From Sparse X-rays to 3D CT: Training-Free Reconstruction with Diffusion Priors

Zhenkai Zhang, Markus Hiller, Krista A. Ehinger, and Tom Drummond

The University of Melbourne, Melbourne, Australia
freddy.zhang@unimelb.edu.au

Abstract. Solving 3D medical inverse problems typically requires training dedicated supervised models for each specific task and measurement setting. To break this dependency, we present TF-PRDiT: a *training-free* conditional sampling framework that converts a frozen voxel-level 3D Diffusion Transformer prior into a versatile inverse medical problem solver. Building on the posterior-sampling view of diffusion inverse solvers, TF-PRDiT enforces measurement consistency during sampling via a task-specific forward operator rather than updating model weights, enabling a *single* pretrained prior to be reused across diverse conditional settings. Our method combines a predictor-corrector sampler with likelihood-based guidance on the denoised prediction, providing stable data-fidelity correction while preserving the underlying 3D anatomical prior. We highlight our framework's capability on the challenging task of X-ray-to-computed tomography (CT) reconstruction by integrating a differentiable digitally reconstructed radiograph (DRR) projector to allow gradients to propagate directly from projection space back to voxels without any retraining. Experiments on the LIDC-IDRI dataset demonstrate that TF-PRDiT provides a strong unified reconstruction framework for varying numbers of input X-rays (1–12). While the single-view setting remains highly under-constrained, reconstruction quality improves consistently as additional views are provided, leading to strong performance in the bi-planar and multi-view settings. Beyond X-ray-to-CT, we show that simply swapping the forward operator extends the same frozen model to 3D super-resolution, volumetric infilling, and deblurring without any task-specific retraining, demonstrating that a single 3D diffusion prior can serve as a universal solver for volumetric medical inverse problems.

Keywords: Training-free Conditional Generation · 3D Medical Inverse Problems · Gradient-Based Diffusion Guidance · Sparse X-ray to CT

1 Introduction

Recovering 3D medical volumes from sparse, incomplete, or degraded measurements is a central inverse problem in computational radiology. Fully sampled CT provides rich anatomical information, but its acquisition is constrained by scanner availability, clinical cost, specialized equipment, and radiation exposure. In contrast, 2D X-rays are widely available and low dose, yet mapping one or a

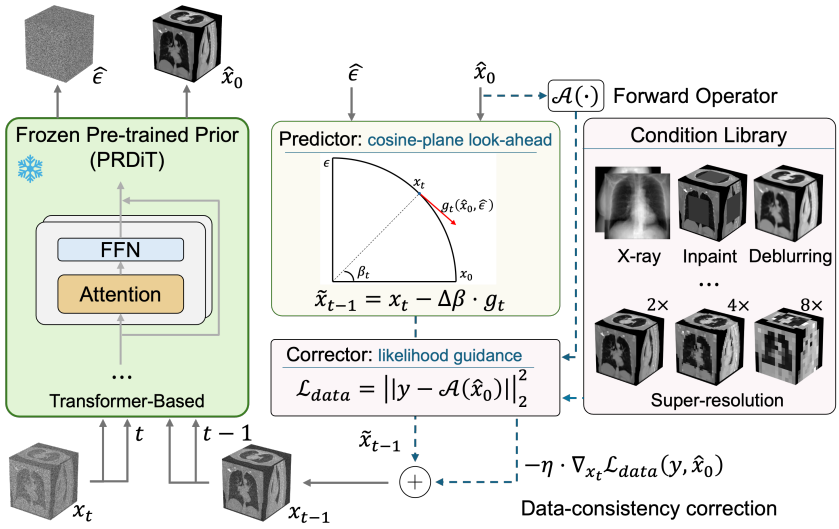

Fig. 1. Overview of TF-PRDiT. A frozen voxel-level 3D Diffusion Transformer prior is guided at inference by measurement-consistency gradients from a task-specific forward operator $\mathcal{A}$. By replacing $\mathcal{A}$, the same frozen model handles X-ray-to-CT reconstruction and volumetric restoration tasks without retraining.

few projections to a full 3D volume is severely ill posed. A useful reconstruction framework should therefore enforce the available measurements while preserving anatomically plausible 3D structure. Most existing X-ray-to-CT methods learn a supervised mapping for a fixed measurement setting [5,13,7,8,4]. These models can work well under the view configuration used during training, but changing the geometry or number of views usually requires architectural changes or retraining. This rigidity is especially limiting in clinical workflows, where imaging protocols vary and paired 3D training data are expensive.

Training-free diffusion-based inverse solvers offer a different route: a pretrained generative prior is kept fixed, and measurement consistency is imposed during sampling through a known forward operator [6,10,2,11]. This posterior-sampling view, notably formalized in Diffusion Posterior Sampling (DPS) [2], allows the same prior to be reused across tasks without retraining. However, most prior work has focused on 2D images or relatively simple restoration operators. Extending this principle to native 3D medical volumes and projection-to-volume reconstruction remains challenging because volumetric generation is computationally costly, sparse X-ray measurements are severely underdetermined, and latent compression can remove fine anatomical details.

In this paper, we present TF-PRDiT, a **t**raining-**f**ree extension of **p**ixel-level **r**esidual **di**ffusion **t**ransformer for conditional sampling at test time to solve medical inverse problems. Our method follows the DPS-style idea of enforcing measurements during sampling rather than retraining the generative model; our

contribution is to adapt this idea to native 3D CT priors and differentiable medical forward models. For X-ray-to-CT, the forward operator is a differentiable DRR projector that maps a sampled 3D volume to one or more 2D projections. For other inverse problems, it can be replaced by a downsampling, masking, or blurring operator. This formulation makes the number of available X-rays a runtime input, rather than a training-time architectural assumption.

Our contributions are: (1) a training-free sampler extending DPS to native voxel-space 3D CT with differentiable volumetric forward operators, (2) denoised-prediction guidance with a cosine-decay schedule and k-step predictor–corrector for stable high-dimensional conditioning, and (3) a single frozen prior that scales to 1–12 X-ray views and transfers to super-resolution, infilling, and deblurring by swapping only $\mathcal{A}$.

2 Related Work

Sparse-view CT reconstruction. Learning 3D CT from sparse X-rays observations is commonly formulated as a supervised image-to-image translation problem. X2CT-GAN [13] learns to fuse 2D X-ray features into a 3D generator, while PerX2CT [7] further incorporate perspective projection geometry. More recent diffusion and Transformer-based approaches, such as DiffuX2CT [8] and DX2CT [4], improve reconstruction quality by learning stronger conditional mapping from fixed input views to CT volumes. However, these methods are tied to the view configuration used during training and cannot handle variable views without retraining. NAF [15] is conceptually related, optimizing a patient-specific neural representation from projections, but assumes sparse-view CBCT with ∼50 views in their main experiments. Our mono-/bi-planar setting (1–2 views) is not measurement-budget aligned, so we omit direct comparison. Our TF-PRDiT instead treats each X-ray as a runtime measurement constraint while using a pretrained 3D diffusion prior to recover anatomically plausible structures.

Diffusion inverse solvers. Score-based inverse problem solvers and Diffusion Posterior Sampling (DPS) [10,2] show pretrained diffusion models can be reused as generative priors by enforcing measurement consistency during sampling. Related training-free methods include DDRM [6], which leverages the singular-value structure of linear degradation operators, DDNM [11], which decomposes the solution into range- and null-space components, and FreeDOM [14], which extends test-time guidance to non-linear operators. These methods are mainly developed for 2D image restoration, where forward operators are lower-dimensional and often linear. Sparse-view X-ray-to-CT is more challenging as the operator maps a high-dimensional 3D volume to 2D projections and becomes increasingly underconstrained as views decrease. TF-PRDiT follows the same posterior-sampling principle but instantiates it for native voxel-space 3D CT, where each additional X-ray view contributes a residual term to the guidance objective.

Native 3D priors. Direct 3D diffusion is expensive because memory and compute scale cubically with resolution. Many 3D generative models rely on latent compression or encoder-decoder architectures, which may weaken fine anatomi-

Algorithm 1 Conditional Predictor-Corrector Sampler

Input: $\mathbf{x}_T$, timestep schedule, frozen prior f_θ, predictor multiplier k, measurement $\mathbf{y}$, forward operator $\mathcal{A}$, guidance scale η
Initialize: $\mathcal{X} \leftarrow \{\mathbf{x}_T\}$, $\beta_t \leftarrow \frac{\pi}{2}\frac{t}{T}$ ▷ Angle for spherical transition
for $t = T, T-1, \ldots, 1$ **do**
  $\mathbf{x}_t \leftarrow \text{last}(\mathcal{X})$, $(\hat{\boldsymbol{\epsilon}}, \hat{\mathbf{x}}_0) \leftarrow f_\theta(\mathbf{x}_t, t)$ ▷ Predict noise and clean signal
  $\mathbf{f}_t \leftarrow \sin(\beta_t)\hat{\mathbf{x}}_0 - \cos(\beta_t)\hat{\boldsymbol{\epsilon}}$ ▷ Prior-induced update direction
  $\mathcal{L}_{\text{data}} \leftarrow \|\mathcal{A}(\hat{\mathbf{x}}_0) - \mathbf{y}\|_2^2$, $\mathbf{g}_t \leftarrow \nabla_{\hat{\mathbf{x}}_0}\mathcal{L}_{\text{data}}$ ▷ Data-consistency guidance
  $\kappa_t \leftarrow \min(k, t)$, $\Delta\beta^{(\kappa_t)} \leftarrow \beta_t - \beta_{t-\kappa_t}$ ▷ Avoid $t - k < 0$
  $\tilde{\mathbf{x}} \leftarrow \mathbf{x}_t - \Delta\beta^{(\kappa_t)}\mathbf{f}_t - \eta\mathbf{g}_t$ ▷ Guided predictor step
  $\Gamma_t^{(\kappa_t)} \leftarrow \cos(\beta_{t-1})/\cos(\beta_{t-\kappa_t})$ ▷ Variance-preserving scale
  $\mathbf{x}_{t-1} \leftarrow \Gamma_t^{(\kappa_t)}\tilde{\mathbf{x}} + \sqrt{1-(\Gamma_t^{(\kappa_t)})^2}\boldsymbol{\epsilon}'$, $\boldsymbol{\epsilon}' \sim \mathcal{N}(0, \mathbf{I})$ ▷ Corrector step
  $\mathcal{X} \leftarrow \mathcal{X} \cup \{\mathbf{x}_{t-1}\}$
end for
Return $\mathcal{X}$

cal boundaries. TF-PRDiT builds on PRDiT [17], a voxel-level residual diffusion Transformer prior, and combines it with differentiable measurement guidance to enable training-free conditional sampling across multiple inverse problems.

3 Method

TF-PRDiT solves inverse problems of the form $\mathbf{y} = \mathcal{A}(\mathbf{x}) + \boldsymbol{\xi}$, where $\mathbf{x}$ is the unknown 3D volume, $\mathbf{y}$ is the measurement, $\mathcal{A}$ is a known forward operator, and $\boldsymbol{\xi}$ is measurement noise. The same formulation covers X-ray projection, downsampling, masking, and deblurring by changing $\mathcal{A}$.

Relation to DPS. Our sampler follows the same posterior-sampling perspective as DPS [2]: an unconditional diffusion prior supplies the generative score, while a task-specific likelihood term enforces agreement with measurements. We do not introduce a new posterior-guidance principle. Instead, TF-PRDiT specializes this principle for native 3D medical volumes by combining a voxel-level CT diffusion prior with differentiable volumetric forward operators. This is important for sparse X-ray-to-CT because the forward model maps a 3D volume to one or more 2D projections, so each additional X-ray can be incorporated by adding another projection-space residual to the same guidance loss.

Frozen 3D diffusion prior. We use a pretrained PRDiT model [17] as a frozen unconditional prior over chest CT volumes. PRDiT is a diffusion transformer trained directly in voxel space on LIDC-IDRI CT volumes [1], avoiding the compressed latent representation used by many latent diffusion models. This voxel-level formulation is important for medical reconstruction because fine anatomical boundaries and small structures may be weakened or lost during latent compression. Given a noisy volume $\mathbf{x}_t$ and timestep t, the frozen network jointly predicts the noise component and the corresponding clean-volume estimate: $(\hat{\boldsymbol{\epsilon}}, \hat{\mathbf{x}}_0) = f_\theta(\mathbf{x}_t, t)$. During downstream reconstruction, all parameters

of f_θ remain fixed. Task adaptation is performed only through measurement-guided sampling, so the same prior can be reused across different view counts and inverse operators without retraining.

Predictor-corrector sampling. Using the cosine-sine parameterization [16], let $\beta_t = \frac{\pi}{2}\frac{t}{T}$. The prior-induced update direction is $\mathbf{f}_t = \sin(\beta_t)\hat{\mathbf{x}}_0 - \cos(\beta_t)\hat{\boldsymbol{\epsilon}}$. We use a k-step predictor,

$$\tilde{\mathbf{x}} = \mathbf{x}_t - \Delta\beta^{(k)}\mathbf{f}_t, \quad \Delta\beta^{(k)} = \beta_t - \beta_{t-k}, \tag{1}$$

which produces $\tilde{\mathbf{x}} \approx \mathbf{x}_{t-k}$ at noise level $t-k$. Because the predictor strides k steps, the noise level of $\tilde{\mathbf{x}}$ is β_{t-k}, which differs from the target β_{t-1} when $k > 1$. A variance-preserving corrector rescales $\tilde{\mathbf{x}}$ and injects fresh noise to reach the correct noise level at timestep $t-1$:

$$\mathbf{x}_{t-1} = \Gamma_t^{(k)}\tilde{\mathbf{x}} + \sqrt{1 - \left(\Gamma_t^{(k)}\right)^2}\,\boldsymbol{\epsilon}', \quad \Gamma_t^{(k)} := \frac{\cos(\beta_{t-1})}{\cos(\beta_{t-k})}, \quad \boldsymbol{\epsilon}' \sim \mathcal{N}(0, \mathbf{I}). \tag{2}$$

This keeps the marginal variance of $\mathbf{x}_{t-1}$ consistent with the forward process, preventing accumulated drift when $k > 1$. Larger k enables broader stochastic exploration, which is useful for ill-posed sparse-view reconstruction.

Likelihood guidance on the denoised estimate. Directly matching measurements on the noisy state can produce unstable gradients. We instead guide sampling through the denoised estimate $\hat{\mathbf{x}}_0$, using the data-consistency loss

$$\mathcal{L}_{\text{data}} = \|\mathcal{A}(\hat{\mathbf{x}}_0) - \mathbf{y}\|_2^2. \tag{3}$$

The gradient $\nabla_{\hat{\mathbf{x}}_0}\mathcal{L}_{\text{data}}$ is computed by backpropagating only through $\mathcal{A}$, treating $\hat{\mathbf{x}}_0$ as the free variable and holding the denoiser f_θ fixed. Following the DPS approximation [2], the Jacobian $\partial\hat{\mathbf{x}}_0/\partial\mathbf{x}_t$ is omitted for tractability; the resulting voxel-space gradient is applied directly as a correction to the $\mathbf{x}_t$ predictor update. For X-ray-to-CT, $\mathcal{A}$ is instantiated as a differentiable DRR projector using DiffDRR [3], allowing gradients to flow from projection space to voxels. With M available X-rays, the loss becomes

$$\mathcal{L}_{\text{xray}} = \sum_{i=1}^{M} \|\mathcal{P}_{g_i}(\hat{\mathbf{x}}_0) - \mathbf{y}_i\|_2^2, \tag{4}$$

where $\mathcal{P}_{g_i}$ denotes projection under view geometry g_i. Thus changing the number or geometry of views changes only the residual terms in the loss, not the network architecture or weights.

Cosine-decay guidance. We use a time-dependent guidance scale $\eta_t = \eta_{\max} \cdot (1 - \cos(\pi t/T))/2$. At high-noise timesteps, $\eta_t \approx \eta_{\max}$ provides strong measurement guidance to shape global structure. As $t \to 1$, η_t decays to zero, reducing over-correction and preserving fine anatomical details. This avoids the trade-off of fixed guidance, which can under-correct at high or over-correct at low noise.

Algorithm 1 assembles the above components into the full conditional sampling procedure. At each timestep, the frozen prior predicts $(\hat{\boldsymbol{\epsilon}}, \hat{\mathbf{x}}_0)$, the predictor applies the prior-induced update together with measurement-consistency guidance, and the corrector restores the target noise level. Task adaptation is entirely controlled by $\mathcal{A}$ and $\mathbf{y}$, while f_θ remains fixed.

Table 1. Quantitative comparison of 1-view and 2-view X-ray-to-CT reconstruction on LIDC-IDRI. Results are mean ± std over three random seeds, where available. †indicates results reported in original papers. MSE is scaled by 10^3.

Method	#V	MSE↓	PSNR↑	SSIM↑	SNR↑
X2CT-GAN [13]*	1	**17.41** ± 0.11	**22.13** ± 0.03	**0.518** ± 0.002	**6.78** ± 0.03
X2CT-GAN [13]	1	18.91 ± 0.41	21.75 ± 0.10	0.509 ± 0.008	6.39 ± 0.10
TF-PRDiT (ours)	1	22.51 ± 1.46	21.34 ± 0.30	0.509 ± 0.012	6.01 ± 0.30
X2CT-GAN [13]*	2	6.55 ± 0.13	26.23 ± 0.09	0.656 ± 0.005	10.88 ± 0.09
X2CT-GAN [13]	2	6.80 ± 0.07	26.05 ± 0.05	0.647 ± 0.004	10.70 ± 0.05
PerX2CT [7]†	2	—	27.45 ± —	0.732 ± —	—
DiffuX2CT [8]†	2	—	26.35 ± —	0.687 ± —	—
DX2CT [4]†	2	—	28.36 ± —	0.763 ± —	—
TF-PRDiT (ours)	2	**3.65** ± 0.12	**29.06** ± 0.12	**0.767** ± 0.002	**13.72** ± 0.12

*X-rays generated with Plastimatch [9]; †MSE/SNR & std not reported; code not publ. available.

4 Experiments

Setup. We evaluate on LIDC-IDRI [1] using the X2CT-GAN [13] split (916 train and 102 held-out test CTs), following its preprocessing protocol (CT intensities clamped to [0,2500] HU) for fair comparison. The frozen PRDiT prior is trained once exclusively on the 916 training cases to match this benchmark setting; the 102 test cases are held out entirely. TF-PRDiT itself remains training-free: no weights are updated for reconstruction, and all task adaptation happens at inference via measurement-guided sampling. Projections are generated with DiffDRR [3], and we report full-volume MSE, PSNR, SSIM [12], and SNR.

4.1 X-ray-to-CT Reconstruction

Baseline comparisons. Table 1 reports single-view and biplanar results on LIDC-IDRI. Single-view is highly under-constrained, yet TF-PRDiT remains competitive with X2CT-GAN despite no task-specific training. In the biplanar setting, TF-PRDiT substantially improves reconstruction quality, outperforming X2CT-GAN, PerX2CT [7], DiffuX2CT [8], and DX2CT [4] across all metrics.
Scaling to arbitrary views. Table 2 shows monotonic gains from 1 to 12 views across all metrics (PSNR 21.34 → 34.61 dB, SSIM 0.509 → 0.880, average HU-MAE 279.3 → 68.8), with the sharpest jump occurring at 2 views where the second projection resolves depth ambiguity.

4.2 Generalization and Ablation

Other inverse problems. Beyond X-ray-to-CT reconstruction, TF-PRDiT also generalizes to 3D super-resolution, volumetric infilling, and deblurring by only replacing the forward operator $\mathcal{A}$. As shown in Figure 2, the same frozen

Table 2. Reconstruction quality vs. number of X-ray views for TF-PRDiT (frozen prior); average HU-MAE across {lung, airway, heart, vessels, bone, esophagus}.

#Views	MSE↓	PSNR↑	SSIM↑	SNR↑	HU-MAE↓
1	22.51 ± 1.46	21.34 ± 0.30	0.509 ± 0.012	6.01 ± 0.30	279.3 ± 125.9
2	3.65 ± 0.12	29.06 ± 0.12	0.767 ± 0.002	13.72 ± 0.12	118.7 ± 36.1
4	1.96 ± 0.02	31.71 ± 0.06	0.824 ± 0.000	16.37 ± 0.06	97.7 ± 27.4
6	1.39 ± 0.02	33.20 ± 0.07	0.852 ± 0.000	17.86 ± 0.07	81.7 ± 19.0
8	1.19 ± 0.01	33.89 ± 0.05	0.865 ± 0.000	18.55 ± 0.05	75.7 ± 16.8
12	**1.02** ± 0.02	**34.61** ± 0.07	**0.880** ± 0.000	**19.27** ± 0.07	**68.8** ± 14.3

Fig. 2. Qualitative results on additional 3D inverse problems. By replacing only the forward operator $\mathcal{A}$, TF-PRDiT handles super-resolution, volumetric infilling, and deblurring with the same frozen prior f_θ.

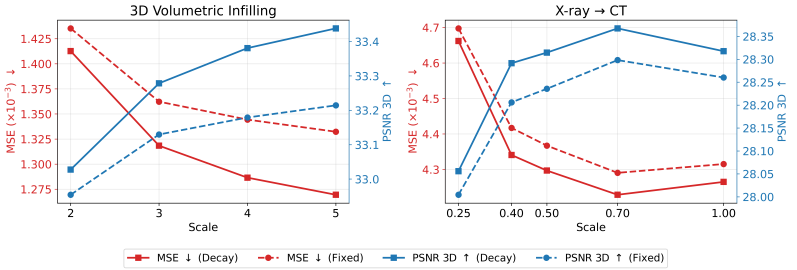

Fig. 3. Ablation study on guidance schedule. Cosine decay vs. fixed guidance on 3D volumetric infilling (left) and 2-view X-ray-to-CT (right).

diffusion prior restores missing or degraded structures across different degradation types without task-specific retraining. This supports the central motivation of TF-PRDiT: once a strong unconditional 3D prior is learned, different inverse problems can be addressed through measurement-consistent sampling rather than separate supervised models.

Ablation. We use $k = 4$ for X-ray-to-CT and $k = 2$ for $2\times$ super resolution, reflecting their differing ambiguity: broader stochastic exploration helps resolve depth uncertainty under sparse views, while the more constrained super-resolution task needs less. Cosine-decayed guidance also outperforms fixed guidance on both tasks: strong early guidance corrects global structure, while decay near $t = 1$ avoids over-correction and preserves fine anatomical detail (Fig. 3).

5 Conclusion

We presented TF-PRDiT, a training-free framework that converts a frozen voxel-level diffusion prior into a versatile solver for 3D medical inverse problems by swapping only the forward operator. Without retraining, it scales to 1–12 X-ray views and transfers to super-resolution, infilling, and deblurring. Limitations include slower inference than feed-forward models and dependence on an accurate forward operator. Broader validation across anatomies, scanners, and clinical settings remains future work. Our code and pre-trained models are publicly available at https://github.com/Fredy-Zhang/TF-PRDiT.

Acknowledgments. This research was supported by The University of Melbourne's Research Computing Services and the Petascale Campus Initiative. MH was supported by the Australian Research Council (ARC) through grant DP230102775.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Armato III, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A., et al.: The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. Medical Physics **38**(2), 915–931 (2011)
2. Chung, H., Kim, J., Mccann, M.T., Klasky, M.L., Ye, J.C.: Diffusion posterior sampling for general noisy inverse problems. arXiv preprint arXiv:2209.14687 (2022)
3. Gopalakrishnan, V., Golland, P.: Fast auto-differentiable digitally reconstructed radiographs for solving inverse problems in intraoperative imaging. In: Workshop on Clinical Image-Based Procedures. pp. 1–11. Springer (2022)
4. Jeong, Y.S., Yoo, H.B., Chun, I.Y.: Dx2ct: Diffusion model for 3d ct reconstruction from bi or mono-planar 2d x-ray (s). In: ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1–5. IEEE (2025)
5. Jin, K.H., McCann, M.T., Froustey, E., Unser, M.: Deep convolutional neural network for inverse problems in imaging. IEEE transactions on image processing **26**(9), 4509–4522 (2017)
6. Kawar, B., Elad, M., Ermon, S., Song, J.: Denoising diffusion restoration models. Advances in neural information processing systems **35**, 23593–23606 (2022)

7. Kyung, D., Jo, K., Choo, J., Lee, J., Choi, E.: Perspective projection-based 3d ct reconstruction from biplanar x-rays. In: 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1–5. IEEE (2023)
8. Liu, X., Qiao, Z., Liu, R., Li, H., Zhang, J., Zhen, X., Qian, Z., Zhang, B.: Diffux2ct: Diffusion learning to reconstruct ct images from biplanar x-rays. In: European Conference on Computer Vision. pp. 458–476. Springer (2024)
9. Sharp, G.C., Li, R., Wolfgang, J., Chen, G., Peroni, M., Spadea, M.F., Mori, S., Zhang, J., Shackleford, J., Kandasamy, N.: Plastimatch: an open source software suite for radiotherapy image processing. In: Proceedings of the XVI'th International Conference on the use of Computers in Radiotherapy (ICCR), Amsterdam, Netherlands. vol. 3 (2010)
10. Song, Y., Shen, L., Xing, L., Ermon, S.: Solving inverse problems in medical imaging with score-based generative models. arXiv preprint arXiv:2111.08005 (2021)
11. Wang, Y., Yu, J., Zhang, J.: Zero-shot image restoration using denoising diffusion null-space model. arXiv preprint arXiv:2212.00490 (2022)
12. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing **13**(4), 600–612 (2004)
13. Ying, X., Guo, H., Ma, K., Wu, J., Weng, Z., Zheng, Y.: X2ct-gan: reconstructing ct from biplanar x-rays with generative adversarial networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10619–10628 (2019)
14. Yu, J., Wang, Y., Zhao, C., Ghanem, B., Zhang, J.: Freedom: Training-free energy-guided conditional diffusion model. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 23174–23184 (2023)
15. Zha, R., Zhang, Y., Li, H.: NAF: Neural attenuation fields for sparse-view CBCT reconstruction. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2022. pp. 442–452. Springer (2022)
16. Zhang, Z., Ehinger, K.A., Drummond, T.: Improving denoising diffusion models via simultaneous estimation of image and noise. arXiv preprint arXiv:2310.17167 (2023)
17. Zhang, Z., Hiller, M., Ehinger, K.A., Drummond, T.: Pixel-level residual diffusion transformer: Scalable 3d CT volume generation. In: The Fourteenth International Conference on Learning Representations (2026)