


Pixel-Dependent Noise Variance Stabilization for Learning-Based Denoising in Imaging Systems with Radially Symmetric Beams

Pablo Corral-Arroyo¹ 

PABLO.CORRALARROYO@SIEMENS-HEALTHINEERS.COM

Fasil Gadjimuradov² 

FASIL.GADJIMURADOV@SIEMENS-HEALTHINEERS.COM

Sai Gokul Hariharan² 

SAL.HARIHARAN@SIEMENS-HEALTHINEERS.COM

¹ *Siemens Healthineers, Tafernstrasse 7, 5405 Baden, Switzerland*

² *Siemens Healthineers, Siemensstrasse 1, 91301 Forchheim, Germany*

Editors: Under Review for MIDL 2026

Abstract

X-ray guided medical procedures may expose the patient to a non-negligible amount of radiation dose. To mitigate the dose and reduce the risk of potentially correlated health issues, it is important to optimize the radiation exposure for both the patients and clinical staff. This means that the applied radiation dose should be as low as reasonably achievable while ensuring that the required image quality is reached. UNETs have become the state-of-the-art denoising algorithms. In this article we show a preprocessing algorithm, which improves the robustness and generalization of denoising models of systems whose images have position-dependent x-ray intensity.

Keywords: Image denoising, UNETs, Convolutional Neural Networks

1. Introduction

Denoising methods based on Convolutional Neural Networks (CNN) are well-established state-of-the-art denoising image algorithms, which deliver optimal denoising levels while preserving fine physical structures and edges. This algorithm is based on a convolutional neural network, specifically a U-NET (Ronneberger et al., 2015), a kind of CNN that consists of a compression part (encoder) and expansion part (decoder). Originally designed for image segmentation, the UNET architecture has also been applied to other image processing operations such as denoising. UNETs comprise an encoder, which is a typical convolutional network using repeated convolutions with ReLU (Rectified Linear Unit) activation and a down-sampling operation that lowers the dimension of the image. This stage compresses the spatial resolution while enriching the feature representation. The decoder then reconstructs the spatial resolution through up-convolutions and concatenating high-resolution features from the contracting path and merging this information in the following channels. Recently, it has been shown that UNET-based denoising methods outperform even the most sophisticated denoising classic methods in the context of medical imaging (Hariharan et al., 2018, 2019, 2022).

We focus on a denoising algorithm based on a UNET (Hariharan et al., 2022). The algorithm is driven by the physics underlying X-ray image formation. The transformation of received X-ray photons, collected by an indirect-detection flat-panel detector, into pixel

gray values can be described as a succession of processing stages. Each stage may involve either a quantum gain or spatial spreading (i.e., blurring due to photon dispersion during the detection process). It is assumed that this process follows a linear model (cite). Thus, an observed noise-corrupted gray value (number of counts) V at row x and column y can be expressed as

$$V[x, y] = \beta(c * k_q)[x, y] + n_{\text{elec}} = kN + n_{\text{elec}}, \quad (1)$$

where c represents the charge signal at the photodiodes, corrupted by quantum noise, convolved with the stochastic spreading function k_q and scaled by the overall detector gain β (Siewerdsen et al., 1997). The term n_{elec} denotes additive electronic noise with zero mean and standard deviation σ_{elec} . The variable N represents the noise-corrupted number of detected X-ray photons, and k is a proportionality constant.

The total noise present in a standard Varian X-ray image can be decomposed into its elemental components, namely electronic noise σ_{elec}^2 and quantum noise σ_{quant}^2 . Accordingly, the total noise variance can be written as

$$\sigma^2 = \sigma_{\text{elec}}^2 + \sigma_{\text{quant}}^2 = \sigma_{\text{elec}}^2 + \alpha (kN)_{\text{avr}}. \quad (2)$$

Aiming to produce images with uniform noise intensity in the input image across all exposure levels, Noise Variance Stabilization algorithms are applied to the input images, such as the Generalized Ascome Transform (GAT) used in (Hariharan et al., 2022). This improves denoiser performance and the learning consistency (avoid noise learning) and shows a faster generalization.

$$V'[x, y] = \frac{2}{\alpha} \sqrt{\alpha V[x, y] + \frac{3}{8} \alpha^2 - \alpha g + \sigma_{\text{elec}}^2} \quad (3)$$

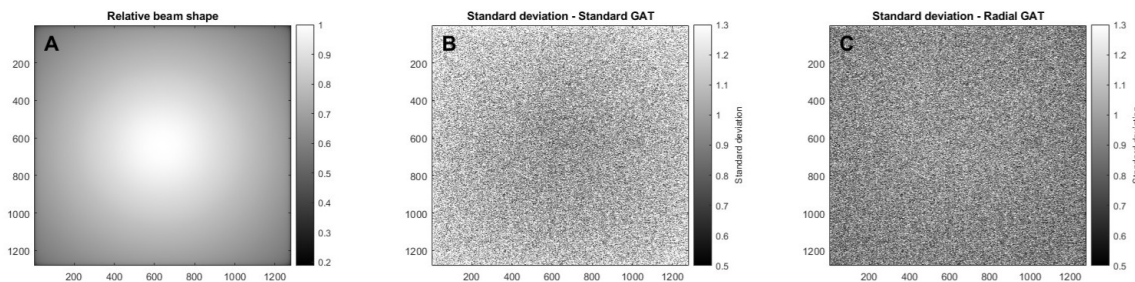


Figure 1: A) Relative beam shape map . B) Standard deviation of image after standard GAT. C) Standard deviation of image after radial GAT.

2. Algorithm

Typical Mega-Volts (MV) x-ray imaging system sources are based on a linear electron accelerator, after which a bending magnet directs the high-energy electrons toward the

target. When the electron beam hits the target, bremsstrahlung x-rays are produced in a forward-peaked distribution; this is a consequence of relativistic bremsstrahlung physics. The MV beam can therefore have a significant radial fall-off in the case of flattening-filter free MV beams. This relative beam shape map for a MV beam is shown in Fig. 1A, where 1 represents the maximum intensity in the central part of the beam.

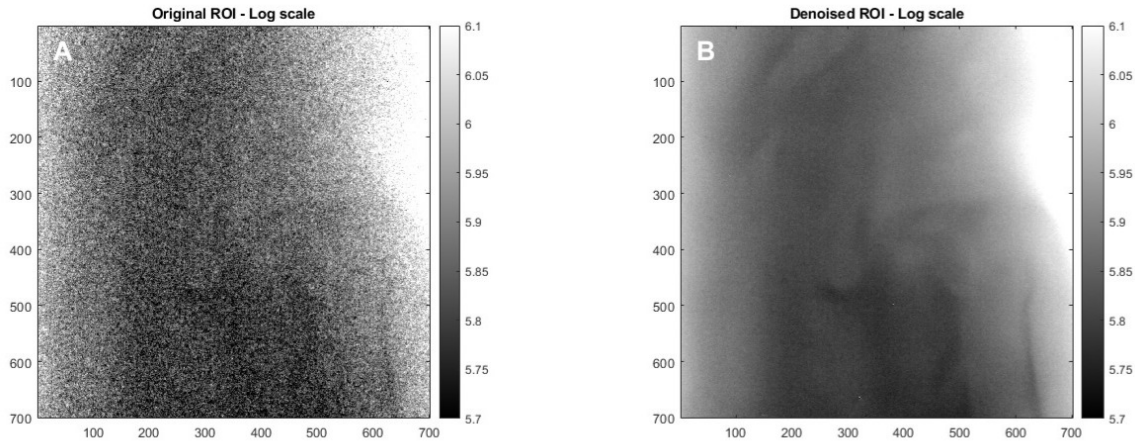


Figure 2: Noisy (A) and denoised (B) MV image of an abdomen phantom.

Due to this pronounced profile, the flat-field correction generates corrected images where the noise is amplified in the low-exposure regions and diminished in the high exposure regions. When applying a standard NVS, it is observed that the standard deviation of the noise is higher in the low-exposure regions (Fig. 1B). This poses a problem for the denoiser performance, since distinct ROIs would have different noise levels. This is solved with the radial-dependent GAT, where the noise parameters are pixel dependent accounting for the beam shape and the corresponding flat-field correction. The standard deviation is similar across the image. The application of such a pre-processing algorithm makes the training of a denoiser based on UNETs more robust for these systems. Figure 2 shows the standard and denoised image for an abdomen phantom. The training and inference with this approach improves the generalization and robustness of the model. This same approach could be used for other similar systems.

References

- S. G. Hariharan, N. Strobel, C. Kaethner, M. Kowarschik, S. Demirci, S. Albarqouni, R. Fahrig, and N. Navab. A photon recycling approach to the denoising of ultra-low dose x-ray sequences. *International Journal of Computer Assisted Radiology and Surgery*, 13 (6):847–854, 2018. doi: 10.1007/s11548-018-1766-4.
- S. G. Hariharan, C. Kaethner, N. Strobel, M. Kowarschik, S. Albarqouni, R. Fahrig, and N. Navab. Learning-based x-ray image denoising utilizing model-based image simulations.

- In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2019*, Lecture Notes in Computer Science, pages 549–557, Berlin, Germany, 2019. Springer. doi: 10.1007/978-3-030-32245-8_61.
- S. G. Hariharan, C. Kaethner, N. Strobel, M. Kowarschik, R. Fahrig, and N. Navab. Robust learning-based x-ray image denoising—potential pitfalls, their analysis and solutions. *Biomedical Physics & Engineering Express*, 8(3):035013, 2022. doi: 10.1088/2057-1976/ac5b7a.
- O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, volume 9351 of *Lecture Notes in Computer Science*, pages 234–241, Cham, Switzerland, 2015. Springer. doi: 10.1007/978-3-319-24574-4_28.
- J. H. Siewerdsen, L. E. Antonuk, Y. El-Mohri, J. Yorkston, W. Huang, J. Boudry, and I. A. Cunningham. Empirical and theoretical investigation of the noise performance of indirect-detection, active matrix flat-panel imagers (amfpis) for diagnostic radiology. *Medical Physics*, 24(1):71–89, 1997. doi: 10.1118/1.597928.