

Dose-Aware Cold Diffusion with Physics Consistency for Generalizable Low-Dose CT Reconstruction

Md Imam Ahasan¹ Guangchao Yang^{1*} A. F. M. Abdun Noor² S. M. Hasan Mahmud² Md Mahfuzur Rahman¹

¹College of Computer Science, Chongqing University, Chongqing, China

²Department of Software Engineering, Daffodil International University, Dhaka, Bangladesh

Abstract—Reducing radiation dose in computed tomography significantly degrades image quality and poses challenges for accurate and clinically reliable reconstruction. While recent approaches have shown promise for low-dose CT, they often struggle to generalize across continuous and previously unseen dose levels, leading to artifacts and loss of anatomical detail. To address these limitations, we propose Dose-Aware Cold Diffusion (DACD), a physics-consistent reconstruction framework that explicitly models radiation dose as a continuous latent factor within a cold diffusion process. The proposed DACD framework integrates image-based dose-aware perception, multi-scale structural prior extraction, and dose-calibrated step allocation to adaptively guide the denoising trajectory. In addition, an iterative forward-backprojection correction is incorporated into the reverse refinement process to enforce projection-domain data consistency. Extensive experiments on three public benchmarks, including Mayo-2020, Mayo-2016, and LoDoPaB-CT, demonstrate that DACD consistently outperforms state-of-the-art diffusion-based and physics-guided methods in both quantitative accuracy and visual fidelity, particularly under ultra-low-dose conditions. The results show that DACD achieves robust generalization across a continuous range of dose levels, including those unseen during training.

Index Terms—Low-dose CT, Cold Diffusion, physics-guided reconstruction, dose embedding, medical image reconstruction

I. INTRODUCTION

Computed tomography (CT) is central to clinical diagnosis, screening, and image-guided intervention. However, X-ray radiation exposure raises persistent safety concerns [1], [2]. In practice, dose reduction is commonly achieved by lowering tube current, which reduces photon statistics and yields increased noise and pronounced streak artifacts in reconstructed images. Fig. 1 illustrates how decreasing tube current induces distinct degradation patterns in low-dose CT (LDCT) reconstructions. Prior LDCT methods use model-based iterative and classical post-processing techniques [3]–[5], operating in either the sinogram domain to denoise projections before FBP or directly in the image domain. Representative techniques include penalized weighted least squares [6], bilateral filtering [7], and structural filtering [8], which aim to preserve structural fidelity under noisy acquisition or limited sampling. Deep learning (DL) has also been widely adopted for LDCT denoising [9], [10]. Most existing DL approaches are trained for a single dose level or a specific anatomical region, limiting their

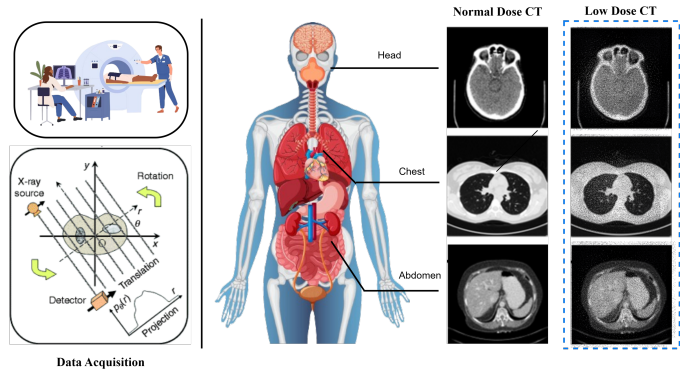


Fig. 1. Overview of CT imaging workflow and dose effects. Varying X-ray tube currents alter photon statistics and noise levels in reconstructed images, highlighting the importance of dose-aware reconstruction.

ability to handle heterogeneous noise characteristics and anatomical variability across protocols and scanners [11]. This is problematic in clinical deployment, where dose settings and reconstruction pipelines vary with patient factors, target anatomy, and manufacturer-specific configurations, even under the same nominal parameters [12]. Training separate models per dose or anatomy is feasible but computationally expensive and difficult to scale. Denoising diffusion probabilistic models (DDPMs) have recently achieved strong performance in image generation [13]–[15], with stable training and high synthesis quality [14], [16]–[18]. DDPMs learn to iteratively remove Gaussian noise through a denoising network, but standard sampling is slow. To accelerate diffusion-based CT reconstruction, fast ordinary differential equation (ODE) solvers have been used to reduce sampling steps [19], [20]. Building on cold diffusion, Gao *et al.* (2023) proposed the contextual error-modulated generalized diffusion (CoreDiff) model for LDCT denoising [21]–[23]. Subsequent work incorporated manifold constraints and prior features to improve score-based reconstruction under ultra-sparse sampling [24]–[26], and Dou *et al.* (2024) introduced a sequential Monte Carlo strategy for asymptotically accurate Bayesian posterior sampling with strong zero-shot performance [27]. Liu *et al.* (2024) learn a general CT denoiser across multiple degradations [28], but it lacks adaptive dose perception and requires costly fine-tuning for unseen data; dose-conditional methods improve generalization [29] yet treat dose as discrete and ignore

*Corresponding author: Guangchao Yang (gchao_yang@cqu.edu.cn).

This work was supported by the National Natural Science Foundation of China under Grant 62572086.

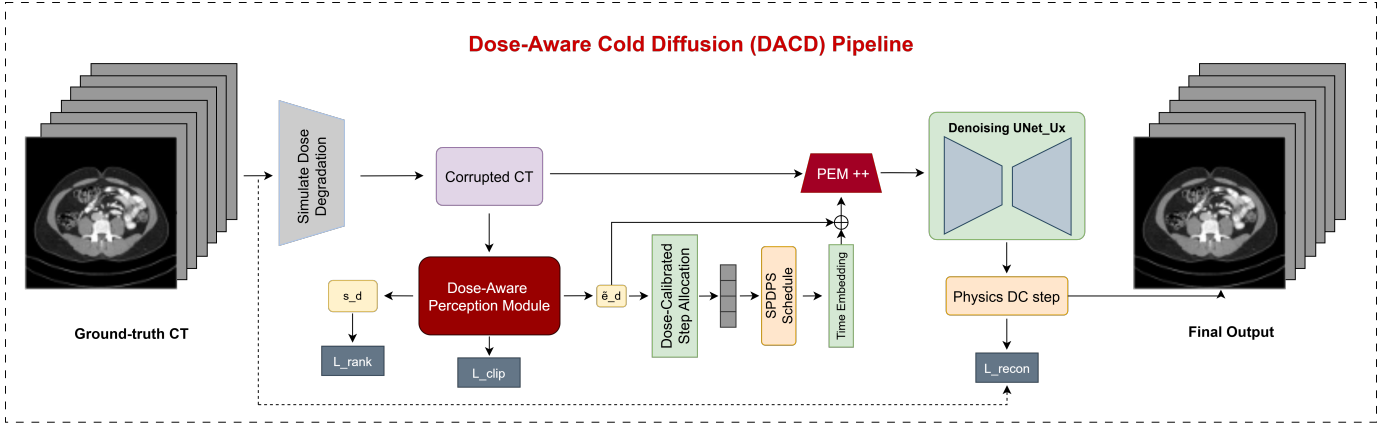


Fig. 2. Overview of the Dose-Aware Cold Diffusion (DACD) framework. The model combines physics-based consistency with adaptive diffusion denoising, guided by dose-aware embeddings that modulate reconstruction according to noise severity.

anatomical variability [30], motivating a continuous, anatomy-aware dose-conditioned approach.

In this work, we propose Dose-Aware Cold Diffusion (DACD), a reconstruction framework for LDCT that combines dose-conditioned diffusion modeling with physics-consistent corrections. DACD explicitly encodes dose effects, incorporates handcrafted structural priors, and enforces fidelity to measured sinograms through iterative forward-backprojection corrections within the reverse refinement process, enabling robust generalization to unseen dose levels. Extensive experiments on three benchmark CT datasets demonstrate the effectiveness of the proposed approach. For reproducibility, the implementation of DACD will be made publicly available at <https://github.com/imamahasane/DACD>. The primary contributions are:

- We propose DACD, a unified framework that models radiation dose as a continuous latent factor for generalizable LDCT reconstruction across a wide range of dose conditions.
- We introduce an in-loop physics consistency mechanism that integrates iterative forward-backprojection correction into each reverse refinement step, improving physical reliability and reducing hallucinated structures.
- We develop a dose-calibrated step allocation strategy that adapts the number of reverse refinement steps to degradation severity, improving inference efficiency while maintaining quality under ultra-low-dose settings.

II. RELATED WORK

Diffusion models have recently achieved strong performance in sparse-view and low-dose CT reconstruction by learning powerful iterative denoisers guided by the CT forward model. Cold diffusion methods replace synthetic Gaussian corruption with physically meaningful degradations, enabling direct inversion of CT-specific operators. Several works tailor diffusion to particular degradation types. CvG-Diff models angular subsampling artifacts for sparse-view CT using a degradation operator and efficient multi-phase sampling [31], while NEED introduces a Poisson-inspired diffusion process and guided

refinement to better match photon noise statistics in low-dose CT [12].

Other approaches pursue broader generalization across acquisition settings. FoundDiff leverages dose- and anatomy-aware pretraining to adapt a single denoiser to multiple anatomical regions and dose categories [32], and CT-SDM supports reconstruction across different sparse-view sampling rates within one model using projection-domain degradation and stochastic sampling [33]. PrideDiff incorporates sinogram-domain priors in a physics-regularized cold diffusion framework to improve stability and reconstruction quality [34]. In contrast to prior work that relies on discrete conditioning or outer-loop data consistency, we model radiation dose as a continuous latent variable that parameterizes the diffusion trajectory and embed forward-backprojection correction inside every reverse step. This tight coupling of dose-aware conditioning and in-loop physics consistency enables robust reconstruction and smooth generalization to unseen dose levels.

III. METHOD

A. Problem Formulation

Let $x_{gt} \in \mathbb{R}^{H \times W}$ denote the ground-truth attenuation map and A the CT forward projection operator. The ideal noiseless sinogram is

$$y_{\text{clean}} = A(x_{gt}). \quad (1)$$

At a reduced radiation dose level $d \in (0, 1]$, photon measurements follow Poisson statistics,

$$y_d \sim \text{Poisson}(d \cdot y_{\text{clean}}), \quad (2)$$

where smaller d yields higher relative noise. A standard analytic reconstruction from noisy projections is obtained by filtered backprojection (FBP),

$$x_{\text{FBP}} = A_{\text{FBP}}^T(y_d), \quad (3)$$

which is fast but highly sensitive to noise and undersampling. LDCT reconstruction seeks an estimate $\hat{x}$ that is both faithful to the measurements and structurally plausible, typically formulated as

$$\hat{x} = \arg \min_x \mathcal{L}(A(x), y_d) + \lambda \mathcal{R}(x), \quad (4)$$

where $\mathcal{L}$ enforces data fidelity to the measured sinogram and $\mathcal{R}$ encodes prior knowledge on image structure. The main challenges are to (i) generalize across continuous dose levels, (ii) preserve fine anatomical details under severe noise, and (iii) maintain consistency with the CT forward model. These requirements motivate a dose-aware, physics-consistent reconstruction framework.

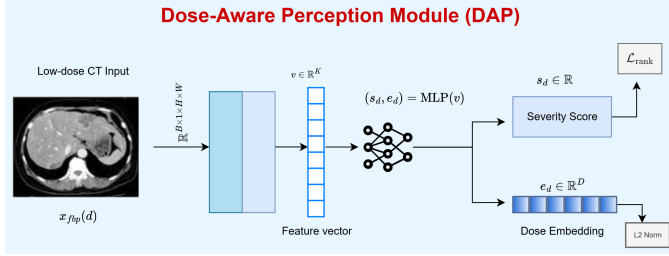


Fig. 3. Overview of the DAP module. DAP extracts feature embeddings from low-dose CTs to estimate dose severity and generate dose-conditioned representations for adaptive reconstruction.

B. Dose-Aware Cold Diffusion Framework

DACD formulates low-dose CT reconstruction as a dose-conditioned, physics-guided iterative refinement process with dose level $d \in (0, 1]$. As illustrated in Fig. 2, the framework combines dose-dependent degradation modeling, dose-aware feature conditioning, structural priors, and projection-domain data consistency within a unified cold diffusion scheme. Instead of adding synthetic Gaussian noise, the forward degradation is generated by the physical low-dose acquisition process.

Forward Process (Training): For a ground-truth attenuation map x_{gt} , the clean sinogram is

$$y_{\text{clean}} = A(x_{gt}). \quad (5)$$

A low-dose measurement is simulated by Poisson thinning,

$$y_d = \text{PoissonThin}(y_{\text{clean}}, d), \quad (6)$$

and reconstructed using filtered backprojection (FBP),

$$x_t(d) = A_{\text{FBP}}^T(y_d). \quad (7)$$

Varying d produces a sequence of deterministic degraded reconstructions $\{x_t(d)\}$, where smaller d yields more severe degradation. A monotonic mapping between dose level d and diffusion index t defines a physically grounded cold diffusion trajectory that the reverse process inverts.

Reverse Process (Inference): Given a measured low-dose sinogram y_d and its FBP reconstruction, DACD iteratively alternates learned denoising and physics-based correction:

$$x_{t-1} = U_\theta(x_t, c_t) - \eta_t A^T(A(\tilde{x}_{t-1}) - y_d), \quad (8)$$

where $\tilde{x}_{t-1} = U_\theta(x_t, c_t)$ is the denoised estimate, U_θ is a U-Net denoiser, η_t controls the correction strength, and c_t concatenates diffusion-time and dose embeddings. The second term enforces projection-domain data fidelity at every step, coupling the learned prior with the CT forward model.

Training Objective: Training combines image fidelity, dose-order consistency, and representation alignment:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{img}} + \lambda_1 \mathcal{L}_{\text{rank}} + \lambda_2 \mathcal{L}_{\text{contrast}}, \quad (9)$$

with

$$\mathcal{L}_{\text{img}} = \|x_{t-1} - x_{gt}\|_1 + \beta \|A(x_{t-1}) - y_{\text{clean}}\|_2^2. \quad (10)$$

This term enforces consistency in both image and projection domains. The ranking and contrastive losses from the dose-aware perception module impose ordinal consistency across dose levels and encourage discriminative dose-aware representations. Together, these constraints promote physically consistent and accurate reconstruction over a continuous range of radiation doses.

C. Dose-to-Step Mapping in Cold Diffusion

DACD adopts a cold diffusion formulation in which each diffusion step corresponds to a deterministic degradation induced by a specific radiation dose, directly reflecting the physics of low-dose CT. Let $d \in (0, 1]$ be the normalized dose fraction ($d = 1$ is full dose). We map dose to a degradation index $t \in \{1, \dots, T_{\text{max}}\}$ by

$$t(d) = \lfloor T_{\text{max}}(1 - d) \rfloor + 1, \quad (11)$$

so lower doses yield larger indices and more severe degradation. Given full-dose projections y_{full} , low-dose measurements and degraded images are generated as

$$y_d \sim \text{Poisson}(d \cdot y_{\text{full}}), \quad x_t = \mathcal{A}^\dagger(y_d), \quad (12)$$

producing a deterministic degradation state for each t . The image x_t initializes the reverse refinement. The degradation index $t(d)$ encodes noise severity and is distinct from the number of reverse refinement iterations. The latter is adaptively set by the dose-calibrated step allocation strategy $T(d)$, which assigns more iterations to lower-dose inputs. Any monotonic mapping between d and t preserves dose ordering and yields stable reconstruction due to dose-aware conditioning and in-loop physics-consistent correction, making this mapping a simple and physically interpretable design choice.

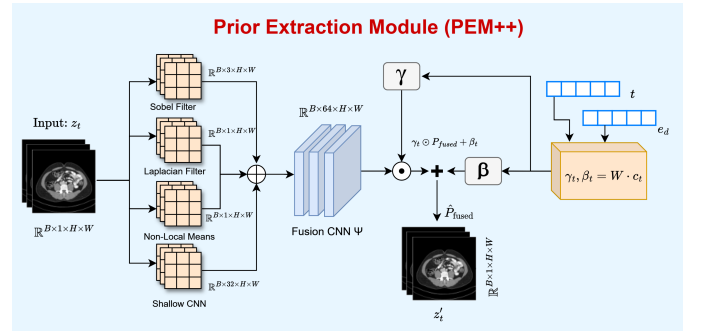


Fig. 4. Architecture of the Prior Extraction Module (PEM++). The module fuses multi-scale spatial priors from edge, texture, and CNN features, modulated by dose and time embeddings for adaptive reconstruction.

D. Core Modules

DACD consists of three components for dose-adaptive reconstruction: the Dose-Aware Perception module (DAP), the Prior Extraction module (PEM++), and the Dose-Calibrated Step Allocation mechanism (DCSA), which respectively enable

dose perception, structural prior modeling, and adaptive control of the reverse process.

1) Dose-Aware Perception (DAP): DAP estimates degradation severity and produces a continuous dose-conditioned embedding from the degraded image (Fig. 3). Given x_t , a lightweight encoder extracts latent features

$$z_t = f_\theta(x_t), \quad (13)$$

which are mapped by an MLP to a severity score and dose embedding,

$$(s_d, e_d) = \text{MLP}(z_t), \quad (14)$$

with ℓ_2 -normalized e_d for stable conditioning. Ordinal consistency across dose levels is enforced by a ranking loss

$$\mathcal{L}_{\text{rank}} = \frac{1}{N} \sum_{i < j} \max(0, m - (s_j - s_i)), \quad d_i < d_j, \quad (15)$$

and representation geometry is regularized by an InfoNCE-style contrastive loss

$$\mathcal{L}_{\text{contrast}} = -\log \frac{\exp(\langle z_i^+, z_i \rangle / \tau)}{\sum_k \exp(\langle z_k, z_i \rangle / \tau)}. \quad (16)$$

The conditioning token for the denoiser is formed by concatenating the timestep and dose embeddings,

$$c_t = [e_t; e_d], \quad e_t = \text{TimeEmbed}(t), \quad (17)$$

enabling continuous conditioning under unseen dose levels.

2) Prior Extraction (PEM++): PEM++ enhances structural fidelity by fusing complementary handcrafted and learned priors (Fig. 4). From x_t , multi-scale priors are computed as

$$\begin{aligned} P_{\text{grad}} &= \nabla x_t, & P_{\text{lap}} &= \Delta x_t, \\ P_{\text{nllm}} &= \text{NLM}(x_t), & P_{\text{cnn}} &= h_\psi(x_t), \end{aligned} \quad (18)$$

capturing edges, curvature, non-local similarity, and semantic context. These are combined by attention-based fusion

$$\alpha_k = \frac{\exp(Q^T K_k)}{\sum_j \exp(Q^T K_j)}, \quad P_{\text{fuse}} = \sum_k \alpha_k V_k, \quad (19)$$

with $Q = f_q(x_t)$ and $(K_k, V_k) = f_{k,v}(P_k)$ implemented by 1×1 convolutions. The fused prior is modulated by dose and timestep through FiLM:

$$P'_{\text{fuse}} = \gamma(d, t) \odot P_{\text{fuse}} + \beta(d, t), \quad (20)$$

$$[\gamma(d, t), \beta(d, t)] = W_{\text{film}}[e_d; e_t] + b_{\text{film}}, \quad (21)$$

and injected into the denoiser,

$$x'_t = U_\theta(x_t \oplus P'_{\text{fuse}}, c_t), \quad (22)$$

providing dose- and severity-adaptive structural guidance.

3) Dose-Calibrated Step Allocation (DCSA): DCSA adaptively sets the number of reverse refinement steps from the learned dose embedding. A bounded severity indicator is predicted as

$$\xi_d = \sigma(W_s e_d + b_s), \quad (23)$$

and the iteration count is

$$T(d) = T_{\text{max}} - \lfloor \xi_d (T_{\text{max}} - T_{\text{min}}) \rfloor. \quad (24)$$

Thus, lower-dose (more degraded) inputs receive more refinement steps, while near full-dose inputs require fewer, improving efficiency without sacrificing quality.

E. Physics Consistency Enforcement

To maintain consistency with the measured projections and suppress hallucinated structures, DACD applies a physics-based correction after each diffusion denoising step, enforcing sinogram-domain data fidelity throughout the reverse process.

Forward-Backprojection Correction: Let A and A^T denote the projection and backprojection operators. Given the intermediate denoised estimate $\tilde{x}_{t-1}$, the projection residual is backprojected to the image domain,

$$\Delta x_{t-1} = A^T(y_d - A(\tilde{x}_{t-1})), \quad (25)$$

and used to update the estimate as

$$x_{t-1} = \tilde{x}_{t-1} + \eta_t \Delta x_{t-1}, \quad (26)$$

where η_t controls the correction strength. This step is analogous to one iteration of algebraic reconstruction (SIRT), but is embedded inside every diffusion step, coupling learned denoising with measurement consistency.

Variational View: The fixed point of the iterative updates approximately minimizes

$$\hat{x} = \arg \min_x \|A(x) - y_d\|_2^2 + \lambda_p \|x - U_\theta(x, c)\|_2^2, \quad (27)$$

which balances projection-domain data fidelity with adherence to the diffusion prior U_θ . The first term enforces agreement with the measured sinogram, while the second keeps the solution close to the learned denoising manifold.

IV. EXPERIMENTAL SETUP AND IMPLEMENTATION

A. Datasets

We evaluate on Mayo-2016 [35], Mayo-2020 [36], and LoDoPaB-CT [37]. Mayo-2020 contains 300 full-dose volumetric scans (head, chest, abdomen). All DICOMs were converted to Hounsfield Units (HU) and clipped to $[-1024, 3072]$ HU. Clean parallel-beam projections were generated from full-dose images, and low-dose data were simulated by Poisson thinning at $d \in \{50\%, 25\%, 12.5\%, 5\%\}$, followed by FBP to obtain inputs $x_{\text{fbp}}(d)$. For Mayo-2016, provided quarter-dose scans were used as low-dose inputs and full-dose scans as ground truth with the same HU preprocessing. LoDoPaB-CT is a large-scale thoracic benchmark (33,525/3,552/3,528 train/val/test slices) with simulated low-dose parallel-beam projections; inputs were obtained via FBP from the provided sinograms and HU-clipped to the same range. Its scale and pulmonary variability provide a challenging test of generalization.

B. Implementation Details

All models were implemented in PyTorch with CUDA acceleration. Forward and backprojection were performed using the GPU-accelerated, differentiable TorchRadon library. Unless stated otherwise, filtered backprojection (FBP) used a Ram-Lak filter. CT slices were cropped to 384×384 patches to balance memory usage and context. Training used the AdamW optimizer with learning rate 1×10^{-4} , $\beta_1 = 0.9$, $\beta_2 = 0.999$, and weight decay 10^{-2} , together with cosine annealing and linear warm-up. The batch size was 8 per GPU. Models were trained for 100 epochs in Stage A and 200 epochs in Stage B on four NVIDIA RTX 4090 GPUs with automatic mixed precision enabled.

TABLE I

QUANTITATIVE COMPARISON ON MAYO-2020 ACROSS ANATOMIES AND DOSE LEVELS. BEST RESULTS ARE IN **BOLD** AND SECOND-BEST ARE UNDERLINED.
* INDICATES A SIGNIFICANT DIFFERENCE FROM THE BEST METHOD ($p < 0.05$, WILCOXON TEST).

Anatomy	Method	50%	25%	12.5%	5%
		PSNR / SSIM / RMSE	PSNR / SSIM / RMSE	PSNR / SSIM / RMSE	PSNR / SSIM / RMSE
Abdomen	PrideDiff*	41.24 / 0.9748 / 0.013	39.83 / 0.970 / 0.016	37.92 / 0.963 / 0.020	35.45 / 0.955 / 0.025
	Cold Diffusion*	39.36 / 0.969 / 0.015	38.07 / 0.964 / 0.018	36.18 / 0.957 / 0.022	34.01 / 0.948 / 0.028
	CoreDiff*	40.31 / 0.972 / 0.014	38.94 / 0.967 / 0.017	37.17 / 0.961 / 0.021	34.83 / 0.952 / 0.026
	RDDM*	39.75 / 0.970 / 0.015	38.32 / 0.965 / 0.018	36.63 / 0.958 / 0.022	34.34 / 0.949 / 0.027
	DDPM-1000*	39.05 / 0.967 / 0.016	37.64 / 0.962 / 0.019	35.82 / 0.955 / 0.023	33.63 / 0.946 / 0.029
	RED-diff*	38.57 / 0.964 / 0.017	37.14 / 0.959 / 0.020	35.36 / 0.952 / 0.024	33.23 / 0.943 / 0.030
	Noise2Sim*	38.08 / 0.962 / 0.018	36.62 / 0.956 / 0.021	34.93 / 0.949 / 0.025	32.85 / 0.940 / 0.031
	DACD (Ours)	42.02 / 0.9773 / 0.012	40.57 / 0.982 / 0.015	38.64 / 0.966 / 0.019	36.23 / 0.958 / 0.024
Chest	PrideDiff*	38.33 / 0.8719 / 0.018	36.95 / 0.862 / 0.021	35.03 / 0.853 / 0.025	32.72 / 0.842 / 0.030
	Cold Diffusion*	36.83 / 0.866 / 0.020	35.41 / 0.857 / 0.023	33.52 / 0.847 / 0.027	31.42 / 0.836 / 0.032
	CoreDiff*	37.63 / 0.869 / 0.019	36.20 / 0.860 / 0.022	34.34 / 0.851 / 0.026	32.03 / 0.840 / 0.031
	RDDM*	37.12 / 0.868 / 0.019	35.74 / 0.858 / 0.022	33.93 / 0.849 / 0.026	31.72 / 0.838 / 0.031
	DDPM-1000*	36.46 / 0.864 / 0.020	35.03 / 0.855 / 0.023	33.14 / 0.846 / 0.027	30.94 / 0.835 / 0.032
	RED-diff*	35.87 / 0.861 / 0.021	34.45 / 0.852 / 0.024	32.62 / 0.843 / 0.028	30.43 / 0.832 / 0.033
	Noise2Sim*	35.31 / 0.858 / 0.022	33.91 / 0.849 / 0.025	32.11 / 0.840 / 0.029	29.84 / 0.829 / 0.034
	DACD (Ours)	39.02 / 0.8768 / 0.017	37.61 / 0.869 / 0.020	35.84 / 0.861 / 0.024	33.53 / 0.850 / 0.029
Head	PrideDiff*	44.83 / 0.9794 / 0.011	43.11 / 0.975 / 0.013	41.03 / 0.970 / 0.016	38.23 / 0.963 / 0.021
	Cold Diffusion*	43.04 / 0.974 / 0.012	41.44 / 0.969 / 0.015	39.32 / 0.964 / 0.018	36.63 / 0.957 / 0.023
	CoreDiff*	44.05 / 0.977 / 0.012	42.35 / 0.972 / 0.014	40.21 / 0.967 / 0.017	37.52 / 0.960 / 0.022
	RDDM*	43.42 / 0.976 / 0.012	41.83 / 0.970 / 0.015	39.84 / 0.965 / 0.018	37.03 / 0.958 / 0.023
	DDPM-1000*	42.61 / 0.972 / 0.013	40.94 / 0.967 / 0.016	38.86 / 0.961 / 0.019	36.15 / 0.954 / 0.024
	RED-diff*	42.04 / 0.969 / 0.014	40.34 / 0.964 / 0.017	38.32 / 0.958 / 0.020	35.54 / 0.951 / 0.025
	Noise2Sim*	41.52 / 0.967 / 0.015	39.83 / 0.962 / 0.018	37.84 / 0.956 / 0.021	35.03 / 0.949 / 0.026
	DACD (Ours)	46.03 / 0.9821 / 0.010	44.35 / 0.978 / 0.012	42.24 / 0.972 / 0.015	39.42 / 0.965 / 0.020

C. Evaluation Metrics

Performance was measured using PSNR, SSIM, and RMSE between reconstruction $\hat{x}$ and ground truth x_{gt} :

$$\text{PSNR} = 10 \log_{10} \left(\frac{L^2}{\|\hat{x} - x_{gt}\|_2^2 / N} \right), \quad (28)$$

$$\text{SSIM} = \frac{(2\mu_{\hat{x}}\mu_{x_{gt}} + C_1)(2\sigma_{\hat{x}x_{gt}} + C_2)}{(\mu_{\hat{x}}^2 + \mu_{x_{gt}}^2 + C_1)(\sigma_{\hat{x}}^2 + \sigma_{x_{gt}}^2 + C_2)}, \quad (29)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \|\hat{x} - x_{gt}\|_2^2}, \quad (30)$$

where L is the HU dynamic range and N the number of pixels. Higher PSNR/SSIM and lower RMSE indicate better quality. Metrics were computed per slice. Statistical significance was assessed using paired Wilcoxon signed-rank tests with FDR control ($q = 0.05$).

V. RESULTS AND ANALYSIS

A. Results on the Mayo-2020 Dataset

DACD achieves the best performance across all anatomies and dose levels. In the abdomen, it improves PSNR by +0.78 dB over the second-best method at both 50% and 5% dose (42.02 vs. 41.24 dB and 36.23 vs. 35.45 dB), corresponding to up to 7.7% lower RMSE. In the chest, gains reach +0.81 dB

at 5% dose (33.53 vs. 32.72 dB) with approximately 5-6% RMSE reduction. The largest improvements are observed in the head region, where DACD provides up to +1.20 dB PSNR at 50% dose and +1.19 dB at 5% dose, together with up to 9.1% lower RMSE. Consistent advantages are also observed at intermediate doses, with SSIM improvements of up to +0.008. These results demonstrate stable and statistically significant gains across anatomies and noise levels, confirming the robustness of the proposed dose-aware and physics-consistent reconstruction framework. And figure. 5 provides a visual comparison across three representative regions abdomen, chest, and head. DACD preserves subtle texture and vessel details with minimal oversmoothing, whereas competing approaches (DDPM-1000 [15], RDDM [28], Cold Diffusion [38]) exhibit visible blurring or loss of low-contrast boundaries.

B. Cross-Dataset Evaluation on LoDoPaB-CT and Mayo-2016

Table II reports quantitative results on two out-of-distribution benchmarks, LoDoPaB-CT and Mayo-2016. On LoDoPaB-CT, DACD achieves the best accuracy with 38.74 dB PSNR and 0.8951 SSIM, surpassing PrideDiff (37.92 dB / 0.891) and CoreDiff (37.40 dB / 0.888), while obtaining the lowest RMSE of 0.0281. This corresponds to PSNR gains of +0.82 dB

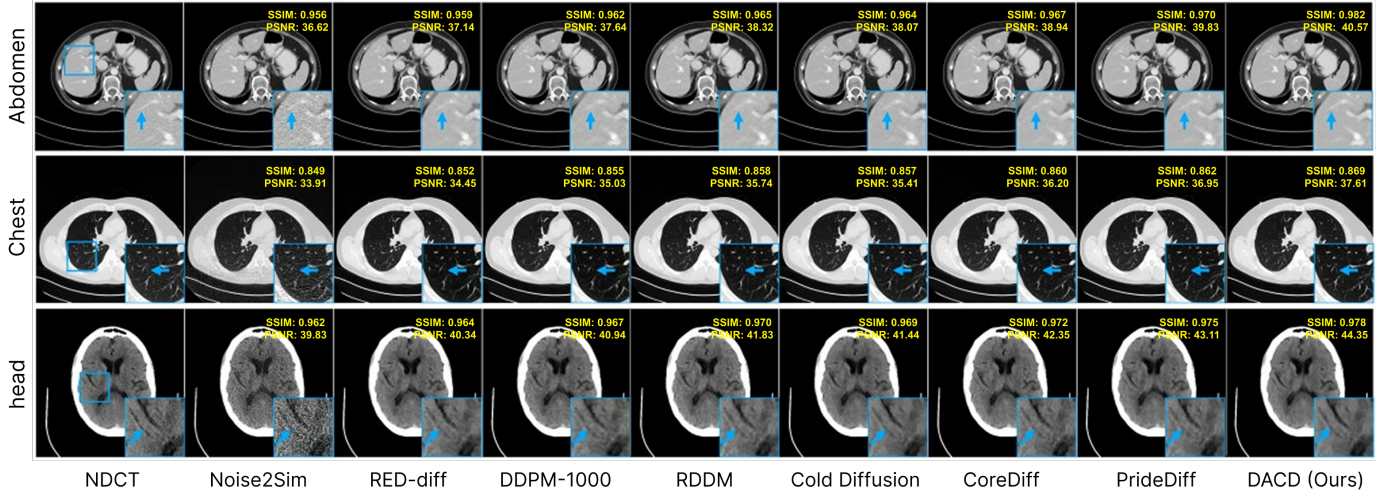


Fig. 5. Qualitative comparison of reconstructed CT images from different methods on the Mayo 2020 dataset. Zoomed-in regions highlight that the proposed DACD framework achieves better texture and noise suppression compared to other approaches.

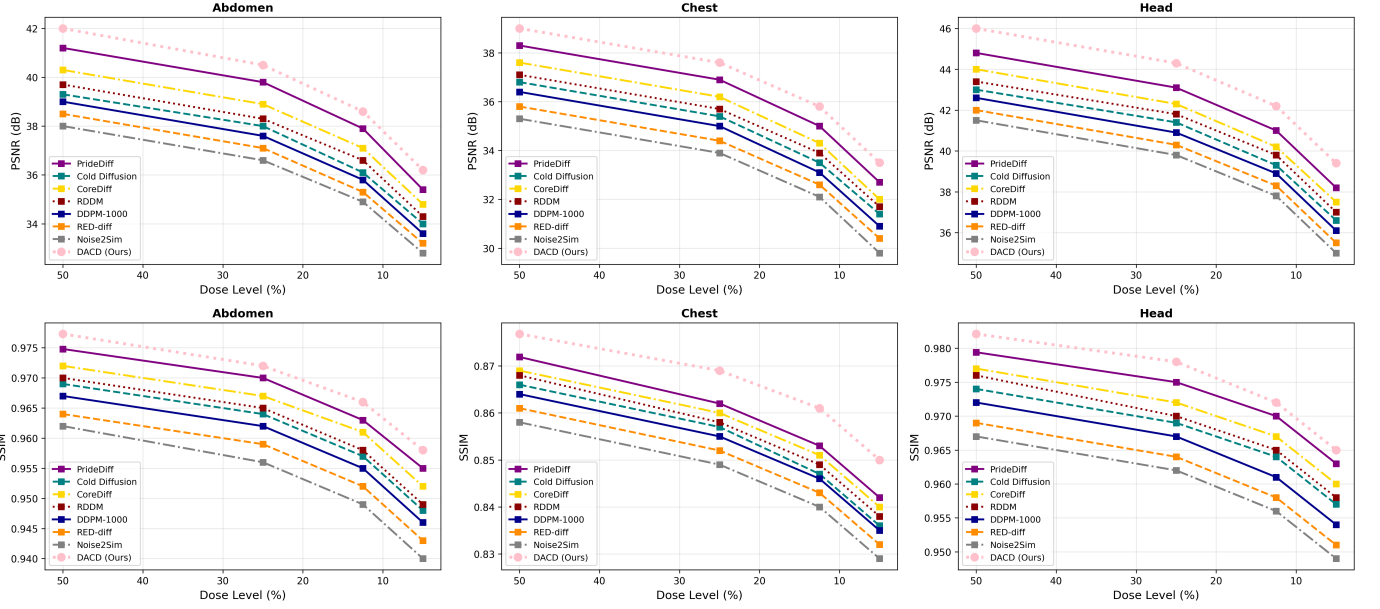


Fig. 6. Quantitative performance across anatomical regions and dose levels on Mayo-2020. The proposed DACD (pink line) consistently achieves the highest PSNR and SSIM across anatomical regions, reflecting robust dose adaptability and structural fidelity under severe noise conditions.

($\approx 2.2\%$) over PrideDiff and $+1.34\text{dB}$ ($\approx 3.6\%$) over CoreDiff, together with RMSE reductions of about 4.1% and 5.7%, demonstrating robustness to unseen dose distributions and acquisition geometries. A similar pattern is observed on Mayo-2016, where DACD reaches 44.36 dB PSNR and 0.9724 SSIM, outperforming PrideDiff (43.72 dB / 0.9693) and CoreDiff (43.25 dB / 0.9670). These correspond to PSNR improvements of $+0.64\text{dB}$ ($\approx 1.5\%$) and $+1.11\text{dB}$ ($\approx 2.6\%$), along with the lowest RMSE of 0.0243, indicating higher fidelity and lower variance across slices. The reduced dispersion suggests more stable and reliable reconstructions under diverse anatomical and noise conditions, consistent with clinical-quality imaging requirements.

C. Results on Unseen Dose Levels

To assess extrapolation to dose levels not seen during training, we evaluate at 30% (abdomen), 10% (chest), and 15% (head). DACD consistently achieves the highest PSNR/SSIM and lowest RMSE across all cases. At the unseen 10% chest dose, DACD attains 37.8 dB PSNR and 0.870 SSIM, clearly surpassing PrideDiff and CoreDiff, indicating strong robustness to dose mismatch. Across regions, DACD preserves fine structures while suppressing noise-induced streaks, as illustrated in Fig. 7. These results show that dose-aware conditioning with in-loop physics consistency enables reliable generalization beyond the trained dose range. While diffusion baselines such as CoreDiff and RDDM remain competitive at

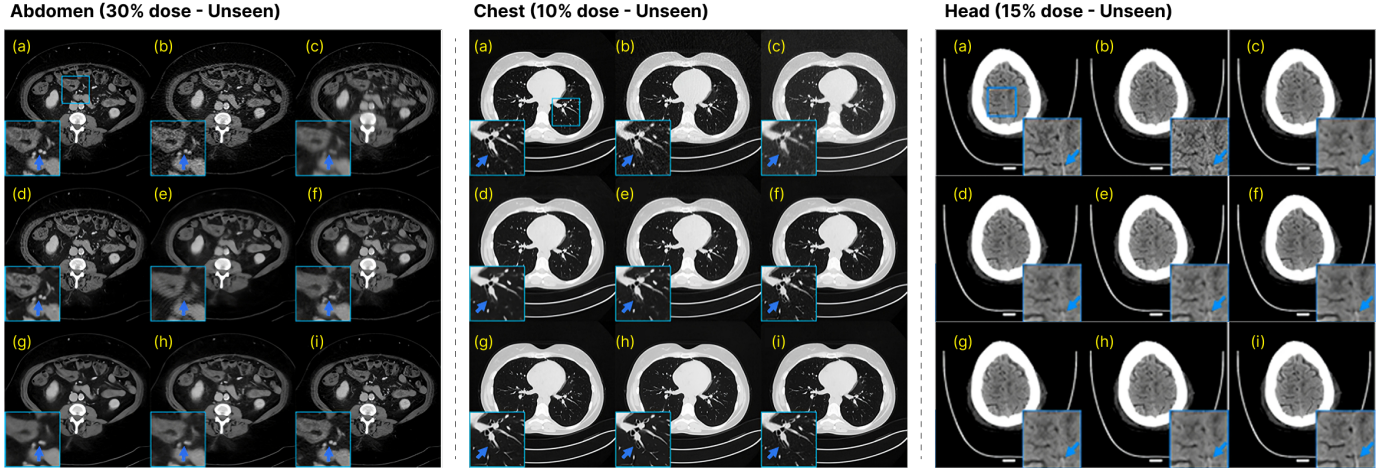


Fig. 7. Visual results on unseen dose levels across abdomen (30%), chest (10%), and head (15%) cases. DACD (h) maintains sharper tissue boundaries and more realistic contrast than competing methods, demonstrating strong generalization to dose conditions not encountered during training.

TABLE II

QUANTITATIVE RESULTS ON LoDoPAB-CT AND MAYO-2016. MEAN $\pm$ STANDARD DEVIATION OF PSNR (dB), SSIM, AND RMSE. BOLD AND UNDERLINED VALUES DENOTE THE BEST AND SECOND-BEST RESULTS. * INDICATES DIFFERENCE FROM THE BEST METHOD ($p < 0.05$, WILCOXON TEST).

Method	LoDoPAB-CT			Mayo-2016		
	PSNR (dB)	SSIM	RMSE	PSNR (dB)	SSIM	RMSE
Cold Diffusion	36.10 $\pm$ 0.41*	0.883 $\pm$ 0.0041*	0.0313 $\pm$ 0.0009*	41.85 $\pm$ 0.42*	0.9632 $\pm$ 0.0031*	0.0278 $\pm$ 0.0009*
RDDM	36.25 $\pm$ 0.38*	0.885 $\pm$ 0.0038*	0.0310 $\pm$ 0.0008*	42.10 $\pm$ 0.38*	0.9641 $\pm$ 0.0028*	0.0273 $\pm$ 0.0008*
DDPM-1000	35.85 $\pm$ 0.45*	0.880 $\pm$ 0.0043*	0.0318 $\pm$ 0.0010*	41.50 $\pm$ 0.47*	0.9610 $\pm$ 0.0035*	0.0281 $\pm$ 0.0010*
RED-diff	35.55 $\pm$ 0.48*	0.876 $\pm$ 0.0046*	0.0324 $\pm$ 0.0011*	40.95 $\pm$ 0.49*	0.9582 $\pm$ 0.0039*	0.0290 $\pm$ 0.0011*
Noise2Sim	35.05 $\pm$ 0.50*	0.871 $\pm$ 0.0048*	0.0333 $\pm$ 0.0012*	40.40 $\pm$ 0.52*	0.9547 $\pm$ 0.0041*	0.0301 $\pm$ 0.0013*
CoreDiff	37.40 $\pm$ 0.34*	0.888 $\pm$ 0.0035*	0.0298 $\pm$ 0.0007*	43.25 $\pm$ 0.33*	0.9670 $\pm$ 0.0025*	0.0256 $\pm$ 0.0007*
PrideDiff	37.92 $\pm$ 0.29*	0.891 $\pm$ 0.0031*	0.0293 $\pm$ 0.0006*	43.72 $\pm$ 0.29*	0.9693 $\pm$ 0.0023*	0.0251 $\pm$ 0.0006*
DACD (Ours)	38.74 $\pm$ 0.26	0.8951 $\pm$ 0.0029	0.0281 $\pm$ 0.0005	44.36 $\pm$ 0.27	0.9724 $\pm$ 0.0021	0.0243 $\pm$ 0.0005

moderate noise, conventional denoisers (e.g., RED-CNN [9] and Noise2Sim [39]) exhibit larger drops, particularly in SSIM, highlighting limited cross-dose generalization.

D. Ablation Study

Table III quantifies the contribution of each component on Mayo-2020 at 12.5% dose. Starting from the cold diffusion baseline (37.10 dB / 0.961), adding the dose-aware perception (DAP) module yields a +0.80 dB PSNR gain ($\approx 2.2\%$) and a +0.004 SSIM increase. Incorporating the structural prior module (PEM++) further improves PSNR to 38.70 dB, providing an additional +0.80 dB ($\approx 2.1\%$) and raising SSIM to 0.968. Introducing dose-calibrated step allocation (DCSA) boosts performance to 39.25 dB and 0.970, corresponding to a cumulative +2.15 dB ($\approx 5.8\%$) improvement over the baseline. Finally, adding the physics-consistency (PC) correction achieves the best result of 39.80 dB / 0.972, giving a total gain of +2.70 dB ($\approx 7.3\%$) and +0.011 SSIM relative to the baseline. These monotonic improvements demonstrate that each module contributes complementary benefits, with the largest cumulative effect arising from the combination of dose awareness, structural priors, adaptive step allocation, and projection-domain consistency.

TABLE III

ABLATION STUDY ON THE MAYO-2020 DATASET AT 12.5% DOSE.

Configuration	PSNR (dB)	SSIM
Cold Diffusion (baseline)	37.10	0.961
+ DAP	37.90	0.965
+ DAP + PEM++	38.70	0.968
+ DAP + PEM++ + DCSA	39.25	0.970
Full DACD (DAP + PEM++ + DCSA + PC)	39.80	0.972

VI. CONCLUSION

This work presented DACD, a physics-consistent framework for generalizable low-dose CT reconstruction. By modeling radiation dose as a continuous latent factor and integrating dose-aware perception, dose-calibrated step allocation, and in-loop forward-backprojection correction, DACD achieves stable and reliable reconstruction across a wide range of dose levels, including previously unseen conditions. Experimental results on multiple public benchmarks demonstrate that the proposed framework consistently improves reconstruction accuracy and structural fidelity under severe noise. These

findings indicate that jointly enforcing dose-aware conditioning and projection-domain consistency within the diffusion process is an effective strategy for robust low-dose CT reconstruction. Future work will explore anatomy-aware representations and further improve cross-scanner generalization and inference efficiency to facilitate practical clinical deployment.

REFERENCES

- [1] R. Smith-Bindman, J. Lipson, R. Marcus, K.-P. Kim, M. Mahesh, R. Gould, A. B. De González, and D. L. Miglioretti, "Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer," *Archives of internal medicine*, vol. 169, no. 22, pp. 2078–2086, 2009.
- [2] A. Sodickson, P. F. Baeyens, K. P. Andriole, L. M. Prevedello, R. D. Nawfel, R. Hanson, and R. Khorasani, "Recurrent ct, cumulative radiation exposure, and associated radiation-induced cancer risks from ct of adults," *Radiology*, vol. 251, no. 1, pp. 175–184, 2009.
- [3] J. Wang, T. Li, H. Lu, and Z. Liang, "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose x-ray computed tomography," *IEEE transactions on medical imaging*, vol. 25, no. 10, pp. 1272–1283, 2006.
- [4] J. Ma, J. Huang, Q. Feng, H. Zhang, H. Lu, Z. Liang, and W. Chen, "Low-dose computed tomography image restoration using previous normal-dose scan," *Medical physics*, vol. 38, no. 10, pp. 5713–5731, 2011.
- [5] Q. Xie, D. Zeng, Q. Zhao, D. Meng, Z. Xu, Z. Liang, and J. Ma, "Robust low-dose ct sinogram preprocessing via exploiting noise-generating mechanism," *IEEE transactions on medical imaging*, vol. 36, no. 12, pp. 2487–2498, 2017.
- [6] K. Zha, P. Cao, J. Son, Y. Yang, and D. Katabi, "Rank-n-contrast: learning continuous representations for regression," *Advances in Neural Information Processing Systems*, vol. 36, pp. 17 882–17 903, 2023.
- [7] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, "Supervised contrastive learning," *Advances in neural information processing systems*, vol. 33, pp. 18 661–18 673, 2020.
- [8] W. Peebles and S. Xie, "Scalable diffusion models with transformers," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2023, pp. 4195–4205.
- [9] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, "Low-dose ct with a residual encoder-decoder convolutional neural network," *IEEE transactions on medical imaging*, vol. 36, no. 12, pp. 2524–2535, 2017.
- [10] J. Chi, Z. Sun, L. Meng, S. Wang, X. Yu, X. Wei, and B. Yang, "Low-dose ct image super-resolution with noise suppression based on prior degradation estimator and self-guidance mechanism," *IEEE Transactions on Medical Imaging*, 2024.
- [11] Z. Chen, Q. Gao, Y. Zhang, and H. Shan, "Ascon: Anatomy-aware supervised contrastive learning framework for low-dose ct denoising," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 355–365.
- [12] Q. Gao, Z. Chen, D. Zeng, J. Zhang, J. Ma, and H. Shan, "Noise-inspired diffusion model for generalizable low-dose ct reconstruction," *arXiv preprint arXiv:2506.22012*, 2025.
- [13] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in neural information processing systems*, vol. 33, pp. 6840–6851, 2020.
- [14] A. Q. Nichol and P. Dhariwal, "Improved denoising diffusion probabilistic models," in *International conference on machine learning*. PMLR, 2021, pp. 8162–8171.
- [15] J. Choi, S. Kim, Y. Jeong, Y. Gwon, and S. Yoon, "Ilvr: Conditioning method for denoising diffusion probabilistic models," *arXiv preprint arXiv:2108.02938*, 2021.
- [16] P. Dhariwal and A. Nichol, "Diffusion models beat gans on image synthesis," *Advances in neural information processing systems*, vol. 34, pp. 8780–8794, 2021.
- [17] G. Müller-Franzes, J. M. Niehues, F. Khader, S. T. Arasteh, C. Haarbuerger, C. Kuhl, T. Wang, T. Han, T. Nolte, S. Nebelung, J. N. Kather, and D. Truhn, "A multimodal comparison of latent denoising diffusion probabilistic models and generative adversarial networks for medical image synthesis," *Scientific Reports*, vol. 13, Jul. 2023.
- [18] C. Saharia, W. Chan, H. Chang, C. Lee, J. Ho, T. Salimans, D. Fleet, and M. Norouzi, "Palette: Image-to-image diffusion models," in *ACM SIGGRAPH 2022 conference proceedings*, 2022, pp. 1–10.
- [19] W. Xia, Q. Lyu, and G. Wang, "Low-dose ct using denoising diffusion probabilistic model for 20 x speedup," *arXiv preprint arXiv:2209.15136*, 2022.
- [20] C. Lu, Y. Zhou, F. Bao, J. Chen, C. Li, and J. Zhu, "Dpm-solver: A fast ode solver for diffusion probabilistic model sampling in around 10 steps," *Advances in neural information processing systems*, vol. 35, pp. 5775–5787, 2022.
- [21] A. Bansal, E. Borgia, H.-M. Chu, J. Li, H. Kazemi, F. Huang, M. Goldblum, J. Geiping, and T. Goldstein, "Cold diffusion: Inverting arbitrary image transforms without noise," *Advances in Neural Information Processing Systems*, vol. 36, pp. 41 259–41 282, 2023.
- [22] Q. Gao, Z. Li, J. Zhang, Y. Zhang, and H. Shan, "Corediff: Contextual error-modulated generalized diffusion model for low-dose ct denoising and generalization," *IEEE Transactions on Medical Imaging*, vol. 43, no. 2, pp. 745–759, 2023.
- [23] M. Mardani, J. Song, J. Kautz, and A. Vahdat, "A variational perspective on solving inverse problems with diffusion models," *arXiv preprint arXiv:2305.04391*, 2023.
- [24] H. Chung, B. Sim, D. Ryu, and J. C. Ye, "Improving diffusion models for inverse problems using manifold constraints," *Advances in Neural Information Processing Systems*, vol. 35, pp. 25 683–25 696, 2022.
- [25] H. Chung, S. Lee, and J. C. Ye, "Decomposed diffusion sampler for accelerating large-scale inverse problems," *arXiv preprint arXiv:2303.05754*, 2023.
- [26] H. Chung, D. Ryu, M. T. McCann, M. L. Klasky, and J. C. Ye, "Solving 3d inverse problems using pre-trained 2d diffusion models," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 22 542–22 551.
- [27] Z. Dou and Y. Song, "Diffusion posterior sampling for linear inverse problem solving: A filtering perspective," in *The Twelfth International Conference on Learning Representations*, 2024.
- [28] J. Liu, Q. Wang, H. Fan, Y. Wang, Y. Tang, and L. Qu, "Residual denoising diffusion models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 2773–2783.
- [29] W. Xia, Z. Lu, Y. Huang, Y. Liu, H. Chen, J. Zhou, and Y. Zhang, "Ct reconstruction with pdf: Parameter-dependent framework for data from multiple geometries and dose levels," *IEEE Transactions on Medical Imaging*, vol. 40, no. 11, pp. 3065–3076, 2021.
- [30] B. R. Mussmann, S. D. Mørup, P. M. Skov, S. Foley, A. S. Brenøe, F. Eldahl, G. M. Jørgensen, and H. Precht, "Organ-based tube current modulation in chest ct: a comparison of three vendors," *Radiography*, vol. 27, no. 1, pp. 1–7, 2021.
- [31] J. Chen, Y. Lin, Y. Qin, H. Wang, and X. Li, "Cross-view generalized diffusion model for sparse-view ct reconstruction," *arXiv preprint arXiv:2508.10313*, 2025.
- [32] Z. Chen, Q. Gao, Z. Li, J. Zhang, Y. Zhang, J. Zhao, and H. Shan, "Founddiff: Foundational diffusion model for generalizable low-dose ct denoising," *arXiv preprint arXiv:2508.17299*, 2025.
- [33] L. Yang, J. Huang, G. Yang, and D. Zhang, "Ct-sdm: A sampling diffusion model for sparse-view ct reconstruction across various sampling rates," *IEEE Transactions on Medical Imaging*, 2025.
- [34] Z. Lu, Q. Gao, T. Wang, Z. Yang, Z. Wang, H. Yu, H. Chen, J. Zhou, H. Shan, and Y. Zhang, "Priddiff: Physics-regularized generalized diffusion model for ct reconstruction," *IEEE Transactions on Radiation and Plasma Medical Sciences*, 2024.
- [35] C. H. McCollough, A. C. Bartley, R. E. Carter, B. Chen, T. A. Drees, P. Edwards, D. R. Holmes III, A. E. Huang, F. Khan, S. Leng *et al.*, "Low-dose ct for the detection and classification of metastatic liver lesions: results of the 2016 low dose ct grand challenge," *Medical physics*, vol. 44, no. 10, pp. e339–e352, 2017.
- [36] T. R. Moen, B. Chen, D. R. Holmes III, X. Duan, Z. Yu, L. Yu, S. Leng, J. G. Fletcher, and C. H. McCollough, "Low-dose ct image and projection dataset," *Medical physics*, vol. 48, no. 2, pp. 902–911, 2021.
- [37] J. Leuschner, M. Schmidt, D. O. Bagger, and P. Maass, "Lodopab-ct, a benchmark dataset for low-dose computed tomography reconstruction," *Scientific Data*, vol. 8, no. 1, p. 109, 2021.
- [38] H. Yen, F. G. Germain, G. Wichern, and J. Le Roux, "Cold diffusion for speech enhancement," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5.
- [39] C. Niu, M. Li, F. Fan, W. Wu, X. Guo, Q. Lyu, and G. Wang, "Noise suppression with similarity-based self-supervised deep learning," *IEEE transactions on medical imaging*, vol. 42, no. 6, pp. 1590–1602, 2022.