engineering proceedings

Proceeding Paper

Fine-Tuning a Denoiser for Low-Dose CT Image Reconstruction †

Tim Selig [1,2,3,*], Thomas März [1,3], Martin Storath [2] and Andreas Weinmann [1,2,3]

[1] Hochschule Darmstadt, Fachbereich Mathematik und Naturwissenschaften, Schöfferstraße 3, 64295 Darmstadt, Germany
[2] Technische Hochschule Würzburg-Schweinfurt, Fakultät Angewandte Natur- und Geisteswissenschaften and Institute of Digital Engineering, Ignaz-Schön-Straße 11, 97421 Schweinfurt, Germany
[3] Data Science Institute, European University of Technology (EUt+), European Union, 64295 Darmstadt, Germany
* Correspondence: tim.selig@h-da.de
† Presented at the 15th International Scientific Conference TechSys 2026—Engineering, Technologies and Systems, Plovdiv, Bulgaria, 14–16 May 2026.

Abstract

A commonly employed imaging modality that relies on ionizing radiation is Computed Tomography (CT). While lowering the radiation dose is beneficial for patient health, it can result in reduced image quality. Therefore, improving low-dose CT (LDCT) reconstruction is a significant area of research. The LoDoPaB-CT benchmark evaluates LDCT reconstruction methods, where many top methods use UNet-type architectures. We explore a two-stage approach for LDCT reconstruction: the first stage employs traditional filtered backprojection (FBP), while the second stage performs CT image enhancement. Our training strategy involves pretraining a neural network to denoise natural grayscale images, which are corrupted by Gaussian noise, followed by fine-tuning the network for CT image enhancement using LDCT and normal-dose CT (NDCT) pairs. Experiments on various small subsets of the LoDoPaB-CT dataset demonstrate the effectiveness of our method, showing that less task-specific data are required for training.

Keywords: computed tomography; image reconstruction; machine learning; UNet

1. Introduction

To provide non-invasive visualization of internal structures, Computed Tomography (CT) is essential in modern medical imaging. Its widespread use is partly due to its versatility, serving various diagnostic purposes [1,2]. However, minimizing ionizing radiation while maintaining image quality remains a critical challenge, particularly for low-dose CT (LDCT) reconstruction [3]. Recent research focuses on LDCT reconstruction, leveraging advanced algorithms and deep learning techniques [4–10].

Historically, iterative variational regularization methods, like total variation [11], total generalized variation [12], and Hessian Schatten norms [13], dominated CT image reconstruction. Recently, deep learning, especially convolutional neural networks (CNNs), has revolutionized medical image reconstruction, effectively reducing artifacts and noise [14]. Approaches range from fully learned methods, which use sinograms or raw projection data, to iterative and multi-stage methods integrating traditional techniques with neural networks. For instance, ISTA UNet [15], the Learned Primal-Dual algorithm [16], and ItNet [17] are notable for their success in LDCT reconstruction, particularly in the LoDoPaB-CT challenge [18].

Academic Editors: Nikola G. Shakev, Valyo N. Nikolov, Nikolay R. Kakanakov and Sevil A. Ahmed-Shieva

Published: 21 July 2026

 https://doi.org/10.3390/engproc2026150038

Building on recent trends of pretraining models on large, diverse datasets from various domains, such as those used in Natural Language Processing (e.g., [19]) and Computer Vision (e.g., [20]), we introduce a two-stage approach for LDCT reconstruction. The first stage involves traditional filtered backprojection (FBP), while the second stage focuses on CT image enhancement. Two-stage methods using UNets or other neural networks are prevalent in CT image processing [9,10,15,21–23]. In the present work, we use the Gaussian denoiser DRUNet [24], which has been pretrained on several natural image datasets, including the Berkeley segmentation, DIV2K, Flick2K, and the Waterloo Exploration dataset. For the subsequent task of LDCT image enhancement, fine-tuning is performed using CT-specific images. We emphasize that our distinctive training strategy involves pretraining on non-CT images with different noise characteristics than LDCT images.

Furthermore, our method demonstrates robustness through experiments on various small subsets of the LoDoPaB-CT dataset, highlighting the advantage of our pretraining strategy over training the network from scratch. These experiments also show that our approach requires less task-specific data to achieve similar results.

2. Problem Formulation

CT images are reconstructed from X-Ray projection data measured at various angles. The forward model defines how the desired image relates to the collected projection data. Let y be the projection data, $x \in \mathbb{R}^{n \times n}$ the true CT image, and R the Radon transform. With these definitions, the forward model can be simplified as follows:

$$y = R(x) + \epsilon,$$

where ϵ represents measurement errors, such as quantum noise, electronic noise, scatter, and artifacts from patient motion. For a detailed model description, we refer to [21]. CT image reconstruction aims to estimate x from the projection data y, utilizing the system matrix R.

This study addresses the issue of reconstructing LDCT images with a reduced radiation dose. The primary challenge is the increased noise level in the projection data resulting from the lower radiation dose. Additionally, the inversion of R is ill-posed, which further amplifies the noise.

2.1. The Two-Stage Approach

The methodology employed for the reconstruction of CT images comprises two stages. In the initial stage, reconstruction is performed utilizing the FBP, which is a conventional analytical method, but the noise is amplified in the measurements. In the subsequent stage, a neural network, designated as G_θ, is employed to enhance the LDCT image. Here, the low-dose FBP-reconstructed image is mapped to an enhanced version. Mathematically, the two-stage method can be expressed in the following way:

$$\hat{x} = G_\theta \circ \mathrm{FBP}(y), \tag{1}$$

where G_θ denotes a neural network with learnable parameters θ, FBP denotes the filtered backprojection, and $\hat{x}$ is the final reconstruction result.

To enhance the performance of G_θ for CT image enhancement, we adopt a fine-tuning strategy. Initially, G_θ is pretrained on a large dataset to denoise natural grayscale images corrupted by Gaussian noise. Subsequently, we fine-tune G_θ for the specific task of LDCT image enhancement using paired data consisting of LDCT and normal-dose CT (NDCT) images. Fine-tuning adapts G_θ to the characteristics of LDCT image enhancement, aiming to improve the quality and fidelity of the FBP-reconstructed images.

2.2. DRUNet: Pretraining for Robust Image Restoration

The DRUNet [24] model is trained for various image restoration tasks, particularly for denoising natural grayscale images corrupted by Gaussian noise. It combines the effectiveness of UNets in image translation tasks, combined with the enhanced modeling capacity of ResNets through stacked residual blocks. The DRUNet model was trained on a dataset comprising 8794 natural images from the Berkeley Segmentation, DIV2K, Flick2K, and Waterloo Exploration datasets. The optimization process utilizes the ADAM algorithm with the objective of minimizing the mean absolute error (MAE) loss between the processed image and the ground truth. This pretrained network is then fine-tuned for the specific purpose of improving CT images using paired LDCT and NDCT images.

3. Implementation Details

Following the methodologies outlined in [10], we adopted a composite loss function to train our neural network architecture. Initially, we employed the MAE loss function, denoted as L_{MAE}, which focuses on ensuring the accuracy of pixel intensity and is known for its robustness in handling outliers [25]. The MAE loss for a single sample is defined as

$$L_{\text{MAE}} = \frac{1}{N} \sum_{i=1}^{N} |\hat{x}_i - y_i|,$$

where N denotes the amount of pixels in the image, $\hat{x}_i$ refers to the i-th pixel of the neural network output, and y_i represents the i-th pixel from the ground truth (NDCT image).

To further enhance reconstruction quality, we integrated Structural Similarity (SSIM) [26] in our loss function. The SSIM metric evaluates the structural similarity of images by analyzing the statistical attributes of local regions, rather than focusing on the precision of individual pixels.

This composite loss function L is formulated as

$$L = L_{\text{MAE}} + \alpha(1 - \text{SSIM}) + \lambda \|\theta\|_2^2,$$

where SSIM is calculated according to the method described in [26], $\alpha > 0$ adjusts the weight of the $(1 - \text{SSIM})$ term, $\lambda > 0$ is the regularization hyperparameter, and $\|\theta\|_2^2$ represents the squared L_2 norm of the parameters of the model.

For optimization, we utilized the ADAM algorithm [27] with default parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The initial learning rate was set to 10^{-4}. In order to enhance the stability of our network, we employed two data augmentation techniques: rotational augmentation and Gaussian noise augmentation.

4. The LoDoPaB-CT Dataset

The LoDoPaB-CT dataset [18] consists of 46,573 pairs of chest CT images obtained from patients and their respective simulated low-dose measurements. The original data were obtained from the LIDC/IDRI dataset [28] and the ground truth images were cropped to 362×362 pixels. This dataset serves as a benchmark for evaluating LDCT reconstruction algorithms, offering 35,820 training, 3522 validation, 3553 test, and 3678 challenge images.

5. Evaluation Metrics

We adopted the evaluation metrics established in the LoDoPaB-CT challenge [18], which include the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). Additionally, we utilize their respective fixed-range (FR) versions of the PSNR and SSIM algorithms, namely, PSNR-FR and SSIM-FR. These metrics standardize the range of values rather than use the full dynamic range of the ground truth image. The PSNR measures the relationship between the highest achievable signal strength and the level of noise distortion [29]. On the other hand, SSIM, introduced by Wang et al. [26], evaluates the structural similarity of two images, as elaborated in the section on implementation details. Higher values for both the PSNR and SSIM indicate better reconstruction quality in our evaluations.

6. Experimental Results

To explore the impact of fine-tuning a pretrained network on our results, we evaluated the performance of the DRUNet in several scenarios. We compared the effectiveness of pretraining the DRUNet on the Gaussian denoising of natural grayscale images against the random initialization of the model weights. In both cases, the models are fine-tuned on various subsets of the LoDoPaB-CT dataset for comparison. We utilized 'drunet_gray.pth' as the pretrained model weights, which is available for download from a GitHub repository (https://github.com/cszn/DPIR/tree/master/model_zoo (accessed on 18 June 2024)).

In the LoDoPaB-CT dataset, we randomly selected 0.2%, 1%, and 5% of the training and validation datasets. Specifically, this selection resulted in 71 training and 7 validation images for the 0.2% case, 358 training and 35 validation images for the 1% case, and 1791 training and 176 validation images for the 5% case. When testing the trained networks, we selected 300 random images from the test set of the LoDoPaB-CT dataset.

Table 1 outlines the results across various scenarios, while Figure 1 provides visual representations of an individual test samples for each scenario. A direct comparison reveals that fine-tuning the pretrained network consistently yields better metrics than training from scratch across all three cases. Additionally, the effectiveness of pretraining becomes more pronounced as the number of images for fine-tuning decreases. Notably, training the network from scratch with 358 images yields similar results to fine-tuning the pretrained network with only 71 images. Similarly, training from scratch with 1791 images shows comparable results to fine-tuning with 358 images. This demonstrates that less task-specific data are needed to achieve similar outcomes.

Table 1. Performance comparison of the DRUNet trained from scratch and fine-tuned using the pretrained weights for Gaussian noise removal. The models are evaluated in terms of PSNR and SSIM across three subsets of the LoDoPaB-CT dataset.

LoDoPaB-CT Dataset	71 Training Images		358 Training Images		1791 Training Images	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DRUNet From Scratch	34.39 ± 4.10	0.8185 ± 0.1508	35.35 ± 4.43	0.8416 ± 0.1505	35.90 ± 4.70	0.8499 ± 0.1485
DRUNet Fine-Tuned	35.30 ± 4.34	0.8374 ± 0.1506	35.87 ± 4.64	0.8468 ± 0.1499	36.14 ± 4.78	0.8516 ± 0.1482

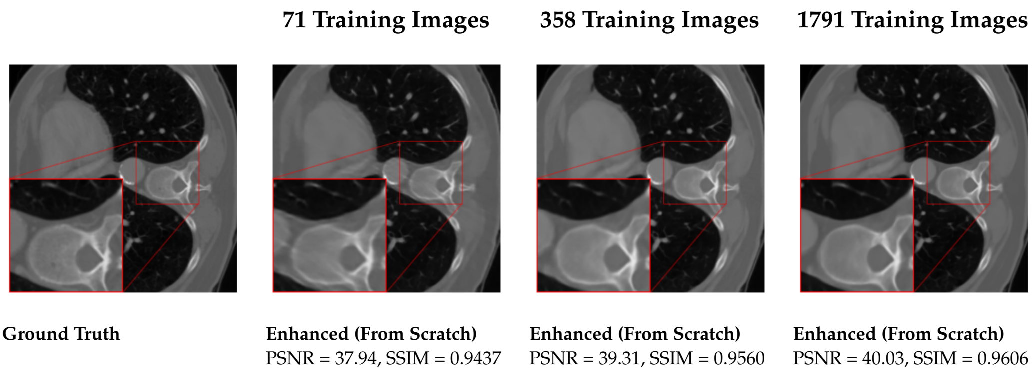
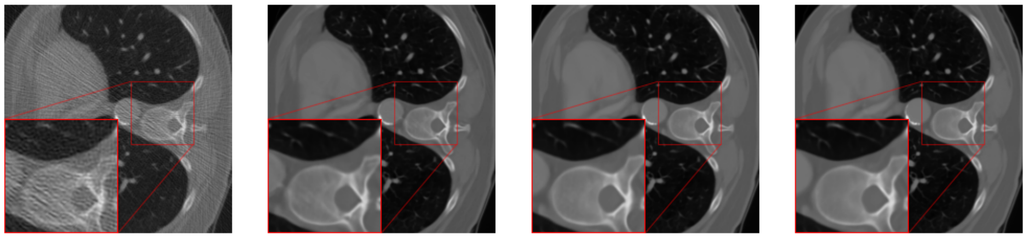


Figure 1. *Cont.*



FBP PSNR = 28.59, SSIM = 0.6290 | **Enhanced (Pretrained)** PSNR = 39.02, SSIM = 0.9530 | **Enhanced (Pretrained)** PSNR = 39.89, SSIM = 0.9584 | **Enhanced (Pretrained)** PSNR = 40.39, SSIM = 0.9619

Figure 1. Reconstruction results for a test sample from the test set of the LoDoPaB-CT dataset. The initial column presents the ground truth, which is followed by the corresponding FBP, which represents the initial stage of our methodology. In the second stage, the DRUNet is employed to enhance the FBP result. Subsequent columns illustrate different amounts of training images (71, 358, and 1791). Each column compares results achieved by training the DRUNet from scratch versus fine-tuning the pretrained DRUNet. PSNR and SSIM metrics quantify the quality improvements, demonstrating better performance when leveraging the pretrained DRUNet.

7. Conclusions

In this study, we presented a two-stage method for reconstructing LDCT images by integrating classical FBP with deep learning techniques, aimed at enhancing image quality. Our training strategy involves the initial pretraining of a neural network on natural grayscale images to remove Gaussian noise, followed by fine-tuning specifically for LDCT image enhancement. This strategy diverges from conventional methods that typically rely on pretraining with solely domain-specific CT data.

Experiments on subsets of the LoDoPaB-CT dataset validated our approach, revealing improvements in both PSNR and SSIM metrics compared to training the network from scratch. Our findings illustrate the benefit of pretraining strategies, highlighting the potential for practical applications in medical imaging where minimizing data requirements is crucial.

For more details, the preprint of our paper can be found on arXiv [30]. Additional results using diverse architectures and datasets will be presented in future work.

Author Contributions: Conceptualization, T.S., T.M., M.S. and A.W.; methodology, T.S., T.M., M.S. and A.W.; software, T.S.; validation, T.S.; formal analysis, T.S.; investigation, T.S.; data curation, T.S.; writing—original draft preparation, T.S.; writing—review and editing, T.S., T.M., M.S. and A.W.; supervision, T.M., M.S. and A.W.; project administration, T.S.; funding acquisition, T.M. and A.W. All authors have read and agreed to the published version of the manuscript.

Funding: T.S., T.M., and A.W. appreciate the funding from the Hessian Ministry of Higher Education, Research, Science and the Arts through the "Programm zum Aufbau eines akademischen Mittelbaus an hessischen Hochschulen" and financial support from the EUt+/h-da (https://h-da.de/hochschule/eut (accessed on 18 June 2024)). M.S. appreciates the support of the DIBCO project financed by the research program "Informations- und Kommunikationstechnik" of the Bavarian State Ministry of Economic Affairs, Regional Development and Energy (DIK-2105-0044/DIK0264). A.W. appreciates the funding provided by the Deutsche Forschungsgemeinschaft (DFG) under project number 514177753.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The LoDoPaB-CT dataset is publicly available from Zenodo at https://zenodo.org/records/3384092 (accessed on 18 June 2024). The pretrained DRUNet weights were obtained from the CSZN DPIR model zoo (https://github.com/cszn/DPIR/tree/master/model_zoo, accessed on 18 June 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lell, M.M.; Wildberger, J.E.; Alkadhi, H.; Damilakis, J.; Kachelriess, M. Evolution in Computed Tomography: The Battle for Speed and Dose. *Investig. Radiol.* **2015**, *50*, 629–644. [CrossRef]
2. Kalender, W.A. *Computed Tomography: Fundamentals, System Technology, Image Quality, Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
3. Yang, Q.; Yan, P.; Zhang, Y.; Yu, H.; Shi, Y.; Mou, X.; Kalra, M.K.; Zhang, Y.; Sun, L.; Wang, G. Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss. *IEEE Trans. Med. Imaging* **2018**, *37*, 1348–1357. [CrossRef] [PubMed]
4. Lu, S.; Yang, B.; Xiao, Y.; Liu, S.; Liu, M.; Yin, L.; Zheng, W. Iterative reconstruction of low-dose CT based on differential sparse. *Biomed. Signal Process. Control* **2023**, *79*, 104204. [CrossRef]
5. Hu, Y.; Zheng, Z.; Yu, H.; Wang, J.; Yang, X.; Shi, H. Ultra-low-dose CT reconstructed with the artificial intelligence iterative reconstruction algorithm (AIIR) in 18F-FDG total-body PET/CT examination: A preliminary study. *EJNMMI Phys.* **2023**, *10*, 1. [CrossRef] [PubMed]
6. Kulathilake, S.; Abdullah, N.; Sabri, A.; Lai, K.W. A review on Deep Learning approaches for low-dose Computed Tomography restoration. *Complex Intell. Syst.* **2021**, *9*, 2713–2745. [CrossRef]
7. Li, H.; Yang, X.; Yang, S.; Wang, D.; Jeon, G. Transformer With Double Enhancement for Low-Dose CT Denoising. *IEEE J. Biomed. Health Inform.* **2023**, *27*, 4660–4671. [CrossRef] [PubMed]
8. Wang, D.; Fan, F.; Wu, Z.; Liu, R.; Wang, F.; Yu, H. CTformer: Convolution-free Token2Token dilated vision transformer for low-dose CT denoising. *Phys. Med. Biol.* **2023**, *68*, 065012. [CrossRef]
9. Mazandarani, F.N.; Babyn, P.; Alirezaie, J. UNeXt: A Low-Dose CT denoising UNet model with the modified ConvNeXt block. In Proceedings of the ICASSP 2023—2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes, Greece, 16–20 April 2023; pp. 1–5.
10. Xiong, L.; Li, N.; Qiu, W.; Zhang, Y. Re-UNet: A Novel Multi-scale Reverse U-shaped Network Architecture for Low-dose CT Image Reconstruction. *SSRN* **2023**, 4426158. Available online: http://dx.doi.org/10.2139/ssrn.4426158 (accessed on 14 June 2026).
11. Liu, Y.; Ma, J.; Fan, Y.; Liang, Z. Adaptive-weighted total variation minimization for sparse data toward low-dose x-ray computed tomography image reconstruction. *Phys. Med. Biol.* **2012**, *57*, 7923–7956. [CrossRef] [PubMed]
12. Niu, S.; Gao, Y.; Bian, Z.; Huang, J.; Chen, W.; Yu, G.; Liang, Z.; Ma, J. Sparse-view x-ray CT reconstruction via total generalized variation regularization. *Phys. Med. Biol.* **2014**, *59*, 2997–3016. [CrossRef] [PubMed]
13. Liu, L.; Li, X.; Xiang, K.; Wang, J.; Tan, S. Low-Dose CBCT Reconstruction Using Hessian Schatten Penalties. *IEEE Trans. Med. Imaging* **2017**, *36*, 2588–2599. [CrossRef] [PubMed]
14. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; Van Der Laak, J.A.W.M.; Van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [CrossRef] [PubMed]
15. Liu, T.; Chaman, A.; Belius, D.; Dokmanic, I. Interpreting U-Nets via Task-Driven Multiscale Dictionary Learning. *arXiv* **2020**, arXiv:2011.12815.
16. Adler, J.; Öktem, O. Learned Primal-Dual Reconstruction. *IEEE Trans. Med. Imaging* **2018**, *37*, 1322–1332. [CrossRef] [PubMed]
17. Genzel, M.; Gühring, I.; Macdonald, J.; März, M. Near-Exact Recovery for Tomographic Inverse Problems via Deep Learning. In Proceedings of the 39th International Conference on Machine Learning, Baltimore, MD, USA, 17–23 July 2022; pp. 7368–7381.
18. Leuschner, J.; Schmidt, M.; Ganguly, P.S.; Andriiashen, V.; Coban, S.B.; Denker, A.; Bauer, D.F.; Hadjifaradji, A.; Batenburg, K.J.; Maass, P.; et al. Quantitative Comparison of Deep Learning-Based Image Reconstruction Methods for Low-Dose and Sparse-Angle CT Applications. *J. Imaging* **2021**, *7*, 44. [CrossRef] [PubMed]
19. Radford, A.; Narasimhan, K.; Salimans, T.; Sutskever, I. Improving Language Understanding by Generative Pre-Training. *OpenAI* 2018. Available online: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf (accessed on 18 March 2026).
20. Liu, Z.; Hu, H.; Lin, Y.; Yao, Z.; Xie, Z.; Wei, Y.; Ning, J.; Cao, Y.; Zhang, Z.; Dong, L.; et al. Swin transformer v2: Scaling up capacity and resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–20 June 2022; pp. 12009–12019.
21. Leuschner, J.; Schmidt, M.; Baguer, D.O.; Maass, P. LoDoPaB-CT, a benchmark dataset for low-dose computed tomography reconstruction. *Sci. Data* **2021**, *8*, 109. [CrossRef] [PubMed]
22. Pelt, D.M.; Batenburg, K.J.; Sethian, J.A. Improving tomographic reconstruction from limited data using mixed-scale dense convolutional neural networks. *J. Imaging* **2018**, *4*, 128. [CrossRef]
23. Liu, T.; Chaman, A.; Belius, D.; Dokmanic, I. Learning Multiscale Convolutional Dictionaries for Image Reconstruction. *IEEE Trans. Comput. Imaging* **2022**, *8*, 425–437. [CrossRef]

24. Zhang, K.; Li, Y.; Zuo, W.; Zhang, L.; Van Gool, L.; Timofte, R. Plug-and-Play Image Restoration with Deep Denoiser Prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 6360–6376. [CrossRef] [PubMed]
25. Bishop, C. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2011.
26. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef] [PubMed]
27. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
28. Armato, S.G., III; 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. *Med. Phys.* **2011**, *38*, 915–931. [CrossRef] [PubMed]
29. Sara, U.; Akter, M.; Uddin, M.S. Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study. *J. Comput. Commun.* **2019**, *7*, 8–18. [CrossRef]
30. Selig, T.; März, T.; Storath, M.; Weinmann, A. Low-dose ct image reconstruction by fine-tuning a unet pretrained for gaussian denoising for the downstream task of image enhancement. *arXiv* **2024**, arXiv:2403.03551.

https://doi.org/10.3390/engproc2026150038