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.

Segmentation-Free Estimation of Fibre, Matrix and Void Fractions from Simulated X-ray CT of Fibre Composites with a 3D Residual Network

Abhinav Swaran Singh
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, fibre- reinforced composites, void content, fibre volume fraction, deep learning, 3D convolutional neural network, synthetic data

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11690358/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.

Segmentation-Free Estimation of Fibre, Matrix and Void Fractions from Simulated X-ray CT of Fibre Composites with a 3D Residual Network

Abhinav Swaran Singh[1] and Mayank Goswami,[1,2]

[1]Department of Physics, IIT Roorkee
[2]MF School of AI and DS, IIT Roorkee
mayank.goswami@ph.iitr.ac.in

***Abstract*—Fibre volume fraction and void content govern the stiffness and interlaminar strength of fibre-reinforced polymers, but measuring them from X-ray computed tomography (CT) normally requires a voxel-wise segmentation that is fragile when carbon fibre and epoxy differ little in attenuation and when fibres are only a few voxels across. We instead regress the three volume fractions of a CT sub-volume directly with a 3D residual network (3.6 M parameters, softmax output so that the fractions sum to one). Training and test data are simulated: 66 synthetic composites with 40–66% fibre and 0–5% voids are generated on a supersampled grid, so every crop has exact ground truth, and are scanned by a physics-based CT simulator (parallel-beam projection, Poisson noise at three doses, detector blur, filtered back-projection). On 216 test crops from 9 held-out composites, the network's mean absolute error (MAE) is 0.84 percentage points (pp) [95% CI 0.66–1.05] for fibre and 0.17 pp [0.11–0.25] for void, against 6.06 pp and 0.28 pp for the best threshold- or histogram-based methods. The network is more accurate on all 9 test composites for fibre and matrix (Wilcoxon $p = 0.0039$, the minimum attainable with $n = 9$), and its fibre error rises only from 0.69 to 1.07 pp as the contrast-to-noise ratio falls from 5.2 to 1.2. For void it does not significantly beat a Gaussian-mixture histogram fit ($p = 0.36$) and over-estimates void below 1%. A mean-grey-value estimator calibrated on validation data comes close (1.09 vs 0.87 pp; $p = 0.20$), but a 2% calibration error raises its error to 8.4 pp, while the network, whose input is normalised per crop, is exactly unaffected.**

***Index Terms*—X-ray computed tomography, fibre-reinforced composites, void content, fibre volume fraction, deep learning, 3D convolutional neural network, synthetic data**

I. Introduction

The fibre volume fraction V_f and the void content V_v are the two composition numbers that most directly control the mechanical performance of fibre-reinforced polymers. Aerospace prepreg laminates are typically made to $V_f \approx 60\%$ for carbon and $\approx 50\%$ for E-glass [1]. Voids are unavoidable in processing, and around 1% is a common acceptance limit. Reported sensitivities include a loss of about 9% of interlaminar shear strength per 1% voids and of about 40% of compressive strength at 4% voids [2]; formation and effects are surveyed in [3].

The reference measurements are destructive. ASTM D3171 removes the matrix by acid digestion or burn-off to obtain fibre, matrix and void content [5], and ASTM D2734 gives void content from densities [6]. Both return a single bulk number and depend on accurate constituent densities; comparisons with micro-CT report errors of no less than $\pm 0.5\%$ for the destructive methods [7]. X-ray computed tomography (CT) is non-destructive and spatially resolved [4], but quantifying phases from it is hard for carbon/epoxy. Fibre and resin have similar chemistry and therefore similar X-ray attenuation, individual 7–8 µm fibres need voxels of about a third of their diameter, and the resulting void content depends on how the image is segmented [8]–[10].

We ask whether a neural network can read the phase fractions directly from the CT image, without first labelling every voxel, and how its accuracy compares with classical estimators under controlled noise. Our contributions are:

- a reproducible simulation pipeline that produces composite CT volumes with *exact* phase-fraction ground truth at three noise levels;
- a compact 3D residual network that predicts the fraction vector directly and is, by construction, invariant to the absolute grey-level scale of the scan;
- a comparison with threshold, histogram and calibrated mean-grey estimators, using a protocol designed against leakage and pseudo-replication (split by composite, cluster bootstrap, paired tests, three seeds, a shuffled-label control and an out-of-distribution set);
- an account of where the network does not win: void content below 1%, and a calibrated mean-grey estimator that comes close in this idealised simulator but fails under a small calibration error.

II. Related Work

Deep learning has been applied to composite CT mainly as *segmentation*: a network labels each voxel and fractions

Corresponding author: mayank.goswami@ph.iitr.ac.in

are then counted. Bertoldo *et al.* [13] segmented three-phase glass-fibre/polyamide CT with a modular U-Net. Tsamos *et al.* [14] and Friemann *et al.* [9] trained segmentation networks on synthetic, automatically labelled CT of complex composites; the latter reached 88% pixel agreement with manual labels on a real scan, and noted that classical segmentation of carbon fibre in epoxy is difficult because of the low contrast. Mehdikhani *et al.* [10] segmented voids in real carbon/epoxy CT with a Gaussian mixture model.

A second line of work regresses physical quantities directly from 3D images without an explicit label map. DeePore predicts 30 pore-scale properties from micro-CT [15], 3D CNNs estimate rock permeability [16], [17] and composite effective stiffness [18], and CNNs have regressed fibre volume content from 2D micrographs with 2.72% MAE [19]. Synthetic composite microstructures are commonly generated with random packing algorithms [20], [21]. We are not aware of work that regresses all three phase fractions jointly from 3D CT and compares the result, with paired statistics, against classical segmentation-and-counting.

III. Background

A. X-ray attenuation and fibre–matrix contrast

A monochromatic X-ray beam crossing a sample is attenuated according to the Beer–Lambert law,

$$I = I_0 \exp\Big(-\int_{\text{ray}} \mu(\mathbf{r})\, dl\Big), \qquad p = -\ln\frac{I}{I_0} = \int_{\text{ray}} \mu\, dl, \tag{1}$$

where $\mu = (\mu/\rho)\,\rho$ is the linear attenuation coefficient. At micro-CT energies, low-Z materials (C, H, O) attenuate mainly by Compton scattering, so carbon and polymer have nearly equal mass attenuation μ/ρ and their contrast comes from density. With NIST mass attenuation coefficients for carbon and Bakelite as an epoxy-like resin [11] and densities of 1.8 and $1.2\,\mathrm{g\,cm^{-3}}$ [1], the matrix/fibre ratio μ_m/μ_f is 0.79 at 15 keV and 0.72 at 40 keV. Voids (air) have $\mu \approx 0$.

B. Partial-volume effect

A voxel that straddles an interface records an attenuation that mixes the phases it contains: $\mu_{\text{voxel}} = \sum_k \phi_k \mu_k$, where ϕ_k are the voxel's sub-volume fractions. Two consequences drive this work. First, any method that assigns each voxel to a single phase misclassifies mixed voxels, and with fibres only a few voxels in radius a large share of voxels are mixed. Second, averaging over a whole crop gives

$$\bar{\mu} = \mu_f y_f + \mu_m y_m + \mu_v y_v, \qquad y_f + y_m + y_v = 1, \tag{2}$$

so in a linear, monochromatic system the mean grey value of a crop encodes its composition, provided the grey scale is calibrated in absolute units.

IV. Synthetic Data

A. Composite phantoms

Each phantom is a cube of $L = 128$ voxels built on a fine grid supersampled by a factor of 2 per axis. Fibres are cylinders parallel to z with radius drawn from $\mathcal{N}(r, (c\,r)^2)$, where the mean radius $r \in [3, 6]$ voxels and the coefficient of variation $c \in [0.02, 0.15]$ bracket measured carbon-fibre diameter scatter ($\approx 4.7\%$ [21]). A sinusoidal in-plane shift of up to 2 voxels, common to all fibres, adds global waviness without creating intersections.

Random sequential addition of discs has difficulty above about 50% area fraction [21], below the 60–66% needed. We therefore use a force-biased scheme [22]. The number of fibres needed for the target $V_f \sim U(0.40, 0.66)$ is placed at random in a domain padded by $1.5r$, so that fibres may cross the boundary as in a crop from a larger part. Overlapping pairs (centre distance below $r_i + r_j + 0.15$ voxels) are then pushed apart iteratively until none remain. Voids are spheres with log-normal radii (median 1.8 voxels, $\sigma_{\ln} = 0.45$, capped at 6) centred in matrix, which replace matrix voxels only, until the target $V_v \sim U(0, 0.05)$ is reached.

Ground truth is counted on the fine grid: for every target voxel we store how many of its 8 sub-voxels are fibre and void, so the fractions of *any* crop are exact sums. The attenuation map is the block average of the fine grid, which reproduces the partial-volume effect at every interface. We use $\mu_f = 1.00$, $\mu_m = 0.80$ and $\mu_v = 0$ (Section III-A): realistic, and slightly lower in contrast than typical laboratory energies.

B. CT simulation

Scans are simulated with the ASTRA Toolbox [23]–[25], slice by slice in parallel-beam geometry. To avoid edge streaks from an isolated cube, each slice is embedded in a cylindrical specimen 188 pixels across, filled with the phantom's mean attenuation, with air outside. We take 360 projections over π, above the $(\pi/2)N_{\text{det}} \approx 302$ required for the 192-pixel detector. Line integrals are scaled so that the longest path has $p \approx 2.5$. Detected counts are drawn as $N \sim \text{Poisson}(I_0 e^{-p})$ with $I_0 = 10^4$, 2×10^3 and 4×10^2 photons per pixel (high, medium and low dose), followed by Gaussian detector blur ($\sigma = 0.7$ pixels) and FBP with a Ram–Lak filter. Image quality is quantified by the fibre–matrix contrast-to-noise ratio

$$\text{CNR} = \frac{\bar{x}_f - \bar{x}_m}{\sqrt{\tfrac{1}{2}(\sigma_f^2 + \sigma_m^2)}}, \tag{3}$$

measured on eroded phase masks so that interface voxels do not count as noise. The mean CNR over the dataset is 5.2, 2.6 and 1.2 for the three doses. Their ratios (2.0 and 2.2) are close to the $\sqrt{5} \approx 2.24$ expected for Poisson noise under linear reconstruction, for which CNR $\propto \sqrt{I_0}$. Fig. 1 shows one slice through the pipeline.

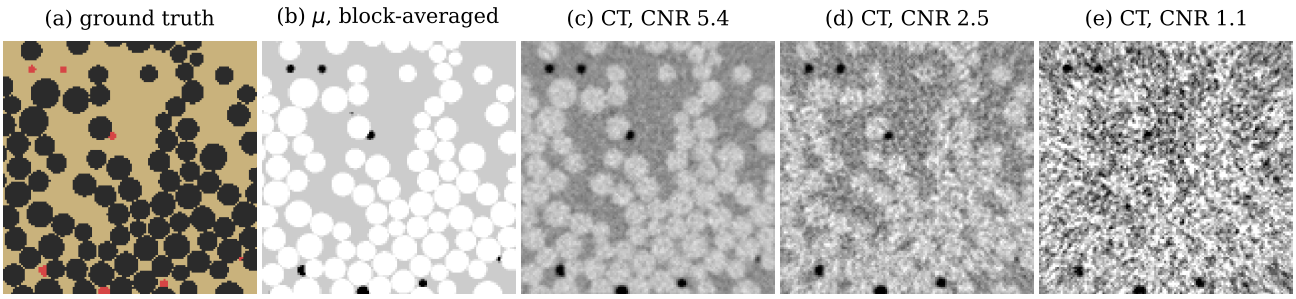


Fig. 1. One slice of a test phantom (fibre 51.3%, void 0.47%, mean fibre radius 5.7 voxels). (a) Fine-grid labels: matrix (tan), fibre (dark), void (red). (b) Block-averaged attenuation, showing partial-volume mixing at every interface. (c–e) Reconstructions at the three doses. Rendered from a CPU re-simulation of the same phantom; the GPU projector used for the dataset gives slightly higher CNR (5.7/2.7/1.2 for this parent).

C. Dataset

Sixty parent phantoms are drawn from one master seed and split 42/9/9 into training, validation and test *by parent*, so overlapping crops of one composite never appear on both sides of a split. Six further out-of-distribution (OOD) parents have fibre radius 6.5–7.5 voxels, outside the training range. Every parent is scanned at all three doses.

V. Methods

A. Network

The regressor is a 3D ResNet [12], [26]. A 3^3 convolutional stem with 32 channels is followed by four residual stages of 32, 64, 128 and 256 channels (one block each, stride-2 downsampling between stages), with instance normalisation [27] and leaky ReLU, global average pooling, and a 256→64→3 head whose softmax output $\hat{y}_k = e^{z_k}/\sum_j e^{z_j}$ gives non-negative fractions that sum to one. The model has 3.6 M parameters. Global average pooling suits the target: a volume fraction is itself an average over voxels, and pooling lets the network accept any cube size. Each input crop is z-scored with its own mean and standard deviation, so the network never sees the absolute grey scale and its prediction is exactly invariant to any gain or offset of the scanner.

B. Training

The loss is a weighted mean squared error,

$$\mathcal{L} = \tfrac{1}{3} \sum_{k\in\{f,m,v\}} w_k(\hat{y}_k - y_k)^2, \qquad w = (1,1,10), \qquad (4)$$

which up-weights void, whose values are an order of magnitude smaller. Each epoch draws 400 random 64^3 crops from the training parents with exact targets, augmented by random 90° rotations about the fibre axis and random flips, which leave the fractions unchanged. We use AdamW [28] (learning rate 10^{-3}, weight decay 10^{-4}) with cosine annealing [29], batch size 16 and 60 epochs (1,500 steps), with mixed-precision training [30] on an NVIDIA T4 GPU (about 10 s per epoch). The checkpoint with the lowest mean validation MAE is kept, and three training seeds are run.

C. Baselines

Multi-Otsu counting. Three-class multi-Otsu thresholding [31], [32] labels the darkest class void, the middle class matrix and the brightest fibre. Void components smaller than 4 voxels are reassigned to matrix as noise, and fractions are voxel counts. It is run on the raw reconstruction and, as the fairer practitioner-style variant, after Gaussian smoothing ($\sigma = 1$ voxel).

GMM histogram. A three-component, tied-variance Gaussian mixture is fitted by expectation–maximisation [33] to the crop's intensity histogram, with means initialised near 0 and at the 25th and 75th percentiles; the mixture weights are the fractions.

Mean-grey estimator. From (2) with our attenuation values, $y_f = 5\bar{\mu} - 4 + 4y_v$. We evaluate this with y_v from the GMM, both with the nominal constants and with (a, b, c) in $y_f = a + b\bar{\mu} + c\,\hat{y}_v$ fitted on the validation crops.

D. Evaluation protocol

Each test parent is tiled into 8 non-overlapping 64^3 crops at each dose, giving 216 crops from 9 parents, and every method is evaluated on the identical crops. We report MAE, RMSE, bias and R^2 in percentage points (pp). Crops from one parent are correlated, so 95% confidence intervals use a cluster bootstrap over parents (2,000 resamples) [34]. Methods are compared with a paired Wilcoxon signed-rank test on per-parent MAE; with $n = 9$ pairs the smallest attainable two-sided p is $2/2^9 = 0.0039$. Agreement is shown with Bland–Altman analysis [35], [36], and seed variability is reported separately from test-sample uncertainty [37]. As a leakage control [38], the network is also trained with targets shuffled within each batch, which must leave it no better than predicting the mean.

VI. Results

A. Accuracy and significance

Table I and Fig. 2 summarise the test set. The network estimates fibre to 0.84 pp ($R^2 = 0.988$), about seven times more accurately than the best classical estimator (smoothed multi-Otsu, 6.06 pp), with negligible

TABLE I
TEST-SET ACCURACY (216 CROPS OF 64^3 FROM 9 HELD-OUT COMPOSITES, ALL DOSES POOLED). MAE IN PERCENTAGE POINTS WITH CLUSTER-BOOTSTRAP 95% CI; BIAS = MEAN(PREDICTED − TRUE); R^2 AGAINST TRUTH. MATRIX RESULTS MIRROR FIBRE (NETWORK: MAE 0.84, $R^2 = 0.989$).

	Fibre fraction			Void fraction		
Method	MAE [95% CI]	bias	R^2	MAE [95% CI]	bias	R^2
3D ResNet (ours)	**0.84** [0.66, 1.05]	−0.13	0.988	**0.17** [0.11, 0.25]	+0.06	0.977
Multi-Otsu, raw	9.79 [7.24, 12.42]	−6.96	−0.745	5.37 [2.48, 8.76]	+5.37	−33.8
Multi-Otsu, smoothed (σ=1)	6.06 [4.72, 7.39]	+2.61	0.535	2.78 [0.50, 5.94]	+2.77	−19.2
GMM histogram	8.02 [6.26, 9.87]	+2.75	0.034	0.28 [0.13, 0.44]	−0.26	0.917

TABLE II
PAIRED WILCOXON SIGNED-RANK TESTS ON PER-COMPOSITE MAE, NETWORK VS BASELINE (9 TEST COMPOSITES). GAIN = MEDIAN MAE OF BASELINE MINUS NETWORK.

Phase	Baseline	Net better	Gain (pp)	p
fibre	Otsu, raw	9/9	+6.23	0.0039
fibre	Otsu, smoothed	9/9	+5.09	0.0039
fibre	GMM	9/9	+6.03	0.0039
matrix	Otsu, raw	9/9	+5.50	0.0039
matrix	Otsu, smoothed	9/9	+5.49	0.0039
matrix	GMM	9/9	+5.92	0.0039
void	Otsu, raw	9/9	+3.97	0.0039
void	Otsu, smoothed	9/9	+0.42	0.0039
void	GMM	5/9	+0.07	0.36

bias (−0.13 pp). Raw multi-Otsu under-counts fibre by 7 pp, mostly because noisy voxels are mislabelled as void (void over-count +5.4 pp). Smoothing and the histogram fit instead over-count fibre by about 2.7 pp, consistent with partial-volume voxels at fibre–matrix interfaces being assigned to the fibre class. For void the network reaches 0.17 pp. Both Otsu variants over-count void as dark noise specks become voids, while the GMM is competitive (0.28 pp) because voids form a separate intensity peak far from both solids.

Paired tests on per-parent MAE (Table II) confirm the fibre and matrix result on every test composite against every baseline ($p = 0.0039$, the floor for $n = 9$). For void, the network beats both Otsu variants on all 9 composites but the GMM on only 5 of 9 ($p = 0.36$). With nine composites the test has little power for an effect of this size, so the void comparison should be read as "no difference detected", not as equivalence.

B. Noise, void level and packing

Fig. 3 and Table III break the error down. As the CNR falls from 5.2 to 1.2, the network's fibre MAE rises only from 0.69 to 1.07 pp, while raw multi-Otsu degrades from 2.84 to 18.99 pp. By void level, the network is best in the 1–3% band (0.10 pp) and above 3% (0.20 vs 0.61 pp for GMM). Below 1% the GMM is better (0.06 vs 0.26 pp): the network's minimum prediction over the test set is 0.58% void, against a smallest true value of 0.17%, giving a +0.25 pp bias there (Fig. 4a). The network's fibre error is flat across packing density (0.78–0.91 pp), whereas the classical methods are worst for the densest composites (> 58% fibre), where interfaces dominate.

C. Comparison with a calibrated mean-grey estimator

Equation (2) makes the crop's mean grey value a strong reference in a linear simulator, and one the network cannot use since its input is z-scored. To evaluate it on identical inputs, we re-simulated the 9 validation and 9 test composites from the same master seed on CPU. Their ground truth matches the original dataset exactly (largest per-crop difference $< 10^{-4}$ pp); only the projector discretisation differs (mean CNR 4.9/2.5/1.1). On these crops the trained network's fibre MAE is 0.87 pp, against 0.84 pp on the original GPU reconstructions.

The calibrated mean-grey estimator comes close (Table IV): 1.09 vs 0.87 pp. The network is better on 7 of 9 test composites, but not significantly (Wilcoxon $p = 0.20$). By dose, the mean-grey estimator is better at medium dose (0.60 vs 0.79 pp) and worse at high dose (1.47 vs 0.70 pp). A plausible cause is the logarithm in $-\ln(N/I_0)$: under Poisson statistics it biases the mean line integral by an amount that depends on dose, so one pooled calibration cannot fit all doses. The fitted slope (485 vs nominal 500) also shows that the reconstructed mean is slightly biased.

The decisive difference is calibration. A 2% error in the absolute attenuation scale raises the mean-grey fibre error from 1.09 to 8.40 pp, and a 5% error to 21.40 pp. The network's predictions do not change at all, because per-crop normalisation makes it exactly invariant to gain and offset.

D. Stability, leakage control and generalisation

Across three training seeds the test MAE is 0.90 ± 0.06 pp for fibre, 0.86 ± 0.04 for matrix and 0.18 ± 0.02 for void (mean ± SD), an order of magnitude smaller than the gap to any threshold or histogram baseline. Trained on shuffled targets, the network's best validation fibre MAE was 8.88 pp, against 0.68 pp for the real model, as expected if no information leaks between splits (Fig. 4b). On the six OOD composites with larger fibres, fibre MAE rises to 1.28 pp, almost entirely a systematic under-estimate (bias −1.28 pp), while void MAE stays at 0.23 pp. Applied to 32^3 and 96^3 tiles without retraining, the 64^3-trained model gave fibre MAE 2.47 pp (GMM: 9.09) and 0.90 pp respectively; this is a scale-transfer test, not a crop-size study.

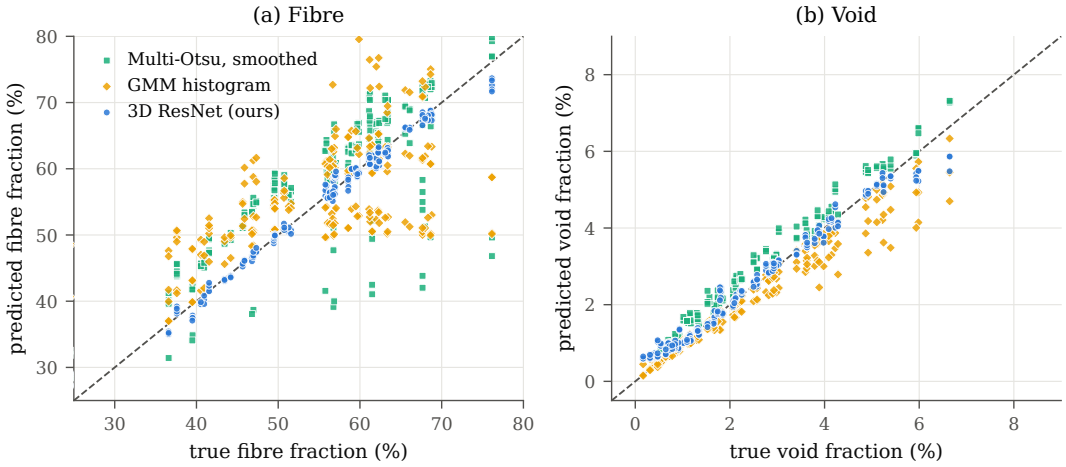


Fig. 2. Predicted versus true fractions for the 216 test crops. The network (blue circles) lies on the identity line for fibre; the classical methods scatter by several percentage points. For void, the network never predicts less than about 0.6%, so it over-estimates nearly void-free crops.

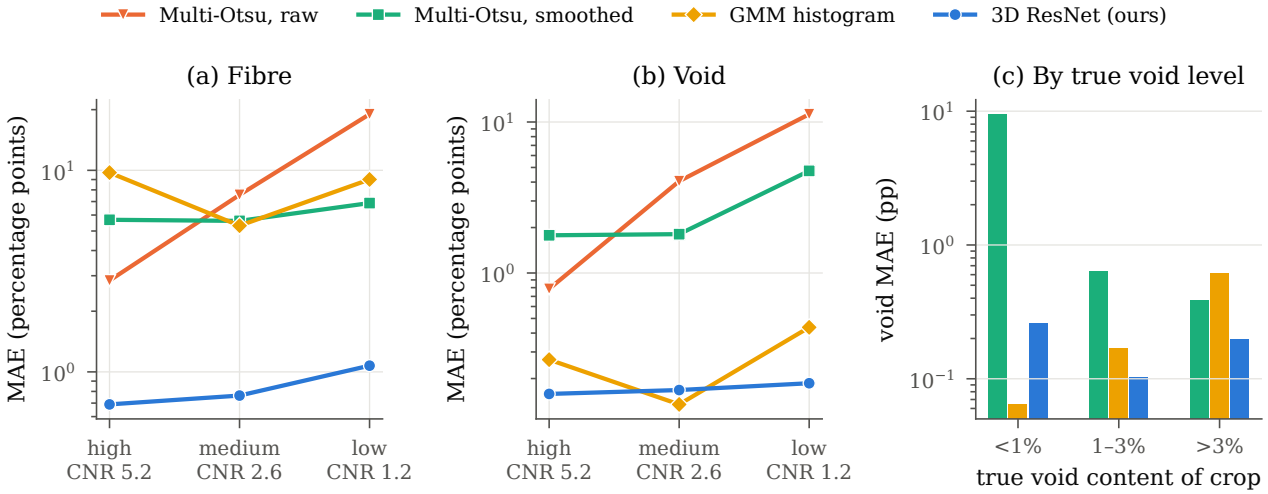


Fig. 3. (a, b) Test MAE against dose (log scale) for fibre and void. (c) Void MAE by true void content of the crop: the GMM histogram fit wins below 1%, the network above.

VII. Discussion

A. Interpretation

Fibre. With fibres 3–6 voxels in radius and a fibre–matrix contrast of 20%, a large share of voxels is mixed and per-voxel classification is dominated by noise. The network instead uses local structure (fibre cross-sections and their edges) to apportion mixed voxels. Its advantage is therefore largest for dense packing and low dose, where interfaces and noise dominate.

Void. Voids are separated from both solids by a large intensity gap, so a mixture fit recovers them well except when almost no void voxels are present. The network's floor near 0.6% is consistent with regression towards the bulk of the training distribution (void uniform on 0–5%) under a weighted squared loss.

Mean grey value. In a linear simulator composition is almost fully encoded in the mean, so a calibrated two-parameter formula is a near-oracle. That the network matches it without access to absolute intensities shows it has learned composition from structure rather than brightness. Real scanners drift in calibration and show beam hardening, which makes the grey-value mapping nonlinear and position-dependent, so calibration-free estimation should matter more on real data.

B. Limitations

Simulation. The beam is monochromatic, so beam hardening and cupping are absent, as are scatter, ring artefacts and phase contrast. The geometry is parallel-beam rather than cone-beam. Fibres are straight and parallel and voids are spheres; ply interfaces, resin-rich regions and fibre misalignment are not modelled. Real carbon/epoxy contrast can be lower still [8], [10].

TABLE III
ERROR BREAKDOWNS ON THE TEST SET. ALL VALUES IN PP; n = NUMBER OF 64^3 CROPS. THE NETWORK'S VOID BIAS IS SHOWN FOR THE VOID BANDS.

	n	Network	Otsu, raw	Otsu, smoothed	GMM	Network bias
Fibre / void MAE by dose (mean CNR)						
high (5.2)	72	0.69 / 0.16	2.84 / 0.78	5.68 / 1.78	9.74 / 0.27	
medium (2.6)	72	0.76 / 0.17	7.55 / 4.06	5.61 / 1.81	5.31 / 0.13	
low (1.2)	72	1.07 / 0.19	18.99 / 11.26	6.88 / 4.75	9.01 / 0.44	
Void MAE by true void content						
$< 1\%$	54	0.26	13.18	9.49	0.06	+0.25
1–3%	96	0.10	4.07	0.64	0.17	+0.05
$> 3\%$	66	0.20	0.86	0.39	0.61	−0.10
Fibre MAE by true fibre content						
$< 50\%$	72	0.91	6.32	6.19	6.80	
50–58%	48	0.87	8.82	5.57	6.15	
$> 58\%$	96	0.78	12.88	6.21	9.88	

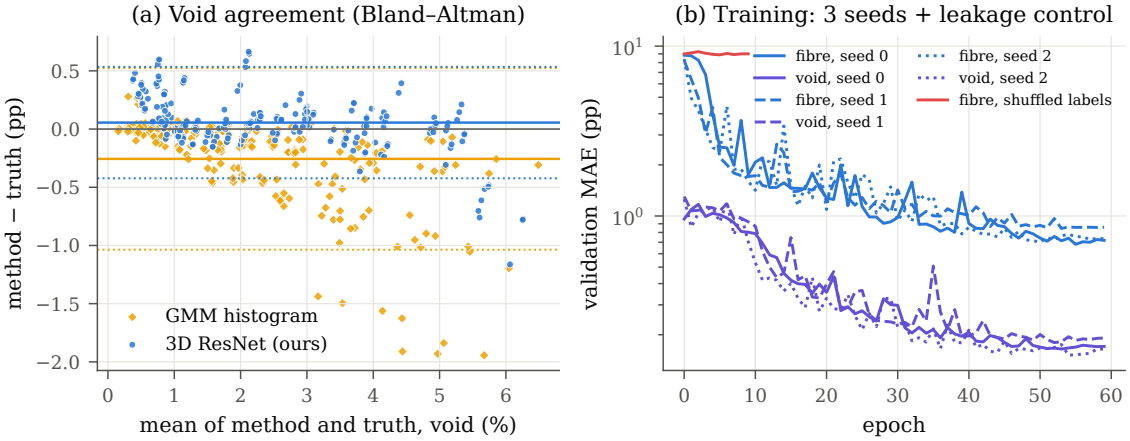


Fig. 4. (a) Bland–Altman agreement for void (solid: bias; dotted: 95% limits of agreement). Network: bias +0.06, limits $[-0.42, +0.53]$ pp; GMM: bias −0.26, limits $[-1.04, +0.52]$ pp. (b) Validation MAE during training for three seeds and for a control trained on shuffled targets, which stays at the predict-the-mean level.

TABLE IV
FIBRE MAE (PP) ON THE 216 RE-SIMULATED TEST CROPS. GAIN COLUMNS MULTIPLY THE RECONSTRUCTION BY 1.02 OR 1.05, EMULATING A 2% OR 5% ERROR IN THE ATTENUATION CALIBRATION.

Estimator	MAE	×1.02	×1.05
3D ResNet (ours)	**0.87**	**0.87**	**0.87**
Mean grey, calibrated, GMM void	1.09	8.40	21.40
Mean grey, nominal, GMM void	1.15	–	–
Mean grey, nominal, true void	1.38	–	–

Data. The network is trained on simulated data only, so a domain gap to real scans is expected [9], [14] and remains untested. Validation on real scans with independent references, such as registered multi-resolution CFRP scans [39] or the carbon/epoxy void dataset of [10], is the next step.

Statistics and scope. Nine test composites give tight intervals for the large fibre effect but cannot resolve small differences such as network versus GMM for void. The model was trained at one crop size (64^3), and it overestimates void below 1%; a dedicated void head, a loss on $\sqrt{y_v}$, or a hybrid that uses the GMM at low void are natural fixes.

Ethical use and bias. The model inherits the biases of its synthetic training distribution, as the systematic −1.28 pp fibre under-estimate on unseen fibre sizes shows. Composite parts are often safety-critical, so the network should support rather than replace qualified inspection, and should not be used for accept/reject decisions without validation against destructive references on the material in question. No personal data are involved, and the compute footprint is small (about 10 GPU-minutes per training run).

VIII. CONCLUSION

A 3D residual network trained only on simulated CT estimates the fibre and matrix fractions of composite subvolumes to within about 0.85 percentage points. That is roughly seven times more accurate than threshold- or histogram-based counting, and the advantage holds consistently across noise levels, packing densities and training seeds. For void content it matches, but does not significantly beat, a Gaussian-mixture histogram fit, and it over-

estimates nearly void-free material. A perfectly calibrated mean-grey estimator comes close (1.09 vs 0.87 pp, not significant) but degrades to 8.4 pp under a 2% calibration error, which leaves the network unchanged. The network's practical value is therefore accurate composition without absolute calibration; whether this holds on real scans is the main open question.

Code and Data Availability

The phantom generator, CT simulator, training and evaluation code, and a Google Colab notebook that regenerates the full dataset and every reported number from one master seed (2026) have been shared with the course instructor, who will make them public after review.

Video

https://youtu.be/8OfwDt2jwuk

References

[1] Hexcel Corporation, "HexPly prepreg technology," Publication FGU 017c, 2013.

[2] X. Liu and F. Chen, "A review of void formation and its effects on the mechanical performance of carbon fiber reinforced plastic," *Eng. Trans.*, vol. 64, no. 1, pp. 33–51, 2016.

[3] M. Mehdikhani, L. Gorbatikh, I. Verpoest, and S. V. Lomov, "Voids in fiber-reinforced polymer composites: A review on their formation, characteristics, and effects on mechanical performance," *J. Compos. Mater.*, vol. 53, no. 12, pp. 1579–1669, 2019, doi: 10.1177/0021998318772152.

[4] S. C. Garcea, Y. Wang, and P. J. Withers, "X-ray computed tomography of polymer composites," *Compos. Sci. Technol.*, vol. 156, pp. 305–319, 2018, doi: 10.1016/j.compscitech.2017.10.023.

[5] *Standard Test Methods for Constituent Content of Composite Materials*, ASTM D3171-22, ASTM International, 2022.

[6] *Standard Test Methods for Void Content of Reinforced Plastics*, ASTM D2734-23, ASTM International, 2023.

[7] M. Elkolali, L. P. Nogueira, P. O. Rønning, and A. Alcocer, "Void content determination of carbon fiber reinforced polymers: A comparison between destructive and non-destructive methods," *Polymers*, vol. 14, no. 6, Art. no. 1212, 2022, doi: 10.3390/polym14061212.

[8] A. Rashidi, T. Olfatbakhsh, B. Crawford, and A. S. Milani, "A review of current challenges and case study toward optimizing micro-computed X-ray tomography of carbon fabric composites," *Materials*, vol. 13, no. 16, Art. no. 3606, 2020, doi: 10.3390/ma13163606.

[9] J. Friemann, L. P. Mikkelsen, C. Oddy, and M. Fagerström, "Synthetic, automatically labelled training data for machine learning based X-ray CT image segmentation: Application to 3D-textile carbon fibre reinforced composites," *Compos. Part B Eng.*, vol. 305, Art. no. 112656, 2025.

[10] M. Mehdikhani, I. Straumit, L. Gorbatikh, and S. V. Lomov, "A dataset of void characteristics in multidirectional carbon fiber/epoxy composite laminates, obtained using X-ray micro-computed tomography," *Data Brief*, vol. 27, Art. no. 104686, 2019, doi: 10.1016/j.dib.2019.104686.

[11] J. H. Hubbell and S. M. Seltzer, "Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients," NIST Standard Reference Database 126. [Online]. Available: https://physics.nist.gov/PhysRefData/XrayMassCoef/

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 770–778.

[13] J. P. C. Bertoldo, E. Decencière, D. Ryckelynck, and H. Proudhon, "A modular U-Net for automated segmentation of X-ray tomography images in composite materials," *Front. Mater.*, vol. 8, Art. no. 761229, 2021, doi: 10.3389/fmats.2021.761229.

[14] A. Tsamos, S. Evsevleev, R. Fioresi, F. Faglioni, and G. Bruno, "Synthetic data generation for automatic segmentation of X-ray computed tomography reconstructions of complex microstructures," *J. Imaging*, vol. 9, no. 2, Art. no. 22, 2023, doi: 10.3390/jimaging9020022.

[15] A. Rabbani, M. Babaei, R. Shams, Y. D. Wang, and T. Chung, "DeePore: A deep learning workflow for rapid and comprehensive characterization of porous materials," *Adv. Water Resour.*, Art. no. 103787, 2020, doi: 10.1016/j.advwatres.2020.103787.

[16] S. Liu, R. Barati, and C. Zhang, "Fast estimation of permeability in sandstones by 3D convolutional neural networks," in *SEG Tech. Program Expanded Abstracts*, 2019, pp. 4833–4838, doi: 10.1190/segam2019-3216569.1.

[17] S. Gärttner, F. O. Alpak, A. Meier, N. Ray, and F. Frank, "Estimating permeability of 3D micro-CT images by physics-informed CNNs based on DNS," arXiv:2109.01818, 2021.

[18] C. Rao and Y. Liu, "Three-dimensional convolutional neural network (3D-CNN) for heterogeneous material homogenization," arXiv:2002.07600, 2020.

[19] F. Rothenhäusler, R. Q. Albuquerque, M. Sticher, C. Kuenneth, and H. Ruckdaeschel, "Application of convolutional neural networks and ensemble methods in the fiber volume content analysis of natural fiber composites," *Mach. Learn. Appl.*, vol. 19, Art. no. 100609, 2025.

[20] A. R. Melro, P. P. Camanho, and S. T. Pinho, "Generation of random distribution of fibres in long-fibre reinforced composites," *Compos. Sci. Technol.*, vol. 68, pp. 2092–2102, 2008, doi: 10.1016/j.compscitech.2008.03.013.

[21] Y. Ismail, D. Yang, and J. Ye, "Discrete element method for generating random fibre distributions in micromechanical models of fibre reinforced composite laminates," *Compos. Part B Eng.*, vol. 90, pp. 485–492, 2016, doi: 10.1016/j.compositesb.2016.01.037.

[22] J. Mościński, M. Bargieł, Z. A. Rycerz, and P. W. M. Jacobs, "The force-biased algorithm for the irregular close packing of equal hard spheres," *Mol. Simul.*, vol. 3, pp. 201–212, 1989, doi: 10.1080/08927028908031373.

[23] W. van Aarle *et al.*, "Fast and flexible X-ray tomography using the ASTRA Toolbox," *Opt. Express*, vol. 24, no. 22, pp. 25129–25147, 2016, doi: 10.1364/OE.24.025129.

[24] W. van Aarle *et al.*, "The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, vol. 157, pp. 35–47, 2015, doi: 10.1016/j.ultramic.2015.05.002.

[25] W. J. Palenstijn, K. J. Batenburg, and J. Sijbers, "Performance improvements for iterative electron tomography reconstruction using graphics processing units (GPUs)," *J. Struct. Biol.*, vol. 176, no. 2, pp. 250–253, 2011, doi: 10.1016/j.jsb.2011.07.017.

[26] K. Hara, H. Kataoka, and Y. Satoh, "Can spatiotemporal 3D CNNs retrace the history of 2D CNNs and ImageNet?" in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 6546–6555.

[27] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," arXiv:1607.08022, 2016.

[28] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2019.

[29] I. Loshchilov and F. Hutter, "SGDR: Stochastic gradient descent with warm restarts," arXiv:1608.03983, 2016.

[30] P. Micikevicius *et al.*, "Mixed precision training," arXiv:1710.03740, 2017.

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

[32] P.-S. Liao, T.-S. Chen, and P.-C. Chung, "A fast algorithm for multilevel thresholding," *J. Inf. Sci. Eng.*, vol. 17, no. 5, pp. 713–727, 2001.

[33] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. Ser. B*, vol. 39, no. 1, pp. 1–38, 1977.

[34] C. A. Field and A. H. Welsh, "Bootstrapping clustered data," *J. R. Stat. Soc. Ser. B*, vol. 69, no. 3, pp. 369–390, 2007, doi: 10.1111/j.1467-9868.2007.00593.x.

[35] J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 327, pp. 307–310, 1986.

[36] D. Giavarina, "Understanding Bland Altman analysis," *Biochem. Med.*, vol. 25, no. 2, pp. 141–151, 2015, doi: 10.11613/BM.2015.015.
[37] X. Bouthillier *et al.*, "Accounting for variance in machine learning benchmarks," in *Proc. Mach. Learn. Syst. (MLSys)*, 2021.
[38] S. Kapoor and A. Narayanan, "Leakage and the reproducibility crisis in machine-learning-based science," *Patterns*, 2023, arXiv:2207.07048.
[39] B. Boos *et al.*, "Multi-voxelsize X-ray computed tomography and optical microscopy image datasets of a thermoplastic CFRP tape," 4TU.ResearchData, 2025, doi: 10.4121/3a864c60-3023-45ab-a6c6-f36a23d67f56.