Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

AI-Based 3D Pore Volume and Morphology Prediction

Raghav Yadav
Indian Institute of Technology Roorkee
Mayank Goswami
mayankfph@iitr.ac.in

Indian Institute of Technology Roorkee https://orcid.org/0000-0001-8829-9980

Method Article

Keywords: X-ray computed tomography, pore metrology, 3D convolution, synthetic benchmark, regression

Posted Date: November 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11729343/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

AI-Based 3D Pore Volume and Morphology Prediction

Raghav Yadavand Mayank Goswami [1, 2, #]
[1] Department of Physics, IIT Roorkee
[2] *Mehta Family School of Data Science and Artificial Intelligence*, IIT Roorkee,
Roorkee, Uttarakhand-247667, India.
[#] mayank.goswami@ph.iitr.ac.in

Abstract. X-ray computed tomography can quantify pores inside a material, but noisy and blurred pore boundaries make measurements sensitive to the segmentation procedure. This work evaluates direct regression of pore volume, surface area proxy, sphericity, and elongation from a centered synthetic grayscale crop. We generated 180 independent sphere, ellipsoid, and irregular pore geometries and rendered each under three degradation conditions, yielding 540 crops of size 24^3. All renderings of each geometry remained in one split. A compact three-dimensional convolutional neural network (CNN) was compared with Otsu thresholding and an Extra Trees regressor using segmentation and intensity features. On 27 held-out geometries at high degradation, volume mean absolute percentage error was 14.76% for the CNN, 67.29% for Otsu, and 3.74% for Extra Trees. Otsu recovered clean discrete volume exactly and had the lowest pooled elongation error. Paired and out-of-domain experiments reveal how size, nearby pores, unresolved features, and sampling affect the results. The contribution is a reproducible synthetic benchmark that identifies conditions under which learned estimation improves on thresholding and conditions where classical measurements remain stronger. Independent industrial scans are needed to assess transfer.

Keywords. X-ray computed tomography, pore metrology, 3D convolution, synthetic benchmark, regression.

1 Introduction

X-ray computed tomography (CT) reconstructs a volumetric representation of the attenuation inside an object. In additively manufactured parts, internal pores can be quantified without cutting the specimen. Micro-CT porosity protocols use voxel size, segmentation and measurement settings to define how pores are reported [1]. Process diagnosis can benefit from measuring both how large a pore is and how strongly it departs from an equiaxed shape. The estimates depend on the quality of the reconstructed image and on the definition of the object to be measured.

Blur mixes intensities near an interface. Noise can add dark voxels to a segmented region or disconnect parts of it. For a small pore, a few wrongly assigned voxels can account for a large fraction of its estimated volume. An estimate of area may be more sensitive still because small boundary irregularities add triangles to the reconstructed surface.

This project asks whether learned methods can estimate four descriptors of a candidate pore from a $24 \times 24 \times 24$ crop. A procedural benchmark holds the underlying geometry constant while image degradation changes. The same held-out inputs are supplied to a CNN, an intensity threshold followed by component measurement, and a tree ensemble. The comparison tests descriptor estimation with the pore center already supplied.

2 Related work and theory

Otsu's method chooses a threshold that maximizes the separation between two groups in an intensity histogram [2]. This gives a deterministic baseline for a dark pore in bright solid. Marching cubes converts a three-dimensional scalar field into a triangulated isosurface [3]. It supports surface measurement for both the clean reference mask and the threshold-derived component.

A 3D convolution combines intensity neighborhoods across depth, height, and width. Shared weights permit the same feature detector to act throughout the crop. This study uses the learned features for scalar regression. Extra Trees instead fits an ensemble of trees with randomized split candidates and averages their outputs [4]. Supplying intensity summaries alongside geometric features lets it learn corrections on degraded examples without training directly on all voxels.

Study boundary. Each main example contains one centered, largely isolated synthetic pore. The reference is a discrete mask and its derived measurements. The setup assesses performance against controlled references and intentionally separates this question from industrial validation, automated localization and scan-level porosity measurement.

3 Methods

3.1 Geometry and image generation

The generator samples 60 spheres, 60 ellipsoids and 60 irregular blobs using seed 42. It varies size and subvoxel center, applies random rotations and perturbs blob boundaries with a smoothed random field. The largest connected mask component is retained. All crops use unit isotropic spacing. For mask M, clean intensity is $1 - 0.8M$, with pore intensity 0.2 and matrix intensity 1. Rendering follows

$$X = \mathrm{clip}_{[0,1]}(G_\sigma * (1 - 0.8M) + \epsilon), \qquad (1)$$

where G_σ is a Gaussian blur and ϵ is independent Gaussian noise. The three (σ, SD) pairs are (0, 0), (0.6, 0.03) and (1.1, 0.10). Blur and noise vary together. The supplemental experiments use occupancy sampling where appropriate and image-domain Poisson-like plus Gaussian noise.

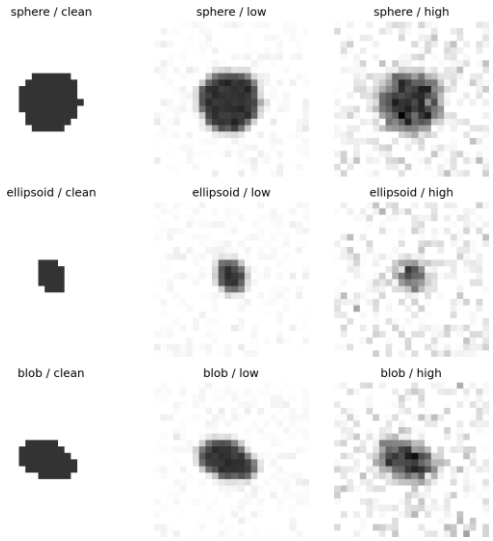


Figure 1: Central slices from the generated sphere, ellipsoid and blob families. Columns show clean, low and high degradation. The predictor receives the full volume.

3.2 Targets and split

Let N_M be the foreground count and h the isotropic spacing. Let A be the total area of the level-0.5 triangulation of the padded binary mask. With $\lambda_{\max}$ and $\lambda_{\min}$ denoting the extreme eigenvalues of the foreground-coordinate covariance, the target vector is

$$(V, A, \psi, E) = \left(N_M h^3, A, \frac{\pi^{1/3}(6V)^{2/3}}{A}, \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}\right). \qquad (2)$$

The implementation floors the eigenvalues used in the ratio at 10^{-8}. Volume and area use voxel^3 and voxel^2. The remaining targets are dimensionless. Discrete mask measurements can differ from continuous analytic ones. The split is by parent volume: 54 training geometries (81 patches) and 27 test geometries (81 patches). Each geometry contributes three renderings. The metadata field parent_volume_id identifies a geometry rather than a physical scan.

3.3 CNN and baseline fitting

The input transform is $x = 2X - 1$. The CNN applies padded 3^3 convolutions with channel counts 8, 16 and 32 and rectified linear activations. Average pooling halves each spatial dimension after the first two blocks. Global average pooling yields 32 features. A dense 32-unit hidden layer maps them to four standardized targets. The network has 18,740 trainable parameters. For target j, training estimates μ_j and s_j from the training labels. With $z_{ij} = (y_{ij} - \mu_j)/s_j$, the learning objective is

$$\mathrm{L} = \frac{1}{4n}\sum_{i=1}^{n}\sum_{j=1}^{4}(\hat{z}_{ij} - z_{ij})^2, \qquad \hat{y}_{ij} = s_j\hat{z}_{ij} + \mu_j. \quad (3)$$

The supplied run uses Adam at 10^{-3}, batch size 16, a 20-epoch cap and patience 5. Validation loss selects the checkpoint.

The classical baseline computes Otsu's threshold and labels the dark components. It measures the center component when present and otherwise the largest. Extra Trees uses these four descriptors, the patch mean and standard deviation, and $\sum_v \max(t - X_v, 0)$ for thresholds $t \in \{0.30, 0.45, 0.60, 0.75\}$. It fits 160 trees with seed 50. Validation standardized mean squared error selects minimum leaf size 3 from candidates 1, 3 and 6.

3.4 Evaluation

For reference y_i, prediction $\hat{y}_i$ and $e_i = y_i - \hat{y}_i$, we compute

$$\mathrm{MAE} = \frac{1}{n}\sum_i |e_i|, \quad \mathrm{MAPE} = \frac{100}{n}\sum_i \left|\frac{e_i}{y_i}\right|. \qquad (4)$$

We use $R^2 = 1 - \sum_i e_i^2 / \sum_i (y_i - \bar{y})^2$ when the denominator is positive. MAPE can be large for tiny references. For each noise level we treat 27 geometries as the paired sampling units. Agreement plots follow the difference-versus-mean approach of Bland and Altman [5].

lytic geometry.

Split assignment is stratified by shape family and grouped by geometry. The archive contains 126 training geometries (378 patches), 27 validation geometries

4 Experimental results

4.1 Training and checkpoint reproduction

The archived training history contains 16 epochs. Validation standardized MSE reached its minimum of 0.3760 at epoch 11. The last logged validation value is 0.3915. The archived checkpoint was evaluated with the original normalization and an independent NumPy implementation. The maximum absolute deviation from the archived CNN volume predictions was below 0.0002 voxel^3. Otsu predictions and the fitted Extra Trees comparison were also recomputed using the supplied split.

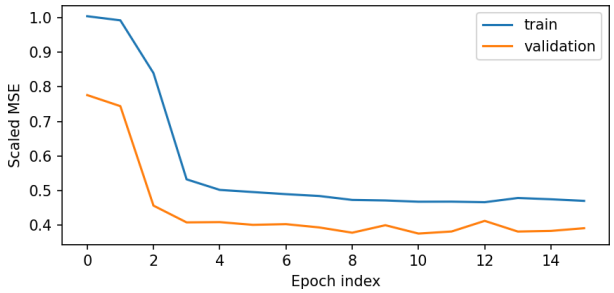


Figure 2: Archived training and validation standardized mean squared error. The exported chart uses zero-based epoch indices, so index 10 is epoch 11 in the training log.

4.2 Pooled and condition-specific errors

Table 1: Pooled test MAE on 81 patches from 27 geometries.

Target	CNN	Otsu	Extra Trees
Volume (voxel^3)	31.81	37.83	**9.84**
Area (voxel^2)	18.09	58.39	**6.01**
Sphericity	.0147	.0811	**.0084**
Elongation	.1898	**.0252**	.0360

The CNN attained pooled R^2 values of 0.947 for volume, 0.942 for surface area, 0.143 for sphericity and −0.143 for elongation. Its strong fit for size coexists with weak elongation estimates. Figure 3 shows the narrower spread of the CNN elongation outputs. Extra Trees had the lowest pooled MAE for volume, area and sphericity. Otsu led for elongation.

Table 2: Volume MAPE by degradation. Each row uses the same 27 held-out geometries.

Degradation	CNN	Otsu	Extra Trees
Clean	10.72%	**0.00%**	5.30%
Low	3.52%	**0.28%**	4.81%
High	14.76%	67.29%	**3.74%**

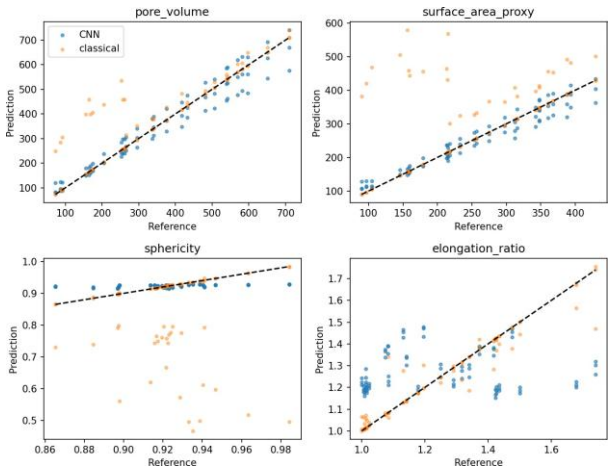


Figure 3: Archived CNN and Otsu estimates against mask references. The diagonal marks equality. Each geometry contributes three points.

Perfect intensity separation on clean synthetic crops lets Otsu recover the discrete volume exactly. The CNN's clean volume error exceeds the project's 5% target. Under high degradation, Otsu's threshold tends to admit excess dark voxels. Pooled Otsu signed volume bias was +37.31 voxel^3. The CNN reduced high-degradation volume MAPE by 52.53 percentage points. The 95% paired percentile interval for CNN minus Otsu was [−81.48, −25.77] percentage points based on 5,000 bootstrap resamples of geometry IDs. The interval describes this generator and fitted model. It does not quantify transfer to other instruments or materials.

4.3 Size and signed agreement

For high degradation, volume MAPE below equivalent radius 5 voxels was 14.13%, 89.28%, and 4.52% for CNN, Otsu, and Extra Trees respectively. At radius at least 5 voxels the errors were 16.56%, 4.47%, and 1.49%. These strata contain 20 and 7 geometries. The pooled Otsu degradation is therefore concentrated in the smaller group. An equal-volume sphere radius provides a common size scale even for non-spherical masks.

The high-degradation difference-versus-mean plots in Fig. 4 use one point per geometry. Signed residuals retain the direction of error that MAE discards. The descriptive limits summarize spread around the mean bias and do not define acceptance tolerances.

5 Robustness experiments

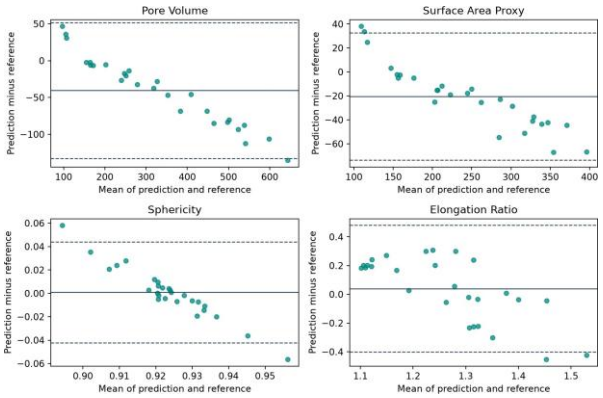


Figure 4: High-degradation CNN agreement on 27 geometries. Solid lines show mean signed difference. Dashed lines show mean difference plus or minus 1.96 sample standard deviations.

5.1 Neighboring and unresolved pores

The pair experiment placed a second sphere beside the designated central sphere and rendered the union. The reference descriptors came from the central sphere before combination. Twelve examples were evaluated for each relationship. Contact describes nominal contact in the generator. Discretization and blur influence whether the threshold-derived component is joined.

Table 3: Volume MAE (voxel^3) for 12 examples per stress condition. Both methods returned 12 measurements per row.

Condition	CNN	Otsu
Separated	38.53	9.08
Touching	91.46	255.08
Overlapping	120.47	228.33
Small sphere, $r = 1.2$	73.12	94.00
Subvoxel sphere, $r = 0.4$	79.62	55.48

Both methods give substantial errors in some cases. In overlap, ownership of the shared region is latent in the generator. The image provides a union, and an individual-pore interpretation requires an additional rule or prior. For tiny spheres, a fine-grid occupancy reference was downsampled to the 24^3 image. A finite result can be large relative to the actual feature.

5.2 Sampling at a fixed field of view

Twelve matched spheres were sampled at 32^3, 64^3, and 96^3 over a 24-unit physical field. Blur width was scaled in voxel coordinates to preserve physical width. Analytic sphere volume provided the reference. Native Otsu estimates were corrected for spacing. CNN inputs were trilinearly resampled to 24^3.

Table 4: Volume MAPE by acquisition sampling. The CNN always receives a 24^3 input.

Sampling grid	CNN	Native Otsu
32^3	17.10%	7.30%
64^3	13.75%	6.19%
96^3	14.80%	5.70%

The test shows the behavior of the fixed CNN under changes in acquisition sampling and resampling. It leaves native-resolution CNN retraining open.

5.3 Joint and single-target regression

The matched ablation trained one four-output network and four single-output networks with the same split, optimizer and epoch cap. Each selected its own best validation checkpoint. The joint and single-target volume MAEs were 40.54 and 31.80 voxel^3. Corresponding area MAEs were 21.89 and 15.27 voxel^2. Sphericity was 0.0146 for both to four decimal places. Elongation MAE was 0.1904 and 0.1871, respectively. The comparison uses one seed and a training run distinct from the main archived checkpoint.

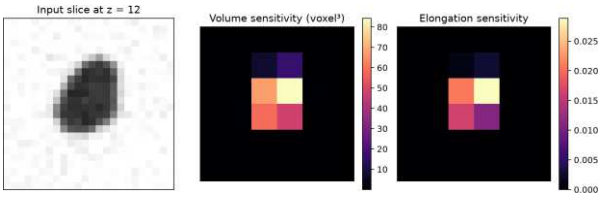


Figure 5: Sensitivity on one demo crop. A 4-voxel cube is replaced with matrix intensity, one position per run. The heatmaps show absolute output change in one depth slab.

6 Discussion and limitations

6.1 Why rankings change

The clean renderer supplies separated pore and matrix intensities, favoring a simple threshold. Under high degradation, blur and fluctuations perturb histogram and component boundaries. Direct regression can exploit distributed evidence across the crop and reduces the high-noise volume error relative to Otsu. The tree ensemble can also use intensity and region features together. Its advantage on several descriptors suggests that these engineered features carry useful information for this small, controlled dataset.

CNN elongation performance was weaker than that of the classical estimator. Global averaging may discard positional relationships helpful for axis ratios, and the training distribution may favor predictions near common shapes. These are hypotheses. The present ablation changes the output and training objective without isolating the effect of pooling. The stronger Otsu result shows that explicit geometric measurement can remain useful when the selected component preserves its principal axes.

6.2 Where estimates become unreliable

The dataset contains only 180 geometries and 27 independent test groups. The three rendered observations

per group support paired comparisons rather than an independent sample of 81 pores. Few examples occupy each size decile. Noise and blur rise together, so their individual causal effects cannot be identified. The main split and synthetic generator limit inference about other reconstruction settings.

The CNN is conditioned on centered crops with normalized intensity and unit isotropic voxel spacing. Scan-level deployment would require an upstream candidate detector, physical spacing metadata, a consistent measurement protocol and checks for ambiguous or truncated objects. Partial-volume effects and discretization can alter both predictions and target definitions. Volume, surface area and sphericity are physically related, while the CNN's separate outputs do not enforce consistency. Occlusion is a local sensitivity test on one crop and depends on the chosen block and replacement intensity.

The stress experiments show substantial failures for tiny and neighboring pores. Overlap poses an identification problem because several latent object configurations can explain a similar grayscale union. Real measurements would need an explicit definition of whether to report a connected cluster or attempt to split it. External validation should specify this choice before evaluating accuracy.

7 Conclusion and future scope

A compact CNN reproduced four reference descriptors from centered synthetic CT crops with performance that varied by target and degradation. The CNN improved high-degradation volume error relative to Otsu, while Otsu retained strong clean volume and pooled elongation accuracy. Extra Trees led pooled volume, area, and sphericity MAE and high-degradation volume MAPE. The synthetic benchmark therefore supports comparing methods condition by condition.

A follow-on study could use full parent volumes and grouped scanner-level splits, compare native-resolution networks at matched physical field of view and test independent industrial scans. A larger benchmark would support repeated-seed estimates, uncertainty by scan, and stronger conclusions about whether shape-aware pooling or physically consistent targets improve performance.

Code and data. The CWS package includes the original generator, training and evaluation code, checkpoint, separated dataset manifests, predictions, and a NumPy inference runner. For the demo, the runner estimates 396.7908 voxel^3 for a held-out synthetic example with a reference of 410 voxel^3. The report uses the archived training history and existing ablation artifacts. The main CNN predictions, Otsu, Extra Trees and diagnostic outputs were checked using the supplied data and checkpoint during preparation. No public repository or preprint identifier is assigned.

Video: https://youtu.be/QF0Hnk2fxYg

References

[1] A. du Plessis et al. Standard method for microCT-based additive manufacturing quality control 1: Porosity analysis. *MethodsX*, 5:1102–1110, 2018. doi:10.1016/j.mex.2018.09.005.

[2] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans. Syst., Man, Cybern.*, 9(1):62–66, 1979. doi:10.1109/TSMC.1979.4310076.

[3] W. E. Lorensen and H. E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. *ACM SIGGRAPH Comput. Graph.*, 21(4):163–169, 1987. doi:10.1145/37402.37422.

[4] P. Geurts, D. Ernst, and L. Wehenkel. Extremely randomized trees. *Machine Learning*, 63:3–42, 2006. doi:10.1007/s10994-006-6226-1.

[5] J. M. Bland and D. G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 327(8476):307–310, 1986. PMID 2868172.