Assessment of Segmentation-Induced Deviations of Porosity Metrics in Powder Bed Fusion Additively Manufactured Components

Peter W. Spaeth[1], Erik L. Frankforter, Samuel J. Hocker, and Joseph N. Zalameda
NASA Langley Research Center, Hampton, VA USA

ABSTRACT

Post processing X-ray computational tomography (CT) inspection data for additively manufactured (AM) components can induce deviations in defect quantification, affecting subsequent fatigue and failure predictions. To assess the influence and potential impact of segmentation-induced measurement deviations, this paper applies several segmentation techniques to X-ray CT data for powder bed fusion Ti-6Al-4V specimens exhibiting porosity conditions. X-ray CT reconstructions were segmented with varying techniques including Otsu's thresholding, random forest, k-nearest neighbors, and the multilayer perceptron. Metrics such as pore size and global porosity were compared for internal validity. Then, top-down X-ray CT measurements of surface-breaking porosity were compared to optical profilometry for cross-validation.

Keywords: Data fusion, X-ray computed tomography, image segmentation, supervised classification, surface profilometry, additive manufacturing, image registration

1. INTRODUCTION

The rapid advancement of metal additive manufacturing (AM) technologies is driven by its strong potential for creating complex, custom parts. However, AM exhibits a large parameter space in processing conditions, creating the potential for AM parts to present varying types of defects while potentially exhibiting high heterogeneity compared to conventionally manufactured metal parts. This variability presents challenges to nondestructive evaluation (NDE) capabilities in establishing quality and safety of AM parts sufficient for safety-critical aerospace applications [1-4]. Two-dimensional imaging techniques such as optical microscopy, surface profilometry, and electron backscatter diffraction (EBSD) provide the highest resolution for accurate part characterization, however they are restricted to 2D top-down measurements [1, 2]. Serial sectioning combined with 2D imaging has become the gold standard for full volumetric measurements in AM specimens [5-7]. However, the applicability of serial sectioning is limited by specimen size and the necessary destruction of the sample. X-Ray CT can provide nondestructive full volumetric assessments of AM defects, which has made it the de facto gold standard amongst AM NDE techniques [1-4, 8]. In practice, though, X-ray CT still presents challenges in porosity detection and quantification.

Various factors affect X-ray CT measurement uncertainty for AM defects including sample positioning, material, X-ray CT system settings, and reconstruction approach [8, 9]. Additional X-ray CT factors such as resolution can affect measurement metrics such as flaw sizing and porosity estimations; beyond resolution effects, X-ray CT measurement uncertainty further translates into deviations in X-ray CT porosity quantifications as compared to more accurate serial sectioning methods [3, 5, 6]. Porosity size estimates are also significantly influenced by choice of segmentation algorithm [1]. The accuracy of porosity measurements further influences probability of detection. For example, Sundar et al. found when comparing X-ray CT to serial sectioning, probability of detection (POD) for X-ray CT was only 50% for pores six times larger than the voxel resolution [2]. Thompson et al. showed in a round robin test of ten X-ray CT facilities that AM

[1]peter.w.spaeth@nasa.gov; phone +1 (757) 864-8622; fax +1 (757) 864-4914

porosity measurements can vary widely in practice due to conditions such as inspection configuration, choice of algorithms, and human factors. In round robin testing of a 10 mm laser powder bed fusion (LPBF) AM Ti-6Al-4V cube specimen, measured maximum pore size varied between facilities by a factor of approximately 4.5, and bulk porosity estimates varied by over a factor of approximately 7 [3].

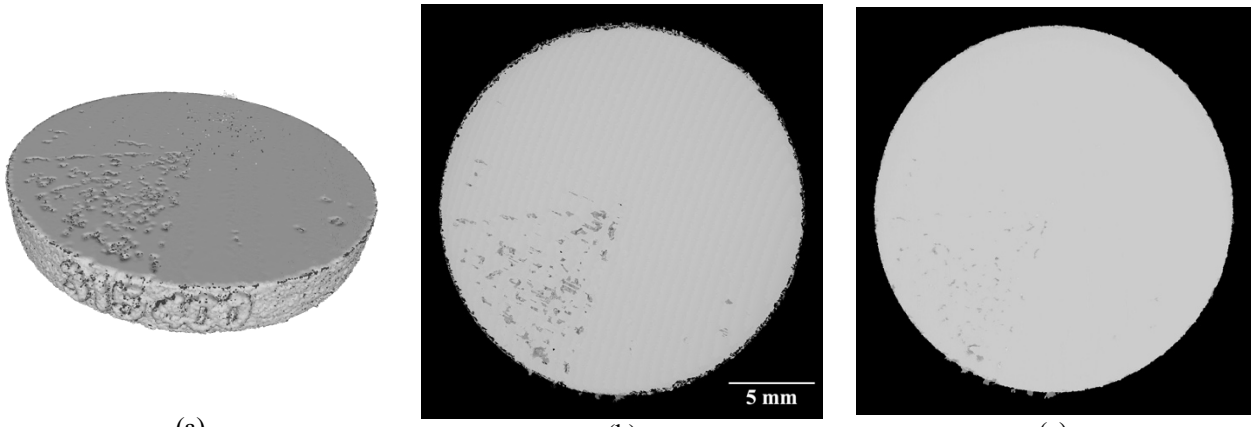
Together these factors indicate a strong need for improving accuracy and reliability of X-ray CT AM porosity measurements. To this end, X-ray CT images were compared to micrographs taken at interior surfaces of cut or polished AM specimens since 2D imaging approaches exhibit higher accuracy. The optical microscopy measurements tended to underpredict pore sizes, likely due to factors such as capturing the pore depth from a depth and an angle that's not along its longest axis [4]. Furthermore, these kinds of assessments are destructive in nature, and to the best of our knowledge no nondestructive techniques have been employed for baselining X-ray CT flaw measurements with more accurate 2D surface imaging.

There are two main goals of this paper. The first is to demonstrate a data-driven method for the selection of hyperparameters in the porosity segmentation pipeline starting with data normalization and feature selection, continuing through the selection of classifier parameters, and ending with a choice scoring function. The second goal is to demonstrate a method to nondestructively quantify the quality of the X-ray CT segmentation against surface profilometry data that measures the lateral size and depth of surface breaking pores.

2. METHODS

2.1 The test specimen

A disk-shaped Ti-6Al-4V titanium alloy test specimen (Figure 1) measuring 21 mm in diameter and 3 mm thick was built using LPBF AM under varying process conditions, which are detailed in [5]. The build was carried out with individual processing conditions across 16 equiangular sectors of the disk in an effort to produce distinct regions with distinct porosity characteristics. Prior to taking X-ray CT and profilometry measurements, the top surface of the disk was polished revealing surface breaking porosity concentrated in an angular section of approximately 67.5 degrees (Figure 1 (a)-(c)).



(a) (b) (c)

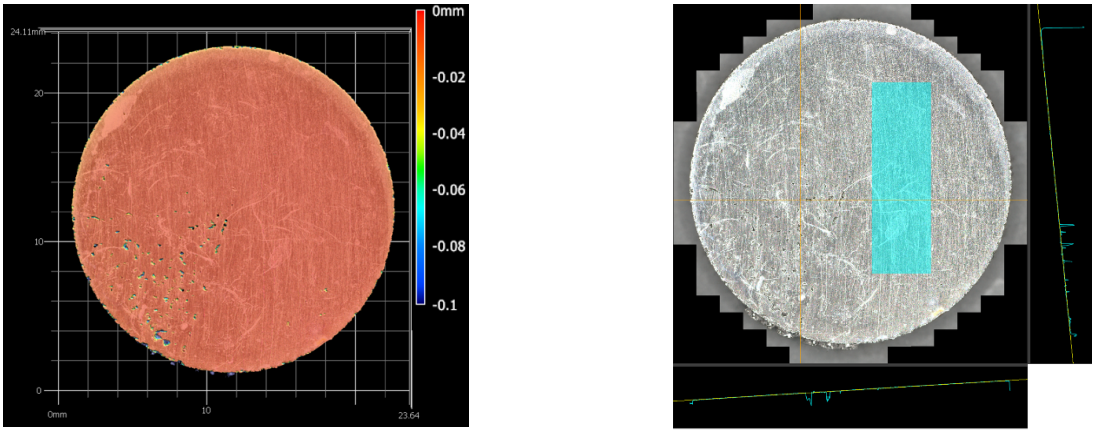
Figure 1. Volumetric representation of a Ti-6Al-4V disk-shaped AM specimen created using X-ray CT data that was used for analysis (a). An image derived from X-ray CT that represents the depth of surface breaking porosity (b). Estimate for depth of surface breaking porosity obtained using surface profilometry (c).

2.2 Surface profilometry measurements

Surface profilometry measurements were used to estimate the depth across the top surface of the disk (Figure 2). Measurements were taken with a Keyence® VR 6200[2], acquired at 160x magnification. The pixel size was 7.4 μm and the

[2] *Specific vendor and manufacturer names are explicitly mentioned only to accurately describe the test hardware. The use of vendor and manufacturer names does not imply an endorsement by the U.S. Government, nor does it imply that the specified equipment is the best available.*

vertical resolution was 10 µm. These surface profilometry scans provided a co-registered height image and a texture image across the surface of the disk. Following data acquisition, a reference plane was set in the Keyence® VR-6000 Analyzer software by specifying planar areas on the non-porous regions of the disk's top surface. This processing step rotates and translates the height data so the top surface has a z-axis normal, while applying the same rotation to the texture data to keep the datasets co-registered.



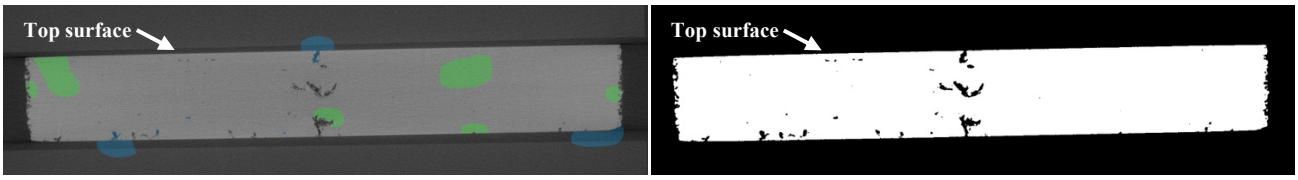
(a) (b)

Figure 2. Surface profilometry height image (a), and texture image showing reference plane alignment (b). The rectangular box on the disk image in (b) shows the area used for alignment; the line plots below and to the right in (b) show the height profile compared to the reference plane fit along the lines specified in the 2D image.

2.3 X-ray CT measurements and porosity segmentation

The disk specimen was interrogated using a Nikon® Metrology HMXST 225™ X-ray system that can resolve to 5 µm. The disk was imaged at an operating voltage of 180 kV, current intensity of 40 µA, and rotational angle of 0.11 degrees. The volumetric reconstruction of the disk was produced from 3,142 projections resulting in a 12.9 µm voxel resolution.

Three supervised binary classification algorithms (random forest, k-nearest neighbors, and the multilayer perceptron) were trained on hand-labeled images in order to estimate porosity in the disk. An example X-ray CT image slice with manual labels and an example segmentation are shown in Figure 3. Across the four training images, the training data contained approximately 1.2 percent of the total number of voxels in the X-ray CT volume, with approximately balanced classes (a 1.1 background-to-foreground voxel ratio).



(a) (b)

Figure 3. Two-dimensional slice of X-ray CT volume of the AM disk orthogonal to the top surface including manual labels of component (shown in green) and background (shown in blue) voxels (a). Example segmentation of the X-ray CT image slice (b).

Image features that were used as inputs for classification were computed using scikit-image [6]. The features included image intensity, averaged local gradients, eigenvalues of the image Hessian, and the three largest eigenvalues of the image structure tensor. Features were computed for each image and its Gaussian blurring, which was computed with standard deviation of 16 pixels, ultimately producing an input feature vector of length 14 at every pixel.

Four training images were used in an 80-20 training-test split to compute, using scikit-learn [7], hyperparameters of the classification pipeline using n-fold cross-validation. During cross-validation, adjusted Rand index [8] was used to score the hyperparameters of the pipeline. The cross-validation process also included feature selection using k-best features

(SelectKBest with k = 2 or 4) and feature scaling (StandardScaler, MinMaxScaler, or RobustScaler). The list of all hyperparameters for each classifier are detailed in Table 1, and the parameters selected through cross validation are highlighted in bold. A fifth labeled image was withheld for additional validation.

Table 1. Modeling pipeline in Scikit-learn for porosity segmentation of the X-ray CT volume of the disk showing preprocessing steps (scaling and feature selection), hyperparameters for each classifier, and cross-validation scores. The best performing hyperparameters and preprocessing steps are shown in bold.

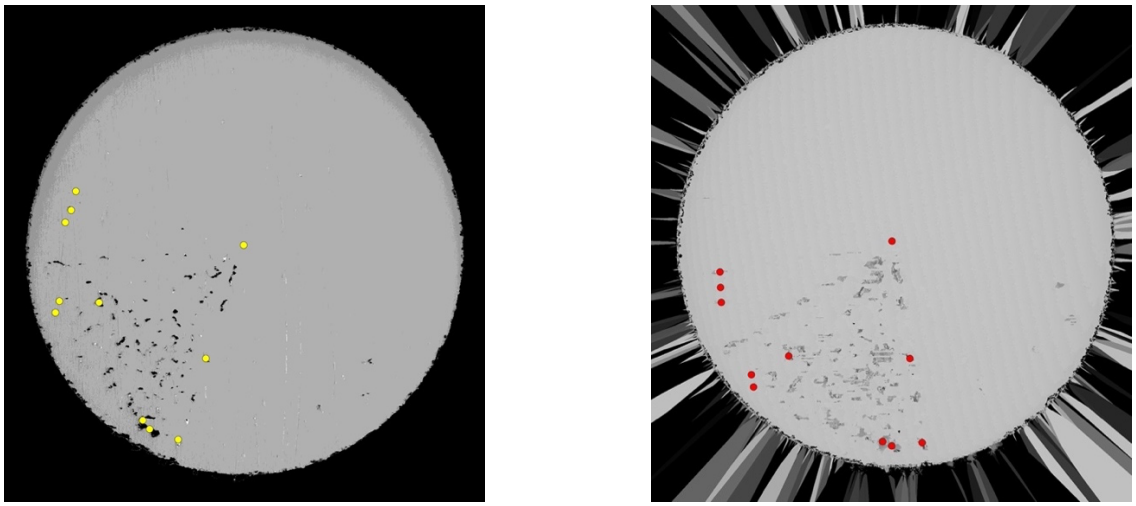
Classifier	Hyperparameters		Scaler	Feature Selector	Mean test score	StDev test score
k-Nearest neighbors	weights	n_neighbors	**MaxMin**, Standard, Robust	SelectKBest	0.9934	0.0010
	uniform, **distance**	**3**, 7, 11		k = 2, **4**		
Multilayer perceptron	hidden_layer_sizes		MaxMin, **Standard**, Robust	SelectKBest	0.9876	0.0007
	50, **100**			k = 2, **4**		
Random forest	max_depth	n_estimators	MaxMin, **Standard**, Robust	SelectKBest	0.9920	0.0011
	5, **10**	50, **100**		k = 2, **4**		

Ultimately, the best performing parameters were applied to segment the entire X-ray CT volume, image-by-image, across the direction parallel to the surface of the disk. For comparison, a fourth segmentation was obtained using Otsu's global thresholding method [9] applied across the entire X-ray CT volume.

2.4 Top-surface depth estimates derived from segmented volumes

In order to use the profilometry data to evaluate the quality of the X-ray CT segmentations, each segmented X-ray CT volume was used to create a depth estimate for surface breaking porosity in the form of a two-dimensional image. Shen et al. demonstrated methods for forming images from top-down projections of X-ray CT data to get a 2D slice for comparison with structured light system metrological measurements [10] (although to our knowledge the techniques has not been assessed for defect evaluation). First, voxels representing the overall surface of the disk were estimated using the Sobel filter [11] after filling interior pores in the binary segmentation with a three-dimensional morphological operation. Next, a plane was aligned using random sample consensus (RANSAC) [12] to the top-surface voxels of the disk, and the orthogonal distance from each top-surface voxel to the fitted plane was computed. Finally, for each pixel in the depth image, its grayscale value was defined to be the radial basis function (RBF) interpolation of the orthogonal distances to the plane for all surface voxels in a neighborhood of the image pixel after its registration to X-ray CT coordinates. In each case the depth was oriented by the outward facing normal to the top surface.

The profilometry depth image was rescaled using bicubic interpolation to match the dimensions of the depth images that were derived from the binary X-ray CT segmentations, and missing data in the foreground of the profilometry depth image was filled using biharmonic interpolation. Each depth image derived from a binary segmentation was registered to the rescaled profilometry depth image using a Euclidean transformation that was computed using RANSAC. Sets of corresponding points used for estimating the Euclidian registration were manually identified after first windowing all images to have pixel values between -500 μm and 150 μm from the top surface. Example registration data are shown in Figure 4.



(a) (b)

Figure 4. Profilometry-based depth estimate for surface breaking porosity of the disk together with manually labeled locations for image registration (shown in yellow) (a). Representative depth estimate derived from X-ray CT segmentation and manually labeled locations for image registration (shown in red) (b).

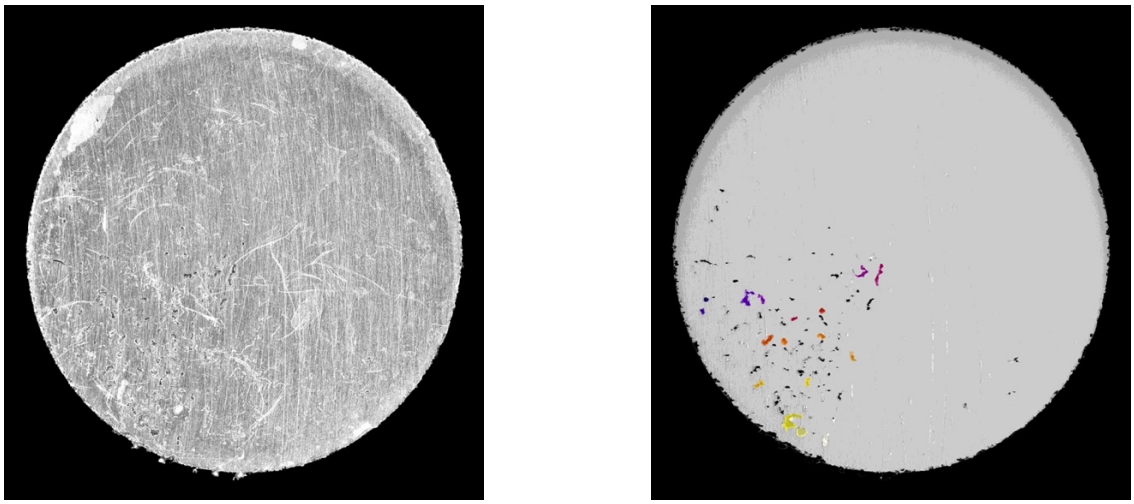
3. RESULTS

The first comparison of the segmentation methods was an internal check for consistency. For each segmentation, porosity volume fraction and equivalent spherical diameter (ESD) were estimated based on internal pores of equivalent spherical diameter larger than five voxels. The results of the porosity analysis in Table 2 show that Otsu's threshold produced the smallest total porosity fraction estimate and smallest estimate of mean ESD for the specimen, while all four methods computed a mean pore diameter to within 8 percent of the mean value across all methods. The total variation of mean ESD was 14 μm, which is similar to the voxel length, 13 μm, of the X-ray CT volume. Global porosity volume fraction among all methods ranged from 0.75% to 1.62%. On the other hand Otsu's thresholding produced an estimate for global porosity fraction that was smallest among all the segmentation methods.

Table 2. Pore statistics obtained from each of the four segmentation methods.

Segmentation method	Mean ESD (mm)	StDev ESD (mm)	Porosity volume fraction (%)
Otsu threshold	0.114	0.065	0.75
k-nearest neighbors	0.125	0.084	1.62
Multilayer perceptron	0.125	0.081	1.40
Random forest	0.128	0.083	1.49

Next the profilometry texture image was used to trace 18 surface breaking pores of varying depth, shape, and aspect ratio. The hand traced pores, shown in Figure 5, were used as domains on which to evaluate the registered depth images (and hence the segmentations from which they were derived) against the profilometry data.



(a) (b)

Figure 5. Texture image obtained from surface profilometry (a). Depth image obtained from profilometry and manually identified pores (shown in color) (b).

The difference in depth between the profilometry data and segmentation-derived estimate was measured at each pore, P_i, $i = 1, \ldots, 18$, using area-normalized mean absolute deviation. Mean absolute deviation (MAD) normalized by area at a pore P is given by

$$|f_1 - f_2|_P^{MAD} = \frac{\int_P |f_1(x,y) - f_2(x,y)|\, dx\, dy}{\int_P dx dy}, \tag{1}$$

where f_1 and f_2 represent the height functions derived from surface profilometry and X-ray CT segmentation, respectively. The mean absolute deviation across all 18 pores is shown as a function of pore aspect ratio in Figure 6. Aspect ratio measures the length of the major axis of the pore in the plane of the top surface divided by the maximum depth of the pore.

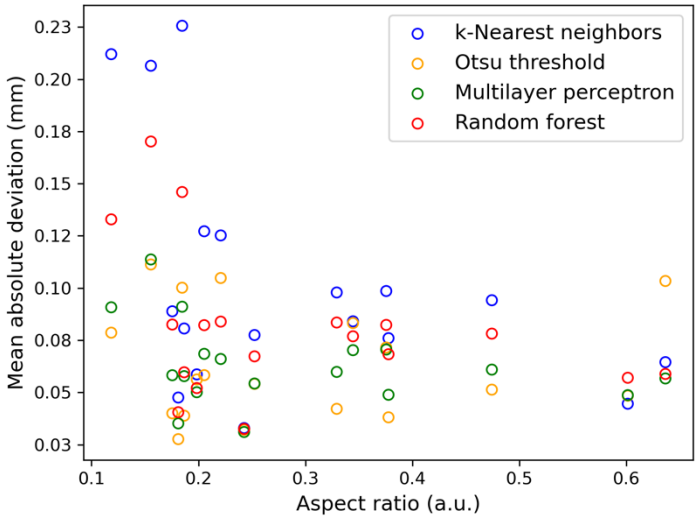


Figure 6. Mean absolute deviation of surface breaking porosity depth estimate versus aspect ratio for each manually identified pore, across all segmentation methods. The data show an increase in error and variation for low aspect ratio pores.

Finally, the (area-normalized) mean absolute deviation was averaged across all 18 pores, with the results summarized in Table 3. Under this metric, the multilayer perceptron classifier and Otsu thresholding segmentations produced the lowest errors.

Table 3. Average and median mean absolute deviation for the depth estimates across all manually labeled pores.

Segmentation method	Average mean absolute deviation (mm)	Median mean absolute deviation (mm)
Otsu threshold	0.063	0.055
k-nearest neighbors	0.102	0.086
Multilayer perceptron	0.063	0.059
Random forest	0.081	0.078

4. DISCUSSION

The results show that for the given training data, the neural network- and Otsu-derived depth estimates best agree with the data provided by surface profilometry. Misclassified voxels in the segmentations, errors derived from the plane fitting procedure, and errors arising from the registration of the depth images to the profilometry coordinates all contribute to the overall error in the comparisons.

However, even in the worst case, the median absolute deviation across all pores is on the same order of magnitude as the quantization error in the profilometry data itself. The analysis of the estimation errors on a pore-by-pore basis also shows that the errors and variability in pore depth estimates are largest for low-aspect ratio pores, i.e. pores that are shallow compared with their area at the surface of the disk.

The best performing classifier compared similarly to Otsu thresholding in estimating depth of surface breaking pores in the disk. To understand why, consider the grayscale values in the X-ray CT representation of the disk that are shown in Figure 7 corresponding to the full X-ray CT volume, and those voxels included in the training set. In both cases the first two peaks in the histogram arise from voxels corresponding to the background and acrylic holder used to secure the disk during acquisition of the X-ray CT data, and the third peak corresponds to voxels in the disk specimen. While Otsu's segmentation is able to successfully separate the specimen from the background it is worth noting that the specific threshold computed by Otsu's method (and hence the pores predicted by Otsu's segmentation) depends on the choice of data used to compute it. Had Otsu's threshold been computed on the training data the threshold would increase by nearly 4 percent resulting in a decrease in porosity, driving the results that were obtained from Otsu's method closer to those derived from the other classifiers.

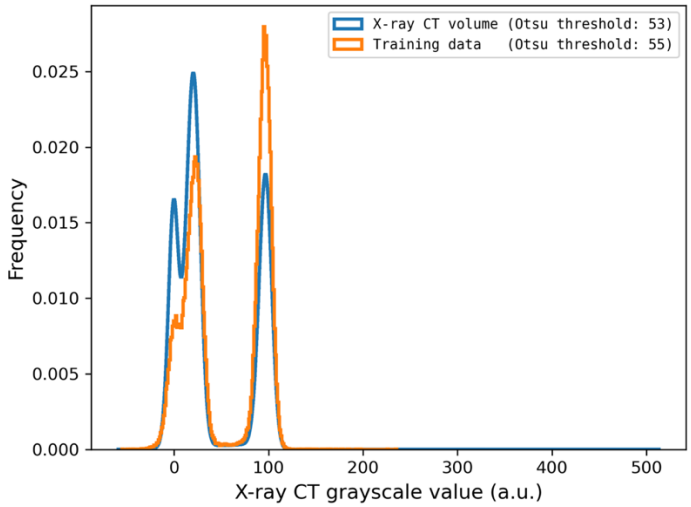
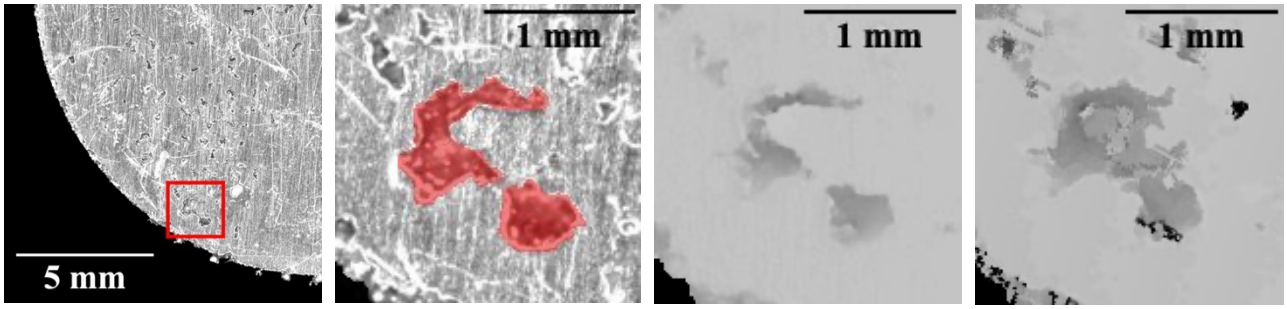


Figure 7. Distribution of X-ray CT grayscale values across the full volume (shown in blue) and across the training data set (shown in orange).

The number of classified voxels in the labeled data that was used to train the three classifiers accounted for approximately 1.2 percent of the overall voxel count in the X-ray CT volume. Although the classifiers were trained using class weighted scoring and sampling methods, despite the similarity of the distribution of grayscale values in the training set to the overall X-ray CT volume, it is possible that the existence of a class imbalance in the training data did impact the overall results. Uncertainty quantification could be applied to estimate the statistical quality of the overall results as a function of the imbalance of the training data through the application of Monte Carlo methods to the training of the individual classifiers.

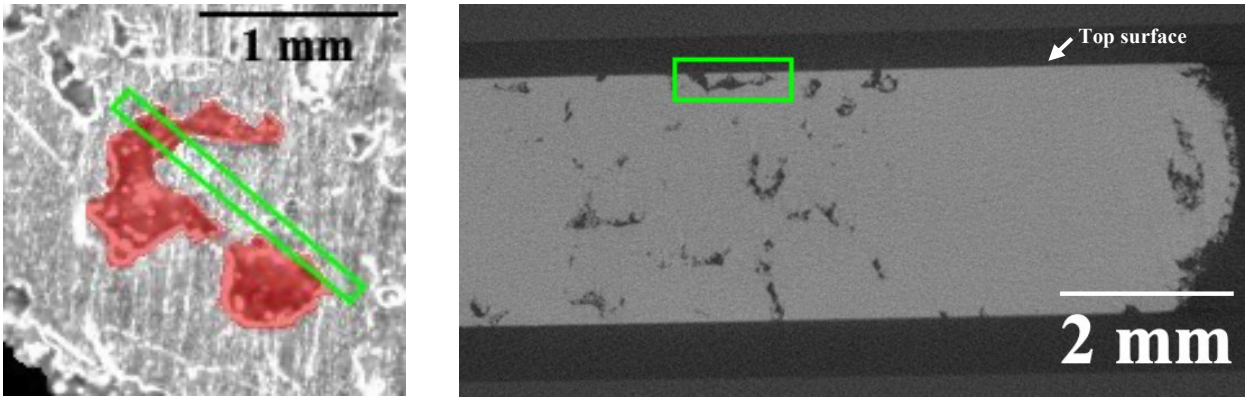
Finally, near surface porosity and its effect on the depth images derived from the X-ray CT segmentation was examined in the neighborhood of a surface breaking pore that is highlighted in Figure 8.



(a) (b) (c) (d)

Figure 8. Texture image obtained from surface profilometry (a) with region of interest (highlighted in red) showing a manually labeled pore in red (b). Depth image in the same region of interest obtained from surface profilometry (c) and estimated from X-ray CT segmentation that was computed with the multilayer perceptron (d).

The dark grayscale values in the depth estimate from the multilayer perceptron to the right of the manually labeled pore shows depth from the top surface that is not visible in the texture image Figure 8 (b) nor the depth measurement obtained using surface profilometry Figure 8 (c). The reason for this can be understood by examining the X-ray CT volume in the corresponding region that is highlighted in green in Figure 9.



(a) (b)

Figure 9. Texture image obtained from surface profilometry that shows a manually labeled pore (red) and region (green) selected for comparison with X-ray CT (a). X-ray CT reveals underlying near surface porosity (green region) that accounts for a difference in depth estimate obtained through profilometry and using X-ray CT segmentation.

This analysis suggests the possibility to extend the method into the component, for example to highlight pores that laterally expand beneath the top surface and cannot be seen by profilometry alone.

5. CONCLUSIONS

This paper demonstrates two capabilities. First, that modeling pipelines can be optimized to segment volumetric X-ray CT inspections of AM components in the context of supervised classification. This optimization is accomplished using training data that is manually derived from the X-ray CT inspection to drive the training and testing framework for hyperparameter selection.

The second capability that this paper demonstrates is the ability to use surface profilometry data to nondestructively evaluate the performance of different porosity segmentation algorithms. To make the comparison, an estimate of surface breaking porosity is first derived from the segmented X-ray CT data, subsequently registered to depth data derived from the profilometry, and finally various metrics of performance are computed. Comparison against the profilometry images ranks the performance of the segmentations and provides an extrinsic assessment of their performance. In this context the profilometry data can be subsequently used to further optimize the hyperparameters in the segmentation pipeline.

DATA AVAILABILITY

The data that support the findings of the study are available for download at the NASA Technical Reports Server [13].

ACKNOWLEDGMENTS

The authors would like to thank William Sommer of the Nondestructive Evaluation Sciences Branch (NASA Langley Research Center) for the X-ray CT measurements, and Brandon Widener of the Durability, Damage Tolerance, and Reliability Branch (NASA Langley Research Center) for the surface profilometry measurements. The authors also recognize the Transformational Tools and Technologies Project of the NASA Transformative Aeronautics Concepts Program for support of this work.

REFERENCES

[1] B. Jolley, M. Uchich, D. Sparkman, M. Chapman and E. Schwalbach, "Application of Serial Sectioning to Evaluate the Performance of x-ray Computed Tomography for Quantitative Porosity Measurements in Additively Manufactured Metals," *Advances in Multi-modal Characterization of Structural Materials,* vol. 73, no. 11, p. 3230–3239, 2021.

[2] V. Sundar, Z. Snow, J. Keist, G. Jones, R. Reed and E. Reutzel, "Flaw Identification in Additively Manufactured Parts Using X-ray Computed Tomography and Destructive Serial Sectioning," *Journal of Materials Engineering and Performance,* vol. 30, no. 7, p. 4958–4964, 2021.

[3] A. du Plessis, S. le Roux, J. Waller, P. Sperling, N. Achilles, A. Beerlink, J.-F. Métayerd, M. Sinico, G. Probst, W. Dewulf, F. Bittnerg, H.-J. Endres, M. Willnerh, Á. Drégelyi-Kiss, T. Zikmund, J. Laznovsky, J. Kaiserj, P. Pinterk, S. Dietrich, O. Fitzek, E. Lopez and P. Konrad, "Laboratory X-ray tomography for metal additive manufacturing: Round robin test," *Additive Manufacturing,* vol. 30, p. 100837, 2019.

[4] C. Chua, Y. Liu, R. Williams, C. Chua and S. Sing, "In-process and post-process strategies for part quality assessment in metal," *Journal of Manufacturing Systems,* vol. 73, pp. 75-105, 2024.

[5] S. J. A. Hocker, B. Richter, P. W. Spaeth, A. R. Kitahara, J. N. Zalameda and E. H. Glaessgen, "A point field driven approach to process metrics based on laser powder bed fusion additive manufacturing models and in situ process monitoring," *Journal of Materials Research,* vol. 38, no. 7, pp. 1866-1881, 2023.

[6] S. van der Walt, J.L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J.D. Warner, N. Yager, E. Gouillart, T. Yu, the scikit-image contributors, "Scikit-image: image processing in Python," *PeerJ,* vol. 2, no. e453, 2014.

[7] F. Pedregosa, et al., "Scikit-learn: Machine Learning in Python," *Journal Machine Learning Research,* vol. 12, pp. 2825-2830, 2011.

[8] W. Rand, "Objective criteria for the evaluation of clustering methods," *Journal American Statistical Association,* vol. 66, no. 336, pp. 846-850, 1971.

[9] N. Otsu, "A threshold selection method from gray-level histograms.," *IEEE Trans. Syst. Man Cybern.,* vol. 9, no. 1, pp. 62-66, 1979.

[10] W. Shen, X. Zhang, X. Jiang, L.-H. Yeh, Z. Zhang, Q. Lia, B. Lib and H. Qin, "Surface extraction from micro-computed tomography data for additive manufacturing," *Procedia Manufacturing,* vol. 53, pp. 568-575, 2021.

[11] I. Sobel and G. Feldman, *A 3x3 Isotropic Gradient Operator for Image Processing,* presented at the Stanford Artificial Intelligence (SAIL) Project, 1968.

[12] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Comm. ACM,* vol. 24, no. 6, pp. 381-395, 1981.

[13] P. Spaeth, E. Frankforter, S. Hocker and J. Zalameda, "Additive Manufacturing Disk Specimen Nondestructive Evaluation Dataset," NASA, 18 March 2024. [Online]. Available: https://ntrs.nasa.gov/citations/20240002767.

[14] C. Mandache, "Overview of non-destructive evaluation techniques for metal-based additive manufacturing," *Materials Science and Technology,* vol. 35, no. 9, p. 1007–1015, 2019.

[15] A. Thompson, I. Maskery and R. Leach, "X-ray computed tomography for additive manufacturing: a review," *Measurement Science and Technology,* vol. 27, no. 7, p. 072001, 2016.

[16] A. du Plessis, I. Yadroitsev, I. Yadroitsava and S. Le Roux, "X-Ray Microcomputed Tomography in Additive Manufacturing:," *3D Printing and Additive Manufacturing,* vol. 5, no. 3, pp. 227-247, 2018.

[17] D. Rowenhorst, L. Nguyen, A. Murphy-Leonard and R. Fonda , "Characterization of Microstructure in Additively Manufactured 316L using Automated Serial Sectioning," *Current Opinion in Solid State and Materials Science,* vol. 24, no. 3, p. 100819, 2020.

[18] M. Khosravani and T. Reinicke, "On the Use of X-ray Computed Tomography in Assessment of 3D-Printed Components," *Journal of Nondestructive Evaluation,* vol. 39, no. 75, pp. 1-17, 2020.

[19] F. Kim, A. Pintar, S. Moylan and E. Garboczi, "The Influence of X-Ray Computed Tomography Acquisition Parameters on Image Quality and Probability of Detection of Additive Manufacturing Defects," *Journal of manufacturing science and engineering,* vol. 141, no. 11, p. 111002, 2019.