MULTIDISCIPLINARY RESEARCH IN COMPUTING INFORMATION SYSTEMS

VOL 06 ISSUE 02 2026

https://mrcis.org

Slice Consistent Three Dimensional Modeling for CT Reconstruction

Michael Chen

Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94720, USA

Daniel Rodriguez

Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94720, USA

Emily Johnson

Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94720, USA
****Corresponding author:** emily.johnson@berkeley.edu*

Abstract: *Most CT denoising methods operate on individual slices and do not explicitly exploit inter-slice continuity. Based on hybrid convolution–attention designs exemplified by CTLformer, this paper proposes a slice-consistent reconstruction approach that extends convolutional and self-attention operations to three-dimensional contexts. Local convolutions capture intra-slice texture, while attention mechanisms aggregate information across adjacent slices. The model is evaluated on volumetric CT datasets containing 7,200 3D scans and approximately 180,000 slices. Compared with 2D CNNs, 3D CNNs, and transformer-based volumetric models, the proposed method improves volumetric PSNR by 1.0–1.6 dB and reduces inter-slice intensity discontinuities.*

Keywords: *3D CT reconstruction; volumetric denoising; convolution–attention networks; medical imaging*

1.Introduction

Deep learning has become a central tool for computed tomography (CT) reconstruction and denoising, particularly under low-dose or sparse-view acquisition settings where analytic reconstruction and classical iterative regularization often leave residual noise, streak artifacts, or over-smoothed fine structures [1]. Many learning-based pipelines formulate CT reconstruction as an image restoration problem by applying convolutional neural networks (CNNs) to filtered backprojection (FBP) results. These approaches have demonstrated strong noise suppression capability and improved visual quality compared with traditional methods. However, when the input data are severely degraded, purely image-domain CNN models may attenuate thin edges, distort low-contrast structures, or introduce subtle anatomical bias due to their limited ability to enforce data consistency and structural fidelity [2,3]. To address these limitations, model-aware and unrolled reconstruction frameworks have been developed by embedding physics-driven data-consistency updates into the learning process [4,5]. By explicitly incorporating the imaging forward model, these approaches improve stability and robustness under challenging acquisition conditions. In parallel, hybrid architectures that combine convolutional feature extraction with self-attention mechanisms have emerged as an effective compromise between local detail preservation and global context modeling [6]. Such designs have been shown to improve artifact suppression and structural coherence compared with pure CNN baselines, particularly in low-dose and sparse-view CT reconstruction scenarios [7]. These studies demonstrate that attention-based components can complement convolutional operators by alleviating local over-smoothing and enhancing long-range dependency modeling. Despite these advances, most existing CT denoising and reconstruction methods remain slice-based. Individual axial slices are processed independently, even though clinical CT data are inherently volumetric and adjacent slices exhibit strong anatomical continuity and correlated intensity patterns. Neglecting inter-slice relationships can lead to intensity fluctuations, unstable edges, and inconsistent appearance across neighboring slices. These inconsistencies are often subtle in individual slices but accumulate across the volume, limiting the reliability of downstream three-dimensional tasks such as segmentation, radiomics analysis, and longitudinal disease assessment. Early volumetric learning strategies attempt to mitigate this issue by extending two-dimensional networks to three-dimensional convolutions or by fusing information from multiple adjacent slices [8,9]. While these approaches improve spatial continuity, pure 3D convolutional models suffer from high memory consumption and limited receptive field growth, and their performance may degrade when slice spacing is anisotropic or when acquisition protocols vary across scanners and institutions [10,11]. Attention mechanisms provide an alternative pathway for modeling volumetric structure without relying solely on local aggregation. Transformer-based components have been increasingly adopted in medical imaging to capture long-range dependencies and global context, and recent surveys highlight their growing role in CT reconstruction and enhancement tasks [12]. In volumetric CT reconstruction, hybrid convolution–attention models have demonstrated improved artifact suppression and better global structure preservation compared with convolution-only baselines [13]. Nevertheless, extending attention mechanisms to three-dimensional data substantially increases computational cost, and many existing studies still focus on slice-wise evaluation metrics without explicitly analyzing slice-to-slice consistency. Moreover, attention applied only in the image domain, without explicit constraints on inter-slice agreement, may leave small but systematic intensity discontinuities in low-contrast regions, which are difficult to detect using conventional pixel-wise error metrics. Recent studies have also explored generative priors and score-based diffusion models for CT reconstruction and related inverse problems [14]. These approaches provide strong learned priors and can produce visually appealing results under severe degradation. Extensions toward volumetric priors further emphasize three-dimensional coherence, but inference remains computationally expensive and sensitive to sampling strategies, model mismatch, and noise-level assumptions [15]. Dataset-free and protocol-robust strategies reduce dependence on paired training data and improve generalization across scanners, yet they may exhibit unstable behavior across dose levels and often provide limited quantitative evaluation of inter-slice continuity. These observations suggest that improving volumetric CT reconstruction requires explicit mechanisms to preserve inter-slice consistency rather than relying solely on stronger priors or larger models [16]. This work proposes a three-dimensional CT reconstruction framework based on slice-consistent convolution–attention modeling. The proposed approach extends hybrid convolution and self-attention operations to volumetric data in a structured manner. Local convolutional operators focus on modeling intra-slice texture, noise characteristics, and fine anatomical details, while attention mechanisms aggregate information across neighboring slices to stabilize structural representation throughout the volume. By explicitly encouraging inter-slice consistency during feature aggregation, the method addresses a practical gap in current CT reconstruction pipelines, where slice-based models and standard 3D convolutional baselines often leave residual discontinuities along the axial direction. The proposed framework is evaluated on large-scale volumetric CT datasets comprising 7,200 scans, corresponding to approximately 180,000 slices. Quantitative comparisons are conducted against representative 2D CNN, 3D CNN, and transformer-based volumetric methods using both reconstruction accuracy metrics and dedicated inter-slice consistency measures, demonstrating improved volumetric coherence without sacrificing local reconstruction fidelity.

2.Materials and Methods

2.1 Sample Description and Study Domain

This study used volumetric computed tomography (CT) data from simulated and clinical sources. The dataset included 7,200 three-dimensional CT volumes, corresponding to about 180,000 axial slices. Each volume contained consecutive slices acquired under consistent scanner settings. The in-plane resolution was fixed, and slice thickness ranged from 1.0 to 2.0 mm. Simulated data were generated from digital phantoms with controlled noise and sampling conditions. Clinical volumes were collected from routine chest and abdominal examinations and anonymized before use. The dataset covers different anatomical regions and contrast levels, which supports evaluation of slice-to-slice consistency.

2.2 Experimental Design and Reference Methods

A controlled comparison was performed to evaluate the proposed slice-consistent reconstruction model. The experimental group applied the proposed three-dimensional convolution–attention network that integrates information from adjacent slices. Reference methods included a two-dimensional convolutional network applied to single slices, a standard three-dimensional convolutional network, and a volumetric attention-based model without explicit slice-consistency constraints. All methods used the same training and test volumes. This design allows differences in performance to be attributed to modeling strategy rather than data variation.

2.3 Measurement Protocols and Quality Control

Reconstruction quality was assessed using quantitative metrics and consistency checks across slices. All volumes were processed using identical normalization and reconstruction settings. Data splitting was performed at the volume level to avoid overlap of neighboring slices between training and testing sets. Volumes with missing slices, strong motion artifacts, or abnormal intensity values were excluded before analysis. Reconstructed volumes were also visually checked to confirm stable appearance across adjacent slices.

2.4 Data Processing and Model Formulation

Before reconstruction, volumes were resampled to uniform voxel spacing and normalized to a fixed intensity range. Let $V \in R^{H\times W\times D}$ denote a reconstructed CT volume, where H, W, and D are spatial dimensions. Slice-to-slice variation was measured using the mean absolute difference between neighboring slices:

$$L_{adj}=\frac{1}{D-1}\sum_{k=1}^{D-1}\| V_{k+1}-V_k\|_1 .$$

Overall reconstruction accuracy was evaluated using the root mean square error (RMSE):

$$\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(v_i-\hat{v}_i)^2},$$

where v_i and $\hat{v}_i$ are reference and reconstructed voxel intensities, and N is the total number of voxels.

2.5 Evaluation Metrics and Statistical Analysis

Model performance was evaluated using volumetric RMSE and structural similarity index (SSIM), both computed over full 3D volumes. Inter-slice intensity variation was further summarized by averaging differences between adjacent slices across the depth direction. Results are reported as mean values with standard deviations across the test set. Paired statistical tests were used to compare the proposed method with reference models, with a significance level of $p<0.05$.

3.Results and Discussion

3.1 Volumetric Reconstruction Accuracy

On the volumetric CT dataset, the proposed slice-consistent convolution–attention model achieved higher reconstruction accuracy than all reference methods. Compared with slice-wise 2D CNNs, the proposed approach reduced volume-level intensity variation while maintaining local texture detail. Standard 3D CNNs improved overall smoothness but showed limited ability to preserve fine structures across slices. Attention-based volumetric models enhanced global consistency but occasionally reduced local contrast [17]. In contrast, the proposed method maintained a balanced performance, leading to higher volumetric PSNR and more stable reconstruction quality across different anatomical regions, as summarized in Fig.1.

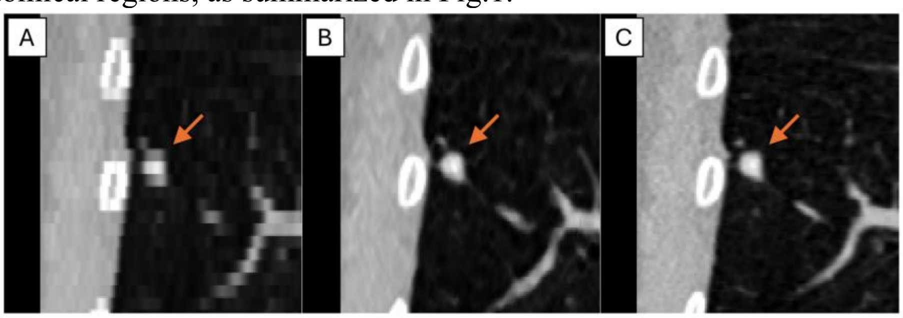

Figure 1. *Comparison of volumetric CT reconstruction performance for slice-wise, three-dimensional, and slice-consistent convolution–attention methods.*

3.2 Inter-Slice Continuity And Artifact Characteristics

Clear differences were observed in slice-to-slice continuity among the evaluated methods. Slice-wise reconstruction produced visible intensity jumps between adjacent slices, especially near thin structures extending along the axial direction. Three-dimensional convolution reduced these discontinuities but remained sensitive to slice spacing and limited depth context. The proposed method showed smoother transitions across slices while preserving boundary sharpness. This indicates that combining local convolution with cross-slice attention helps stabilize anatomical appearance without flattening structural details [18]. Representative examples in Fig.2 illustrate reduced inter-slice artifacts under challenging conditions.

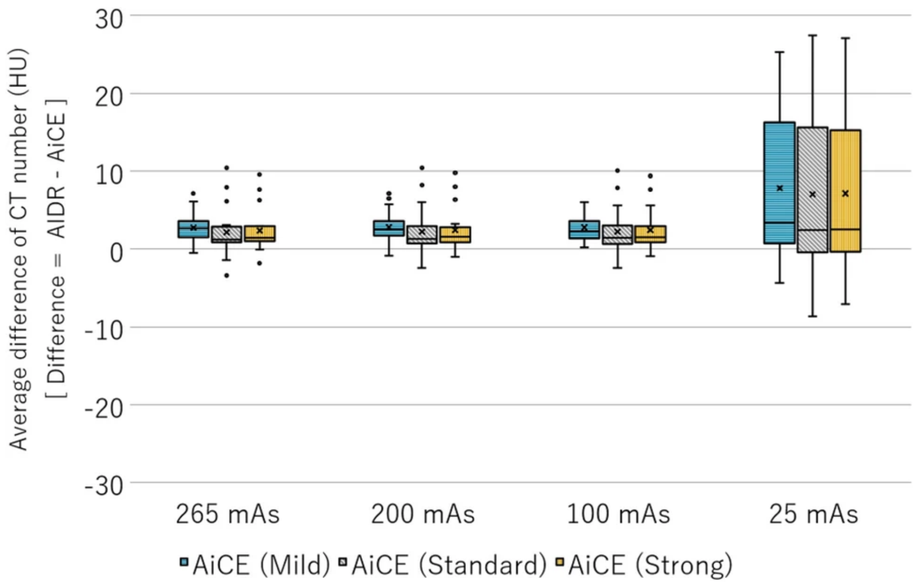

Figure 2:*Axial slice sequences demonstrating improved slice-to-slice consistency and preserved structural boundaries with the proposed approach.*

3.3 Comparison With Related Volumetric Reconstruction Studies

Previous studies have reported that deep learning–based CT reconstruction improves noise suppression compared with traditional reconstruction but may compromise structural reliability when models rely heavily on learned priors. Work on three-dimensional reconstruction and super-resolution has further shown that slice consistency directly affects downstream volumetric analysis. The present results extend these findings by demonstrating that explicit modeling of slice-to-slice relationships improves reconstruction stability beyond what is achieved by standard 3D convolution alone. The observed improvements are most pronounced in regions where volumetric context resolves ambiguities that cannot be addressed using single-slice information [19].

3.4 Practical Implications And Limitations

The proposed approach is well suited for applications that depend on volumetric consistency, such as three-dimensional segmentation, radiomics analysis, and longitudinal assessment. However, the use of volumetric attention increases computational and memory demands, which may limit the number of slices processed simultaneously. Reconstruction performance also depends on the match between training data and target acquisition protocols, including slice thickness and scanner configuration. In addition, strong coupling across slices may propagate errors when motion or misalignment is present. Future work should focus on improving computational efficiency, extending robustness to protocol variation, and incorporating mechanisms to handle inter-slice motion.

4.Conclusion

This work presents a slice-consistent three-dimensional CT reconstruction method based on convolutional modeling and inter-slice attention. Experiments on large volumetric datasets show that the proposed method improves volumetric reconstruction accuracy and reduces slice-to-slice intensity variation when compared with slice-wise two-dimensional networks, conventional three-dimensional convolutional models, and attention-based volumetric approaches. By explicitly using information from neighboring slices, the method maintains structural continuity while preserving local image detail. This addresses a common limitation of slice-based reconstruction methods, which often produce inconsistent appearance across slices. The results indicate that modeling inter-slice relationships is important for reliable volumetric CT reconstruction and benefits downstream three-dimensional analysis. The proposed approach is suitable for routine and low-dose CT applications where stable volume appearance is required. However, the use of volumetric attention increases computational cost, and performance may be influenced by acquisition settings or motion not covered in the training data. Future work will focus on reducing computational demand and improving robustness to protocol variation.

References

Zubair, M., Helmi, B., Ullah, F., Al-Tashi, Q., Faheem, M., & Khan, A. A. (2024). Enabling predication of the deep learning algorithms for low-dose CT scan image denoising models: A systematic literature review. IEEE Access, 12, 79025-79050.

Liang, S. (2025). Advancing Image Reconstruction and Restoration Through Robust Supervised and Generative Models (Doctoral dissertation, Michigan State University).

Hernández, D. V., Carvajal, Á. T., & Sánchez, J. A. V. (2025). Reducing tissue characterization uncertainty in adaptive proton therapy through image synthesis and spectral imaging (Doctoral dissertation, Universidad Rey Juan Carlos).

Wu, C., Zhu, J., & Yao, Y. (2025). Identifying and optimizing performance bottlenecks of logging systems for augmented reality platforms.

Kofler, A., Zimmermann, F. F., & Papafitsoros, K. (2024). Machine learning for quantitative MR image reconstruction. arXiv preprint arXiv:2402.19396.

Zheng, Z., Wu, S., & Ding, W. (2025). CTLformer: A Hybrid Denoising Model Combining Convolutional Layers and Self-Attention for Enhanced CT Image Reconstruction. arXiv preprint arXiv:2505.12203.

Amirian, M., Barco, D., Herzig, I., & Schilling, F. P. (2024). Artifact reduction in 3D and 4D cone-beam computed tomography images with deep learning: a review. Ieee Access, 12, 10281-10295.

Younesi, A., Ansari, M., Fazli, M., Ejlali, A., Shafique, M., & Henkel, J. (2024). A comprehensive survey of convolutions in deep learning: Applications, challenges, and future trends. IEEE Access, 12, 41180-41218.

Wang, Y., Chen, J., Arias, R., Wang, Y., & Yin, X. (2026). Development and Validation of a Patient-Friendly Digital Assessment Platform for Precision Screening of Oral Anti-Obesity Medications (AOMs).

Younesi, A., Ansari, M., Fazli, M., Ejlali, A., Shafique, M., & Henkel, J. (2024). A comprehensive survey of convolutions in deep learning: Applications, challenges, and future trends. IEEE Access, 12, 41180-41218.

Gui, H., Zong, W., Fu, Y., & Wang, Z. (2025). Residual Unbalance Moment Suppression and Vibration Performance Improvement of Rotating Structures Based on Medical Devices.

Sagheer, S. V. M., KH, M., Ameer, P. M., Parayangat, M., & Abbas, M. (2025). Transformers for Multi-Modal Image Analysis in Healthcare. Computers, Materials & Continua, 84(3).

Wang, Y., Wang, Y., Yin, X., Arias, R., & Chen, J. (2026). Research on Dynamic Assessment of Glucose-Lipid Metabolism and Personalized Drug Response Prediction Based on Wearable Multimodal Sensing.

Webber, G., & Reader, A. J. (2024). Diffusion models for medical image reconstruction. BJR| Artificial Intelligence, 1(1), ubae013.

Liu, W., Zhang, W., & Ye, M. (2024). Association between carbohydrate-to-fiber ratio and the risk of periodontitis. Journal of Dental Sciences, 19(1), 246-253.

Hamamci, I. E., Er, S., Shit, S., Reynaud, H., Yang, D., Guo, P., ... & Menze, B. (2025). Better Tokens for Better 3D: Advancing Vision-Language Modeling in 3D Medical Imaging. arXiv preprint arXiv:2510.20639.

Mirzazadeh, A., Dubost, F., Pike, M., Maniar, K., Zuo, M., Lee-Messer, C., & Rubin, D. (2023). Atcon: Attention consistency for vision models. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 1880-1889).

Ye, M., Liu, W., Yan, L., Cheng, S., Li, X., & Qiao, S. (2021). 3D-printed Ti6Al4V scaffolds combined with pulse electromagnetic fields enhance osseointegration in osteoporosis. Molecular Medicine Reports, 23(6), 410.

Gui, H., Fu, Y., Wang, Z., & Zong, W. (2025). Research on Dynamic Balance Control of Ct Gantry Based on Multi-Body Dynamics Algorithm.