NASD-Diff: A Noise-Assisted Structural Difference Diffusion Model for CT Angiography Generation from NCCT

Zeng Xie[1], Lijun Guo[1⋆], Bang Chen[1], Yuning Pan[2], Wenming He[3], Xianwang Ye[2], and Rong Zhang[1]

[1] Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China
guolijun@nbu.edu.cn
[2] Department of Radiology, The First Affiliated Hospital of Ningbo University, Ningbo 315010, China
[3] Department of Cardiology, The First Affiliated Hospital of Ningbo University, Ningbo 315010, China

Abstract. Generating CT Angiography (CTA) from the latent vascular structural information in Non-Contrast CT (NCCT) has significant clinical value and remains highly challenging. With the successful application of diffusion models in medical image modality conversion tasks, research on NCCT-to-CTA generation has been promoted. However, most diffusion-based methods still use random noise to model the entire CTA image, lacking explicit modeling of vascular enhancement changes caused by contrast agent injection, which limits the quality of vascular visualization in generated images. To address this limitation, we propose a Noise-Assisted Structural Difference Diffusion Model (NASD-Diff) for NCCT-to-CTA generation. In the forward process, a noise-assisted structural difference degradation operator explicitly models the attenuation of contrast-enhanced vascular signals while incorporating controllable noise for high-frequency texture modeling. In the reverse process, a dual-decoder network decouples the prediction of vascular structural differences and noise components. Furthermore, a cross-slice conditional embedding module leverages contextual information to enhance the quality of image generation, and a global perception refinement module improves the overall visual perception of the generated images. Experiments on a real clinical dataset demonstrate that NASD-Diff significantly outperforms existing methods in PSNR, SSIM, RMSE, and multiple perceptual metrics, highlighting its effectiveness and clinical potential for high-quality CTA synthesis without contrast agents.

Keywords: CT Angiography · Non-Contrast CT · Medical image translation · Degradation operator · Diffusion model.

⋆ Corresponding author

1 Introduction

CT Angiography (CTA) is a technique crucial for assessing vascular diseases such as atherosclerosis, aneurysms, aortic dissection, and stenosis[11]. However, CTA requires intravenous iodinated contrast agents[17], which pose risks for patients with iodine allergy, renal impairment, or multiple myeloma[2], and repeated use may further increase the risk of adverse reactions[4]. The high cost and procedural complexity of CTA examinations also limit their accessibility and clinical adoption. In contrast, Non-Contrast CT (NCCT) is simpler and safer but provides limited vascular visualization, which restricts its application in vascular disease assessment. Nevertheless, NCCT contains latent vascular information, and generating CTA-like images from NCCT using generative models could reduce reliance on contrast agents, providing a safer and more cost-effective alternative for high-risk patients and enabling contrast-free CTA assessment.

With the advancement of generative models, medical image modality translation tasks have gradually attracted increasing attention. Lyu et al.[14] employed a conditional generative adversarial network to synthesize CTA images from NCCT images, but due to training instability and mode collapse, the generated images still exhibited significant flaws in structural consistency and detail stability. With the rise of diffusion models, Jiang et al.[10] proposed Fast-DDPM, which aligns training and sampling processes and adopts a non-uniform noise scheduling strategy to reduce sampling steps while maintaining high generation quality, achieving strong performance in MRI modality translation. Gao et al.[6] introduced CoreDiff, a contextual error-modulated generalized diffusion model that models the degradation from normal-dose CT to low-dose CT for denoising and improved generalization. Nevertheless, most diffusion-based medical image modality conversion methods still rely primarily on global random noise in the forward degradation process, which confines the modeling to the global image distribution[16,12]. In the NCCT-to-CTA conversion process, contrast agent injection induces varying degrees of distributional changes across different tissue regions, with vascular regions exhibiting the most pronounced variation. Therefore, it is worth considering a model that prioritizes vascular modeling while leveraging noise to assist in modeling other texture details.

To address these limitations, we propose a novel Noise-Assisted Structural Difference Diffusion Model (NASD-Diff) for generating CTA images from NCCT, as shown in Fig. 1. Specifically, we propose a noise-assisted structural difference degradation operator (NASD), which transforms the diffusion process into a contrast signal attenuation process while introducing Gaussian noise to guide the modeling of high-frequency textures, thereby preventing excessive smoothing in the generated images. To support this, we construct a Decoupled Structural-Texture Network (DeST-Net) that decouples structural difference modeling from texture restoration through a dual-branch decoder, enabling parallel optimization of vascular morphology and image texture. Furthermore, DeST-Net incorporates multi-slice contextual information to improve image generation quality, and the introduction of visual perceptual loss regulates global contrast and semantic consistency, leading to improved global structural fidelity and texture quality

of the generated images. The main contributions of our work are: (1)**A noise-assisted structural difference modeling strategy:** This strategy explicitly captures contrast-induced vascular structural changes between NCCT and CTA in the forward diffusion process and incorporates a noise-assisted texture enhancement strategy, thereby improving the generation of vascular structures and global texture details. (2)**A generative network for decoupling the prediction of structural differences and noise:** A dual-branch decoder performs parallel optimization of vascular structures and texture details, enabling independent prediction of structural differences and texture-related noise components, thereby enhancing both structural accuracy and detail clarity in the generated CTA images. (3)**Multi-slice contextual embedding and perceptual loss:** The conditional embedding leverages multiple adjacent slices to improve image generation quality, and a perceptual loss further enhances global structural consistency and texture realism. (4)**Experimental validation:** Experiments on a clinical dataset demonstrate that NASD-Diff outperforms existing methods in terms of vascular accuracy and texture fidelity, confirming its effectiveness in high-quality cross-modal image generation and clinical potential.

2 Methods

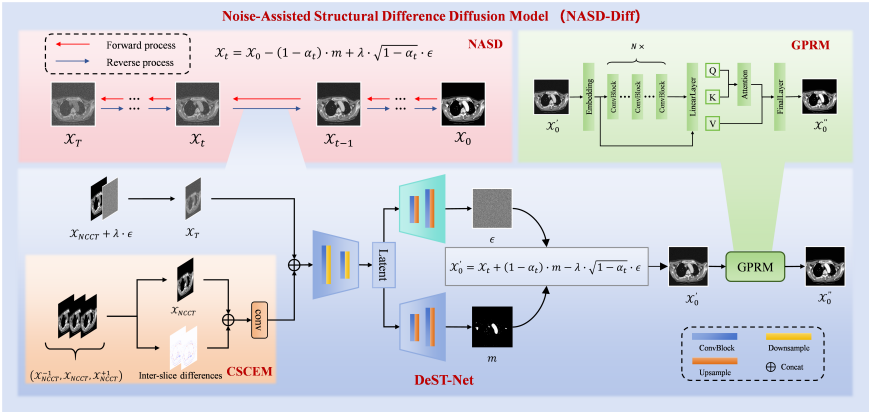

Fig. 1. Overview of the proposed NASD-Diff for generating CTA from NCCT.

2.1 Preliminaries

We first provide a brief introduction to the generalized diffusion framework underlying our approach. This framework provides a unified formulation of the forward degradation and reverse restoration processes[3]. Specifically, given an initial image x_0, it is progressively transformed into x_T, which follows a known distribution P, via a customizable operator $D(\cdot)$. In a standard diffusion model, the

forward process is defined as: $x_t = D(x_0, x_T, t) = \sqrt{\alpha_t} \cdot x_0 + \sqrt{1-\alpha_t} \cdot x_T$, where T is the total number of degradation steps, t denotes the current step, α_t is the timestep scheduling parameter, and x_T denotes random noise. In the reverse process, the model samples x_T from P, and reconstructs the image through a restoration operator $R(\cdot)$ parameterized by a neural network. To prevent blurring caused by single-step prediction, [3] employed a progressive "restoration-redegradation" algorithm, and an improved sampling strategy is applied to mitigate error accumulation during iterative restoration: $x_{t-1} = x_t - D(\widehat{x}_0, x_T, t) + D(\widehat{x}_0, x_T, t-1)$, where $\widehat{x}_0$ is obtained from the restoration operator $R(\cdot)$.

2.2 Noise-assisted structural difference degradation operator

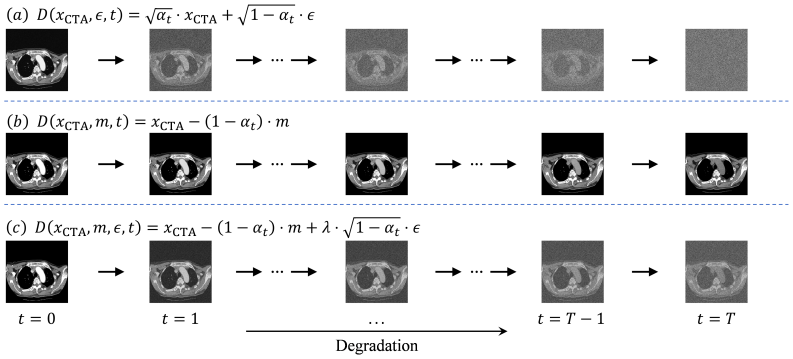

Fig. 2. Comparison of different degradation processes.

As shown in Fig. 2(a), previous diffusion-based models for modality conversion progressively corrupt the target image into pure noise and use the source image to guide target image generation. To better align the degradation with the CTA-to-NCCT transformation, we design a structural difference degradation operator inspired by [6], as shown in Fig. 2(b). Specifically, given an NCCT image x_{NCCT} and its corresponding CTA image x_{CTA}, the structural difference is defined as: $m = x_{\text{CTA}} - x_{\text{NCCT}}$, where m primarily consists of vascular enhancement information introduced by contrast agent injection. Based on this definition, we construct the degradation operator as:

$$x_t = D(x_{\text{CTA}}, m, t) = x_{\text{CTA}} - (1-\alpha_t) \cdot m \tag{1}$$

this operator leverages the inter-modality structural difference m to reformulate degradation process as a gradual modality conversion from CTA to NCCT, simulating the progressive attenuation of the contrast agent signal.

However, this operator relies solely on linearly scaling m without noise, limiting supervision for high-frequency details and resulting in over-smoothed image

details. To address this, we propose a noise-assisted structural difference degradation operator (NASD), as shown in Fig. 2(c), which adds a random noise term in Eq.(1) to model both vascular structures and high-frequency information:

$$x_t = D(x_{\mathrm{CTA}}, m, \epsilon, t) = x_{\mathrm{CTA}} - (1 - \alpha_t) \cdot m + \lambda \cdot \sqrt{1 - \alpha_t} \cdot \epsilon \tag{2}$$

where ϵ denotes random Gaussian noise, and λ is the noise control coefficient. To investigate the role of noise in frequency modeling, we further analyze Eq.(2) in the frequency domain. As the Power Spectral Density of general images generally decreases with increasing spatial frequency, while Gaussian noise exhibits a flat spectrum, high-frequency components are more easily affected during the diffusion process[12]. Consequently, by adjusting λ, the final Signal-to-Noise Ratio can be controlled to induce the degradation of high-frequency components for texture supervision while relatively preserving low-frequency structures, allowing the degradation operator to perform texture modeling while relatively preserving the dominant role of the degradation of structural differences.

2.3 Reverse restoration process

In this subsection, we introduce our proposed restoration network architecture.

Decoupled Structural-Texture Network: In the forward degradation process, both the structural difference term m and the noise term ϵ are involved. Therefore, we design a DeST-Net to predict the structural difference m and the texture-related noise ϵ in parallel during the reverse restoration process. Based on the predicted m and ϵ, the preliminary CTA image x_0' is obtained by reformulating Eq.(2). As shown in Fig. 1, DeST-Net adopts an encoder–decoder network composed solely of 3×3 convolutions following the design philosophy of [20], with multistage downsampling and upsampling to enlarge the receptive field and sparse skip connections to reduce redundant feature transmission.

Cross-Slice Conditional Embedding Module: To leverage inter-slice context[18], we introduce the CSCEM, which inputs the current and adjacent slices to compute inter-slice difference maps. These maps are concatenated with the current slice for convolution, enriching the conditional representation to enable the network to better perceive inter-slice variations within a 2D framework, thereby generating higher-quality images.

Global Perception Refinement Module: To enhance the visual quality of generated CTA images, we employ GPRM to project the initial estimate x_0' into a high-dimensional feature space and refines it through multiple fixed-resolution 3×3 convolutional layers. A lightweight cross-attention fuses these features with the initial ones to produce x_0''. The perceptual loss[1], computed using a pre-trained LPIPS(VGG) network on replicated three-channel images, is applied to x_0'' to update only the parameters of GPRM while freezing the backbone network. This design enables GPRM to focus on perceptual optimization while ensuring stable training of the backbone network.

Training and Sampling: In the training process, we decompose the overall training objective into three subtasks, and impose separate constraints on each:

$$\begin{aligned} &\mathcal{L}_\epsilon = \|\widehat{\epsilon} - \epsilon\|^2, \quad \mathcal{L}_m = \|\widehat{m} - m\|^2, \\ &\mathcal{L}_{\text{CTA}} = \|x_0'' - x_{\text{CTA}}\|^2, \quad \mathcal{L}_{\text{lpips}} = d_{\text{LPIPS}}(x_0'',\ x_{\text{CTA}}), \\ &\mathcal{L}_{\text{GPRM}} = \lambda_{\text{CTA}} \cdot \mathcal{L}_{\text{CTA}} + \lambda_{\text{lpips}} \cdot \mathcal{L}_{\text{lpips}} \end{aligned} \tag{3}$$

where $\mathcal{L}\epsilon$, $\mathcal{L}_m$, and $\mathcal{L}_{\text{CTA}}$ are mean squared errors for noise prediction, structural difference prediction, and CTA image prediction, respectively; $\mathcal{L}_{\text{GPRM}}$ is the joint loss of GPRM, with λ_{CTA} and λ_{lpips} as the corresponding weighting coefficients; $\mathcal{L}_{\text{lpips}}$ is the perceptual loss on CTA images; $d_{\text{LPIPS}}(\cdot)$ denotes a deep-feature perceptual similarity metric measuring visual consistency between generated and real CTA images. In the sampling process,we first obtain the initial state x_T by adding λ-scaled noise to the current NCCT image. Then, we employ the restoration-redegradation generation strategy and apply the algorithm proposed in [3] at each step. Following observations that later timesteps are more critical for perceptual quality[5], we adopt a non-uniform schedule[10] with sparse early and dense later steps to improve generation quality.

3 Experiments

3.1 Datasets and Experimental Setup

To the best of our knowledge, no publicly available paired NCCT–CTA dataset exists. We collected a carotid artery dataset from The First Affiliated Hospital of Ningbo University, comprising 492 patients with paired NCCT and CTA scans (200 slices per modality). Among them, 10 patients were used for testing (2,000 pairs), 10 patients for validation (2,000 pairs), and the remaining patients for training (94,400 pairs). All images were 256 × 256 with a slice thickness of 0.5 mm, registered using SyNQuick algorithm. All procedures were conducted in accordance with the ethical standards approved by the Institutional Review Board (IRB).

Implementation Details: The NASD-Diff is implemented in PyTorch and trained on four NVIDIA RTX 4090 GPUs for 300K iterations with a batch size of 8, using the Adam optimizer (learning rate 2×10^{-5}). The noise control coefficient λ is set to 3 through experiments, and the joint loss uses weights $\lambda_{\text{CTA}} = 1.0$ and $\lambda_{\text{lpips}} = 0.1$[1]. The diffusion process uses 100 forward and reverse sampling steps, following the parameter settings and sampling strategy in [10].

Evaluation measures: We employed five objective image quality evaluation metrics: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Root Mean Square Error (RMSE) , Visual Information Fidelity (VIF)[19] , and Gradient Magnitude Similarity Deviation (GMSD)[21]. Previous studies have shown that these metrics exhibit good consistency with the subjective ratings of radiologists in medical image assessment[15].

3.2 Results

We adopt the same experimental settings to compare NASD-Diff with several image generation models, including Pix2Pix[9], CycleGAN[22], CTA-GAN[14], UNIT[13], MUNIT[8], DDPM[7], SynDiff[16], Fast-DDPM[10], and CoreDiff[6]. Table 1 presents the quantitative comparison of all methods on the test set.

Experimental results show that NASD-Diff achieves the best performance across all metrics, with a PSNR of 28.92, SSIM of 0.9372, and RMSE of 9.75, indicating high pixel-wise fidelity and accurate vascular generation. Its VIF (0.5224) and GMSD (0.0588) further confirm superior recovery of high-frequency details. As shown in Fig. 3, NASD-Diff generates clearer vascular contours and more complete structures, especially in the red-circled regions where key vessels are accurately reconstructed. In contrast, other models exhibit more pronounced morphological deviations, as evidenced by the error maps, which may lead to inaccurate representation of vessel positions or shapes. Although NASD-Diff shows minor differences in certain areas, the overall error level remains within clinically acceptable limits, and vessel morphology is largely preserved.

Table 1. Quantitative evaluation results of different models. The table reports the mean values of each metric, with arrows indicating the direction of improvement.

Methods	PSNR↑	SSIM↑	RMSE↓	VIF↑	GMSD↓
Pix2Pix	25.67	0.9000	13.92	0.3974	0.0863
CycleGAN	27.65	0.9355	11.32	0.4973	0.0681
CTA-GAN	28.14	0.8761	11.04	0.4908	0.0659
UNIT	28.05	0.9323	10.67	0.4986	0.0621
MUNIT	23.67	0.8943	17.10	0.3509	0.1110
DDPM	28.14	0.8715	10.70	0.4974	0.0629
SynDiff	28.22	0.9063	10.67	0.4928	0.0667
Fast-DDPM	28.24	0.8948	10.60	0.4981	0.0621
CoreDiff	28.26	0.9345	10.23	0.5086	0.0640
NASD-Diff (Ours)	**28.92**	**0.9372**	**9.75**	**0.5224**	**0.0588**

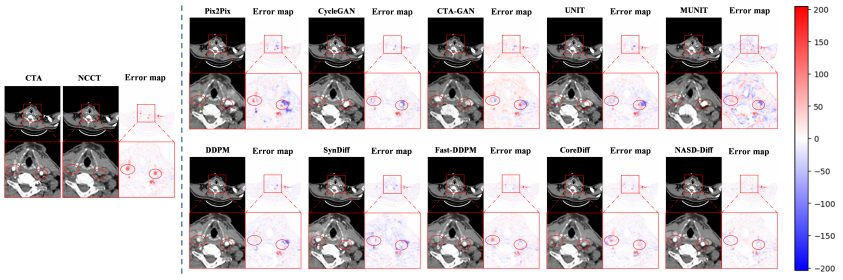

Fig. 3. Visual comparison of different methods for the NCCT-to-CTA task. Each generated result is accompanied by an error map (Errormap = RealCTA − GeneratedCTA), with the color bar showing pixel differences: red for positive, blue for negative residuals.

3.3 Ablation study

In this subsection, we conduct model component ablation and noise control coefficient ablation experiments on the validation set.

Table 2. Ablation studies of NASD-Diff.

Ablation a: Contribution of each module							
Modules			PSNR↑	SSIM↑	RMSE↓	VIF↑	GMSD↓
NASD	CSCEM	GPRM					
-	-	-	29.05	0.8628	9.93	0.5163	0.0575
✓	-	-	29.78	0.9317	8.92	0.5282	0.0557
✓	✓	-	30.11	0.9341	8.47	0.5335	0.0529
✓	-	✓	30.08	0.9357	8.62	0.5336	0.0537
✓	✓	✓	**30.33**	**0.9379**	**8.26**	**0.5374**	**0.0518**
Ablation b: Impact of noise control coefficient λ							
λ			PSNR↑	SSIM↑	RMSE↓	VIF↑	GMSD↓
$\lambda = 0$			27.80	0.9223	11.57	0.5126	0.0761
$\lambda = 1$			28.97	0.9288	9.79	0.5036	0.0599
$\lambda = 2$			29.34	0.9292	9.36	0.5112	0.0577
$\lambda = 3$			**29.78**	**0.9317**	**8.92**	**0.5282**	**0.0557**
$\lambda = 4$			29.66	0.9229	9.06	0.5267	0.0567
$\lambda = 5$			29.21	0.8804	9.45	0.5031	0.0585

Model component ablation: The ablation study (Table 2, Ablation a) shows that gradually adding NASD, CSCEM, and GPRM consistently improves performance. NASD alone significantly enhances the quality of the generated images (PSNR 29.78, SSIM 0.9317, RMSE 8.92). On this basis, adding CSCEM alone further improves performance, and incorporating GPRM independently also leads to additional gains. When all three modules are combined, all metrics reach their optimal values (PSNR 30.33, SSIM 0.9379, RMSE 8.26, VIF 0.5374, GMSD 0.0518), demonstrating the synergistic effect and comprehensive improvement in vascular structures, global textures, and overall visual quality.

Noise control coefficient ablation: Table 2(Ablation b) shows the effect of different λ values with only NASD introduced. When $\lambda = 0$, relying solely on the structural differences degradation operator(Fig. 2(b)), the generated CTA images capture vascular positions but show blurred details and unclear edges(Fig. 4). Increasing λ to 1–3 progressively improves performance, with consistent improvements observed in PSNR, SSIM, RMSE, VIF, and GMSD, indicating that moderate noise effectively restores high-frequency textures. Further increasing λ to 4–5 degrades performance, as excessive noise begins to disrupt low-frequency structural information. Accordingly, $\lambda = 3$ is chosen as a trade-off between texture enhancement and low-frequency structural preservation.

4 Conclusion

This paper proposes NASD-Diff for CTA generation from NCCT, achieving consistent improvements in vascular imaging, texture preservation, and over-

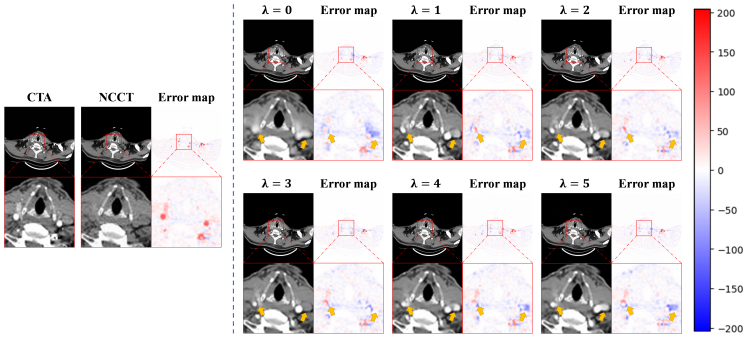

Fig. 4. Visualization of results for different noise control coefficient λ values.

all visual quality. Experimental results demonstrate its superior performance. Despite its strengths, NASD-Diff relies on paired NCCT-CTA data and has limited 3D vascular modeling capability. Future work will investigate stronger 3D constraints and unpaired training strategies to further improve generalization. Overall, NASD-Diff provides a promising solution for high-quality vascular imaging without contrast agents, offering substantial potential for future clinical translation and continued progress in non-contrast CTA research.

Acknowledgments. This work was supported by the Ningbo International Sci-tech Cooperation Projects (No. 2024H010), and the Guangxi Key Technologies R&D Program (No. FN2504240023).

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

References

1. An, J., Yang, Z., Wang, J., Li, L., Liu, Z., Wang, L., Luo, J.: Bring metric functions into diffusion models. pp. 578–586. International Joint Conferences on Artificial Intelligence Organization (8 2024). https://doi.org/10.24963/ijcai.2024/64, main Track
2. Bae, K.T.: Intravenous contrast medium administration and scan timing at ct: considerations and approaches. Radiology **256**(1), 32–61 (2010). https://doi.org/10.1148/radiol.10090908
3. Bansal, A., Borgnia, E., Chu, H.M., Li, J., Kazemi, H., Huang, F., Goldblum, M., Geiping, J., Goldstein, T.: Cold diffusion: Inverting arbitrary image transforms without noise. In: Advances in Neural Information Processing Systems. vol. 36, pp. 41259–41282. Curran Associates, Inc. (2023)
4. Beckett, K.R., Moriarity, A.K., Langer, J.M.: Safe use of contrast media: What the radiologist needs to know. Radiographics **35**(6), 1738–50 (2015). https://doi.org/10.1148/rg.2015150033
5. Chen, T.: On the importance of noise scheduling for diffusion models (2023), https://arxiv.org/abs/2301.10972

6. Gao, Q., Li, Z., Zhang, J., Zhang, Y., Shan, H.: Corediff: Contextual error-modulated generalized diffusion model for low-dose ct denoising and generalization. IEEE Transactions on Medical Imaging **43**(2), 745–759 (2024). https://doi.org/10.1109/TMI.2023.3320812
7. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. In: Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., Lin, H. (eds.) Advances in Neural Information Processing Systems. vol. 33, pp. 6840–6851. Curran Associates, Inc. (2020)
8. Huang, X., Liu, M.Y., Belongie, S., Kautz, J.: Multimodal unsupervised image-to-image translation. In: Computer Vision – ECCV 2018: 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part III. p. 179–196. Springer-Verlag, Berlin, Heidelberg (2018). https://doi.org/10.1007/978-3-030-01219-9_11
9. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5967–5976 (2017). https://doi.org/10.1109/CVPR.2017.632
10. Jiang, H., Imran, M., Zhang, T., Zhou, Y., Liang, M., Gong, K., Shao, W.: Fast-ddpm: Fast denoising diffusion probabilistic models for medical image-to-image generation. IEEE Journal of Biomedical and Health Informatics **29**(10), 7326–7335 (2025). https://doi.org/10.1109/JBHI.2025.3565183
11. Kumamaru, K.K., Hoppel, B.E., Mather, R.T., Rybicki, F.J.: Ct angiography: current technology and clinical use. Radiol Clin North Am **48**(2), 213–35, vii (2010). https://doi.org/10.1016/j.rcl.2010.02.006
12. Li, Y., Shao, H.C., Liang, X., Chen, L., Li, R., Jiang, S., Wang, J., Zhang, Y.: Zero-shot medical image translation via frequency-guided diffusion models. IEEE Transactions on Medical Imaging **43**(3), 980–993 (2024). https://doi.org/10.1109/TMI.2023.3325703
13. Liu, M.Y., Breuel, T., Kautz, J.: Unsupervised image-to-image translation networks. In: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (eds.) Advances in Neural Information Processing Systems. vol. 30. Curran Associates, Inc. (2017)
14. Lyu, J., Fu, Y., Yang, M., Xiong, Y., Duan, Q., Duan, C., Wang, X., Xing, X., Zhang, D., Lin, J., Luo, C., Ma, X., Bian, X., Hu, J., Li, C., Huang, J., Zhang, W., Zhang, Y., Su, S., Lou, X.: Generative adversarial network-based noncontrast ct angiography for aorta and carotid arteries. Radiology **309**(2), e230681 (2023). https://doi.org/10.1148/radiol.230681
15. Mason, A., Rioux, J., Clarke, S.E., Costa, A., Schmidt, M., Keough, V., Huynh, T., Beyea, S.: Comparison of objective image quality metrics to expert radiologists' scoring of diagnostic quality of mr images. IEEE Transactions on Medical Imaging **39**(4), 1064–1072 (2020). https://doi.org/10.1109/TMI.2019.2930338
16. Özbey, M., Dalmaz, O., Dar, S.U.H., Bedel, H.A., Özturk, Ş., Güngör, A., Çukur, T.: Unsupervised medical image translation with adversarial diffusion models. IEEE Transactions on Medical Imaging **42**(12), 3524–3539 (2023). https://doi.org/10.1109/TMI.2023.3290149
17. Pasternak, J.J., Williamson, E.E.: Clinical pharmacology, uses, and adverse reactions of iodinated contrast agents: A primer for the non-radiologist. Mayo Clinic Proceedings **87**(4), 390–402 (2012). https://doi.org/https://doi.org/10.1016/j.mayocp.2012.01.012
18. Shan, H., Zhang, Y., Yang, Q., Kruger, U., Kalra, M.K., Sun, L., Cong, W., Wang, G.: 3-d convolutional encoder-decoder network for low-dose ct via transfer learning

from a 2-d trained network. IEEE Transactions on Medical Imaging **37**(6), 1522–1534 (2018). https://doi.org/10.1109/TMI.2018.2832217

19. Sheikh, H., Bovik, A.: Image information and visual quality. IEEE Transactions on Image Processing **15**(2), 430–444 (2006). https://doi.org/10.1109/TIP.2005.859378
20. Tian, Y., Han, J., Wang, C., Liang, Y., Xu, C., Chen, H.: Dic: Rethinking conv3x3 designs in diffusion models. In: 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2469–2478 (2025). https://doi.org/10.1109/CVPR52734.2025.00236
21. Xue, W., Zhang, L., Mou, X., Bovik, A.C.: Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. IEEE Transactions on Image Processing **23**(2), 684–695 (2014). https://doi.org/10.1109/TIP.2013.2293423
22. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV). pp. 2242–2251 (2017). https://doi.org/10.1109/ICCV.2017.244