

# UC Berkeley

## UC Berkeley Previously Published Works

### Title

Materials Data Science for Microstructural Characterization of Archaeological Concrete

### Permalink

<https://escholarship.org/uc/item/5jc1v56p>

### Journal

MRS Advances, 5(7)

### ISSN

0272-9172

### Authors

Ushizima, Daniela  
Xu, Ke  
Monteiro, Paulo JM

### Publication Date

2020-02-01

### DOI

10.1557/adv.2020.131

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Latest full publication:

MRS Advances © 2020 Materials Research Society DOI:  
10.1557/adv.2020.131

# Materials Data Science for Microstructural Characterization of Archaeological Concrete

Daniela Ushizima<sup>1,2</sup>, Ke Xu<sup>1,2</sup> and Paulo J.M. Monteiro<sup>1,2</sup>

<sup>1</sup>University of California Berkeley, Berkeley, CA 94720

<sup>2</sup>Lawrence Berkeley National Laboratory, Berkeley, CA 94720

*Ancient Roman concrete presents exceptional durability, low-carbon footprint, and interlocking minerals that add cohesion to the final composition. Understanding of the structural characteristics of these materials using X-ray tomography (XRT) is of paramount importance in the process of designing future materials with similar complex heterogeneous structures. We introduce Materials Data Science algorithms centered on image analysis of XRT that support inspection and quantification of microstructure from ancient Roman concrete samples. By using XRT imaging, we access properties of two concrete samples in terms of three different material phases as well as estimation of materials fraction, visualization of the porous network and density gradients. These samples present remarkable durability in comparison with the concrete using Portland cement and nonreactive aggregates. Internal structures and respective organization might be the key to construction durability as these samples come from ocean-submersed archeological findings dated from about two thousand years ago. These are preliminary results that highlight the advantages of using non-destructive 3D XRT combined with computer vision and machine learning methods for systematic characterization of complex and irreproducible materials such as archeological samples. One significant impact of this work is the ability to reduce the amount of data for several computations to be held at minimalistic computational infrastructure, near real-time, and potentially during beamtime while materials scientists are still at the imaging facilities.*

**INTRODUCTION:**

Concrete, the second most-consumed resource in the world, is currently responsible for 8 percent of all carbon dioxide emissions [1, 2] due to one of its key ingredients, cement, with an yearly global production of more than 4 billion tons. Designers and contractors have been urged to modify the concrete formulations [3] due to the cement environmental impact, driving several scientists to look into alternative technologies to minimize this source of anthropogenic greenhouse gas emissions. One of these efforts is the research on ancient Roman concrete [4, 5], its volcanic rock minerals and microscale structures in the presence of seawater, which can lead to the growth of interlocking minerals, bringing added cohesion to the concrete.

For example, poorly crystalline, calcium-aluminum-silicate-hydrate (C-A-S-H binder) in the cementing matrix of the mortar contributed to the long-term durability and pervasive crystallization of zeolite and Al-tobermorite minerals; in addition, the sequestration of chloride and sulfate ions in discrete microstructures of concrete further improves the crack resistance [4]. These hydration and crystallization products compose specific microstructures in ancient Roman concrete.

A major challenge is to characterize the composition of ancient Roman concrete while identifying the structure of the cementing fabric; this paper reports on methods for the investigation of internal structures of such samples using non-destructive 3D imaging by means of X-ray tomography (XRT). From industry to national laboratories, XRT has become fundamental to measure the function and resilience of new materials and for probing dynamic properties. The growth of X-ray brilliance and extremely quick snapshots allied to advances in machine learning create new opportunities to streamline the description of materials structures as part of the design of new compounds. However, the analysis of these rich datasets at scale depends on systematic analysis and automation that combines computational algorithms and experimental outputs at increasing data rate experiments; what is needed are new data science algorithms in support of quantitative image inspection to accelerate the scientific discovery.

Recent efforts in deep learning [6-8] applied to image representation and structural fingerprints have made sample sorting and ranking possible, allowing automated identification of special materials configurations from million-sized databases. These complex networks recognize events from data gathered in two regimes: by experimental testing and by simulation. While such methods successfully bypass hand-engineered features, their full extension to three-dimensional imagery from Materials Science seldom meets standards that are comparable to manual curation. Additionally, complete labelling of gigabyte volumes of experimental data is practically impossible. For example, the inspection of material deformation using X-ray attenuation contrast data from XRT often generates over billions of voxels per time step [9] for one experimental setting.

In order to keep track of these sheer data sizes, we propose a set of tools based on computer vision and machine learning algorithms addressing fundamental problems, such as multiresolution

representation for image segmentation, stereological analysis, and enumeration of material phases within microtomography imagery.

The main contributions of this paper include:

- Design of algorithms to analyse XRT using attenuation contrast that stem from physical experiments using concrete samples from an Italian archaeological site;
- Construction of computational schemes to enable exploratory Materials Data Science (MDS), combining nonlinear filters to gradient ascent methods that iteratively refine pixel clusters to form superpixels using unsupervised methods (simple linear iterative clustering or SLIC), as well as supervised schemes based on random forests, support vector machine and neural nets, whose trained models are used to segment concrete XRT images;
- Development of scripts and curation of XRT datasets to improve experimental reproducibility, integrating both software artefacts<sup>1</sup> as well as the data sets<sup>2</sup> to the content of this paper.

---

<sup>1</sup> [github.com/dani\\_ibnl/mrs2020concrete](https://github.com/dani_ibnl/mrs2020concrete)

<sup>2</sup> K. Xu, P. Monteiro, Y. Su, D. Parkinson, D. Ushizima, *3D Microtomographic Images of Ancient Roman Concrete: Baia samples, v3*, UC Berkeley, Dataset, <https://doi.org/10.6078/D1Q698>

## MATERIALS & METHODS:

This section describes details about the material specimens as well as supervised and unsupervised machine learning (ML) methods tested during the segmentation of different material phases, e.g., matrix, aggregate and interstitial space (pores) from the tomograms.

### Archaeological samples

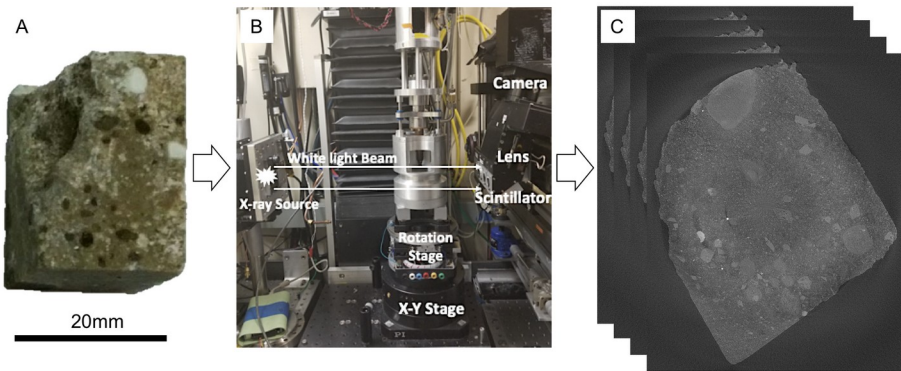


Figure 1. Image acquisition process: (a) Example of a Baiae concrete specimen, (b) XRT setup with X-ray source, the detector and sample stage, and (c) slices from XRT volume after reconstruction using filtered back projection [10], which enables pixel resolution of 8.4  $\mu\text{m}$ .

Ancient Roman concrete has been recognized for its exceptional durability, low-carbon footprint, and the potential to aid the development of the next generation of sustainable high performance concrete [7,8,11]. Unlike the modern concrete and its tendency to decay in a comparatively short time, the ancient Roman concrete has shown higher resiliency and even the ability to improve its strength over 2,000 years. The stability of microcracks and the void system of ancient Roman concrete are some of the reasons it outperforms modern concrete durability, which is possible due to the incorporation of chunks (aggregates) of volcanic rock into a matrix formed by volcanic ash, lime and seawater. Those raw materials have low-carbon footprint and are accessible around the volcanic areas. The hydration and crystallization product of the raw materials formed a durable matrix and stable void network. Besides, the embedded aggregates further contributed the crack resistance and durability.

Due to the pozzolanic reaction, named after the city of Pozzuoli near Naples, Romans noticed that naturally occurring tuff or cemented volcanic ash deposits common in the area were ideal for construction, and then they perfected the art by developing methods to mix ash, water, and heated limestone to be used in many architectural structures [4,5]. Baiae concrete was constructed about 55 BCE to 115 CE, and samples were obtained by the Roman Maritime Concrete

Survey (ROMACONS) [12] drilling program from 2002 to 2006, which has been the most comprehensive study of Roman marine concrete. The field program extracted core samples from eleven locations throughout the Mediterranean.

This paper describes specimens acquired from ancient Roman concrete seawater structures in the bay of Pozzuoli (Baiae), which gives the name to our samples. Our samples were around 0.85 m below the surface of the breakwater and now lies 3.45 m below present sea level, and they are composed of volcanic ash, lime, seawater, and volcanic rocks. In order to unveil the secrets of ancient Roman concrete structures, we investigate these archaeological (67 B.C. to A.D. 64) samples using XRT images to visualize and quantify the different phases of Roman concrete.

### **X-ray tomography (XRT)**

Using advanced X-ray light sources generated by particle accelerators, synchrotron X-ray microtomography enables non-destructive inspection of microscale internal microstructures, especially mineral composition and voids. More specifically, in-situ micro X-ray Computed Tomography tests at LBNL Advanced Light Source (ALS) beamline 8.3.2 [13] have allowed progressive compressive loading (Figure 1) to support both visualization and quantification of materials deformation. Before those steps can proceed, a full characterization of the pristine samples is necessary, which is one of the goals of this paper. Each projection was acquired on a 2560px CCD camera (PCO.edge sCMOS) equipped with an 8.3 mm field of view and a 1X Mitutoyo magnification optical objective.

### **Materials Data Science**

This section describes computer vision and machine learning (ML) algorithms utilized to explore XRT data, to perform image transformation into more suitable data representations to improve discrimination of composition elements, and quantification of detected patterns associated with different components of the cementitious specimens. The proposed end-to-end workflow allows constructing the baseline for comparison with more sophisticated ML algorithms.

After watching several materials scientists analyzing XRT, we noticed that the image understanding follows three main MDS steps with the next sections discussing the main algorithms within each processing module.

1. **Observatory**: volume visualization and assessment of key characteristics of the imaged sample, such as sample conditions, image quality and signal-to-noise ratio given results from straightforward reconstruction algorithms;
2. **Discriminatory**: estimation of region of interest (ROI) to enable lossless data reduction, evaluation of image intensity heterogeneity that may arise from the sample and/or reconstruction artifacts and/or composition interactions, including

iterative algorithms that replace the rigid structure of the pixel grid;

3. **Explanatory:** design of segmentation algorithms that enable the separation of different components, also considering prior information such as pixel resolution, expected number of phases, and interactive process that explore parameter space of algorithms for segmentation proposals, auditing and proper metrology.

These analyses steps are available through reusable functions in Jupyter Notebooks, and are potentially applicable to diverse XRT images that go beyond Roman concrete samples. Several of the proposed algorithms leverage functions from python packages such as *numpy*, *scipy*, *matplotlib*, *pandas*, *skimage*, *sklearn*, *itkwidgets*, which combined integrate a set of calculations often mandatory as part of automated reports about the experiments at the beamline.

## EXPERIMENTAL RESULTS

This paper describes a dataset containing two pristine (Table 1) Baiae specimens imaged using XRT, generating two separate stacks of 32-bit after XRT reconstruction of sinograms using the filtered back projection algorithm in Tomopy [14], with default parameters as prescribed by LBNL ALS.

Table 1. Ancient concrete samples: XRT stacks with 2560\*2560\*504 as illustrated in Figure 7.

Sample ID	Maximum Diameter (mm)	Area of largest cross-section (mm <sup>2</sup> )	Volume of convex polytope (mm <sup>3</sup> )
S1	15	117	456.16
S2	22	365	1615.56

Each stack contains 2560\*2560\*504 voxels, and occupy about 9GB each, with individual images being about 26MB. Our goal is to investigate the complex 3D characteristics of the material by means of the attenuation coefficient, which is proportional to the materials composition density. Spurious intensity variations often arise due to ringing artefacts inherent to the reconstruction algorithm, therefore we design several modules to support MSD for real XRT experiments as it follows.

### Observatory module

As the first step in handling image analysis, we set up scripts to automatically perform the following observatory functions for qualitative assessment of the reconstructed stack:

**Slicing** for accessing each image of the stack individually, including a x-y-z scaling parameter for handling large datasets and allowing fast stack exploration (Figure 2 (a)); in addition, we create an animation generator that accepts user-defined resolution as the single parameter to generate .gif files of a XRT stack;

**Binarization** schemes including isodata, Li, mean, minimum, Otsu, triangle and Yen methods for thresholding [15,16] for exploration of alternative masking schemes (Figure 2(c)) using the image histogram; we also expose interactive histogram functions for user-defined indexing of the stack, including fast nonlinear smoothing and bit-depth transformation for x-axis over [0,255];

**Maximum projection** over z-axis (Figure 2(b)) for global view as well as the creation of image mask [17], while the image enhancement function combines rescaling of the range of intensities with anisotropic image smoothing [15] to preserve borders and eliminate spurious gray level fluctuations that arise from XRT imaging;

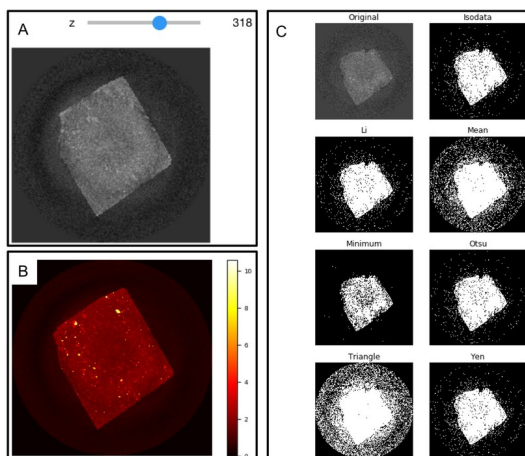


Figure 2. S1 sample exploration: (a) volume slicing, (b) maximum projection used for masking, (c) binarization schemes applied to the top stack slice.

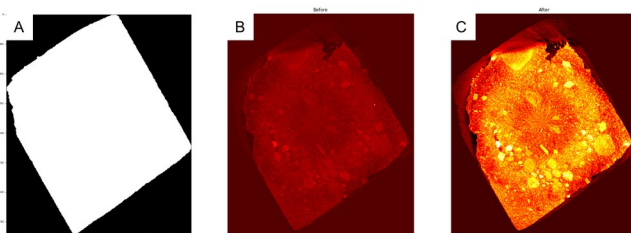


Figure 3. S1 image reduction and enhancement: a) Unsupervised mask calculation; b) cropped image with original intensities and c) enhanced image using edge-preserving bilateral filter.



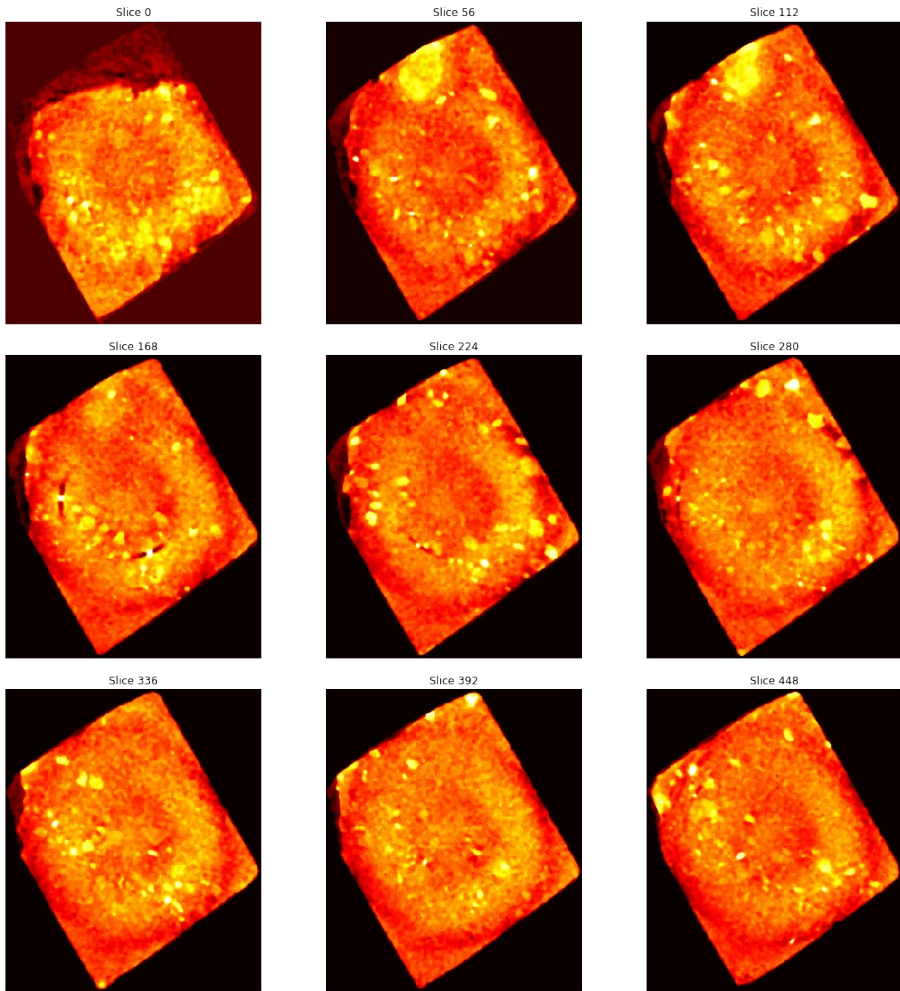


Figure 4. Image montage using equispaced stack slices after automated masking (S1).

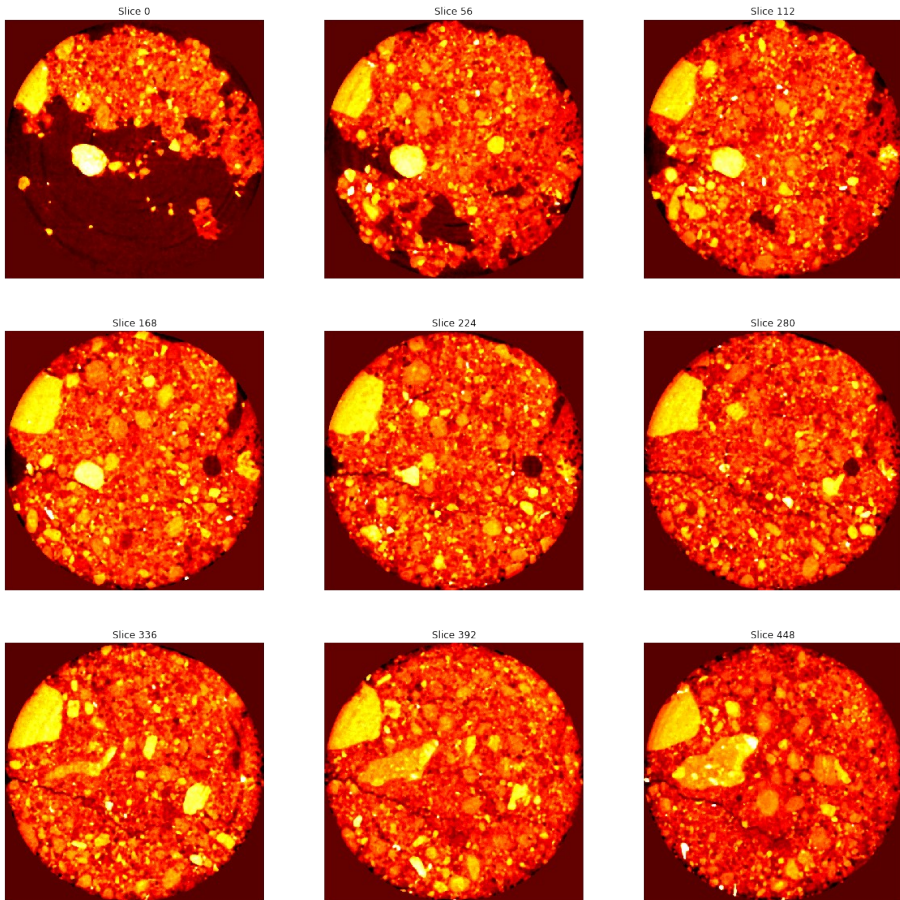


Figure 5. Image montage using equispaced stack slices after automated masking (S2).

### Discriminatory module

This step targets on defining the ROI based on the maximum projection to focus computations at the specimen, so the scripts will automatically calculate several image representations that will be key for materials phase segmentation. These transformations are:

**Masking:** lossless data reduction of the stack by using an unsupervised algorithm combining maximum projection (Figure 3) allied to isodata (Figure 2), which supports creation of a mask to quickly separate background from foreground across the full stack [17], as illustrated in Figure 4-5; it works for most of XRT samples imaged within cylindrical containers

**Supervoxel** using Simple Linear Iterative Clustering (SLIC) [18], an unsupervised machine learning algorithm that controls the XRT intensity variation and can efficiently decompose XRT slices into homogeneous regions that are likely to correspond to the mortar components. SLIC is based on a spatially localized version of the k-

means clustering, similarly to the mean shift. The clustering procedure begins with an initialization step where  $k$  initial cluster centers  $C_i = \{x_i, y_i, g_i\}^T$ , for location  $x, y$  and gray level  $g$ , which are sampled on a regular grid spaced  $S$  pixels apart to produce roughly equally sized superpixels, for  $S = \sqrt{N}/k$ , for a target image of  $N$  pixels.

**Vesselness** using Sato et. al [19] emphasizes tubular structures, therefore it enhances the interfaces between different components of the materials by combining the second order differential quantity with a filter function with asymmetrical characteristics and dependent on two eigenvalues. Figure 6 (d) illustrates a vesselness feature map for sample S1.

**Texture descriptors** [20] or local binary patterns are efficiently calculate texture descriptors that consider a user-defined neighborhood around each pixel. Here, we use radius 10, which is approximately 15  $\mu\text{m}$  or the diameter of an aggregate; parametrization variation might lead to largely different results.

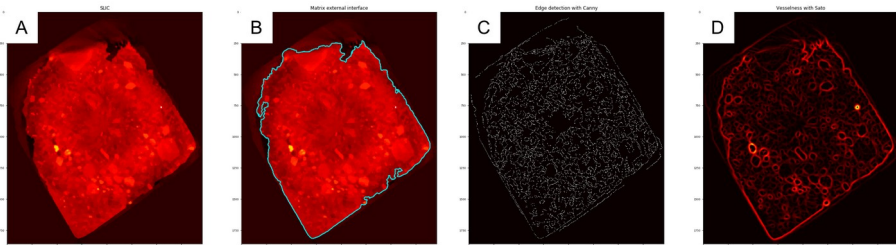


Figure 6. From controlled heterogeneity to feature maps: (a) superpixel with SLIC [17], (b) ROI adjustments based on SLIC, (c) Canny edge detection [14], (d) vesselness using Sato [18] for local estimation of components interfaces (S1).

## Explanatory module

High classification accuracy obtained through the creation of ensemble of trees [21-24] has motivated the use of random forest as one of the main supervised machine learning strategies, which here is exploited to segment XRT data. This technique generates and selects feature vectors that govern the growth of each tree in the ensemble, and it is known for preventing data overfitting even with relatively small datasets.

This paper describes machine learning experiments using two methods from different software stacks:

(a) the *FastRandomForest* java-based algorithm (**method A**) available in the trainable Weka segmentation plugin within Fiji for 2D image segmentation, dependent on feature vectors extracted from: Gaussian blur, Hessian, membrane projections, Sobel filter, difference of gaussians, membrane projection, variance, mean, maximum, minimum, median, anisotropic diffusion, bilateral, entropy and neighbors filters, applied to each slice. For the model training, batch size, number of trees and number of threads were set as 100, 200 and 12 respectively, and includes 10-fold cross-validation. Tables 2-3 show

the classification accuracy and Table 10 describes volume fraction estimates.

(b) a set of python-based classification algorithms (**method B**), including random forest (*RandomForestClassifier*), a support vector machine using radial basis functions (*SVC*) and a feedforward neural network (*MLPClassifier*), available in the *sklearn* package, which we adapted for 2D image segmentation, here designed to depend on feature vectors derived from image transformations [25], such as median, SLIC, vesseness and texture. Tables 4-9 summarize the classification accuracy for these methods.

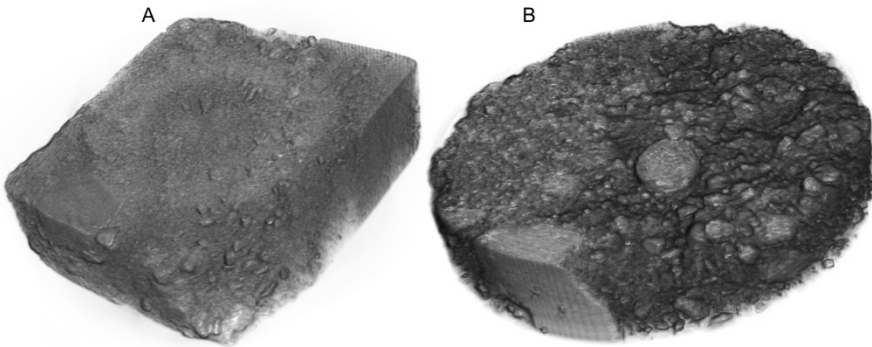


Figure 7. Volume rendering of XRT stacks of cementitious specimens (a) S1 and (b) S2 before classification of phase components.

Tables 2-9 show the confusion matrices that highlights the accuracy in pixel classification, and also allows the comparison of results using different random forest methods with inputs obtained from manually segmented samples.

Table 2. Confusion matrix for S1 using method A - random forest.

%	Background	Aggregate	Matrix	Pore
Background	99.39	0	0.04	0.57
Aggregate	0	99.66	0.34	0
Matrix	0.01	0.07	99.61	0.31
Pore	6.76	0.02	3.73	89.49

Table 3. Confusion matrix for S2 using method A - random forest.

%	Background	Aggregate	Matrix	Pore
Background	99.96	0	0	0.04
Aggregate	0	97.70	2.30	0
Matrix	0.01	0.37	96.14	3.48
Pore	0	0	3.50	96.50

Table 4. Confusion matrix for S1 using method B - random forest.

%	Background	Aggregate	Matrix	Pore
---	------------	-----------	--------	------

<b>Background</b>	100	0	0	0
<b>Aggregate</b>	0	97.97	1.74	0.29
<b>Matrix</b>	0	65.79	34.21	0
<b>Pore</b>	0	0	0	100

Table 5. Confusion matrix for S2 using method B - random forest.

%	<b>Background</b>	<b>Aggregate</b>	<b>Matrix</b>	<b>Pore</b>
<b>Background</b>	100	0	0	0
<b>Aggregate</b>	0	98.81	1.19	0
<b>Matrix</b>	0	54.17	41.67	4.16
<b>Pore</b>	0	0	2.78	97.22

Table 6. Confusion matrix for S1 using method B - SVM.

%	<b>Background</b>	<b>Aggregate</b>	<b>Matrix</b>	<b>Pore</b>
<b>Background</b>	100	0	0	0
<b>Aggregate</b>	0	99.71	0	0.29
<b>Matrix</b>	0	0	100	0
<b>Pore</b>	0	0	0	100

Table 7. Confusion matrix for S2 using method B - SVM.

%	<b>Background</b>	<b>Aggregate</b>	<b>Matrix</b>	<b>Pore</b>
<b>Background</b>	100	0	0	0
<b>Aggregate</b>	0	99.49	0.51	0
<b>Matrix</b>	0	72.92	20.83	6.25
<b>Pore</b>	0	0	0	100

Table 8. Confusion matrix for S1 using method B - neural net.

%	<b>Background</b>	<b>Aggregate</b>	<b>Matrix</b>	<b>Pore</b>
<b>Background</b>	100	0	0	0
<b>Aggregate</b>	0	98.56	0.56	0.28
<b>Matrix</b>	0	94.74	2.63	2.63
<b>Pore</b>	0	0	0	100

Table 9. Confusion matrix for S2 using method B - neural net.

%	<b>Background</b>	<b>Aggregate</b>	<b>Matrix</b>	<b>Pore</b>
<b>Background</b>	100	0	0	0
<b>Aggregate</b>	0	99.83	0	0.17
<b>Matrix</b>	0	89.58	0	10.41
<b>Pore</b>	0	0	0	100

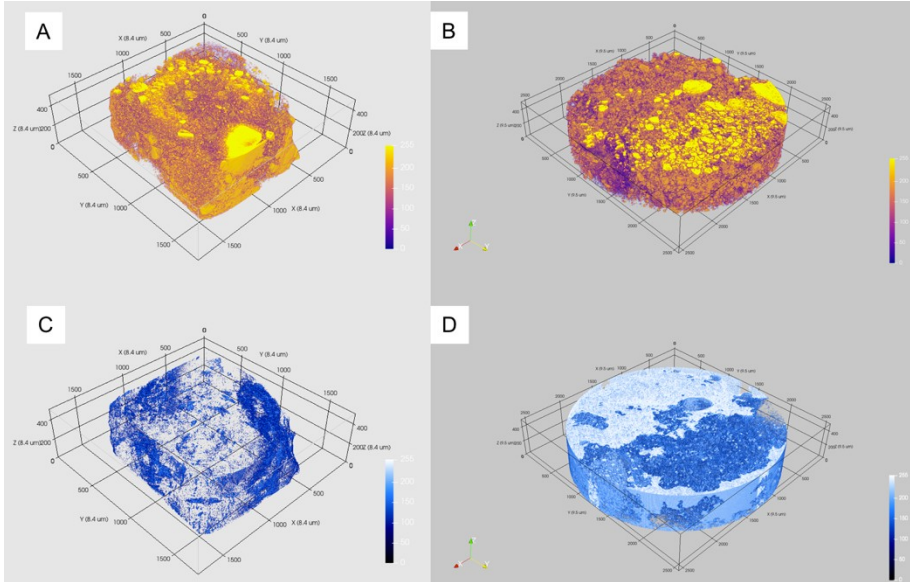


Figure 8. Phase segmentation estimates using method A on full stacks S1 (left) and S2 (right): (a,b) aggregates and (c,d) pore space, which includes cracks and remaining ringing artifacts.

Figure 7-8 show insightful visualizations which help to elucidate the roughness of the specimens as well as their heterogeneity [26]. While volume rendering provides valuable qualitative information, quantification of partial areas across slices helps to pinpoint potential segmentation errors. In order to audit the segmentation results for different phases, we calculate the cross-sectional pore dimensions, i.e., area variation across stack height as in Figure 9 (a) and (b). Notice that S1 shows large pore area variation, however within a small range while S2 has a sharp pore area variation due to large cavities as illustrated in Figure 7. Figure 9 (c-d) illustrates the pore size distribution as a surrogate measurement to access the porosity of the samples, suggesting that S1 presents smaller pores than S2, as expected.

Table 10. Phase segmentation estimates using method A on full stacks S1 and S2

%Volume	<b>Matrix</b>	<b>Aggregate</b>	<b>Pore</b>	<b>Background</b>
S1	35.80	15.88	0.82	47.5
S2	35.89	29.60	13.10	21.40

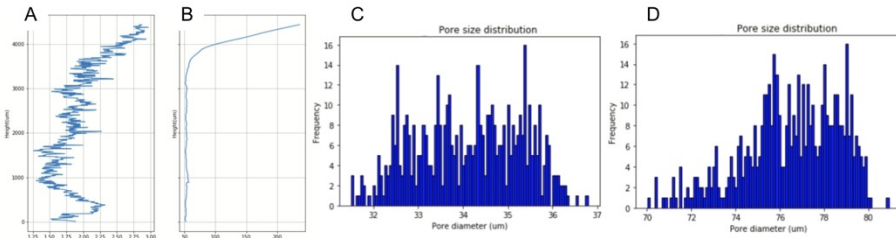


Figure 9. Cross-sectional pore statistics: area vs height for sample (a) S1 and (b) S2 and pore size distribution for (c) S1 and (d) S2.

## CONCLUSIONS

Concrete contributes to climate change in the world because the processes of cement manufacturing are energy intensive and dependent on high fuel consumption. Modern cement is a major cause of global warming and the third biggest industrial source of pollution. Alternative solutions based on research on ancient Roman cementing fabric may conceal answers for new advances in sustainable and improved more energy-efficient constructions, particularly regarding its microstructures and interplay between cement and aggregates.

This paper explained key methods for Materials Data Science to inspect XRT images of archaeological concrete samples, based on python scripts for investigation of microtomography data from observation, to discrimination of components up to explanation of different concrete phases, which is also compared with Weka segmentation. Most of the algorithms were designed in order to enable materials scientists to quickly explore their samples using Jupyter notebooks with classes for higher level command access, so that students can run code and reproduce the investigation on standard computational infrastructure, e.g., laptops, as codes explore virtual memory as opposed to RAM, or commodity clusters, such as Google Cloud.

In this preliminary investigation, we evaluated the classification results during segmentations based on the random forest, SVM and neural nets to estimate the volume of different phases (matrix, aggregate and pore) and audited results using spatial analysis. We noticed that S2 is more heterogeneous than S1, despite the fact that proposed methods lack ability to individualize aggregates, which are sintered to the matrix both in the physical space as well as in the XRT images. A few cracks were detected as part of the pore space, but segmentation is highly inaccurate in those areas due to the image resolution and cracks to be within 1 to 3 pixels width. While the accuracy of classification for the different phases varies considerably across the proposed methods, it is speculative to state that method A performs best since the human curation of ground-truth data is still in its early stages.

We illustrate an end-to-end and systematic characterization process during inspection of hierarchical materials, particularly offering

baseline methods for future insights about mechanical properties and durability. In order to achieve better characterization of the pore structure, further reconstruction algorithms should be considered as well as larger datasets including materials deformations under compressive and tensile forces. Future work will extend feature extraction algorithms to include viewpoint invariance and scale invariance, as well as more extensive tests using the current and new classification approaches. Also, we intend to explore graph-based tracking of objects to improve the characterization of the pore network.

## ACKNOWLEDGMENTS

This work was supported by the Office of Science, of the U.S. Department of Energy (DOE) under Contract No.DE-AC02-05CH11231. P.J.M.M. acknowledges the financial support given by the Roy W. Carlson Chair. We acknowledge the support of D. Parkinson from the Lawrence Berkeley National Laboratory Advanced Light Source, a DOE Office of Science User Facility.

## REFERENCES:

- [1] R. Andrew, Global CO<sub>2</sub> emissions from cement production, *Earth System Science Data*, 10, 195–217, (2018).
- [2] F. Williams, Concrete responsible for 8 percent of all CO<sub>2</sub> emissions, <https://www.architectsjournal.co.uk/news/concrete-responsible-for-8-per-cent-of-all-co2-emissions-says-report/10038404.article> (accessed 10 December 2019).
- [3] E. Mittelman, The Cement Industry, One of the World's Largest CO<sub>2</sub> Emitters, Pledges to Cut Greenhouse Gases, <https://e360.yale.edu/digest/the-cement-industry-one-of-the-worlds-largest-co2-emitters-pledges-to-cut-greenhouse-gases> (accessed 10 December 2019).
- [4] M. Jackson, C. Sejung, S. Mulcahy, C. Meral Akgul, R. Taylor, P. Li, A. Emwas, J. Moon, S. Yoon, G. Vola, H. Wenk, P. Monteiro, Unlocking the secrets of Al-tobermorite in Roman seawater concrete. *American Mineralogist*. 98. 1669-1687. 10.2138/am.2013.4484, (2013).
- [5] M. D. Jackson, S. R. Mulcahy, H. Chen, Y. Li, Q. Li, P. Cappelletti, H. Rudolf Wenk, *American Mineralogist* 102 (7): 1435-1450, (2017).
- [6] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press (2016).
- [7] D. M. Ushizima, H. A. Bale, E. W. Bethel, P. Ercius, B. Helms, H. Krishnan, L. T. Grinberg, M. Haranczyk, A. A. Macdowell, K. Odziomek, D. Y. Parkinson, T. Terciano, R. O. Ritchie, C. Yang. IDEAL: Images across domains, experiments, algorithms and learning. *The Journal of The Minerals, Metals & Materials Society*, pages 1-10, Sep (2016).
- [8] S. Liu, C. N. Melton, S. Venkatakrishnan, R. Pandolfi, G. Freychet, D. Kumar, H. Tang, A. Hexemer, D. M. Ushizima, Convolutional neural networks for grazing incidence x-ray scattering patterns: thin film structure identification, (9)2: 586-592, (2019).
- [9] H. A. Bale, A. Haboub, A. A. MacDowell, J. R. Nasiatka, D. Y. Parkinson, B. N. Cox, D. B. Marshall and R. O. Ritchie, *Nature Materials* 12: 40-46, (2013).
- [10] M. Willeminck and P. Noel, The evolution of image reconstruction for CT— from filtered back projection to artificial intelligence, *European Radiology*, 29:10, (2018).
- [11] P. J. M. Monteiro, C. Y. Pichot, and K. Belkebir. Computer tomography of reinforced concrete. In *Materials Science of Concrete*, American Ceramics Society, volume 5: pages 537-572, (1998).



- [12] C. Brandon, R. L. Hohlfelder, M. D. Jackson, and J. P. Oleson, *Building for Eternity: The History and Technology of Roman Concrete Engineering in the Sea*. Oxbow Books, Oxford, page 327, (2014).
- [13] Lawrence Berkeley National Laboratory Advanced Light Source Beamline 8.3.2. <http://microct.lbl.gov/> (accessed 10 December 2019).
- [14] D. Gürsoy, F. De Carlo, X. Xiao, and C. Jacobsen, Tomopy: a framework for the analysis of synchrotron tomographic data, *Journal of Synchrotron Radiation*, 21(5):1188–1193, (2014).
- [15] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Guillard and T. Yu, scikit-image: image processing in Python.” *PeerJ* vol. 2: e453, (2014).
- [16] E. Guillard, Segmentation of 3-D tomography images with Python and scikit-image, September (2015). <http://emmanuelle.github.io/segmentation-of-3-d-tomography-images-with-python-and-scikit-image.html> (accessed 10 December 2019).
- [17] S. Miramontes-Lizarraga, D. Ushizima, D. Parkinson, Evaluating fiber detection models using Neural Networks, *ISVC’19*, (2019).
- [18] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua and S. Süsstrunk, SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11): 2274-2282, (2012).
- [19] Y. Sato, S. Nakajima, H. Atsumi, T. Roller, G. Gerig, S. Yoshida, R. Kikinis, 3D multi-scale line filler for segmentation and visualization of curvilinear structures in medical images. In J. Troccaz, E. Grimson, and R. Mösges, eds., *Proc. CVRMEd-MRCAS’97*, LNCS, pages 213-222, (1997).
- [20] T. Ojala, M. Pietikäinen, and D. Harwood, Performance evaluation of texture measures with classification based on Kullback discrimination of distributions, *Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR 1994)*, vol. 1, pp. 582 - 585, (1994).
- [21] L. Breiman, Random Forests, *Machine Learning*, 45(1): 5-32, (2001).
- [22] I. Witten, E. Frank, and M. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3 edition, (2011).
- [23] N. Kumar, S. Khatri, Implementing WEKA for medical data classification and early disease prediction, *IEEE International Conference on Computational Intelligence & Communication Technology*, (2017).
- [24] K. Xu, P. Monteiro, D. Ushizima, Unveiling the secrets of Roman Concrete with Computer Vision, 2018 SSR/LCLS Users' Meeting, (2018).
- [25] D. Ushizima and A. de Siqueira, “scikit-image: 3D Image Processing”, *Image Analysis across Domains, ImageXD’2019*, Berkeley, CA. <https://github.com/imagexd/2019-tutorial-skimage>. (accessed 10 December 2019).
- [26] S. Brisard, M. Serdar, P. J. M. Monteiro, “Multiscale X-ray tomography of cementitious materials: A review”, *Cement and Concrete Research* 128 (2020).