Distilling Photon-Counting CT into Routine Chest CT through Clinically Validated Degradation Modeling

Junqi Liu[1], Xinze Zhou[1], Wenxuan Li[1], Scott Ye[2], Arkadiusz Sitek[3], Xiaofeng Yang[4], Yucheng Tang[5], Daguang Xu[5], Kai Ding[6], Kang Wang[2], Yang Yang[2], Alan L. Yuille[1], and Zongwei Zhou[1,6⋆]

[1] Johns Hopkins University
[2] University of California, San Francisco
[3] Harvard Medical School
[4] Emory University
[5] Nvidia
[6] Johns Hopkins Medicine

Abstract. Photon-counting CT (PCCT) delivers higher spatial resolution and lower noise than conventional energy-integrating CT (EICT). However, PCCT scanners are scarce in clinical practice. This scarcity limits large-scale research and broad deployment. We propose SUMI, a degradation-to-enhancement method that distills PCCT image quality into routine EICT scans. The central idea is to model realistic acquisition degradations that turn high-quality PCCT into low-quality counterparts, and then learn to invert this process. A board-certified radiologist verified that the simulated degradations are clinically realistic. This design provides faithful supervision without paired PCCT and EICT acquisitions. We make three contributions. **First**, we release an open-source autoencoder pretrained on 1,046 PCCT scans and 405,379 EICT scans from 145 hospitals. To our knowledge, this is the largest and most geographically diverse CT autoencoder released to date. We then train an enhancement network in this latent space. **Second**, we release a large-scale dataset of over 17,316 public EICT scans enhanced to PCCT-like quality. The dataset includes radiologist-validated voxel-wise annotations of airway trees, arteries, veins, lungs, and lobes. **Third**, on external data, SUMI outperforms state-of-the-art enhancement methods by 15% in SSIM and 20% in PSNR. SUMI preserves anatomical structure and tissue density, as verified by a radiologist. Applied to off-the-shelf detectors without retraining, our enhanced scans improve lesion detection sensitivity by up to 15% and F1 score by up to 10%. Our results show that imaging advances from specialized hardware can be distilled into routine EICT using a limited number of high-quality reference scans. All datasets, code, and models are available at https://github.com/KumaKuma2002/OpenVAE.

Keywords: photon-counting CT · CT enhancement · generative model.

⋆ Correspondence to: Zongwei Zhou (zzhou82@jh.edu)

1 Introduction

Computed tomography (CT) is a frontline imaging modality for screening, diagnosis, and longitudinal monitoring of thoracic diseases. In routine clinical practice, CT systems with energy-integrating detectors (EICT) suffer from noise and limited spatial resolution. These limitations introduce artifacts such as partial volume effects. As a result, small anatomical structures (e.g., distal airways and fine pulmonary vessels) and subtle pathological findings are often obscured.

Photon-counting CT (PCCT) is a recent hardware innovation that improves on EICT [23]. PCCT scanners register individual x-ray photons and estimate their energy. This design provides spectral sensitivity and improves dose efficiency and spatial resolution [3]. These advances enhance the visibility of small anatomical structures and reduce artifacts. They benefit airway analysis, vascular assessment, and lesion characterization in chest CT [26]. However, PCCT is difficult to deploy at scale. PCCT scanners are far more expensive than EICT scanners. They are concentrated in a small number of well-resourced academic and tertiary-care centers. This concentration restricts clinical access and disproportionately limits availability in underserved communities [3].

Limited PCCT penetration also impedes large-scale research. High-quality PCCT scans are not available in the volume or diversity needed for robust population studies, external validation, and development of generalizable models [3]. Most clinical workflows and most public datasets remain dominated by EICT [11]. This heterogeneity creates a second barrier. Systematic domain shifts between PCCT and EICT, and across EICT platforms, complicate direct comparison and model transfer. They reduce reproducibility and limit the generalizability of findings in multi-institutional research [3]. The landscape thus creates a translational gap. The clinical benefits of PCCT are increasingly recognized, but no scalable strategy exists to extend comparable image quality to the much larger population imaged with EICT.

A direct solution would train an enhancement method that maps EICT to PCCT quality. However, this approach requires paired PCCT and EICT scans from the same patient. Acquiring such pairs at scale is impractical. It adds cost, scanner time, and operational complexity. More importantly, two scans introduce avoidable radiation exposure and raise ethical and regulatory concerns. Purely data-driven enhancement also risks producing unrealistic textures, misleading structures, or subtle anatomical distortions. Even minor artifacts are unacceptable in clinical imaging and can undermine trust [9,24,18]. We therefore ask a different question. *Can we learn a clinically plausible forward degradation process from PCCT, and then train an enhancement method to invert that process in a controlled and auditable way?*

To address these challenges, we introduce SUMI, a degradation-to-enhancement method that distills PCCT quality into EICT. We use 1,046 PCCT scans as the high-quality reference and 405,379 EICT scans as examples of routine clinical data. The central idea is to model realistic acquisition degradations that map PCCT to a degraded version. These degradations emulate the physical and statistical properties of EICT. A board-certified radiologist verified that the simulated

degradations reflect authentic acquisition limitations rather than artificial artifacts. We then train an enhancement method that maps degraded PCCT back to high-quality PCCT. Once trained, the same method enhances any EICT scan toward PCCT-like quality without retraining.

Our contributions are threefold. **First**, we introduce a PCCT-to-EICT degradation simulator that captures realistic acquisition artifacts (Figure 1). A board-certified radiologist verified its clinical realism, providing controlled supervision for enhancement. **Second**, we release an open-source autoencoder pretrained on 1,046 PCCT scans and 405,379 EICT scans from 145 hospitals across 19 countries (Figure 1). To our knowledge, this is the largest and most geographically diverse CT autoencoder released to date. We use it as a backbone for our enhancement network. **Third**, we curate a dataset of 17,316 enhanced CT scans with PCCT-like image quality and radiologist-validated voxel-wise annotations of airway trees, pulmonary arteries, pulmonary veins, lungs, and lobes. We demonstrate improved image quality (Table 1, Figure 4), higher anatomical and HU fidelity (Figure 2), and improved off-the-shelf lesion detection across three external datasets (Table 3, Figure 3).

Overall, this work outlines a pragmatic strategy for distilling imaging gains from specialized hardware into routine CT workflows using a limited number of high-quality reference scans. It can facilitate large-scale multi-institutional research and clinical impact across diverse care settings.

2 Method

Our method has two components. We first pretrain a CT autoencoder on 1,046 PCCT scans and 405,379 EICT scans from 145 hospitals (§2.1). The autoencoder learns general CT latent features and is released as a reusable backbone for medical generative tasks [20]. We then train an enhancement network in this latent space that maps degraded inputs to PCCT-quality outputs (§2.2).

2.1 Autoencoder Pretraining

We adopt the autoencoder architecture from latent diffusion models [20] and pretrain it on 405,379 EICT scans and 1,046 PCCT scans from 145 hospitals across 19 countries. Training uses the standard pixel-wise reconstruction loss and an adversarial loss as in Rombach et al. [20]. The technical components are standard. Our contribution at this stage is the scale and geographic diversity of the pretraining data, which yields a CT backbone that we release for reuse.

2.2 Degradation-to-Enhancement Method

To ensure robust enhancement across heterogeneous CT scans, we design a degradation-to-enhancement method. The method explicitly models real-world acquisition artifacts and learns to reverse them in a controlled manner.

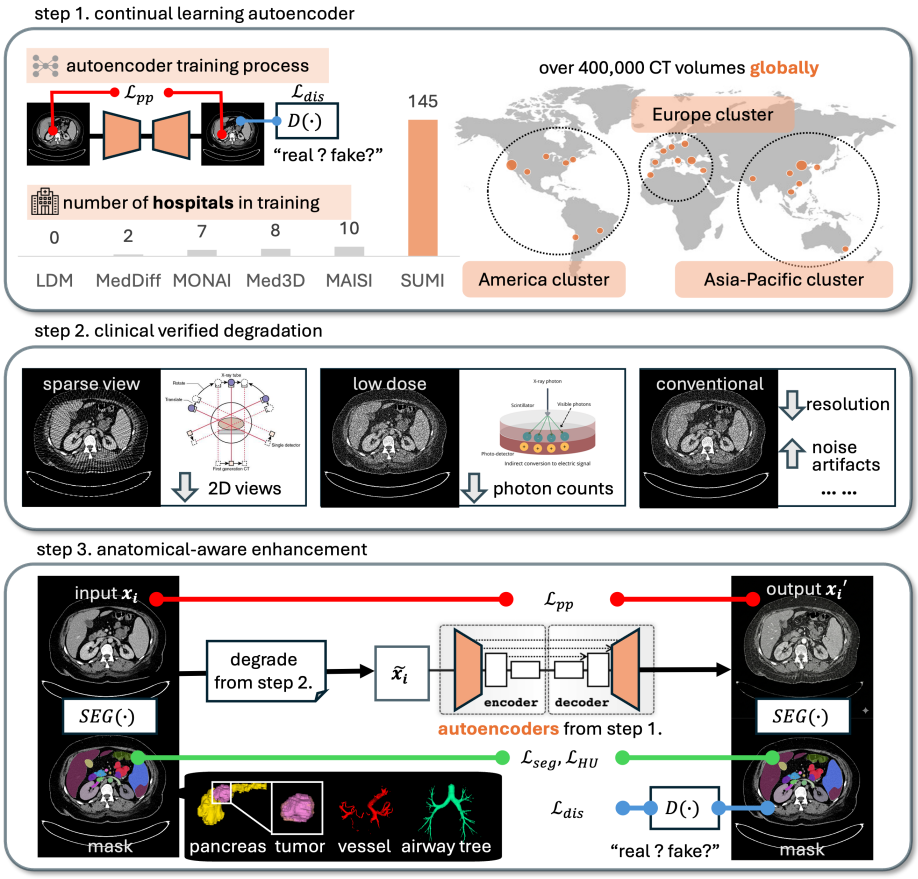

Fig. 1. Overview of SUMI. (1) A CT autoencoder is pretrained on over 400,000 CT scans from 145 hospitals across 19 countries with pixel-wise loss $\mathcal{L}_{pp}$ and adversarial loss $\mathcal{L}_{dis}$. To our knowledge, this is the largest and most geographically diverse open-source medical CT autoencoder to date, surpassing prior released backbones (LDM [20]: 0, MedDiff [27]: 2, MONAI [5]: 7, Med3D [6]: 8, MAISI [29]: 10). **(2)** A radiologist-verified degradation simulator transforms high-quality PCCT into three realistic low-quality counterparts: sparse-view (fewer 2D projections), low-dose (fewer photon counts), and conventional (lower resolution with more noise and artifacts). These cover the primary sources of EICT degradation. **(3)** SUMI takes a degraded input $\tilde{x}_i$ from step (2) and passes it through the autoencoder from step (1) to produce an enhanced output x_i'. Training uses the original PCCT x_i as ground truth under four losses: pixel-wise loss $\mathcal{L}_{pp}$ for structural fidelity; segmentation loss $\mathcal{L}_{seg}$ and HU consistency loss $\mathcal{L}_{HU}$ to preserve organ boundaries and tissue densities (such as pancreas, tumor, vessel, and airway tree) using pre-computed masks $SEG(\cdot)$; and adversarial loss $\mathcal{L}_{dis}$ via discriminator $D(\cdot)$ for image realism. Once trained, SUMI enhances any EICT scan to PCCT-like quality without retraining.

CT Degradation. We degrade high-quality PCCT scans to simulate common clinical artifacts and generate EICT-like appearances. A board-certified radiologist with thoracic imaging experience verified the realism of each degradation type on a sample of PCCT and matched degraded scans. We apply three degradation strategies. **Sparse View** reduces the number of projections in Radon space and induces streak artifacts:

$$\mathbf{p}_{\text{sparse}} = \mathcal{S}(\mathbf{p}), \quad \hat{\mathbf{x}}_{\text{sparse}} = \mathcal{R}(\mathbf{p}_{\text{sparse}}).$$

Low Dose decreases photon counts and models signal-dependent Poisson noise before reconstruction:

$$\hat{\mathbf{p}} \sim \text{Poisson}(\alpha \mathbf{p}), \quad \hat{\mathbf{x}}_{\text{low}} = \mathcal{R}(\hat{\mathbf{p}}).$$

Conventional Degradation applies spatial downsampling with Gaussian and Poisson noise injection. This mimics reduced resolution and electronic noise in standard CT scanners.

CT Enhancement. We train an enhancement network in the autoencoder's latent space. A degraded input $\tilde{x}$ is encoded into the latent space, refined by the network, and decoded into an enhanced output x'. The paired PCCT x serves as ground truth (step 3 of Figure 1). After training, the same network enhances routine EICT toward PCCT-like quality without retraining.

Training uses four losses (Figure 1). The **pixel-wise loss** $\mathcal{L}_{pp} = \|x' - x\|_1$ encourages structural fidelity. The **segmentation loss** $\mathcal{L}_{seg} = \text{CE}(\text{SEG}(x'), \text{SEG}(x))$ preserves organ boundaries using pre-computed masks $\text{SEG}(\cdot)$. The **HU consistency loss** $\mathcal{L}_{HU} = \sum_c \|\mu_c(x') - \mu_c(x)\|_1$ matches mean Hounsfield units μ_c within each region c. The **adversarial loss** $\mathcal{L}_{dis} = -\mathbb{E}[\log D(x')]$ with discriminator $D(\cdot)$ improves realism.

3 Experiment

Datasets and evaluation. Our autoencoder is pre-trained on 405,379 EICT scans from 145 hospitals. SUMI is then trained on 1,046 high-quality PCCT scans. We evaluate image quality (SSIM and PSNR) on 446 held-out private PCCT cases. We evaluate downstream detection on Luna16 [22], LNDb19 [19], and DSB17 [15]. We further release 17,316 public EICT scans enhanced to PCCT-like quality (Table 2). The release includes radiologist-validated voxel-wise annotations of airways, arteries, veins, lungs, and lobes.

Baseline and implementation. We select representative methods from four categories: traditional filtering (NLM [4]), Vision Transformers (Swin2SR [8]), GAN-based methods (Pix2Pix [14], SRGAN [16]), and diffusion methods (SR3 [21], NEED [10]). These methods were developed for natural images. We adapt and retrain each on our PCCT data using their official codebases. Full training scripts and configurations for all baselines and for SUMI are released in our GitHub repository. Many advanced medical enhancement methods remain closed-source, which limits reproducible comparison.

Table 1. SUMI improves image quality across degradations. On 446 external patients, SUMI surpasses the second-best method by up to +35.9% SSIM and +19.5% PSNR under sparse-view, low-dose, conventional, and mixed settings. Baselines are ordered by ascending average performance.

method	sparse view		low dose		conventional		mixed	
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
SR3 [21]	46.2	8.7	42.0	8.8	44.4	8.6	47.0	8.7
Pix2Pix [14]	43.6	11.9	60.0	22.6	60.8	18.1	62.9	22.7
Swin2SR [8]	62.1	26.7	59.9	28.0	29.0	24.9	66.8	26.4
NLM [4]	73.2	28.4	72.3	30.2	40.9	26.1	76.9	27.0
SRGAN [16]	56.8	23.2	85.5	33.0	82.8	32.4	63.1	26.5
NEED [10]	67.4	27.8	83.1	34.1	80.3	33.1	69.5	28.2
SUMI	**91.6**	**33.0**	**92.8**	**37.3**	**88.2**	**35.0**	**90.2**	**33.7**
Δ	**+35.9%**	**+18.7%**	**+8.5%**	**+9.4%**	**+6.5%**	**+5.7%**	**+17.3%**	**+19.5%**

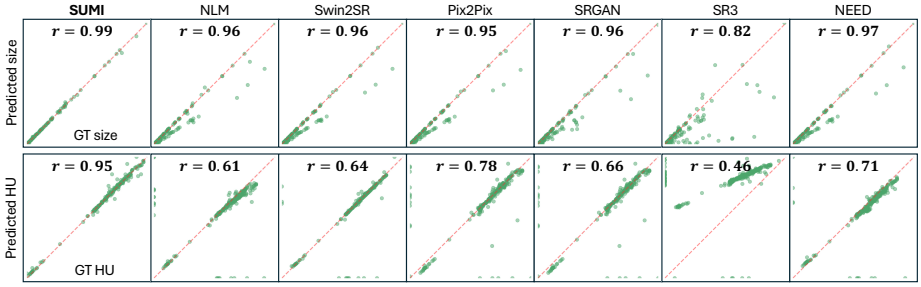

Fig. 2. SUMI preserves anatomical structure and tissue density. Pearson correlation (r) between ground truth and enhanced CT measurements across all organs shows that SUMI maintains anatomical accuracy and HU consistency. Organ masks are obtained with VISTA3D [13].

Image Quality and Enhancement. We compare SUMI with diverse baselines (Table 1). Across sparse-view, low-dose, conventional, and mixed degradations, SUMI consistently outperforms all competitors. It exceeds the second-best method by up to +35.9% SSIM and +19.5% PSNR. As shown in Figure 3, SUMI suppresses noise and streak artifacts. The output is sharp and closely resembles the ground truth.

Anatomical Fidelity and Tissue Consistency. Medical generative methods must preserve physical fidelity. Figure 2 reports the Pearson correlation (r) of organ size and Hounsfield units (HU) between ground truth and enhanced CT scans. Organ masks come from VISTA3D [12]. The results show that SUMI preserves anatomical boundaries and tissue densities without structural hallucination.

Downstream Clinical Performance. We test whether enhancement improves off-the-shelf detectors without modification. We use MONAI [5] for Luna16 and LNDb19, and grt123 [17] for DSB17. Both are released checkpoints pretrained on regular CT scans. We apply each detector to raw test scans and again to

Table 2. SUMI processes and enhances a multi-source cohort of 17,316 public chest CT scans. Cohorts are listed in alphabetical order.

dataset	DSB17 [15]	LIDC-IDRI [1]	LNDb19 [19]	LTRC [2]
scans (N)	1,596	1,018	324	1,496
dataset	Luna16 [22]	MIDRC [28]	NLST [25]	RSNA-STR [7]
scans (N)	854	10,496	422	1,110

Table 3. SUMI improves off-the-shelf detection across three independent chest CT cohorts. The detectors (MONAI, grt123) are off-the-shelf models pre-trained on regular CT scans and applied without any fine-tuning. The "baseline" row applies each detector to raw test scans. The "+SUMI" row applies the same frozen detector to the same test scans after SUMI enhancement. Improvements are therefore attributable to the input scans, not to the detector. Enhancement consistently increases performance on Luna16 [22], LNDb19 [19], and DSB17 [15], with gains up to +15.2% F1 and +10.5% AUC.

	Luna16 [22] (N=222) lung nodule, *MONAI*[1] [5]				**LNDb19 [19] (N=224)** lung nodule, *MONAI*[1] [5]				**DSB17 [15] (N=198)** lung tumor, *grt123*[2] [17]			
method	Sen.	Spec.	F1	AUC	Sen.	Spec.	F1	AUC	Sen.	Spec.	F1	AUC
baseline	75.8	88.4	81.6	88.0	61.5	86.9	55.2	74.0	78.4	82.1	80.2	87.0
+ SUMI	84.5	88.4	87.2	92.5	75.0	89.2	70.4	84.5	88.2	86.4	87.3	91.8
Δ	**+8.7**	+0.0	**+5.6**	**+4.5**	**+13.5**	**+2.3**	**+15.2**	**+10.5**	**+9.8**	**+4.3**	**+7.1**	**+4.8**

[1]MONAI: First-place open-source SOTA model for LUNA16/LNDb19 challenges.
[2]grt123: First-place SOTA model for DSB2017 challenge.

the same scans after SUMI enhancement. As shown in Table 3, enhancement increases F1 by up to +15.2% and AUC by up to +10.5% without any detector retraining. Quality score distributions further show a consistent shift toward the real PCCT standard.

Ablation Study. A leave-one-out ablation study (Table 4) validates our design. Omitting any single degradation strategy during training causes a clear performance drop in the corresponding scenario. All three simulations are essential for robust clinical generalization.

Table 4. Ablation of degradation simulations in SUMI. Leave-one-out experiments under identical settings show that removing any single degradation simulation degrades performance on an external test set (446 scans). All three simulations contribute to robust generalization.

method	sparse view		low dose		conventional		mixed	
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
w/o. sparse view	71.9	30.2	90.0	35.2	87.2	35.3	73.2	30.1
w/o. low dose	72.7	30.2	82.6	28.0	84.3	32.9	76.2	31.5
w/o. conventional	75.8	27.0	86.0	33.0	85.3	32.0	81.9	32.4
all degrades	**91.6**	**33.0**	**92.8**	**37.3**	**88.2**	**35.0**	**90.2**	**33.7**

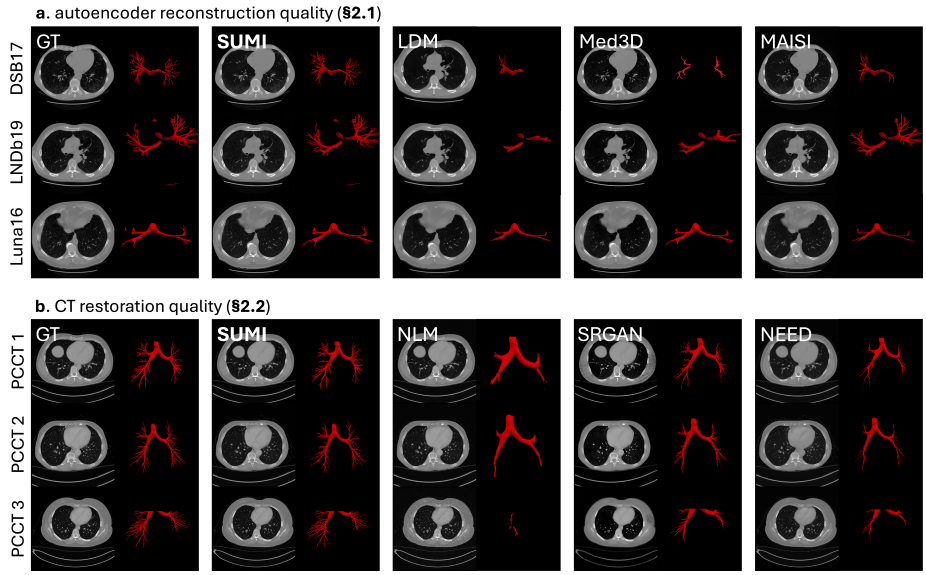

Fig. 3. SUMI delivers superior generalization and enhancement quality. For each example, the left image shows the CT slice and the right shows the airway tree segmentation. (a) Cross-dataset evaluation: compared with autoencoders trained on limited medical data (LDM, Med3D, MAISI), SUMI better preserves fine airway topology under domain shift. (b) PCCT enhancement with ground truth: compared with strong baselines (NLM [4], SRGAN [16], NEED [10]), SUMI maintains structural integrity and small airway branches more faithfully.

4 Discussion and Conclusion

Limitation. Our study has three limitations. First, evaluation is limited to chest CT, though extension to other anatomies is planned; chest CT is prioritized as its quality degradation is more pronounced than abdominal or pelvic CT. Second, the method operates on 2D slices rather than full 3D volumes; z-dimension continuity was validated across consecutive slices, with full 3D implementation deferred due to computational constraints. Third, enhancement may suppress low-contrast cues radiologists rely on to distinguish small nodules from noise. Segmentation and HU consistency losses partially mitigate this by anchoring anatomical boundaries and tissue densities to ground-truth PCCT, but a larger reader study is needed to quantify the risk.

Conclusion. We present an AI method that learns from real PCCT scans to enhance routine EICT scans toward PCCT-like quality by simulating realistic degradations and learning to reverse them. Results show substantial improvements in image quality and lesion detection, with a board-certified radiologist confirming clinical realism and anatomical fidelity. This paradigm distills the benefits of expensive specialized hardware into routine CT scanners using only

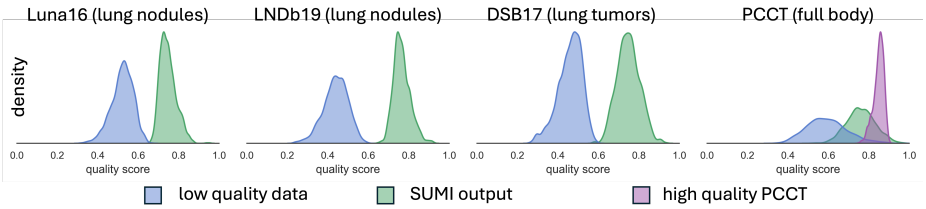

Fig. 4. SUMI improves chest CT quality across datasets. Quality scores on Luna16 [22], LNDb19 [19], DSB17 [15], and a private PCCT dataset. The scorer is an independent CNN-based image-quality classifier trained on radiologist-rated CT scans, outputs a value in $[0, 1]$ (higher is better), and was not used during SUMI training. On public datasets, enhanced scans consistently shift toward higher scores, indicating robust quality gains. On PCCT, a controlled degradation-to-enhancement benchmark shows that SUMI narrows the gap to real high-quality PCCT standards.

a limited number of high-quality reference scans — no costly upgrades required. Datasets, code, and models will be publicly available.

Acknowledgments. This work was supported by the Lustgarten Foundation for Pancreatic Cancer Research and the National Institutes of Health (NIH) under Award Number R01EB037669. We would like to thank the Johns Hopkins Research IT team in IT@JH for their support and infrastructure resources where some of these analyses were conducted; especially DISCOVERY HPC. We thank Jaimie Patterson for writing a news article about this project. Paper content is covered by patents pending.

Disclosure of Interests. The authors declare no competing interests.

References

1. Armato III, S.G., 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. Medical physics **38**(2), 915–931 (2011)
2. Bartholmai, B., Karwoski, R., Zavaletta, V., Robb, R., Holmes, D.: The lung tissue research consortium: An extensive open database containing histological, clinical, and radiological data to study chronic lung disease. The Insight Journal (2006)
3. van der Bie, J., van der Laan, T., van Straten, M., Booij, R., Bos, D., Dijkshoorn, M.L., Hirsch, A., Oei, E.H., Budde, R.P.: Photon-counting ct: an updated review of clinical results. European Journal of Radiology **190**, 112189 (2025)
4. Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). vol. 2, pp. 60–65. Ieee (2005)
5. Cardoso, M.J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., et al.: Monai: An open-source framework for deep learning in healthcare. arXiv preprint arXiv:2211.02701 (2022)
6. Chen, S., Ma, K., Zheng, Y.: Med3d: Transfer learning for 3d medical image analysis. arXiv preprint arXiv:1904.00625 (2019)
7. Colak, E., Kitamura, F.C., Hobbs, S.B., Wu, C.C., Lungren, M.P., Prevedello, L.M., Kalpathy-Cramer, J., Ball, R.L., Shih, G., Stein, A., et al.: The rsna pulmonary embolism ct dataset. Radiology: Artificial Intelligence **3**(2), e200254 (2021)
8. Conde, M.V., Choi, U.J., Burchi, M., Timofte, R.: Swin2sr: Swinv2 transformer for compressed image super-resolution and restoration. In: European conference on computer vision. pp. 669–687. Springer (2022)
9. Eulig, E., Ommer, B., Kachelrieß, M.: Benchmarking deep learning-based low-dose ct image denoising algorithms. Medical physics **51**(12), 8776–8788 (2024)
10. Gao, Q., Chen, Z., Zeng, D., Zhang, J., Ma, J., Shan, H.: Noise-inspired diffusion model for generalizable low-dose ct reconstruction. Medical image analysis **105**, 103710 (2025)
11. Hamamci, I.E., Er, S., Wang, C., Almas, F., Simsek, A.G., Esirgun, S.N., Dogan, I., Durugol, O.F., Hou, B., Shit, S., et al.: Developing generalist foundation models from a multimodal dataset for 3d computed tomography. arXiv preprint arXiv:2403.17834 (2024)
12. He, Y., Guo, P., Tang, Y., Myronenko, A., Nath, V., Xu, Z., Yang, D., Zhao, C., Simon, B., Belue, M., et al.: Vista3d: Versatile imaging segmentation and annotation model for 3d computed tomography. arXiv preprint arXiv:2406.05285 (2024)
13. He, Y., Guo, P., Tang, Y., Myronenko, A., Nath, V., Xu, Z., Yang, D., Zhao, C., Simon, B., Belue, M., et al.: Vista3d: A unified segmentation foundation model for 3d medical imaging. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 20863–20873 (2025)
14. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1125–1134 (2017)
15. Kuan, K., Ravaut, M., Manek, G., Chen, H., Lin, J., Nazir, B., Chen, C., Howe, T.C., Zeng, Z., Chandrasekhar, V.: Deep learning for lung cancer detection: tackling the kaggle data science bowl 2017 challenge. arXiv preprint arXiv:1705.09435 (2017)

16. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A.P., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: CVPR. vol. 2, p. 4 (2017)
17. Liao, F., Liang, M., Li, Z., Hu, X., Song, S.: Evaluate the malignancy of pulmonary nodules using the 3-d deep leaky noisy-or network. IEEE transactions on neural networks and learning systems **30**(11), 3484–3495 (2019)
18. Liu, J., Wu, Z., Bassi, P.R., Zhou, X., Li, W., Hamamci, I.E., Er, S., Lin, T., Luo, Y., Menze, B., Xu, D., Ding, K., Wang, K., Yang, Y., Tang, Y., Yuille, A.L., Zhou, Z.: See more, change less: Anatomy-aware diffusion for contrast enhancement. arXiv preprint arXiv:2512.07251 (2025), https://github.com/MrGiovanni/SMILE
19. Pedrosa, J., Aresta, G., Ferreira, C., Rodrigues, M., Leitão, P., Carvalho, A.S., Rebelo, J., Negrão, E., Ramos, I., Cunha, A., et al.: Lndb: a lung nodule database on computed tomography. arXiv preprint arXiv:1911.08434 (2019)
20. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684–10695 (2022)
21. Saharia, C., Ho, J., Chan, W., Salimans, T., Fleet, D.J., Norouzi, M.: Image super-resolution via iterative refinement. IEEE transactions on pattern analysis and machine intelligence **45**(4), 4713–4726 (2022)
22. Setio, A.A.A., Traverso, A., De Bel, T., Berens, M.S., Van Den Bogaard, C., Cerello, P., Chen, H., Dou, Q., Fantacci, M.E., Geurts, B., et al.: Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge. Medical image analysis **42**, 1–13 (2017)
23. Shah, K.D., Zhou, J., Roper, J., Dhabaan, A., Al-Hallaq, H., Pourmorteza, A., Yang, X.: Photon-counting ct in cancer radiotherapy: technological advances and clinical benefits. Physics in Medicine & Biology **70**(10), 10TR01 (2025)
24. Tavakoli, N., Shakeri, Z., Gowda, V., Samsel, K., Bedayat, A., Ghasemiesfe, A., Bagci, U., Hsiao, A., Leiner, T., Carr, J., et al.: Generative ai and foundation models in radiology: Applications, opportunities, and potential challenges. Radiology **317**(2), e242961 (2025)
25. Team, N.L.S.T.R.: The national lung screening trial: overview and study design. Radiology **258**(1), 243–253 (2011)
26. Tóth, A., Chetta, J.A., Yazdani, M., Matheus, M.G., O'Doherty, J., Tipnis, S.V., Spampinato, M.V.: Neurovascular imaging with ultra-high-resolution photon-counting ct: preliminary findings on image-quality evaluation. American Journal of Neuroradiology **45**(10), 1450–1457 (2024)
27. Wang, H., Liu, Z., Sun, K., Wang, X., Shen, D., Cui, Z.: 3d meddiffusion: A 3d medical latent diffusion model for controllable and high-quality medical image generation. IEEE Transactions on Medical Imaging (2025)
28. Whitney, H.M., Baughan, N., Myers, K.J., Drukker, K., Gichoya, J., Bower, B., Chen, W., Gruszauskas, N., Kalpathy-Cramer, J., Koyejo, S., et al.: Longitudinal assessment of demographic representativeness in the medical imaging and data resource center open data commons. Journal of Medical Imaging **10**(6), 061105–061105 (2023)
29. Zhao, C., Guo, P., Yang, D., He, Y., Tang, Y., Simon, B., Belue, M., Harmon, S., Turkbey, B., Xu, D.: Maisi-v2: Accelerated 3d high-resolution medical image synthesis with rectified flow and region-specific contrastive loss. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 40, pp. 13088–13098 (2026)