Restoration of X-ray detector responses using a diffusion-inspired deblur model

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

*Keywords: X-ray detectors, diffusion-inspired model, image restoration, industrial inspection

1. Introduction

The demand for reliable inspection of advanced electronic components, including multilayer printed circuit boards (PCBs) and lithium-ion batteries, has grown with the expansion of high-tech industries. Detecting micro-defects in complex internal structures is critical for safety but remains challenging. X-ray imaging is widely used in non-destructive testing (NDT) due to its ability to visualize interiors without damage. In indirect detectors, however, a trade-off arises: thick phosphors improve X-ray absorption efficiency but blur images due to lateral scattering, while thin phosphors provide sharper resolution but lower absorption efficiency, reducing industrial throughput.

Given these physical constraints, post-processing approaches have been investigated to restore image quality from degraded detector outputs. Deep learning has shown promise in image restoration tasks such as super-resolution, deblurring, and denoising. However, most existing models are optimized for natural images and do not account for the detector-specific blur characteristics in X-ray systems. In indirect imaging, spatial resolution is fundamentally determined by the modulation transfer function (MTF), which describes how different spatial frequencies are transmitted through the scintillator. Conventional regression-based networks often fail to recover fine high-frequency details, while generative approaches can introduce inconsistencies that reduce their reliability in safety-critical inspection.

Recently, diffusion models [1] have emerged as powerful tools for image restoration, demonstrating strong performance in tasks such as super-resolution, inpainting, and deblurring. Their iterative refinement process allows for the recovery of details beyond conventional regression methods. However, the stochastic nature of standard diffusion models can also generate spurious information, raising concerns for applications in industrial defect detection where fidelity to the true signal is essential.

In this work, we propose a physics-informed diffusion framework that integrates the measured MTFs of low-resolution (LR) and high-resolution (HR) detectors into the degradation and restoration process. By modeling blur as a deterministic operator in the frequency domain, the network is guided to restore HR-like responses from LR inputs. The proposed method is benchmarked against Wiener filtering [2], a U-Net [3] variant regression baseline, and a conditional diffusion model (CDM) [4]. Performance is evaluated in both the spatial domain and the frequency domain, enabling a comprehensive assessment of restoration quality.

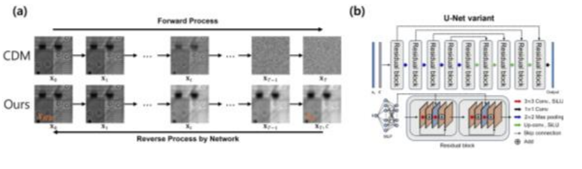

Fig. 1. (a) Forward and reverse process of CDM and Ours. (b) Network model used in this study.

2. Methods and Materials

2.1. Diffusion-inspired network

We formulated detector blur as a diffusion-like process in which spatial resolution gradually deteriorates with increasing phosphor thickness. Fig. 1 (a) shows the comparison of CDM and our approach, and the network used in this study. Inspired by denoising diffusion probabilistic models (DDPM) [1], we defined a deterministic forward operator $D_t(\cdot)$ that progressively degrades the HR image x_0 into a blurred image x_t over T steps. Unlike conventional stochastic diffusion, where Gaussian noise is added at each iteration, our degradation operator is constructed directly from the measured MTFs of the detectors:

$$x_b = \mathcal{F}^{-1}\left[\frac{\mathrm{MTF_{LR}}(u)}{\mathrm{MTF_{HR}}(u)} \cdot \mathcal{F}\{x_0\}\right], \quad (1)$$

where x_b denotes the LR-degraded version of x_0, and $\mathcal{F}, \mathcal{F}^{-1}$ are the Fourier and inverse Fourier transforms, respectively. Intermediate degradations are generated as linear blends of x_0 and x_b:

$$D_t(x_0) = \frac{T-t}{T} x_0 + \frac{t}{T} x_b, \quad t \in \{1, \ldots, T\}. \quad (2)$$

This design ensures that the degradation is physically grounded, reflecting the measured detector response.

2.2. Dataset preparation

Training and evaluation were performed on X-ray images of PCBs. Data were acquired using a laboratory-scale radiographic system with a tungsten-target X-ray tube (70 kV, 42.56 μGy·s^{-1}) coupled to phosphor-based CMOS detectors. Detector configurations simulated both HR and LR responses by varying the scintillator thickness (33.91 mg·cm^{-2} for HR, 134.55 mg·cm^{-2} for LR).

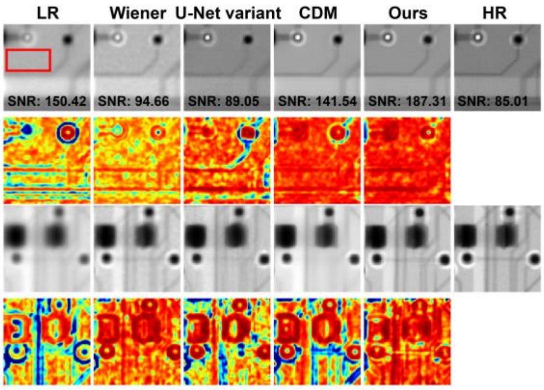

Fig. 2. Comparison of deblurring results. The first and third rows show the deblurring results of each method, while the second and fourth rows present the corresponding SSIM maps with respect to the HR images. The bounding boxes in the LR images indicate the regions where SNR was calculated.

The HR reference image x_0 was obtained by averaging 20 repeated measurements to the relative reduction in the noise-to-signal ratio of the averaged flood field image between successive averages fell below 1%. From these, 128×128-pixel patches were extracted and augmented via random flips and rotations, yielding 15,000 training samples. A separate test set of 600 samples was used.

2.3. Training details and sampling

The network was trained using the Adam optimizer with an initial learning rate of $2 \times 10{-}5$ and a batch size of 16. At each iteration, a random time step t was selected, enabling the network to learn restoration over different levels of blur. For sampling, our method employed transformation-agnostic cold sampling [5], while the CDM baseline was implemented using the denoising diffusion implicit models (DDIM) framework [6].

2.4. Performance evaluation

To comprehensively assess restoration performance, both spatial and frequency domain metrics were considered. In the spatial domain, peak signal-to-noise ratio (PSNR) and structural similarity index measurement (SSIM) were computed against HR references. In the frequency domain, the MTF, normalized noise power spectrum (NNPS), and contrast transfer function (CTF) were measured.

For comparison, three baseline methods were implemented: Wiener filtering, a regression baseline, and a CDM. All networks employed the same U-Net variant, as shown in Fig.1 (b), to ensure fair benchmarking.

3. Preliminary Results

Fig. 2 presents qualitative comparisons of input LR images, deblurring results from each method, and the HR

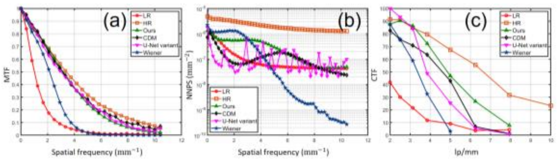

Fig. 3. Comparison of Fourier metrics for various deblurring approaches: (a) MTF, (b) NNPS, and (c) CTF.

ground truth. All methods yielded improved image quality compared to the LR input. Among them, our approach achieved the highest SSIM values, as shown in the SSIM maps, while also demonstrating superior SNR performance. Over the entire test set, the Wiener filter achieved the highest PSNR of 23.301, whereas our method provided the best SSIM of 0.890, indicating more faithful structural restoration.

Fig. 3 compares Fourier-domain metrics, including MTF, NNPS, and CTF. All deblurring methods improved MTF relative to LR, though Wiener filtering showed the weakest performance. In terms of NNPS, all methods yielded lower values compared to the HR reference. The regression baseline U-Net exhibited fluctuations across frequencies, reflecting instability in frequency response. For CTF, our method consistently outperformed all other approaches, demonstrating superior preservation of contrast across spatial frequencies.

4. Conclusion

In this study, the blurring process in X-ray images was modeled as a progressive loss of MTF, analogous to a thermodynamic diffusion process. The proposed method consistently outperformed conventional Wiener filtering, a regression baseline U-Net, and the CDM approach. As shown in Fig. 2, our model achieved the highest SSIM values and superior SNR compared to other methods, while avoiding noise amplification. Furthermore, the Fourier-domain analysis in Fig. 3 confirmed that our method yielded the most stable frequency response and the best contrast transfer across spatial frequencies. These results suggest that diffusion-inspired restoration models can be tailored for X-ray image deblurring, effectively recovering 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] J. Dong, S. Roth, and B. Schiele, Deep Wiener Deconvolution: Wiener Meets Deep Learning for Image Deblurring, Advances in Neural Information Processing Systems, Vol.33, p.1048, 2020.
[3] 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.
[4] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi, Image Super-Resolution via Iterative Refinement, 2022.
[5] A. Bansal, E. Borgnia, H.-M. Chu, J. Li, H. Kazemi, F. Huang, et al., Cold Diffusion: Inverting Arbitrary Image Transforms Without Noise, Advances in Neural Information Processing Systems, Vol.36, p.41259, 2023.
[6] J. Song, C. Meng, and S. Ermon, Denoising Diffusion Implicit Models, International Conference on Learning Representations (ICLR), May 3–7, 2021.