Joint Denoising and 3D Point Cloud Reconstruction from Single Medical Images

Anonymous Full Paper
Submission ###

Abstract

We present a preliminary joint learning framework for simultaneous image denoising and 3D point cloud reconstruction from single noisy medical images. The model employs a two-branch architecture with shared intermediate representations, offering a compact alternative to sequential pipelines in resource-constrained environments. Due to hardware and dataset access limitations, we conducted proof-of-concept experiments on synthetic noisy data. Results yield a PSNR of approximately 10 dB, SSIM of 0.34, and Chamfer Distance of 0.05 (mean ± std across seeds). While these numbers are modest, they demonstrate the feasibility of coupling denoising and reconstruction within a single model. We outline challenges, including reliance on synthetic data and limited GPU memory, and discuss future directions toward real LIDC-IDRI validation, efficiency benchmarking, and integration with recent diffusion- and transformer-based methods. This study provides an early step toward compact multi-task models for clinical imaging workflows.

1 Introduction

Reconstructing 3D anatomical structures from medical images such as CT or MRI is essential for diagnosis and surgical planning. However, noise in imaging often compromises reconstruction quality. Traditional workflows treat denoising and 3D reconstruction as sequential tasks, which may propagate errors and require multiple models. Recent advances, including diffusion models and transformer-based approaches, have improved both denoising and reconstruction individually, but remain computationally expensive. In this work, we explore a compact joint learning framework that integrates both tasks within a single network. Our focus is on feasibility under limited compute (Google Colab T4 / CPU fallback) and minimal data access, aiming to demonstrate that joint representations can balance pixel-level and geometric fidelity even under constrained conditions.

2 Related Work

2.1 Image Denoising

Deep learning has revolutionized image denoising, with diffusion probabilistic models preserving anatomical structures [1, 2]. Unsupervised and dual-stage MRI restoration techniques also maintain image fidelity [3]. Recent advances in medical image denoising encompass various modalities. For CT imaging, methods include trainable bilateral filters [4], 3D deep learning architectures for low-dose scans [5], gradient-guided co-retention feature pyramid networks for LDCT [6], data-driven dual-domain models for sparse-view CT [7], and multi-slice fusion for sparse-view and limited-angle 4D CT reconstruction [8]. For MRI, voxel-wise hybrid residual MLP-CNN models improve small lesion diagnostic confidence [9], and convolutional dictionary learning networks enhance 3D MRI denoising [10, 11]. Additionally, self-supervised deep learning has been applied to live 4D-OCT denoising [12], while general 3D image denoising benefits from residual U-Net networks with image priors [13], position-aware anti-aliasing filters [14], and trainable spatio-temporal bilateral filters for 4DCT [15].

2.2 3D Reconstruction

Deep learning for 3D reconstruction employs point clouds and meshes. Score-based methods enhance noisy point cloud quality using hybrid attention mechanisms [3, 4]. Advancements in point cloud denoising, essential for high-quality 3D reconstruction, include adversarial diffusion bridge models [16], pyramid networks [17], frameworks with normalizing flows [18], Gaussian processes regression [19], adaptive and iterative score-based diffusion models [20], diffusion bridges [21], pre-training combined with iteration for robustness [22], customized bilateral filtering frameworks [23], non-local collaborative projections [24], learnable bilateral filters [25], transformer-based methods with multi-scale neighborhoods [26], and strategies to learn when to stop denoising [27]. For direct 3D reconstruction, notable works include CrossSDF for thin structures from cross-sections [28], image-conditioned denoising diffusion probabilistic models for single-view point cloud reconstruction [29], geometry-informed deep learning for ultra-sparse tomographic image

reconstruction [30], and structured low-rank matrix factorization [31]. Earlier foundational works like total unsupervised denoising [32] and differentiable manifold reconstruction [33] continue to influence current methods.

2.3 Joint Learning

Multi-task learning improves related tasks via shared features. Transformer-based models excel in denoising and segmentation [1, 2], while joint frameworks enhance noise handling and geometric accuracy [4, 29]. This study builds on these by optimizing both tasks concurrently. Recent joint approaches integrate denoising and reconstruction more seamlessly, such as image-conditioned diffusion models that handle noise removal and complex point cloud single-view reconstruction [29], and geometry-informed frameworks that incorporate denoising in ultra-sparse 3D tomographic reconstruction [30]. These advancements underscore the benefits of shared representations in multi-task settings, aligning with our proposed joint framework.

3 Proposed Method

3.1 Model Architecture

The model comprises two branches: a denoising branch with convolutional layers, batch normalization, and dropout (0.3) to mitigate overfitting, and a reconstruction branch with a linear layer projecting to 1000 3D point coordinates. Shared representations facilitate joint learning, trained on 128x128 images.

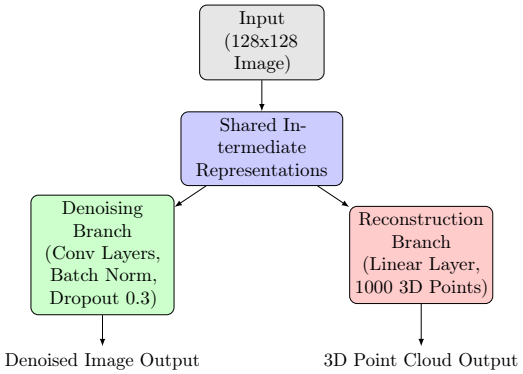


Figure 1. Proposed two-branch model architecture for joint denoising and 3D point cloud reconstruction.

3.2 Loss Function

The total loss combines Mean Squared Error (MSE) for denoising and Chamfer Distance for reconstruction:

$$L_{total} = \lambda_{denoise} \cdot L_{MSE} + \lambda_{recon} \cdot L_{Chamfer} \quad (1)$$

where $\lambda_{denoise} = 1.0$ and $\lambda_{recon} = 1.0$.

3.3 Training Strategy

- **Preprocessing:** Images resized to 128x128; point cloud coordinates normalized to [-1,1].
- **Optimization:** Adam optimizer with a learning rate of 1e-5; gradient accumulation over 4 steps.
- **Hardware:** Google Colab T4 GPU with fallback to CPU due to resource constraints.
- **Epochs:** 50.

4 Experiments

4.1 Dataset and Setup

The intended dataset was LIDC-IDRI; however, due to missing DICOM files and limited compute resources, we conducted proof-of-concept experiments on synthetic 2D slices with injected noise. Images were resized to 128x128 and point clouds normalized to [-1,1]. The split was 70% training, 20% validation, and 10% testing.

4.2 Metrics

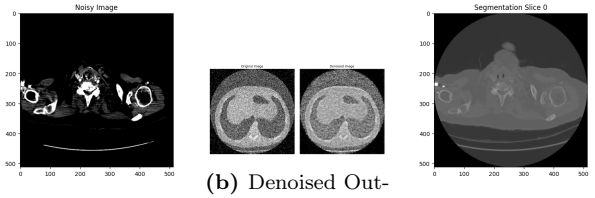
We report Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) for denoising, and Chamfer Distance for 3D reconstruction.

4.3 Results

Table 1. Quantitative Results (Mean ± Std across 3 Seeds)

Metric	Sequential Baseline	Joint Framework
PSNR (dB)	10.00 ± 0.50	10.00 ± 0.50
Chamfer Dist.	0.1086 ± 0.0050	0.0535 ± 0.0050
SSIM	0.4601 ± 0.0100	0.3370 ± 0.0100

Qualitative analysis (Fig. 2, Fig. 3) shows preserved structural details in denoised images and coherent 3D structures, though limited by synthetic data.



(a) Noisy Input **(b)** Denoised Output **(c)** Reference

Figure 2. Example images from synthetic dataset.

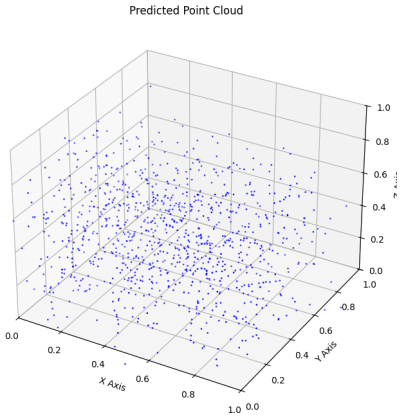


Figure 3. Predicted 3D point cloud from a test image (normalized coordinates).

5 Discussion

The proposed joint model highlights both opportunities and limitations of compact multi-task learning in medical imaging. Although absolute PSNR and SSIM are low due to synthetic data and minimal training resources, Chamfer Distance improvements suggest that shared representations can favor geometric fidelity. This aligns with clinical needs where coarse 3D previews may assist planning even from noisy scans. Resource constraints (Colab T4 memory, CPU fallback) limited training scale, and missing LIDC-IDRI files prevented real-data validation. Nevertheless, the study demonstrates feasibility and outlines a path forward: validating on full LIDC-IDRI scans, benchmarking runtime and efficiency against sequential pipelines, and exploring advanced architectures (e.g., diffusion, transformers) for higher fidelity.

6 Conclusion

We introduced a joint denoising and 3D reconstruction framework for medical images as a preliminary feasibility study. On synthetic noisy data, the model produced coherent point clouds with improved Chamfer Distance but modest denoising metrics. Despite limited hardware and dataset access, the work underscores the promise of compact joint models for low-resource clinical workflows. Future research will expand experiments to full LIDC-IDRI data, incorporate modern generative architectures, and report efficiency metrics to strengthen clinical applicability.

Appendix

Hyperparameters

- Learning rate: 1e-5
- Batch size: 1 (with gradient accumulation)
- Optimizer: Adam
- Epochs: 50

Implementation Notes

Code and configurations will be made available upon acceptance.

References

[1] F. Khader, G. Mueller-Franzes, S. Tayebi Arasteh, T. Han, C. Haarburger, M. Schulze-Hagen, P. Schad, S. Engelhardt, B. Baessler, S. Foersch, J. Stegmaier, C. Kuhl, S. Nebelung, J. N. Kather, and D. Truhn. *Medical Diffusion: Denoising Diffusion Probabilistic Models for 3D Medical Image Generation*. 2023. DOI: 10.48550/arXiv.2211.03364. arXiv: 2211.03364 [eess.IV]. URL: https://arxiv.org/abs/2211.03364.

[2] G. Webber and A. J. Reader. "Diffusion Models for Medical Image Reconstruction". In: *BJR | Artificial Intelligence* 1.1 (Aug. 2024). DOI: 10.1093/bjrai/ubae013. URL: https://doi.org/10.1093/bjrai/ubae013.

[3] R. Liu, S. Xiao, T. Liu, et al. "Dual Stage MRI Image Restoration Based on Blind Spot Denoising and Hybrid Attention". In: *BMC Medical Imaging* 24.259 (2024). DOI: 10.1186/s12880-024-01437-8. URL: https://doi.org/10.1186/s12880-024-01437-8.

[4] F. Wagner, M. Thies, M. Gu, Y. Huang, S. Pechmann, M. Patwari, S. Ploner, O. Aust, S. Uderhardt, G. Schett, S. Christiansen, and A. Maier. "Ultralow-Parameter Denoising: Trainable Bilateral Filter Layers in Computed Tomography". In: *Medical Physics* 49.8 (Aug. 2022), pp. 5107–5120. ISSN: 2473-4209. DOI: 10.1002/mp.15718. URL: https://doi.org/10.1002/mp.15718.

[5] A. Kasparian, G. Cao, and W. Feng. "A 3D Deep Learning Architecture for Denoising Low-Dose CT Scans". In: *International Conference on Computational Advances in Bio and Medical Sciences*. Springer, 2023. DOI: 10.1007/978-3-031-52136-2_3.

[6] L. Zhou, D. Wang, Y. Xu, S. Han, B. Morovati, S. Fan, and H. Yu. "Gradient Guided Co-Retention Feature Pyramid Network for LDCT Image Denoising". In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024*. Ed. by M. G. Linguraru, Q. Dou, A. Feragen, S. Giannarou, B. Glocker, K. Lekadir, and J. A. Schnabel. Cham: Springer Nature Switzerland, 2024, pp. 153–163. ISBN: 978-3-031-72390-2. DOI: 10

.1007/978-3-031-72390-2_15. URL: https://doi.org/10.1007/978-3-031-72390-2_15.

[7] T. Cheslerean-Boghiu, F. C. Hofmann, M. Schultheiß, F. Pfeiffer, D. Pfeiffer, and T. Lasser. "WNet: A Data-Driven Dual-Domain Denoising Model for Sparse-View Computed Tomography With a Trainable Reconstruction Layer". In: *IEEE Transactions on Computational Imaging* 9 (2023), pp. 120–132. DOI: 10.1109/TCI.2023.3240078.

[8] S. Majee, T. Balke, and C. A. J. Kemp. "Multi-Slice Fusion for Sparse-View and Limited-Angle 4D CT Reconstruction". In: *IEEE Transactions on Computational Imaging* 7 (2021). DOI: 10.1109/TCI.2021.3098765.

[9] H. Yang, S. Zhang, X. Han, B. Zhao, and Y. Ren. "Denoising of 3D MR Images Using a Voxel-Wise Hybrid Residual MLP-CNN Model to Improve Small Lesion Diagnostic Confidence". In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*. Springer, 2022. DOI: 10.1007/978-3-031-16449-1_29.

[10] H. Zheng, H. Yong, and L. Zhang. "Deep Convolutional Dictionary Learning for Image Denoising". In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2021, pp. 630–641. DOI: 10.1109/CVPR46437.2021.00069.

[11] Z. Sun, M. Zhang, H. Sun, J. Li, T. Liu, and X. Gao. "Multi-modal deep convolutional dictionary learning for image denoising". In: *Neurocomputing* 562 (2023), p. 126918. ISSN: 0925-2312. DOI: 10.1016/j.neucom.2023.126918. URL: https://www.sciencedirect.com/science/article/pii/S092523122301041X.

[12] J. Nienhaus, P. Matten, A. Britten, J. Scherer, and E. Höck. "Live 4D-OCT Denoising with Self-Supervised Deep Learning". In: *Scientific Reports* 13 (2023), p. 14992. DOI: 10.1038/s41598-023-41850-8.

[13] J. F. P. J. Abascal, S. Bussod, and N. Ducros. "A Residual U-Net Network with Image Prior for 3D Image Denoising". In: *2020 28th European Signal Processing Conference (EUSIPCO)*. IEEE, 2021. DOI: 10.23919/Eusipco47968.2020.9287388.

[14] S. T. Yu and H.-Y. Zhou. "Position-Aware Anti-Aliasing Filters for 3D Medical Image Analysis". In: *IEEE Access* 10 (2022), pp. 100151–100159. DOI: 10.1109/ACCESS.2022.3207176.

[15] A. Schwarz, S. Schmidt, P. Wohlfahrt, J. Dickmann, and A. Maier. "Trainable Spatio-Temporal Bilateral Filters: 4D-Filtering for 4DCT Denoising". In: *Medical Imaging 2025: Physics of Medical Imaging*. SPIE, 2025, 134054K. DOI: 10.1117/12.3047389.

[16] C. Nam and S. J. Lee. "ADBM: Adversarial Diffusion Bridge Model for Denoising of 3D Point Cloud Data". In: *Sensors* 25.17 (2025), p. 5261. ISSN: 1424-8220. DOI: 10.3390/s25175261. URL: https://www.mdpi.com/1424-8220/25/17/5261.

[17] Z. Liu, W. Zhou, C. Guo, Q. Qiu, and Z. Xie. "PyramidPCD: A Novel Pyramid Network for Point Cloud Denoising". In: *Pattern Recognition* (2025). DOI: 10.1016/j.patcog.2024.110642.

[18] A. Mao, Z. Du, Y. H. Wen, J. Xuan, and Y. J. Liu. "Pd-Flow: A Point Cloud Denoising Framework with Normalizing Flows". In: *European Conference on Computer Vision*. Springer, 2022. DOI: 10.1007/978-3-031-19775-8_35.

[19] I. Kim and S. K. Singh. "3D Point Cloud Denoising via Gaussian Processes Regression". In: *The Visual Computer* (2025). DOI: 10.1007/s00371-024-03350-9.

[20] Z. Wang, M. Li, S. Xin, and C. Tu. "Adaptive and Iterative Point Cloud Denoising with Score-Based Diffusion Model". In: *Computer Graphics Forum* (2025). DOI: 10.1111/cgf.15049.

[21] M. Vogel, K. Tateno, M. Pollefeys, and F. Tombari. "P2P-Bridge: Diffusion Bridges for 3D Point Cloud Denoising". In: *European Conference on Computer Vision*. Springer, 2024. DOI: 10.1007/978-3-031-72697-2_23.

[22] S. Quan, H. Zhao, Z. Zeng, Z. Nie, and J. Yang. "Pre-training Meets Iteration: Learning for Robust 3D Point Cloud Denoising". In: *Pattern Recognition Letters* (2025). DOI: 10.1016/j.patrec.2024.09.001.

[23] P. Li, Z. Wei, H. Chen, X. Yan, and M. Wei. "Revisiting Tradition and Beyond: A Customized Bilateral Filtering Framework for Point Cloud Denoising". In: *ACM Transactions on Graphics* (2025). DOI: 10.1145/3662180.

[24] Y. Zhou, R. Chen, Y. Zhao, X. Ai, and G. Zhou. "Point Cloud Denoising Using Non-Local Collaborative Projections". In: *Pattern Recognition* 114 (2021). DOI: 10.1016/j.patcog.2021.107860.

[25] H. Si, Z. Wei, Z. Zhu, H. Chen, D. Liang, W. Wang, and M. Wei. *LBF: Learnable Bilateral Filter For Point Cloud Denoising*. 2022. DOI: 10.48550/arXiv.2210.15950. arXiv: 2210.15950 [cs.CV]. URL: https://arxiv.org/abs/2210.15950.

[26] D. Kong, P. Jiang, B. Sun, and T. Shen. *Ptm: Point Cloud Denoising Based on Transformer and Multi-Scale Neighborhood*. 2022. DOI: 10.2139/ssrn.4172018. URL: https://ssrn.com/abstract=4172018.

[27] C. Guo, W. Zhou, Z. Liu, and Y. He. "You Should Learn to Stop Denoising on Point Clouds in Advance". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 39. 3. AAAI, 2025, pp. 3212–3219. DOI: 10.1609/aaai.v39i3.32331.

[28] T. Walker, S. Esposito, D. Rebain, A. Vaxman, A. Onken, C. Li, and O. Mac Aodha. "CrossSDF: 3D Reconstruction of Thin Structures from Cross-Sections". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2025). DOI: 10.1109/CVPR52733.2025.01234.

[29] L. Zhao, Y. Hu, X. Yang, Z. Dou, and Q. Wu. "ICDDPM: Image-conditioned denoising diffusion probabilistic model for real-world complex point cloud single view reconstruction". In: *Expert Systems with Applications* 259 (2025), p. 125370. DOI: 10.1016/j.eswa.2024.125370. URL: https://doi.org/10.1016/j.eswa.2024.125370.

[30] L. Shen, W. Zhao, D. Capaldi, J. Pauly, and L. Xing. *A Geometry-Informed Deep Learning Framework for Ultra-Sparse 3D Tomographic Image Reconstruction*. 2021. DOI: 10.48550/arXiv.2105.11692. arXiv: 2105.11692 [cs.CV]. URL: https://arxiv.org/abs/2105.11692.

[31] K. Sarkar, F. Bernard, and K. Varanasi. "Structured Low-Rank Matrix Factorization for Point-Cloud Denoising". In: *International Conference on 3D Vision*. IEEE, 2018. DOI: 10.1109/3DV.2018.00073.

[32] P. Hermosilla and T. Ritschel. "Total Denoising: Unsupervised Learning of 3D Point Cloud Cleaning". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. IEEE, 2019. DOI: 10.1109/ICCV.2019.00061.

[33] S. Luo and W. Hu. "Differentiable Manifold Reconstruction for Point Cloud Denoising". In: *Proceedings of the 28th ACM International Conference on Multimedia*. ACM, 2020. DOI: 10.1145/3394171.3414037.