrotron light source. Nowadays, the biggest bottleneck for data analysis in this type of experiment is the image segmentation task, given that MOGNO may generate one *3.6 gigavoxels* 3D image in *1-5s*. To achieve near real-time image segmentation in the future and reduce manual processing, we propose to convert the image segmentation task into *superpixel classification*. We have evaluated different combinations of superpixel estimation algorithms, feature extraction filters, and pattern classifiers aiming to automatically segment the pore space in 3D micro-CT rock grain images.**

***Index Terms*—Pore spaces; rock analysis; image segmentation; micro-tomography**

I. Introduction

With the development of modern X-ray generation methods, such as *synchrotron radiation sources*, micro and nanotomography have recently emerged as prime techniques for high-resolution analysis of small structures in 3D images. In particular, 4th (latest) generation synchrotron light sources are pushing forward yet another innovation: 4D micro and nanotomography. This will allow researchers to study how time-resolved phenomena affect different materials.

The Brazilian Synchrotron Light Laboratory (LNLS) is currently engaged in developing Sirius, a new 4th generation 3 GeV synchrotron light source designed to be a leading machine in its energy class. Sirius will replace the current 2nd generation light source at LNLS (UVX), representing an improvement of orders of magnitude in image acquisition. The proposed X-ray tomography beamline at Sirius, named MOGNO (MicrO and NanO Tomography), is being designed to be a micro and nano imaging beamline focused towards multi-scale analysis of the internal 3D structures of different materials and objects. The beamline will be primarily devoted and specialized in zoom-tomography, where a specimen can be studied at low and high-resolution, and 4D tomography – a series of 3D images, each one containing nearly *3.6 billion voxels* and obtained in the order of *1-5s* [1], [2]. Currently, a single 3D experiment at the IMX microtomography beamline of the UVX takes up to *4 hours*.

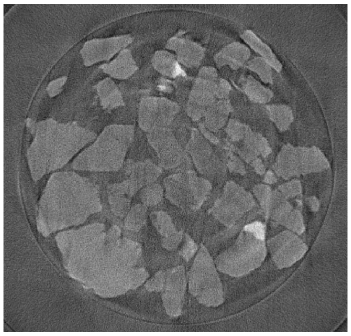

Figure 1: Slice from a 3D tomography experiment of a rock sample acquired on the IMX beamline at LNLS.

Nowadays, the biggest bottleneck in the entire process, from tomography acquisition to final data analysis, is the image segmentation task and, due to MOGNO's real time characteristic, this problem is expected to be even worse with 4D tomography. The segmentation process is currently done interactively through conventional image segmentation techniques, using methods such as the seeded watershed transform [3]. This process relies on the user's expertise in how to use the tool, requiring a long invested time to learn and, therefore, making it difficult to immerse new users in the technique of X-ray tomography.

This paper focuses on the design of image segmentation tools to address the aforementioned challenges that will be imposed by the MOGNO beamline upon its completion. Our goal is to evaluate machine learning techniques that can be

The authors thank MCTIC, CNPq, CAPES, and the Serrapilheira Institute (Serra-1708-16161) for funding. Corresponding authors: thiago.spina@lnls.br and helio@ic.unicamp.br

978-1-5090-1897-0/16/$31.00 ©2017 IEEE

implemented in the future to run at near real-time speeds, aiming to segment 4D micro/nano-tomography images with several billion voxels in a few minutes. In particular, we are interested in segmenting images of rocks and soil to perform *pore network analysis*. The high energy and flux of the MOGNO beamline makes it a leading equipment for pore analysis of rocks an soil [4], [5]. 4D pore scale imaging and analysis may open the door for significant improvements in agriculture and the oil industry, by allowing fast fluid flow experiments to study how water, oil, gas, and other particles interact inside the pore space of soil and rocks.

Motivated by this challenging scenario, our main contribution is an evaluation of image segmentation methods focused on the needs of the IMX and MOGNO beamlines. To achieve near real-time image segmentation, we propose to convert the image segmentation task into a *superpixel classification* process. Superpixels are formed by grouping neighboring pixels with similar intensities. In this way, superpixels carry more information than pixels and reduce the dimensionality of the image segmentation problem, since less data has to be processed to achieve the final segmentation. We have evaluated different combinations of superpixel estimation algorithms, feature extraction filters, and pattern classifiers aiming to segment the pore space in 3D micro-CT rock grain images. Our goal was to select those methods keeping in mind that they can usually be efficiently implemented later using High Performance Computing (HPC) – i.e., parallel programming via GPUs. We compare those algorithms with state-of-the-art deep learning approaches to image segmentation [6], which usually require a much higher computational time for training and prediction.

This paper is organized as follows. Section II details the proposed methodology for image segmentation using superpixels, as well as the evaluated methods for feature extraction and pattern classification. Section III presents the experiments conducted to evaluate our methodology. We then state our conclusions in Section IV.

II. Superpixel-based Image Segmentation

In computer vision and image processing, segmentation is the process of partitioning a digital image into multiple segments, in order to precisely define the spatial location objects and boundaries in images. The goal is to assign a label $L(p) \in \{0, 1, \ldots, c\}$ to every pixel p in an image I [3] corresponding to one out of c semantic objects of interest or the background ($L(p) = 0$). We achieve this task by evaluating the machine learning pipeline for superpixel classification.

We first estimate superpixel labels $S(p) \in \{1, 2, \ldots, k\}$ for the pixels of the original image $p \in I$, computed in 2D for the z-slices for efficiency. Afterwards, a feature vector $\vec{F}_S(s) = \{F_s^1, F_s^2, \ldots, F_s^m\}$ is computed to represent every superpixel, with the mean value F_s^i of the corresponding pixels in the ith feature map s.t. $S(p) = s$ and $i = 1, 2, \ldots, m$. The m feature maps F^i are filtered versions of the image I (Section II-B).

Each sample s (superpixel) is assigned a label $L_S(s)$ from the ground truth image $\lambda(p) \in \{0, 1, \ldots, c\}$, generated by an area expert (Figure 2b), via majority voting of the pixels $p \in I$ with $S(p) = s$. A classification model M is trained by considering the superpixel feature vectors $\vec{F}_S$ in a training set, along with the corresponding ground truth labels L_S. Afterwards, prediction occurs by computing the superpixel labels S' of a test image I', extracting the feature vectors $\vec{F}'_S$, and applying classifier M over the result. The final label assignment $L'(p)$ for every pixel $p \in I'$ is simply obtained from the result of the superpixel classification.

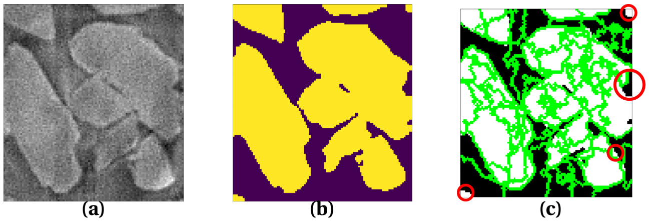

Figure 2: (a) Central part of a slice. (b) Ground truth segmentation. (c) Examples of object edge errors in superpixel estimation.

A. Superpixel estimation

The choice of the superpixel algorithm and its parameters for the specific application is crucial. The main concern with superpixels for the segmentation process through supervised classification is to ensure the correct delimitation of the object edges in order to avoid errors (Figure 2c).

a) SLIC: The Simple Linear Iterative Clustering algorithm [7] produces compact and nearly uniform superpixels by using k-means to cluster pixels surrounding candidate *seed pixels* selected on a regularly spaced grid.

b) Felzenszwalb: This algorithm computes superpixels with varying shape and great boundary adherence by merging regions obtained by oversegmenting the image using the minimum spanning tree of the pixels [8].

c) Compact watershed: This method is a regularized version of the seeded watershed transform [3] to compute regions with a somewhat uniform and compact shape [9].

B. Feature extraction

Feature extraction should ensure that relevant properties of the objects of interest be captured by the superpixels. We adopt a strategy of locally filtering the original image I using operators on patches with radius ρ surrounding each pixel, and then summarizing the values as previously described to form a superpixel feature vectors $\vec{F}_S$. Afterwards, we performed a feature selection procedure to select the most relevant features for posterior classification.

The following filters were applied: Auto-level, Local bottom-hat, Equalize, Gradient, Maximum, Mean, Subtract Mean, Median, Mode, Enhance Contrast, Pop Bilateral, Sum, Sum Bilateral, Threshold, Local Tophat, Entropy, Otsu, Percentile. They can be found on the Scikit-image library.

C. Machine learning models for supervised classification

We have evaluated the Random Forests, Support Vector Machines, and Logistic Regression classifiers. They were selected particularly due to their low computational complexity and parallelization capability during inference.

1) Random Forests [10]: creates an ensemble of random uncorrelated decision trees, which partition the feature space $\mathbb{R}^m$ into classes by learning thresholds for each feature $F^i_S(s)$, with $i = 1, 2, \ldots, m$. During testing, the feature vector $\vec{F}'_S(s)$ of a test sample $s \in S'$ is fed to the decision trees in order to decide the class $\lambda(s) \in \{1, 2, \ldots, c\}$ to which sample s belongs.

2) Support Vector Machines [11] (SVM): attempts to separate the training samples of two classes by finding a hyperplane $\vec{H} \in \mathbb{R}^m$, with the widest possible margin. A test sample s' is classified according to the sign of the dot product $\vec{H} \cdot \vec{F}'_S(s')$. The inputs into are mapped into a higher dimension feature space using the RBF kernel.

3) Logistic Regression [12]: estimates the probability $p(s') \in [0, 1]$ of a test sample s' to belong to one out of two classes by applying a sigmoid function over the weighted sum of the training feature vectors $F_S(s)$, plus a bias term. Training involves determining the coefficient parameters of the equation system that best fits the training samples to the expected classes, using maximum likelihood estimation.

D. Deep learning baseline

The U-net [6] is a fully convolutional neural network consisting of a contracting and an expansive path. The former applies convolutions followed by ReLU and max-pooling operations to reduce spatial information while increasing feature information. The latter combines the feature and spatial information through a sequence of up-convolutions and concatenations with high resolution features from the contracting path to achieve image segmentation.

III. Experimental Evaluation

In our experiments, we first evaluated which superpixel estimation algorithm best performed for the task. Then, the best filtering parameters were selected prior to feature selection. Finally, we compared different classifiers using the previous experiments results versus the U-Net baseline. We considered the Dice Similarity Coefficient (DSC) and the Average Symmetric Surface Distance (ASSD) to evaluate the accuracy of segmentation and boundary error, respectively.

A. Datasets

Our experiments involved three datasets of 3D micro-CT images. The first dataset is composed of one 3D micro-CT image of dolomite rock acquired using a benchtop micro-CT scanner. The image is sized 1024 × 1024 × 845 voxels, and has been fully segmented into background, grains, and pores by two experts in oil engineering. We subdivided the image into non-overlapping tiles with size 200 × 200 × 20 voxels. The tiles presenting the background were excluded, since our goal is to evaluate binary segmentation of pores, and the remaining ones were randomly separated into a training and a testing dataset with size of 70% and 30%, respectively. 5% of the training images were selected for performing the grid search parameter optimization, in each of the three experiments. The final comparison considered the entire training/testing sets.

Our second dataset contains one 1024 × 1024 × 150 voxels micro-CT image acquired at the IMX beamline, displaying glass beads that are used to simulate rock grains in fluid flow experiments. The same previously stated procedure of ground truth and dataset generation was performed. The third dataset contains two 3D tomography images from different experiments of carbonate rocks performed on the IMX beamline at LNLS. The first image used as a training set contains 20 slices of 137 × 127, while the second image used as a test dataset contains 35 slices of 116 × 106. They were acquired from the center of two larger images that were used in fast fluid flow experiments (the IMX image acquisition quality was degraded to reduce the imaging time to a few minutes).

B. Superpixel estimation

In this experiment, we assessed the best possible accuracy that the selected superpixel estimation algorithms may achieve by assigning a segmentation label based on the majority voting of pixels from the ground truth (SP). We also evaluated the classification result obtained by training the Random Forest (RF) using the mean pixel intensity as feature vector. Table I, presents the segmentation accuracy metrics obtained with the best parameters found via grid search on the parameter evaluation subsets.

Table I: Metrics for best parameters found on superpixel grid search in Dataset 1.

Dataset	Method	Metric	Felzenszwalb	SLIC	Comp. WS
1	SP	DICE	0.895	0.886	0.783
		ASSD	1.100	1.094	2.081
	RF	DICE	0.886	0.878	0.777
		ASSD	1.195	1.194	2.256
2	SP	DICE	0.988	0.988	0.964
		ASSD	0.247	0.256	0.852
	RF	DICE	0.917	0.970	0.930
		ASSD	3.839	1.145	2.133
3	SP	DICE	0.966	0.969	0.949
		ASSD	0.594	0.424	0.897
	RF	DICE	0.885	0.881	0.871
		ASSD	1.710	1.771	1.851

In general, the Felzenszwalb and SLIC algorithms achieve similiar metrics with scores higher than Compact Watershed (Comp. WS), and good boundary adhesion. The SLIC superpixels provide better classification accuracy for the RF classifier on Dataset 2, which can be explained by the uniform superpixel segment size. The Compact Watershed algorithm has an intermediate characteristic between the two previous ones, however, it is observed a failure in the generation of the superpixels in the edge of the image.

C. Feature vector extraction comparison

In this section we analyze the ability of the filters mentioned in subsection II-B to extract relevant features to train the machine learning model. The tests were performed using the Random Forest model with 100 trees and the SLIC oversegmentation algorithm with parameters `compactness` = 0.1 and `sigma` = 0.1 for Dataset 1, `compactness` = 0.01 and `sigma` = 0.1 for Dataset 2 and `compactness` = 0.01 and `sigma` = 1 for Dataset 3, also `number_of_segments` = 1200 for all datasets. A grid-search was performed generating maps from filters kernels varying from disks of radius $\rho = 3$ to $\rho = 31$, being a critical parameters according to our experiments. The superpixel mean intensity for each map generated was used as input for the model classification and the best size for each kernel was determined. Afterwards, classifier-based feature selection was performed to select the best 6 features for each dataset.

For Dataset 1, the filters that produced the best results were the Autolevel, Mean, Median, Enhance contrast and Tophat, with corresponding radii $\rho = 31, 3, 3, 3, 15$. For Dataset 2, the filters that produced the best results were the Mean, Median, Mode, Enhance contrast and Otsu, with corresponding radii $\rho = 3, 3, 5, 3, 3$. For Dataset 3 the best filters were Equalize, Subtract mean, Median, Enhance contrast and Threshold, with corresponding radii $\rho = 31, 31, 3, 3, 31$. Also, for all datasets the original image intensity was selected.

D. Classifier comparison with the baseline

Lastly, we have analyzed the segmentation results of the classifier using the features selected for each dataset in the previous section. The models were trained with different numbers of images, to analyze their ability to generalize with less data (Table II). For datasets 1 and 2, the models were trained with $1, 10, 100$ and 200 images, while for the Dataset 3, which is smaller, they considered 1 and 10 images.

Table II: Segmentation metrics.

		Dataset 1		Dataset 2		Dataset 3	
#	Models	DICE	ASSD	DICE	ASSD	DICE	ASSD
1	RF	0.873	1.308	0.985	0.632	0.887	1.539
	SVM	0.875	1.236	0.986	0.627	0.889	1.633
	LogReg	0.873	1.311	0.985	0.633	0.887	1.674
10	RF	0.875	1.220	0.987	0.606	0.904	1.423
	SVM	0.875	1.188	0.987	0.623	0.908	1.385
	LogReg	0.873	1.296	0.986	0.625	0.906	1.460
100	RF	0.875	1.220	0.988	0.584	-	-
	SVM	0.875	1.204	0.987	0.615	-	-
	LogReg	0.875	1.245	0.986	0.590	-	-
200	RF	0.875	1.214	0.988	0.574	-	-
	SVM	0.876	1.204	0.987	0.612	-	-
	LogReg	0.875	1.248	0.986	0.594	-	-

It can be observed from Table II that among the models the SVM classifier usually obtained the best results. For comparison, the U-Net network trained using the whole training dataset and data-augmentation, achieved for Dataset 1 DICE of 0.906 and ASSD of 0.644, for Dataset 2 DICE of 0.982 and ASSD of 0.991 and for Dataset 3 DICE of 0.881 and ASSD of 1.445. Hence, for the datasets considered in this work, the results using the superpixels with roughly 1000 segments are comparable with the state-of-the-art, with much less computational effort to perform the training and also requiring much less annotated data. We can thus affirm, for the datasets used in this paper, that the bottleneck to improve the segmentation is in the superpixel algorithms and not in the models or features.

IV. Conclusions

We have evaluated a strategy to segment 3D microtomography images, aiming to address the real-time requirements of the MOGNO beamline being developed in the Sirius synchrotron light source. Our goal was to evaluate if superpixel-based classification could achieve high accuracy when compared to state-of-the-art techniques for image segmentation, since the former can usually be implemented very efficiently on the GPU. Hence, we have analyzed:

- The quality of the superpixels generated;
- The relevance of filters for superpixel feature extraction;
- The accuracy of different standard supervised classifiers versus deep learning approaches.

Our findings suggest that the combination of the SLIC superpixel generation method, the SVM classifier, and some patch-based filters may achieve accuracy comparable with the state-of-the-art in image segmentation. Hence, our approach seems to be quite promising. Some questions that should be addressed in future work are:

- Is the mean intensity of the maps the best method for feature summarization?
- What other methods of feature extraction can produce better results?
- Where should pixel-level classification be applied to increase segmentation accuracy?

References

[1] G. Costa, N. L. Archilha, F. O'Dowd, and G. Vasconcelos, "Automation Solutions and Prototypes for the X-Ray Tomography Beamline of Sirius, the New Brazilian Synchrotron Light Source," in *ICALEPCS*, no. 16. Barcelona, Spain: JACoW, 2018, pp. 923–927.

[2] E. X. Miqueles, N. Koshev, and E. S.Helou, "A backprojection slice theorem for tomographic reconstruction," *IEEE Trans. Image Process.*, vol. 27, pp. 894 – 906, 2017.

[3] A. X. Falcão, J. Stolfi, and R. A. Lotufo, "The image foresting transform: theory, algorithms, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26(1), pp. 19–29, 2004.

[4] N. L. Archilha, "Quantifição de parâmetros geométricos do sistema poroso por tomografia de raios x e análise da influência em propriedades físicas de rochas carbonáticas," Ph.D. dissertation, Universidade Estadual do Norte Fluminense, 2015.

[5] C. Hollis, V. Vahrenkamp, S. Tull, A. Mookerjee, C. Taberner, and Y. Huang, "Pore system characterisation in heterogeneous carbonates: An alternative approach to widely-used rock-typing methodologies," *Mar. Pet. Geol.*, vol. 27, no. 4, pp. 772–793, 2010.

[6] O. Ronneberger, P.Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *MICCAI*, ser. LNCS, vol. 9351. Springer, 2015, pp. 234–241. [Online]. Available: http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a

[7] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[8] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *IJCV*, vol. 59, no. 2, pp. 167–181, Sep 2004.

[9] P. Neubert and P. Protzel, "Compact watershed and preemptive slic: On improving trade-offs of superpixel segmentation algorithms," in *ICPR*, Stockholm, Sweden, Aug 2014, pp. 996–1001.

[10] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832–844, Aug 1998.

[11] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep 1995.

[12] D. R. Cox, "The regression analysis of binary sequences," *J. R. Stat. Soc. Series B Stat. Methodol.*, vol. 20, no. 2, pp. 215–242, 1958.