
Physics-aware Discrete Reparameterization with Symmetry-aware Bayesian Fusion for Multimodal Radiography

Nga T. T. Nguyen-Fotiadis
Computing and Artificial Intelligence
Los Alamos, National Laboratory
Los Alamos, NM 87545, USA
nga.nguyen@lanl.gov

Bradley T. Wolfe
Physics Division
Los Alamos, National Laboratory
Los Alamos, NM 87545, USA
bwolfe@lanl.gov

David P. Broughton
Nuclear Engineering & Nonproliferation
Los Alamos, National Laboratory
Los Alamos, NM 87545, USA
david.broughton@lanl.gov

Zhehui Wang
Physics Division
Los Alamos, National Laboratory
Los Alamos, NM 87545, USA
zwang@lanl.gov

Nathan A. DeBardeleben
High Performance Computing
Los Alamos, National Laboratory
Los Alamos, NM 87545, USA
ndebarde@lanl.gov

Earl C. Lawrence *
Computing and Artificial Intelligence
Los Alamos, National Laboratory
Los Alamos, NM 87545, USA
earl@lanl.gov

Abstract

Radiographic inspection must identify materials from X-ray and neutron transmissions; single-probe signatures are often ambiguous, whereas multi-modal fingerprints are complementary because photons and neutrons have distinct interaction physics. Data acquisitions (shots) contain *discontinuities* in density as voids, interfaces, and inclusions. Modeling density as a smooth field, via classical inversion or image-to-image regression, blurs those jumps and yields non-physical values. In practice, each voxel of density field is drawn from a small, known material set; thus many inverse problems that *appear* continuous are in fact discrete, and standard regression fatally assumes continuity where none exists. We perform a *physics-aware discrete reparameterization* of the inverse problem: dense per-voxel *discrete material classification* with calibrated uncertainty, rather than continuous regression. We introduce BayesFuse, a symmetry-aware Bayesian model that fuses five probes (neutrons 14 MeV and 2.5 MeV, plus three Bremsstrahlung-like exponential X-ray spectra) and encodes in-plane rotational/view symmetries. On a multimodal dataset with 64^3 voxel and 1D targets reparameterized to four materials (air, tungsten, copper, polystyrene; densities $\{0, 19.3, 8.96, 0.5\}$ g/cm³), BayesFuse attains near-perfect fidelity: mean Intersection over Union mIoU $\approx 98\text{--}99\%$ and best Expected Calibration Error ECE $\approx (3\text{--}5) \times 10^{-4}$, with uncertainty concentrated at interfaces and rising gracefully under corruptions. The result is a well-posed, physics-aware decision system fit-for-purpose in high-stakes, low-repetition experiments.

*Corresponding author.

1 Introduction: MultiModal Physics in Material Inspection

Radiographic inspection studies “what material is where” from line-integrated transmissions. Two facts collide: discontinuities (interfaces, voids) are the signal, yet smoothness priors blur them. Single probes are underdetermined, X-rays sense electrons ($\propto Z^n$), neutrons sense nuclei(1). Classical methods e.g. backprojection(2), Abel inversion(3), and deep regression(4; 5; 6) impose fatal smoothness: density jumps ($0 \rightarrow 19.3 \text{ g/cm}^3$) become non-physical intermediates (9.6 g/cm^3). With experimental shots costing \$200K+ and irreversible, calibrated uncertainty quantification (UQ) via Bayesian Neural Networks (BNNs) enables triage: auto-accept high-confidence voxels, flag interfaces for review. X-ray attenuation ($\mu \approx \rho Z^{3.8}/E^3$ photoelectric + $\rho Z/E$ Compton) and neutron cross-sections (non-monotonic with Z (1)) provide complementary signatures enabling material disambiguation impossible with single probes. In manufactured components and controlled experiments, materials

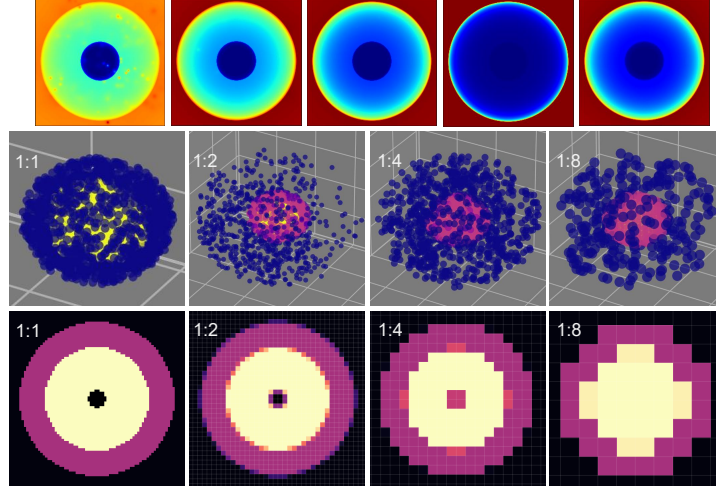


Figure 1: Top: five probe (modality) input images (left to right, X-ray exponential channels with energy 10 MeV, 2 MeV, and 1 MeV, neutrons 2.5 MeV and 14 MeV). Middle: Original 3D density target 256^3 , downsized 128^3 , 64^3 , and 32^3 voxel. Bottom: corresponding cross-section at plane $z = 0$ of Middle. A sample of the second target (1D density) is seen in Figure 2 (rightmost column).

are drawn from a small, known set: air (0), polystyrene (0.5), copper (8.96), or tungsten (19.3 g/cm^3). A voxel value like 4.5 g/cm^3 is not a “half copper” state but ambiguity at an interface. By recasting the task as *dense voxel-wise discrete material classification with calibrated uncertainty*, we align learning with physical reality: definitive decisions in homogeneous regions and principled uncertainty at boundaries. **BayesFuse** is a Bayesian, symmetry-aware instantiation of this principle, but critically, *the reparameterization itself is model-agnostic and represents the core innovation*. Unlike traditional semantic segmentation that groups regions, we perform *dense per-voxel material inference*: each of 64^3 voxels independently classified into discrete materials, fundamentally different from both continuous regression and coarse segmentation. This framework transforms an ill-posed inverse problem into $\sim 262\text{K}$ well-posed local decisions, each with calibrated uncertainty.

We demonstrate this principle through BayesFuse, advancing beyond efficient BNNs(7; 8; 9) via: **(i) Physics-aware design:** focused architecture with shallow *per-modality* encoders (6–8 layers) vs. deep WideResNets (28+ layers)(7; 8; 9; 10) preserve multi-scale features critical at interfaces, explicitly separating 2D modality processing from 3D reconstruction. **(ii) Symmetry-constrained Bayesian layers:** Rank-1 factorization with rotation-equivariant kernels achieves 99.55% rotation consistency (vs. 87.42% no-symmetry BNN). **(iii) Multi-scale fusion:** Attention-weighted combination across resolutions (32×32 , 64×64 , 3D) adapts to probe availability.

Related work. Classical inversion methods [Beer–Lambert(11), backprojection(2)] assume smoothness and blur discontinuities. Neural approaches, e.g., CNNs/U-Nets for tomography(12; 13; 14) and neutron/X-ray applications (tomography reconstruction(15), scattering/spectroscopy(16), and our previous work(17)). For example, Chu et al. generated(17) synthetic neutron and X-ray datasets on the same FTO platform used here, while Nguyen-Fotiadis et al. introduced foundational ViXReg for detail recovery at scales in fusion radiographs(18), demonstrated deep regression surpass generative neural networks of conditional GANs on ICF double-shell reconstructions(21). Complementary

advances in sparse tomographic fusion(19) and in learning shockwave dynamics from radiograph images(20) further illustrated the progression toward physics-aware, uncertainty-calibrated multimodal inference. In addition, toolbox accelerators like TIGRE(22) largely target single-output segmentation or image-to-image learning. To our knowledge, BayesFuse is the first to reparameterize multimodal radiographic inversion as *physics-aware discrete material inference* (voxel-wise classification) rather than continuous regression, and the first to enable *simultaneous* 1D and 3D reconstruction with calibrated uncertainty via symmetry-aware Bayesian fusion.

Our contribution: (1) A *physics-aware discrete reparameterization* transforming ill-posed continuous radiographic inversion into well-posed per-voxel decisions. (2) A *decision-theoretic* justification [Maximum A Posterior (MAP) under 0–1 loss and additive log-evidence across probes] explaining multimodal fusion advantages. (3) A *symmetry-aware Bayesian* instantiation performing empirical validation with high mIoU, low ECE, interface-localized uncertainty, and corruption robustness enhancing reproducibility.

2 BayesFuse: A Physics-integrated UQ Dense Classification

Data. We simulated a multi-probe dataset with $\sim 1\text{K}$ shots per modality that mimic experimental conditions. Simulations were performed using MCNP6.2 (Monte Carlo N-Particle version 6.2)(23; 24; 25; 26), with the “FIR” radiographic tally option. Runtime for *each source* was ≈ 24 hours on both the source-to-object and object-to-detector distances were set to 1 m, resulting in a magnification of 1. The simulated imaging grid measured $110\text{ cm} \times 110\text{ cm}$ and consisted of 256×256 pixels. Each shot has five 256×256 inputs: neutron 14MeV, neutron 2.5MeV, and three “exponential imaging” channels of energy 1 MeV, 2 MeV, 10 MeV. The neutron energies respectively correspond to those of Deuterium-Tritium and Deuterium-Deuterium fusion, and the X-ray spectra are representative of Bremsstrahlung sources with different incident electron energies. Example of these multi-modal (*continuous*) input is shown in Fig. 1 (top row). Targets are (i) a 256^3 voxel grid discretely parametrized with labels as in design $\mathcal{C} = \{0, 19.3, 8.96, 0.5\} \text{ g/cm}^3$ (air, tungsten, copper, polystyrene) and (ii) 1D lineout of length 1K encoded in $\{0, 1, 2, 3\}$ values. We downscale the 3D target to 64^3 for efficiency while preserving class topology (bottom rows in Fig. 1). We predict voxel-wise class logits for both 3D & 1D densities with calibrated UQ. We obtained 5K + 2K shots (inputs + targets) and split the entire data to train/validation/test set with the ratio 80/10/10. Results shown hereafter are on the test set.

Table 1: 3D voxel dense classification. Mean IoU (mIoU, %, \uparrow) / ECE (\downarrow) vs. #modalities.

Model	2 mod	3 mod	5 mod
BayesFuse no-symm	97.07 / 0.0006	98.56 / 0.00059	98.67 / 0.00037
Full BayesFuse	97.46 / 0.00053	98.92 / 0.00039	99.65 / 0.00028

Table 2: 1D dense classification. Mean IoU (mIoU, %, \uparrow) / ECE (\downarrow) vs. #modalities.

Model	2 mod	3 mod	5 mod
BayesFuse no-symm	95.30 / 0.013	97.51 / 0.0062	95.96 / 0.0057
Full BayesFuse	95.63 / 0.013	97.47 / 0.0056	97.48 / 0.0053

BayesFuse: Why Per-Pixel Classification Succeeds Where Regression Fails. Consider the forward radiographic model: $I_m = I_{0,m} \exp(-\int \mu_m(\rho(s))ds)$ with I_m being measured intensity, $I_{0,m}$ incident intensity, and $\mu_m(\rho)$ the mass attenuation coefficient for probe m at density $\rho(s)$ along ray path s . Regression seeks $\rho \in \mathbb{R}^+$, an infinite space where regularization, not physics, determines solutions. Dense voxel-wise classification over K materials $\{\rho_1, \dots, \rho_K\}$ provides: (i) **well-posedness** (finite solution space), (ii) **physical fidelity** (only real materials predicted), and (iii) **uncertainty localization** where $p(c_i|x) \approx p(c_j|x)$ at interfaces. With M distinct probes, log-evidence accumulates, $\log p(c|I_1, \dots, I_M) = \sum_m \log p(I_m|c) + \log p(c)$, making class margins improved with M (28). Critically, while the 3D target uses physical density values $\{0, 19.3, 8.96, 0.5\} \text{ g/cm}^3$, the 1D lineout is independently encoded as $\{0, 1, 2, 3\}$ without explicit correspondence. This dual-target design creates a self-supervised learning signal: BayesFuse must discover the material mapping between 3D densities and 1D encodings purely from data, learning robust material representations that generalize across both geometric projections and label spaces. BayesFuse implements this via: (i) 2D Bayesian encoders with rank-1 parameterization(9) for epistemic UQ, (ii) 3D reconstruction

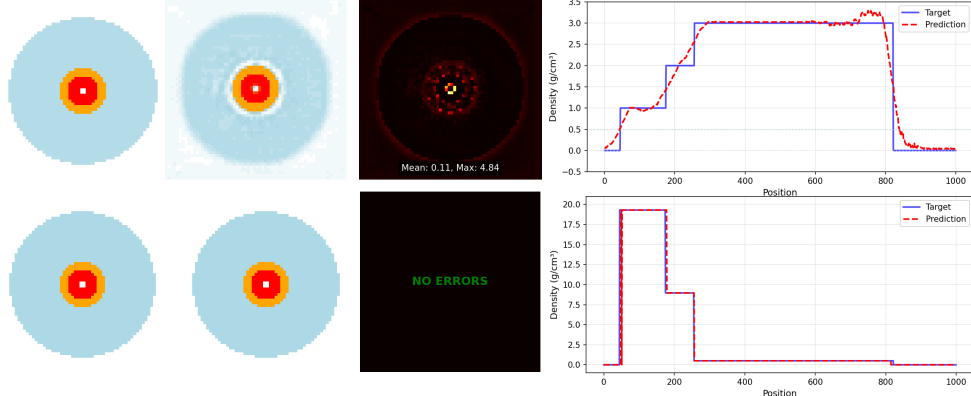


Figure 2: Regression (top) vs. BayesFuse (bottom) for one random test case. First 3 left columns are true 3D voxel (target 1), the prediction, and $\|\text{error}\|$. Rightmost column: 1D lineout (target 2). The axial slice shown (bottom) has zero errors and coronal/sagittal slices (not-shown) were also error-free. BayesFuse preserves sharp, discrete transitions in 1D profile, allowing accurate 3D density mapping.

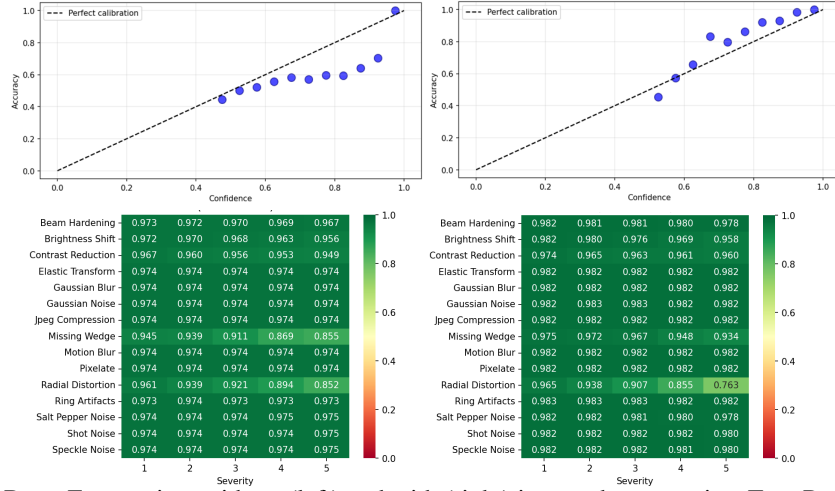


Figure 3: BayesFuse variant without (left) and with (right) imposed symmetries. Top: Representative ECE landscapes. Bottom: Corruption robustness (in mIoU) examined at 5 levels of corruption severity for input perturbations to validate domain shift adaptability of BayesFuse.

branch, (iii) cross-scale fusion lifting 2D \rightarrow 3D features. We enforce crystallographic symmetries through equivariant operations(27) and learn modality weights to suppress corrupted channels. Dense voxel classifications parallelize efficiently on GPUs. Training uses cross-entropy, KL regularization, and diversity terms; test-time Monte-Carlo sampling estimates predictive uncertainty and calibration.

3 Results: Density Prediction and Domain Shift Evaluation

We evaluate three probe combinations $M \in \{2, 3, 5\}$ (X-Rays 10 MeV and 1 MeV $\in M=2$ and neutron 14 MeV added for $M=3$). Metrics include mIoU-like [mean Intersection over Union, mathematically equivalent to the Jaccard Index(29) averaged across 4 material classes], voxel-wise accuracy, ECE, and robustness to physics-inspired corruptions. BayesFuse achieves mIoU ≈ 98 –99% (3D) (voxel-wise accuracy $\geq 99\%$) with lowest ECE $\approx (3$ –5) $\times 10^{-4}$, improving with probe count M (Tables 1, 2). Against regression baselines having identical architecture except output heads (continuous density vs. dense classes, yielding non-physical intermediates) and against BayesFuse variant without symmetry, full BayesFuse concentrates uncertainty precisely at material interfaces. For 1D lineouts, the projection’s inherent asymmetry was restrained; performance on this branch improved from 88–89% to 96–98% mIoU by preserving multi-scale features (8^3 , 4^3 , 2^3 , global) rather than global pooling. Figure 2 demonstrates crucial classification’s advantage: discrete material predictions (bottom) in 1D density profiles enable accurate 1D \rightarrow 3D density mapping, while regression produces non-physical density smearing across the spectrum (top). The discrete formulation preserves material

boundaries that regression destroys through interpolation. BayesFuse maintains physical validity by restricting predictions to real materials [aligned with statistical pattern recognition principles(30)]. We trained all models on 8×NVIDIA TITAN-V 16 GB GPUs (96 GB total) for 300 epochs.

Regression limitations: Figure 2 shows regression (top) producing non-physical densities (e.g., 4.5 g/cm³) at boundaries, misleading when averaged values mask voids/inclusions. This occurs at every test case. Our control runs confirmed sophisticated regression backbones don’t improve on 64³ inputs, revealing problem formulation matters more than model complexity for discrete materials.

Ablations: (i) modality count 2 → 3 → 5 increases mIoU/margins; (ii) symmetry priors improve rotation consistency; (iii) learned weights suppress corrupted channels, effectively addressing domain shift encountered in real experiments; (iv) deeper architectures degrade performance. Figure 3 demonstrates robustness under 15 physics-inspired corruptions at 5 severity levels, BayesFuse concentrates uncertainty at interfaces, remains stable under noise/blur/compression. BayesFuse’s UQ prevents silent misclassifications by flagging uncertain predictions for review. We ran a baseline no-UQ protocol and obtained similar corruption matrix on with accuracy declined on average ≈20-40%.

4 Discussion and Future Work

To our knowledge, this is the first physics-aware discrete reparameterization for multimodal radiographic material characterization achieving voxel accuracy $\geq 99\%$. *The dense classification paradigm, not the specific BNN, drives the improvement from regression’s non-physical intermediates to near-perfect material identification.* BayesFuse illustrates physics-informed ML: reframing continuous inversion as discrete inference aligned with probe physics. It generalizes to discrete materials (security, manufacturing) or continuous settings via reject options/hierarchical dictionaries. By converting ill-posed regression into well-posed calibrated decisions, BayesFuse enables fit-for-purpose automation: high-confidence voxels proceed automatically while uncertain interfaces route for expert review, critical when shots are scarce.

The dual-target formulation, 3D with physical densities, 1D with abstract encodings, enforces implicit consistency without explicit supervision, requiring BayesFuse to learn material-invariant representations across projections. This design proved crucial: the 1D branch improved from 88% to 96-98% mIoU by preserving multi-scale features at 8³, 4³, 2³ resolutions, revealing that material boundaries manifest at specific spatial frequencies lost in global pooling. The framework’s robustness to 15 physics-inspired corruptions suggests readiness for experimental radiographs where ground truth contains inherent ambiguity from partial volume effects and quantum noise. Extending to experimental data requires only calibration of probe-specific attenuation models, not architectural changes. Dense per-voxel decisions scale to 128³–256³ on modern GPUs with much less overhead versus regression. Larger material libraries require per-class forward models and additional data. For $K \gg 4$ materials, core theory holds where hierarchical heads at scales control complexity while maintaining the discrete formulation’s advantages.

Limitations. Our current models assume modest domain shift in probe energy/spectrum, approximate conditional independence across probes, and can inherit a symmetry bias in the 1D head. *Future work:* (i) broaden symmetry groups, inference experimental data, add energy-conditioned probes, (ii) integrate open-set recognition for unseen materials, (iii) add adapters for fast probe/energy finetuning. Code will be released at GitHub.

Acknowledgments

NTTNF, NAD, and ECL were supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory (LANL) under the Director Initiative, Artificial Intelligence for Mission (ArtIMis). BTW, DPB, and ZW were supported by the U.S. Department of Energy, National Nuclear Security Administration, through LANL under Contract No. 89233218CNA000001. We thank Bob Reinovsky and Steve Batha for strategic advice and insightful discussion. Computational resources were provided by the LANL High Performance Computing Division and the Darwin team. Darwin is a testbed at LANL which is funded by the Computational Systems and Software Environments subprogram of Advanced Simulation and Computing program (National Nuclear Security Administration /Department of Energy). We thank Joseph Hickey for critical proofreading and the anonymous reviewers for useful comments that helped us refine the final version of this work.

References

- [1] D. A. Brown, M. B. Chadwick, R. Capote, A. C. Kahler, A. Trkov, M. W. Herman, A. A. Sonzogni, Y. Danon, A. D. Carlson, M. Dunn, D. L. Smith, G. M. Hale, G. Arbanas, R. Arcilla, C. R. Bates, B. Beck, B. Becker, F. Brown, R. J. Casperson, J. Conlin, D. E. Cullen, M. A. Descalle, R. Firestone, T. Gaines, K. H. Guber, A. I. Hawari, J. Holmes, T. D. Johnson, T. Kawano, B. C. Kiedrowski, A. J. Koning, S. Kopecky, L. Leal, J. P. Lestone, C. Lubitz, J. I. Márquez Damián, C. M. Mattoon, E. A. McCutchan, S. Mughabghab, P. Navratil, D. Neudecker, G. P. A. Nobre, G. Noguere, M. Paris, M. T. Pigni, A. J. Plompen, B. Pritychenko, V. G. Pronyaev, D. Roubtsov, D. Rochman, P. Romano, P. Schillebeeckx, S. Simakov, M. Sin, I. Sirakov, B. Sleaford, V. Sobes, E. S. Soukhovitskii, I. Stetcu, P. Talou, I. Thompson, S. van der Marck, L. Welser-Sherrill, D. Wiarda, M. White, J. L. Wormald, R. Q. Wright, M. Zerkle, G. Žerovnik, Y. Zhu. ENDF/B-VIII.0: The 8th major release of the nuclear reaction data library. *Nuclear Data Sheets*, 148:1–142, 2018.
- [2] A. C. Kak, M. Slaney. *Principles of Computerized Tomographic Imaging*. SIAM, 2001.
- [3] R. Bracewell. *The Fourier Transform and Its Applications*. McGraw-Hill, 3rd edition, 1999.
- [4] K. H. Jin, M. T. McCann, E. Froustey, M. Unser. Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, 26(9):4509–4522, 2017.
- [5] G. Wang, J. C. Ye, B. De Man. Deep learning for tomographic image reconstruction. *Nature Machine Intelligence*, 2(12):737–748, 2020.
- [6] S. Arridge, P. Maass, O. Öktem, C. B. Schönlieb. Solving inverse problems using data-driven models. *Acta Numerica*, 28:1–174, 2019.
- [7] D. Krueger, C.-W. Huang, R. Islam, R. Turner, A. Lacoste, A. Courville. Bayesian hypernetworks. arXiv preprint arXiv:1710.04759, 2017.
- [8] D. Hendrycks, T. Dietterich. Benchmarking neural network robustness to common corruptions and perturbations. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2019.
- [9] M. Dusenberry, G. Jerfel, Y. Wen, Y. Ma, J. Snoek, K. Heller, B. Lakshminarayanan, D. Tran. Efficient and scalable Bayesian neural networks with rank-1 factors. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, 2020.
- [10] Y. Wen, D. Tran, J. Ba. BatchEnsemble: An alternative approach to efficient ensemble and lifelong learning. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2020.
- [11] J. Als-Nielsen, D. McMorrow. *Elements of Modern X-ray Physics*. Wiley, 2nd edition, 2011.
- [12] O. Ronneberger, P. Fischer, T. Brox. U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 234–241. Springer, 2015.
- [13] A. R. Sinchuk, P. Piat, A. Krywonos, T. Romanova. A modular U-Net for automated segmentation of X-ray tomography images in composite materials. *Frontiers in Materials*, 8:761229, 2021.
- [14] D. M. Pelt, J. A. Sethian. A mixed-scale dense convolutional neural network for image analysis. *Proceedings of the National Academy of Sciences*, 115(2):254–259, 2018.
- [15] D. Miceli, T. Minniti, L. M. Evans, G. Gorini. Accelerating neutron tomography experiments through artificial neural network based reconstruction. *Scientific Reports*, 9:2450, 2019.
- [16] C. Chang, J. Rosen, J. Bjorkman. Machine learning on neutron and x-ray scattering and spectroscopy. *Chemical Physics Reviews*, 2(3):031301, 2021.
- [17] P. Chu, B. T. Wolfe, D. P. Broughton, R. E. Reinovsky, S. K. Sjue, Z. Wang. Synthetic data generation for machine learning of 3D features using neutrons and X-rays. In *Optica Imaging Congress (3D, COSI, DH, FFlatOptics, IS, pCAOP)*, Technical Digest, paper DM3A.4, 2023.
- [18] N. T. T. Nguyen-Fotiadis, B. T. Wolfe, Z. Wang. Enhancing detail recovery in ICF radiographs: A transformer-based approach with ViXReg. In *NeurIPS 2024 Workshop on Foundation Models for Science*, 2024.
- [19] B. T. Wolfe, M. J. Falato, X. Zhang, N. T. T. Nguyen-Fotiadis, J. P. Sauppe, P. M. Kozlowski, P. A. Keiter, R. E. Reinovsky, S. A. Batha, Z. Wang. Machine learning for detection of 3D features using sparse X-ray tomographic reconstruction. *Review of Scientific Instruments*, 94(2):023502, 2023.
- [20] N. T. T. Nguyen-Fotiadis, G. R. Maskaly, A. S. Liao, K. S. Hickmann. Predicting shockwaves in radiograph images using different deep learning models. In *Proceedings of SPIE Conference on Applications of Machine Learning*, volume 11843, 2021.
- [21] N. T. T. Nguyen-Fotiadis, B. T. Wolfe, Z. Wang. Deep regression outperforms conditional GAN mapping on reconstructing double shells of ICF images. In *Optica Imaging Congress (3D, COSI, DH, FFlatOptics, IS, pCAOP)*, Technical Digest, paper DTh2A.5, 2023.
- [22] A. Biguri, M. Dosanjh, S. Hancock, M. Soleimani. TIGRE: A MATLAB-GPU toolbox for CBCT image reconstruction. *Biomedical Physics & Engineering Express*, 2(5):055010, 2016.
- [23] C. J. Werner, J. S. Bull, C. J. Solomon, F. B. Brown, G. W. McKinney, M. E. Rising, D. A. Dixon, R. L. Martz, H. G. Hughes, L. J. Cox, A. J. Zukaitis, J. C. Armstrong, R. A. Forster, L. Casswell. MCNP user’s manual, code version 6.2. Technical Report LA-UR-17-29981, Los Alamos National Laboratory, 2017.
- [24] C. J. Werner, J. S. Bull, C. J. Solomon, F. B. Brown, G. W. McKinney, M. E. Rising, D. A. Dixon, R. L. Martz, H. G. Hughes, L. J. Cox, A. J. Zukaitis, J. C. Armstrong, R. A. Forster, L. Casswell. MCNP version 6.2 release notes. Technical Report LA-UR-18-20808, Los Alamos National Laboratory, 2018.

- [25] E. F. Shores, C. J. Flamig, T. R. Hill, R. C. Johns. Flux image radiograph (FIR) tally in MCNP: Overview and use. Technical Report LA-UR-11-04757, Los Alamos National Laboratory, 2011.
- [26] E. F. Shores, C. J. Flamig. Radiographic test problem for MCNP and other mesh-based transport codes. *Progress in Nuclear Science and Technology*, 4:502–506, 2014.
- [27] T. Cohen, M. Welling. Group equivariant convolutional networks. In *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, pages 2990–2999, 2016.
- [28] E. T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, 2003.
- [29] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, S. Savarese. Generalized intersection over union: A metric and a loss for bounding box regression. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 658–666, 2019.
- [30] R. O. Duda, P. E. Hart, D. G. Stork. *Pattern Classification*. Wiley, 2nd edition, 2001.

Supplementary Material

Implementation Details

Code BayesFuse implementation structure: `models/` (low-rank BNN layers, symmetry-aware convolutions, multi-scale 1D head), `data/` (multi-probe loader, 15 physics corruptions), `training/` (cross-entropy for classification vs MSE for regression baseline, mIoU/ECE metrics), `evaluation/` (corruption robustness, visualization). Key parameters: Cauchy priors ($\sigma=1.0$), KL scaling $1/N_{\text{voxels}}$, $1.5\times$ weight on 1D loss. Regression baseline: identical architecture except final layer outputs continuous densities. Code release pending institutional approval.

Extended Results

Figures S1–S2 show additional random test cases from the held-out set, comparing continuous regression (top) with BayesFuse (bottom). Left-to-right panels: target 3D slice, prediction, absolute error; the rightmost panel shows the 1D lineout (second target) and prediction.

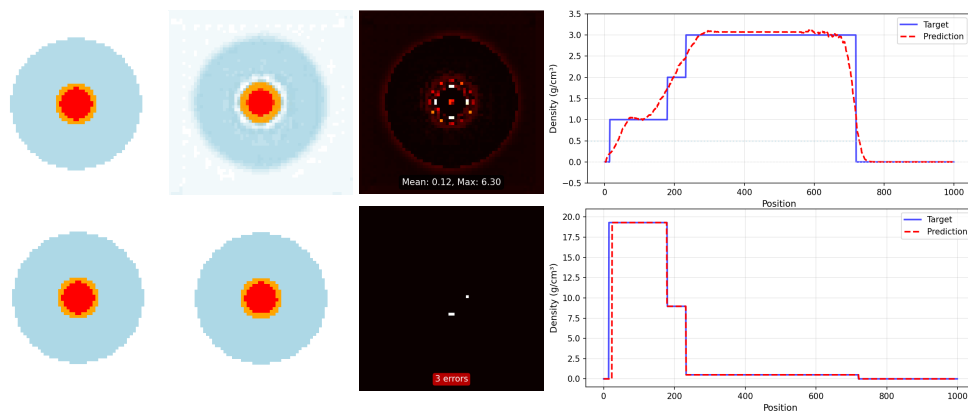


Figure S1: **Supplementary Fig. S1.** Regression (top) vs. BayesFuse (bottom) for a random test sample. The first three columns show the target 3D slice, the prediction, and the absolute error; the rightmost panel shows the 1D lineout (target) and prediction. BayesFuse preserves sharp class boundaries. Errors of $\approx 0.2\%$ (bottom) shown for axial view; similar low errors $0.2\% - 0.4\%$ obtained for coronal and sagittal view (not-shown). Magnification is suggested for better visualization.

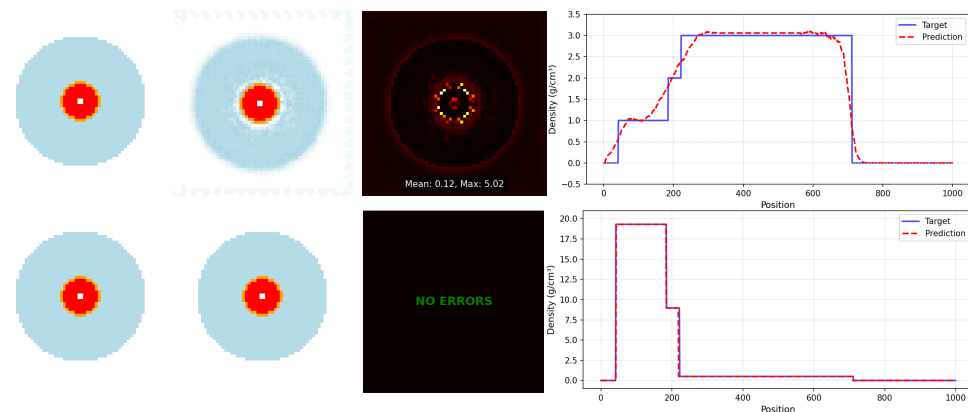


Figure S2: **Supplementary Fig. S2.** Another random test case. Regression (top) vs. BayesFuse (bottom). The first three columns show the target 3D slice, the prediction, and the absolute error; the rightmost panel shows the 1D lineout (target) and prediction. The axial slice predicted by BayesFuse (bottom) has zero errors; coronal and sagittal slices (not shown) were also error-free for this sample. The 1D lineout in this case matches all step transitions implied by the 3D density profile.