Any-to-One CT Kernel Conversion without Source-Kernel Data

Ji-Hoon Jung[1,2][0009−0002−7416−6727],
Sang Min Lee, MD, PhD[3][0000−0002−2173−2193], and
June-Goo Lee, PhD[1,2](✉)[0000−0002−1380−6682]

[1] Biomedical Engineering Research Center, Asan Institute for Life Sciences, Asan Medical Center, Seoul, South Korea
[2] Department of Biomedical Engineering, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea
[3] Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea
junegoo.lee@gmail.com

Abstract. The reconstruction kernel of a CT scan governs its spatial resolution and noise texture, and mismatched kernels degrade quantitative and downstream analysis. We present a method to convert an image of *any* reconstruction kernel to a single *reference* kernel (any-to-one) – spanning both intra-vendor and cross-vendor kernels – while requiring *only* images reconstructed with that reference kernel to build the model. Instead of paired multi-kernel scans or any source-kernel data, we synthesize (soft input, sharp target) training pairs by applying a *radial spectral degradation* to reference-kernel images: a frequency-domain radial low-pass that reshapes the noise power spectrum to follow real soft kernels more closely than a Gaussian blur. Concentrating the synthetic severity on the regime of clinically used soft kernels (hard-biased sampling) yields a single generator that maps a broad range of input softness to the reference kernel. On a paired Br40f→Br60f chest-CT test set of 238 volumes, the model reaches SSIM **0.929** / RMSE **58.2** HU, competitive with the strongest unpaired baseline that additionally uses source-kernel images; an ablation isolates the contribution of each design choice. By removing the per-kernel data requirement, the method offers a practical route to kernel harmonization when only reference-kernel images are available.

Keywords: CT kernel conversion · Image harmonization · Frequency-domain degradation · Noise power spectrum · Generative adversarial network.

1 Introduction

The reconstruction kernel sets a resolution–noise trade-off: sharp kernels resolve fine structures but amplify noise, whereas soft kernels suppress noise at the cost of resolution [4, 5]. Because the kernel varies across institutions, scanners, and vendors, identical anatomy appears with different texture—a known confounder

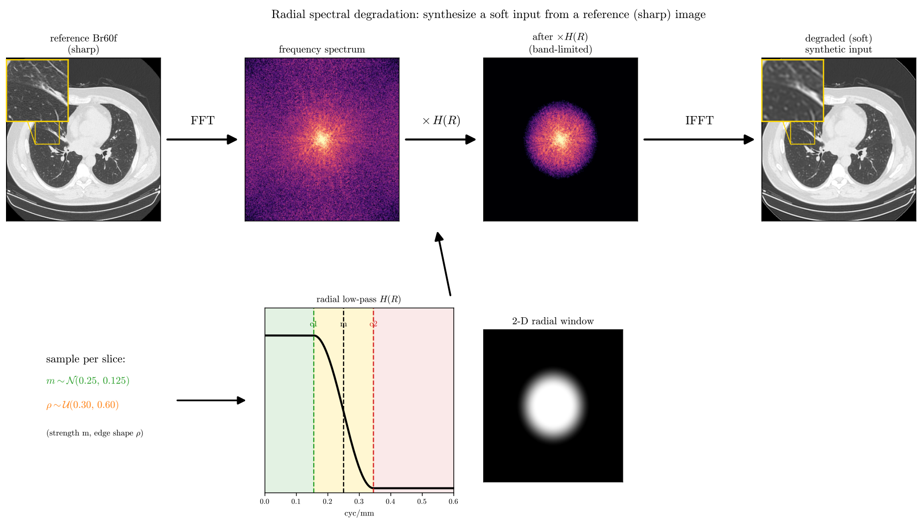

Fig. 1. Radial spectral degradation pipeline. A reference (sharp) Br60f image is multiplied in the frequency domain by a radial low-pass window $H(R)$ and inverse-transformed to a soft synthetic input; the sampled cutoff strength m and edge-shape ρ set the keep/fade/cut boundaries c_1, c_2 (bottom).

for radiomics and longitudinal comparison, where features track the kernel rather than the patient [1]. Harmonizing images to a common reference kernel is therefore a prerequisite for reproducible analysis [1, 10].

Most kernel-conversion methods cast the problem as image-to-image translation [7, 17], in paired [1, 14, 11] or, increasingly, unpaired formulations [16, 3, 10]. Regardless of formulation, they require training data of every kernel to be converted: a new input kernel—especially from another vendor—needs new images and retraining, so these methods cannot realize any-to-one conversion.

We avoid this data dependence by treating the kernel not as an unknown mapping to be learned from data, but through its *known physical effect*. In filtered back-projection, the kernel acts as an apodization window on spatial frequency, setting how much high-frequency content—of both signal and noise—enters the image [4, 5]. We model this effect with a *radial spectral degradation*—a frequency-domain low-pass—which, unlike a spatial Gaussian blur, attenuates signal and noise jointly and therefore reshapes the noise power spectrum (NPS) as well as the resolution (MTF), keeping the synthetic inputs closer to real soft-kernel images.

Because the synthesis uses reference-kernel images alone, training needs no paired scans and no source-kernel data. A single generator trained on such pairs maps any input kernel onto the reference in one forward pass. Since no source kernel or vendor is ever seen during training, the model is not tied to any source domain; we evaluate intra-vendor conversion with full-reference metrics, and cross-vendor conversion with an NPS-based measure.

Although much prior work converts toward soft kernels, we deliberately choose a sharp reference (Br60f): soft-to-sharp conversion must restore high-frequency detail and noise texture that the input lacks, and we adopt this most demanding direction as the strongest test of our method.

Contributions.

1. **Physics-grounded synthesis:** a radial spectral degradation that attenuates signal and noise jointly, reshaping both MTF and NPS with a kernel-like spectral shape that a spatial Gaussian blur does not provide.
2. **Reference-kernel-only training:** the converter is built from target-kernel images alone—no paired multi-kernel scans and no source-kernel data.
3. **Any-to-one conversion** with a single model, evaluated with full-reference metrics on intra-vendor Br40f→Br60f and with an NPS-based measure on cross-vendor kernels.

2 Method

2.1 Overview

We address *any-to-one* reconstruction-kernel conversion: given a chest CT image reconstructed with an arbitrary kernel, produce its appearance under a single fixed *reference* kernel (Br60f). Our central design choice is that **training uses only reference-kernel images**. As illustrated in Fig. 1, from each reference image we synthesize a degraded counterpart with a physically-motivated *radial spectral degradation* (Sec. 2.2), drawing its severity from a *hard-biased* distribution (Sec. 2.3); the resulting (degraded input, reference target) pairs train a single feed-forward generator adversarially (Sec. 2.4). At inference the generator maps an image directly, with no kernel identification and no paired data; inputs sharper than the reference are first brought into the trained range by the same radial low-pass (Sec. 2.5), so one model spans the full range of input kernels, including other vendors (Sec. 3.4).

All images are 512×512 axial slices, clipped to $[-1000, 1000]$ HU and rescaled to $[-1, 1]$; the in-plane pixel spacing Δ (mm/pixel), read from DICOM, converts pixel frequencies to cyc/mm.

2.2 Radial Spectral Degradation

A softer kernel attenuates high frequencies of both the anatomical signal and the correlated noise. We approximate this with an isotropic low-pass window applied in the 2-D Fourier domain to the *entire* image, so signal and noise are attenuated jointly; directional anisotropy of the true kernel is ignored.

Let $x \in \mathbb{R}^{512\times512}$ be a reference-kernel image and $X = \mathcal{F}\{x\}$ its 2-D DFT. Define per-axis frequencies $f_i = \mathrm{fftfreq}(N, \Delta)$ (cyc/mm) with $N = 512$, and the radial frequency $R(u,v)=\sqrt{f_u^2+f_v^2}$ (cyc/mm). The degradation is parameterized by a cutoff strength m (cyc/mm)—the frequency at which the window has faded to half—and a roll-off shape parameter $\rho \in (0,1)$. We set the transition band $[c_1, c_2]$ as

$$c_2 = \frac{2m}{1+\rho}, \qquad c_1 = \rho\, c_2, \tag{1}$$

so that the window equals $\frac{1}{2}$ at the midpoint $\frac{c_1+c_2}{2} = m$; a smaller m removes more high-frequency content (a softer result). The radial window is a raised-cosine (Hann) low-pass:

$$H(R) = \begin{cases} 1, & R < c_1, \\ \frac{1}{2}\left[1+\cos\left(\pi \frac{R-c_1}{c_2-c_1}\right)\right], & c_1 \le R < c_2, \\ 0, & R \ge c_2. \end{cases} \tag{2}$$

The degraded image is $\tilde{x} = \mathrm{Re}\,\mathcal{F}^{-1}\{X \odot H\}$. Because H multiplies the spectrum of the full image, the degraded noise power spectrum satisfies $\mathrm{NPS}_{\tilde{x}} = |H|^2\,\mathrm{NPS}_x$, so the noise texture is shaped by the same window as the resolution. A spatial Gaussian blur is also a low-pass but with a fixed roll-off; our window instead lets (m,ρ) control the cutoff and edge shape to better follow a real soft kernel (Sec. 3.3).

2.3 Hard-Biased Severity Sampling

While the degradation *operator* is physically motivated, the severity *distribution* is a design choice for input coverage. For every training sample we draw the severity independently:

$$m \sim \max\bigl(\mathcal{N}(\mu{=}0.25,\ \sigma{=}0.125),\ 0.02\bigr)\ \text{cyc/mm}, \qquad \rho \sim \mathcal{U}(0.10,\ 0.60). \tag{3}$$

The mean $\mu = 0.25$ cyc/mm lies toward the strong-softening end, so most samples are **heavily degraded (hard) inputs**, focusing capacity on the information-poor regime, while the unclipped upper tail ($m > 0.5$) still covers mild, near-identity cases. The lower clip at 0.02 avoids a blank image, and randomizing ρ prevents overfitting a single roll-off shape.

2.4 Restoration Network and Training Objective

Network. Because our degradation yields exactly-aligned (soft, sharp) pairs, we train a single generator G and PatchGAN discriminator D [7] as a paired translator (conditional-GAN, not cycle-consistency; Fig. 2). For G we reuse the adaptive-normalization residual U-Net of the strongest baseline (Sec. 3.2) [16] unchanged; its invertible polyphase downsampling discards no high frequencies and its adaptive-normalization decoder reproduces kernel-specific texture well [6, 8, 13]. The conditional-GAN translator itself is well established—pix2pix-style

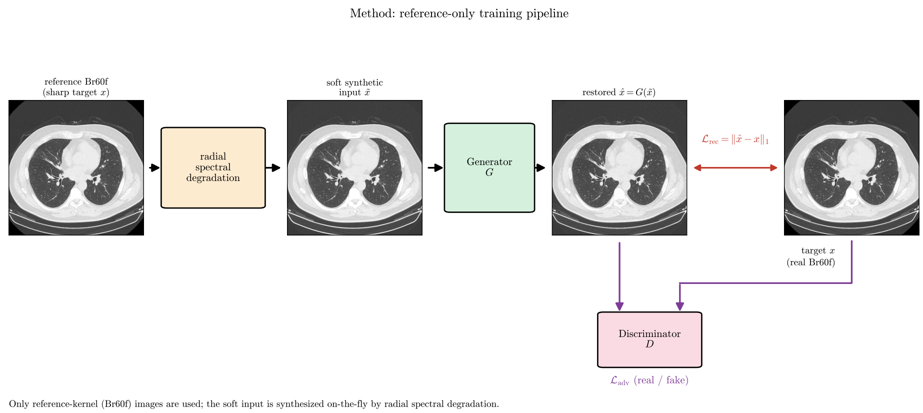

Fig. 2. Reference-only training. A Br60f slice x is degraded to a soft input $\tilde{x}$ (Sec. 2.2) and restored by G; training uses an L_1 loss (red) and a PatchGAN adversarial loss.

networks have been applied to kernel conversion before [14, 11]—and our focus is instead the frequency-domain synthesis of the training pairs.

Objective. With degraded input $\tilde{x}$, target x, and $\hat{x} = G(\tilde{x})$, we use a least-squares GAN [12] with a pixel-wise L_1 term:

$$\mathcal{L}_G = \big\|D(\hat{x}) - 1\big\|_2^2 + \lambda_{\text{rec}}\big\|\hat{x} - x\big\|_1, \qquad \mathcal{L}_D = \big\|D(x) - 1\big\|_2^2 + \big\|D(\hat{x}) - 0\big\|_2^2. \tag{4}$$

We optimize with Adam (β_1=0.5, β_2=0.999, constant lr 1×10^{-5}, batch size 48, $\lambda_{\text{rec}} = 10$), updating G and D alternately with $\hat{x}$ detached for D. Each epoch draws one random axial slice from each of the 4,157 reference volumes and degrades it on-the-fly; we train from scratch for 100 epochs. Every model in this paper follows this identical protocol, so comparisons isolate the studied design choice.

2.5 Inference

At test time an image is mapped in a single forward pass, $\hat{x} = G(x_{\text{in}})$, and rescaled to HU. Because training only ever exposes the generator to inputs at or below the reference resolution, inputs *sharper* than the reference are out-of-distribution. We handle them with a single fixed pre-filter: the same radial low-pass (Sec. 2.2) with $m = 0.37$ cyc/mm and $\rho = 0.13$, applied once before the generator. These values were fixed once by fitting the operator to the NPS of real Br40f images in CT-RATE. As a low-pass only removes high frequencies, it caps a sharper input at the trained resolution while leaving already-softer inputs essentially unchanged, so the *same* pre-filter can be applied uniformly. Inference thus requires no knowledge of the input kernel, no severity estimation, no conditioning input, and no paired data—the same generator handles any input kernel, realizing any-to-one conversion.

Table 1. Comparison with unpaired methods on the paired Br40f→Br60f test set (238 volumes; mean ± std over volumes). Our model trains on Br60f only.

Method	Train data	SSIM ↑	PSNR ↑	RMSE (HU) ↓
Switchable CycleGAN [16]	Br40f/Br60f	0.937 ± 0.037	37.97 ± 1.78	53.2 ± 27.3
Multipath CycleGAN [10]	Br40f/Br60f	0.912 ± 0.037	35.99 ± 1.52	66.2 ± 26.4
GGCL [3]	Br40f/Br60f	0.690 ± 0.057	29.75 ± 1.23	134.1 ± 26.2
Ours	Br60f only	0.929 ± 0.035	37.14 ± 1.54	58.2 ± 26.7

3 Experiments

3.1 Dataset and Evaluation Protocol

We use chest CT from CT-RATE with two Siemens kernels, Br40f (input) and Br60f (reference), reconstructed from the *same* acquisition, hence spatially registered and allowing per-pixel comparison against a ground-truth reference. Training uses all Br60f volumes of the official CT-RATE training split (4,157 volumes from 3,715 patients); no other kernel or split is seen during training. The test set is the 238 validation-split acquisitions that provide both a Br40f and a Br60f reconstruction of the same chest scan. The CT-RATE splits are patient-level, so no patient appears in both training and test. We report 3-D SSIM, PSNR, and RMSE (HU) on lung slices, identified with TotalSegmentator lung labels [15]. Crucially, our model is trained on Br60f *only*, whereas the compared methods use both Br40f and Br60f.

3.2 Comparison with Prior Methods

We compare against the publicly-available *unpaired* kernel-conversion methods Switchable CycleGAN [16], GGCL [3], and Multipath CycleGAN [10] (authors' public implementation [9]), each trained on Br40f/Br60f under its own protocol. Our emphasis is on the frequency-domain synthesis rather than the network itself: Switchable CycleGAN is the strongest (Table 1), so we adopt its generator, retrained as a paired conditional GAN to fit our setting (Sec. 2.4). Even so, our model—using Br60f alone and never seeing the source domain—matches it (SSIM 0.929 vs. 0.937, RMSE 58.2 vs. 53.2 HU) and clearly surpasses the others (Table 1, Fig. 3).

3.3 Ablation Study

We ablate the degradation design with the network architecture and training protocol held fixed, varying only how the soft input is synthesized. The *blur baseline* uses a 2σ Gaussian blur in place of our radial operator (trained as blur(Br60f) → Br60f); we then swap in the *radial spectral degradation* (Sec. 2.2); and finally add *hard-biased sampling* (Sec. 2.3), giving our full model. Table 2 and Fig. 3 show that both changes help: replacing the Gaussian blur with the

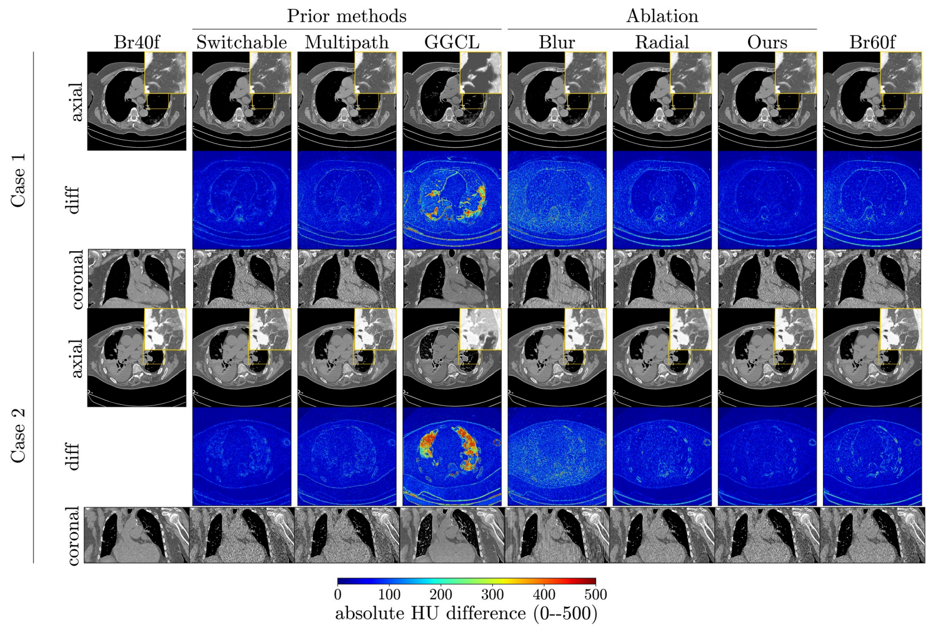

Fig. 3. Qualitative results (Cases 1–2). Rows: axial with a lung zoom (WL/WW = −700/1500 HU), error map (0–500 HU), coronal view. Columns: *prior methods* (Switchable/Multipath/GGCL) and *ablation* (Blur/Radial/Ours); error is |GT−prediction| (GT column: |input−GT|, the difference to remove). Ours has the smallest error and no z-banding.

radial operator—so that the synthesized noise texture (NPS), not just resolution, matches a real kernel—gives by far the largest gain (SSIM $0.735 \rightarrow 0.902$), and hard-biased sampling instead of uniform sampling adds a further improvement ($\rightarrow 0.929$). Although the generator runs per axial slice, the coronal views in Fig. 3 show no z-direction banding.

3.4 Cross-Vendor Generalization

Since training never sees any other vendor, we ask whether the same model converts other-vendor kernels toward Br60f. As no paired reference exists, we evaluate the noise power spectrum (NPS), an anatomy-independent kernel descriptor [10]: from radially-averaged NPS [4] on soft-tissue patches we take its centroid $\bar{f}$ (cyc/mm; higher = sharper) and compare to the full Br60f test set ($\bar{f} = 0.218 \pm 0.043$). On two soft kernels from other vendors—Philips "B" and PNMS "SA" (30 volumes each), each passed through the inference pre-filter (Sec. 2.5)—the converter restores the Br60f noise level: $\bar{f}\, 0.094 \rightarrow 0.198$ (B) and $0.097 \rightarrow 0.237$ (SA), both close to the 0.218 target (Table 3).

B and SA are both soft kernels. Kernels *sharper* than Br60f exist in CT-RATE (e.g. Siemens Bl64f/Br64f, PNMS LungB) but are almost all reconstructed at

Table 2. Ablation of the degradation design (same architecture and training protocol; 238-volume test set, mean $\pm$ std).

Configuration	SSIM ↑	PSNR ↑	RMSE (HU) ↓
Gaussian blur (2σ)	0.735 ± 0.065	31.56 ± 1.41	109.4 ± 27.2
Radial (uniform)	0.902 ± 0.039	35.69 ± 1.53	68.5 ± 26.4
Radial + hard sampling (**Ours**)	0.929 ± 0.035	37.14 ± 1.54	58.2 ± 26.7

1024×1024, leaving fewer than seven 512×512 volumes per kernel—too few for the distributional analysis, so we check them qualitatively on the few 512×512 Siemens Bl64f. Figure 4 shows one representative conversion per source kernel next to a real Br60f case. Owing to this dataset constraint, a broader cross-vendor validation is left to future work.

4 Conclusion

We presented an any-to-one CT kernel converter trained on reference-kernel images alone, using a physically-motivated radial spectral degradation to synthesize soft inputs from Br60f. Without source-kernel or paired data it matches the strongest unpaired baseline (SSIM 0.929, RMSE 58.2 HU) and, seeing no source domain in training, also converts other-vendor kernels toward the Br60f noise level.

Limitations and future work. (i) Cross-vendor conversion matches the Br60f noise level ($\bar{f}$) but not the full NPS shape, and the restored noise texture was not evaluated at scale. (ii) We report image- and noise-level metrics (SSIM, RMSE, NPS), not downstream tasks; downstream clinical evaluation [2, 1] is planned as future work. (iii) Results are limited to chest CT, one dataset, one target kernel, and few cross-vendor kernels; broader validation is planned as future work, and per-vendor kernel conditioning is a natural extension.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Table 3. Cross-vendor conversion by NPS centroid $\bar{f}$ (cyc/mm, mean±std); no paired GT. Target Br60f: $\bar{f} = 0.218 \pm 0.043$ (238 vol); each source 30 vol.

Source	Vendor	$\bar{f}$ source	$\bar{f}$ converted (target 0.218 ± 0.043)
B	Philips	0.094 ± 0.009	0.198 ± 0.042
SA	PNMS	0.097 ± 0.015	0.237 ± 0.043

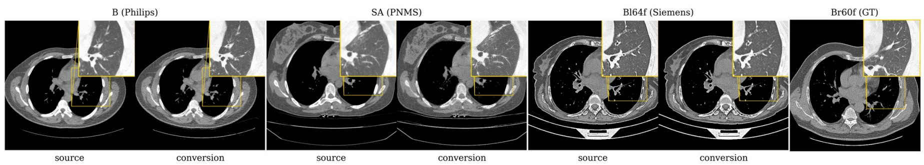

Fig. 4. Cross-vendor conversion. For each source kernel, the source and its conversion are boxed together; the rightmost panel is a real Br60f case (GT) for reference. B and SA reach the Br60f centroid, and the sharper-than-Br60f Bl64f is softened toward it. Soft-tissue window with a lung zoom inset (gold), as in Fig. 3.

References

1. Choe, J., Lee, S.M., Do, K.H., Lee, G., Lee, J.G., Lee, S.M., Seo, J.B.: Deep learning–based image conversion of CT reconstruction kernels improves radiomics reproducibility for pulmonary nodules or masses. Radiology **292**(2), 365–373 (2019). https://doi.org/10.1148/radiol.2019181960
2. Choe, J., Yun, J., Kim, M.J., Oh, Y.J., Bae, S., Yu, D., Seo, J.B., Lee, S.M., Lee, H.Y.: Leveraging deep learning-based kernel conversion for more precise airway quantification on CT. European Radiology **35**(11), 7185–7198 (2025). https://doi.org/10.1007/s00330-025-11696-w
3. Choi, C., Jeong, J., Lee, S., Lee, S.M., Kim, N.: CT kernel conversion using multi-domain image-to-image translation with generator-guided contrastive learning. In: Medical Image Computing and Computer Assisted Intervention (MICCAI). LNCS, vol. 14229 (2023). https://doi.org/10.1007/978-3-031-43999-5_33
4. Friedman, S.N., Fung, G.S.K., Siewerdsen, J.H., Tsui, B.M.W.: A simple approach to measure computed tomography (CT) modulation transfer function (MTF) and noise-power spectrum (NPS) using the American College of Radiology (ACR) accreditation phantom. Medical Physics **40**(5), 051907 (2013). https://doi.org/10.1118/1.4800795
5. Gang, G.J., Stayman, J.W., Zbijewski, W., Siewerdsen, J.H.: Task-based detectability in CT image reconstruction by filtered backprojection and penalized likelihood estimation. Medical Physics **41**(8), 081902 (2014). https://doi.org/10.1118/1.4883816
6. Huang, X., Belongie, S.: Arbitrary style transfer in real-time with adaptive instance normalization. In: IEEE International Conference on Computer Vision (ICCV) (2017)
7. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017)
8. Karras, T., Laine, S., Aila, T.: A style-based generator architecture for generative adversarial networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
9. Krishnan, A.R., Li, T.Z., Remedios, L.W., Xu, K., Zuo, L., Sandler, K.L., Maldonado, F., Landman, B.A.: Anatomy-guided multi-path cycleGAN for lung CT kernel harmonization. Medical Imaging with Deep Learning (MIDL), OpenReview (2025), `https://openreview.net/forum?id=w3p7GddsQ8`, code: `https://github.com/MASILab/AnatomyconstrainedMultipathGAN`
10. Krishnan, A.R., Xu, K., Li, T., Gao, C., Remedios, L.W., Kanakaraj, P., Lee, H.H., Bao, S., Sandler, K.L., Maldonado, F., Išgum, I., Landman, B.A.: Multipath cycleGAN for harmonization of paired and unpaired low-dose lung computed tomography reconstruction kernels. Medical Physics (2025). https://doi.org/10.1002/mp.70120
11. Krishnan, A.R., Xu, K., Li, T.Z., Remedios, L.W., Sandler, K.L., Maldonado, F., Landman, B.A.: Lung CT harmonization of paired reconstruction kernel images using generative adversarial networks. Medical Physics **51**(8), 5510–5523 (2024). https://doi.org/10.1002/mp.17028
12. Mao, X., Li, Q., Xie, H., Lau, R.Y.K., Wang, Z., Smolley, S.P.: Least squares generative adversarial networks. In: IEEE International Conference on Computer Vision (ICCV) (2017)

13. Park, T., Liu, M.Y., Wang, T.C., Zhu, J.Y.: Semantic image synthesis with spatially-adaptive normalization. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
14. Tanabe, N., Kaji, S., Shima, H., Shiraishi, Y., Maetani, T., Oguma, T., Sato, S., Hirai, T.: Kernel conversion for robust quantitative measurements of archived chest computed tomography using deep learning-based image-to-image translation. Frontiers in Artificial Intelligence **4**, 769557 (2022). https://doi.org/10.3389/frai.2021.769557
15. Wasserthal, J., Breit, H.C., Meyer, M.T., Pradella, M., Hinck, D., Sauter, A.W., Heye, T., Boll, D.T., Cyriac, J., Yang, S., Bach, M., Segeroth, M.: TotalSegmentator: Robust segmentation of 104 anatomic structures in CT images. Radiology: Artificial Intelligence **5**(5), e230024 (2023). https://doi.org/10.1148/ryai.230024
16. Yang, S., Kim, E.Y., Ye, J.C.: Continuous conversion of CT kernel using switchable CycleGAN with AdaIN. IEEE Transactions on Medical Imaging **40**(11), 3015–3029 (2021). https://doi.org/10.1109/TMI.2021.3077615
17. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: IEEE International Conference on Computer Vision (ICCV) (2017)