Physics-Informed SDE-Based Diffusion for X-ray Image Deblurring

Seokwon Oh, Seungjun Yoo, Junho Lee, Yoonsang Hong, Jeonghyeon Kim, and Ho Kyung Kim*
Computational X-ray Imaging Laboratory, School of Mechanical Engineering, Pusan Nat'l Univ., 2, Busandaehakro 63beon-gil, Busan 46241
**Corresponding author: hokyung@pusan.ac.kr*

*Keywords: Physics-informed diffusion, X-ray image deblurring, Non-destructive testing

1. Introduction

The fundamental spatial resolution of indirect X-ray imaging systems is inherently constrained by the physical properties of the scintillator layer. While a thicker phosphor screen is desirable to maximize X-ray absorption efficiency and reduce exposure time, it inevitably induces severe spatial blurring due to the lateral scattering of optical photons. This irreversible degradation of spatial frequencies is mathematically characterized by the modulation transfer function (MTF), accompanied by inherent quantum mottle described by the noise power spectrum (NPS). Consequently, recovering sharp, high-resolution (HR) responses from highly efficient but blurred low-resolution (LR) measurements has been a longstanding and severely ill-posed inverse problem in X-ray imaging.

Conventional attempts to reverse this physical blur, such as Wiener deconvolution, are fundamentally limited by the noise-amplification problem. Recently, deep learning-based regression models have been extensively deployed for image restoration, However, these regression approaches are notoriously prone to the regression-to-the-mean problem, penalizing high-frequency details and resulting in over-smoothed textures that fail to recover micro-structural defects critical for non-destructive testing (NDT).

To overcome the limitations of deterministic regression, generative diffusion models, including Denoising Diffusion Probabilistic Models [1, 2] and score-based Stochastic Differential Equations (SDEs) [3], have recently emerged as state-of-the-art priors for complex image restoration tasks. These models learn to reverse a stochastic degradation process, offering capabilities in recovering realistic high-frequency details through iterative refinement. Inspired by these advancements, we propose a physics-informed diffusion framework tailored to address the deterministic physical blur inherent in X-ray systems.

In standard diffusion formulations, the reverse sampling phase typically relies on estimating the injected noise to gradually recover the clean signal. While highly effective for general image generation, we take a different architectural approach to optimally align with the physics of X-ray deblurring. Specifically, we formulate the reverse diffusion process using direct target prediction ($\hat{\mathbf{x}}_0$). In X-ray imaging systems, the MTF naturally decays to near-zero at high spatial frequencies. Consequently, recovering the signal through frequency-domain inversion can lead to numerical instability. By training the network to directly estimate the uncorrupted HR signal at each integration step, the proposed sampling algorithm operates without requiring this inversion.

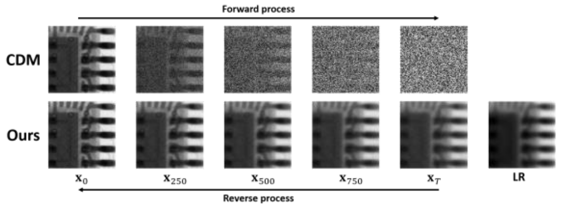

Fig. 1. Forward and reverse process of CDM and Ours.

2. Methods and Materials

2.1. Physics-Informed Forward SDE based on Detector Characteristics

Standard diffusion models typically simulate degradation through isotropic Gaussian noise injection. In this study, to explicitly account for the physical characteristics of X-ray detectors, we formulate the degradation as a physics-informed SDE as Fig. 1. The degradation encompasses frequency-dependent spatial blurring dictated by the MTF and the accumulation of quantum noise governed by the normalized NPS (NNPS). Let $\mathbf{X}_0$ be the HR ground-truth image in the frequency domain. The forward diffusion process at a continuous time variable $t \in [0, 1]$ is defined directly in the Fourier domain as:

$$\mathbf{X}_t = \mathbf{B}_t\mathbf{X}_0 + \mathbf{Z}\sqrt{\mathbf{\Sigma}_t}, \qquad (1)$$

where $\mathbf{B}_t$ represents the time-dependent blur operator denoting the ratio of the MTFs between the low-resolution (LR) and HR detectors, explicitly defined as $((\mathrm{MTF}_{\mathrm{LR}}/\mathrm{MTF}_{\mathrm{LR}})^{t/T})$. $\mathbf{Z}$ is standard complex Gaussian noise in the frequency domain.

We model the initial state $\mathbf{X}_0$ as a physical realization that contains baseline quantum mottle NNPS_0 corresponding to the HR detector. To maintain thermodynamic consistency and avoid non-physical noise overlap at the initial boundary condition, the noise variance schedule $\mathbf{\Sigma}_t$ is derived as:

$$\mathbf{\Sigma}_t = \mathbf{q}(1 - \mathbf{B}_t^2), \qquad (2)$$

where $\mathbf{q}$ is the target noise spectrum map calibrated from the LR detector. This formulation ensures that the degradation trajectory naturally evolves from the intrinsic HR response to the degraded LR condition.

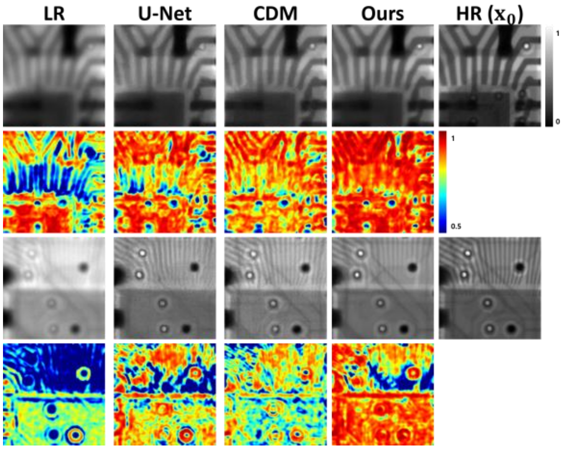

Fig. 2. Comparison of deblurring results. The first and third rows show the deblurring results of each method, while the second and fourth rows present the corresponding SSIM maps with respect to the HR images. The bounding boxes in the LR images indicate the regions where SNR was calculated.

2.2. Reverse Process

In standard diffusion models, the reverse process is often parameterized to predict the injected noise. While conventional diffusion models can be successfully adapted for deblurring tasks, recovering the uncorrupted signal from noise predictions typically involves frequency-domain inversion by the blur operator $\mathbf{B}_t$. Because the X-ray MTF naturally decays to near-zero at high spatial frequencies, this inversion step can introduce numerical instability during the sampling process.

To alleviate this, our model is formulated to directly estimate the uncorrupted spatial domain image ($\mathbf{x}_0$). A conditional U-Net [4] architecture, $f_\theta(\mathbf{x}_t, \mathbf{y}, t)$, predicts the HR image $\hat{\mathbf{x}}_0$ directly from the degraded state $\mathbf{x}_t$ and the LR condition $\mathbf{y}$. During the reverse sampling within the denoising diffusion implicit models (DDIM) framework [5].

2.3. Data Preparation and Training Details

Training and evaluation were performed on X-ray images of PCBs. Data were obtained using a tungsten-target X-ray tube (70 kV, 42.56 μGy) and coupled to phosphor-based CMOS detectors. HR and LR images were acquired by varying the scintillator thickness (33.91 mg·cm^{-2} for HR, 134.55 mg·cm^{-2} for LR).

To construct the HR ground-truth, 20 repeated measurements were averaged. This procedure reduced transient noise while preserving the base structural features and the fundamental NNPS_0

2.4. Performance evaluation

To evaluate the restoration performance, metrics from both the spatial and frequency domains were analyzed. In the spatial domain, peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were calculated with respect to the HR reference images. In the frequency domain, the MTF, NNPS, and contrast transfer function (CTF) were evaluated.

For benchmarking, three baseline approaches were implemented: regression-based model (U-Net), and CDM. To ensure a fair comparison, all learning-based methods adopted the same U-Net architecture.

3. Preliminary Results

Preliminary experiments on PCB X-ray images showed that the proposed physics-informed SDE diffusion model achieved the best overall restoration performance among the compared methods. In quantitative evaluation, our method yielded the highest SSIM (0.87 ± 0.06) and PSNR (22.65 ± 3.80 dB), outperforming the LR input, CDM, and the regression-based U-Net. Although U-Net achieved a comparable PSNR, its lower SSIM indicates inferior structural preservation. Visual comparisons in Fig. 2 further show that the proposed method more effectively restored fine circuit patterns and edge details while suppressing over-smoothing and residual artifacts, and the SSIM maps confirmed that structural similarity to the HR reference was consistently improved across complex high-frequency regions.

4. Conclusion

In this study, we proposed a physics-informed diffusion framework for X-ray image deblurring that incorporates detector-dependent blur and noise characteristics into the forward SDE and employs direct clean-image prediction in the reverse process. By avoiding unstable frequency-domain inversion and explicitly reflecting the physical imaging process, the proposed method provided more accurate and stable restoration than conventional regression and diffusion-based baselines. These results suggest that physics-informed generative restoration is a promising solution for recovering high-resolution details in indirect X-ray imaging and may be extended to practical non-destructive testing applications requiring both high sensitivity and spatial resolution.

ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00340520). S. Oh was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. RS-2024-00408137).

REFERENCES

[1] J. Ho, A. Jain, and P. Abbeel, Denoising Diffusion Probabilistic Models, Advances in Neural Information Processing Systems, Vol.33, p.6840, 2020.

[2] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi, Image Super-Resolution via Iterative Refinement, 2022.

[3] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, Score-Based Generative Modeling through Stochastic Differential Equations, International Conference on Learning Representations (ICLR), May 3–7, 2021.

[4] O. Ronneberger, P. Fischer, and T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015, p.234, 2015.

[5] J. Song, C. Meng, and S. Ermon, Denoising Diffusion Implicit Models, International Conference on Learning Representations (ICLR), May 3–7, 2021.