A Dual-Domain Diffusion Model for Sparse-View CT Reconstruction

Chun Yang, Dian Sheng, Bo Yang, Wenfeng Zheng, Chao Liu

A Dual-domain Diffusion Model for Sparse-view CT Reconstruction

Chun Yang*, Dian Sheng*, Bo Yang, *Member, IEEE*, Wenfeng Zheng, and Chao Liu, *Senior Member, IEEE*

***Abstract*—A new deep-learning approach, dual-domain diffusion model (DDDM), is proposed for sparse-view CT reconstruction, which is composed of a sinogram upgrading module (SUM) and an image refining module (IRM) connected in series. In the sinogram domain, a novel degrading and upgrading framework is defined, in which SUM is trained to upgrade sparse-view sinograms step by step to reverse the degradation process of CT images caused by successive down-sampling of scanning views. In the image domain, IRM adopts an improved denoising diffusion framework to further reduce remaining artifacts and restore image details, where a skip connection from the original sparse-view sinogram is introduced to constrain the generation of details. Our DDDM shows significant improvement over deep-learning baseline models in both classical similarity metrics and perceptual loss, and has good generalization to untrained organs. We release our code at: https://github.com/YC-Markus/code-for-DDDM.**

***Index Terms*—Computed tomography (CT), sparse-view CT, dual domain, diffusion models, low dose**

I. Introduction

THE widespread use of CT scans has raised concerns about the potential risk of excessive X-ray radiation to human health. Low-dose CT can be achieved by lowering the tube current or reducing the scanning views. This work focuses on the CT reconstruction using sparse scanning views to achieve high-quality CT images for clinical diagnosis.

Deep learning technology has received much attention for its great potential in sparse-view CT reconstruction. Early deep-learning methods [1]-[4], inspired from conventional computer vision tasks, were usually designed in the image domain to post-process low-dose reconstructed CT images to remove artifacts. However, it is difficult for the image-domain models to exploit the potential correlations within the raw projection data (especially between neighboring scanning views). Recent works [5]-[16] have focused on the dual domains, i.e., both the sinogram and image domains. DuDoTrans [5] introduced a sinogram restoration transformer module that emphasizes the global nature of sinogram sampling. To incorporate multi-domain features for sparse-view CT reconstruction, MIST-net [6] employed three sub-networks of different backbones for initial recovery, data consistency correction, and high-fidelity reconstruction, respectively. Although the introduction of *pre-processing* in the sinogram domain can effectively suppress radial artifacts, the dual-domain methods still face the challenges from blurring of the internal structure and additional false information. In addition, iterative reconstruction algorithms [17]-[22] incorporating neural networks have recently been proposed. However, existing iterative pipelines typically have large memory and computational costs and long reconstruction times due to the need to compute and save the gradients at all iteration steps.

As a state-of-the-art generative technique, diffusion models have broken previous records in many computer vision tasks [23]-[26]. However, the diffusion models that perform amazingly well on traditional vision tasks are difficult to be directly transplanted to medical imaging tasks. The denoising generation framework typically adopted by existing diffusion models often fabricates structural details that seem to fit anatomical features but do not actually exist due to the lack of necessary constraints in the inversion process. In addition, the slow sampling speed and excessive number of timesteps also make existing diffusion models difficult to meet real-time clinical tasks.

To address these challenges, we propose a new deep-learning pipeline for sparse-view CT, called dual-domain diffusion model (DDDM). As shown in Fig. 1, our DDDM has a very simple architecture, and is comprised of two separately trainable modules, the sinogram upgrading module (SUM) and the image refining module (IRM), both of which have explicit modeling mechanisms, work in the sinogram and image domains, respectively, and are connected in series by the untrainable filtered back-projection (FBP) layers.

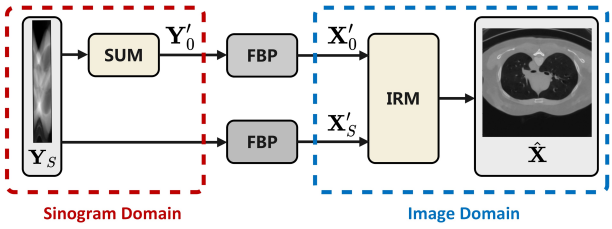

Fig. 1. Architecture of the proposed dual-domain diffusion model (DDDM).

To the best of the authors' knowledge, this is the *first* sparse-view CT pipeline applying diffusion techniques in the dual domains. Although diffusion models have been used for under-sampled MRI reconstruction [27]-[28], the related works mainly focused on the image domain and still fol-

This work was supported by the Science and Technology Planning Project of Sichuan Province (2021YFQ0003, 2023YFH0004, 2023YFSY0026). (Corresponding author: Bo Yang. *Co-first authors.)

Chun Yang and Dian Sheng are with the School of Automation Engineering, and also with the Glasgow College, University of Electronic Science and Technology of China, Chengdu, 610054 China (e-mail: 2020190504033@std.uestc.edu.cn; 2020190504008@std.uestc.edu.cn).

Bo Yang and Wenfeng Zheng are with the School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu, 610054 China (e-mail: boyang@uestc.edu.cn; winfirms@uestc.edu.cn).

Chao Liu is with LIRMM, University of Montpellier-CNRS, Montpellier, 34095 France (e-mail: liu@lirmm.fr).

lowed the denoising-based diffusion framework used for the conventional vision tasks. In this work, the idea of diffusion is introduced not only to the image domain but also to the sinogram domain. In the sinogram domain, a novel down-sampling-based diffusion framework suitable for processing sparse-view sinograms is designed, in which for the first time, the progressively degraded CT image chain formed by successive down-sampling of the views of a sinogram is modeled as a forward diffusion process, and SUM is trained to upgrade the sparse-view sinogram to reverse the image-domain degradation process. In the image domain, IRM is trained to refine the reconstructed images, which employs an improved denoising diffusion framework with a skip connection from the FBP of the sparse-view sinogram to constrain the *generation* of image details. The proposed method is finally validated on a benchmark database widely-used for low-dose CT reconstruction and compared with three deep-learning baseline models.

II. Methods

A. CT imaging modeling

Assume that $\mathbf{X}$ is an underlying CT image, and $\mathbf{Y}$ is the $M \times N$ sinogram acquired by CT scan, where M is the number of X-ray detector and N is the number of rotation angles (views). The CT projection can be expressed as a linear mapping from the image domain to the sinogram domain,

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \epsilon, \tag{1}$$

where $\mathbf{x}$ and $\mathbf{y}$ denote the column vectors reshaped from $\mathbf{X}$ and $\mathbf{Y}$ respectively, ϵ is the system noise during the CT scanning, and $\mathbf{A}$ is the projection matrix determined by the geometry of the CT device and the scanning protocol used.

The CT reconstruction is the inverse problem of the projection process defined in Eq. (1), i.e., recovering $\mathbf{x}$ from $\mathbf{y}$. The classical FBP method can be represented as a linear mapping from the sinogram domain to the image domain, $\hat{\mathbf{x}} = \mathbf{A}^*\mathbf{y}$. Ideally, $\mathbf{A}$ is column full rank for a full-view scan, and $\mathbf{A}^*$ can be viewed as the pseudo-inverse of $\mathbf{A}$, i.e., $\hat{\mathbf{x}}$ is the least-squares solution of the over-determined problem defined in Eq. (1). However, Eq. (1) degenerates to under-determined for a sparse-view scan, and the linear mapping defined by FBP will inevitably introduce artifacts in the reconstructed image.

B. DDDM architecture

The two modules in our DDDM are complementary. SUM can fully exploit the global features embedded in the sinogram-domain projection data to upgrade the reconstructed images in big steps; while IRM in the image domain can compensate for the deficiency of the sinogram-domain process in recovering local details.

1) Sinogram upgrading module: Unlike ordinary diffusion models that use Gaussian noise as the degradation operator [23], our SUM uses the view down-sampling in the sinogram domain as the degradation operator, which better aligns the sparse-view CT task. As shown in Fig. 2, let $\mathbf{Y}_0$ be a full-view sinogram, and $\{\mathbf{Y}_s\}_{s=1}^{S}$ denote its sparse-view versions obtained by successively down-sampling the views of $\mathbf{Y}_0$ by the factor 2 (denoted as $2\downarrow$). Reconstructing these sinograms with FBP forms a chain of CT images $\{\mathbf{X}_s\}_{s=0}^{S}$ of decreasing quality in the image domain. In order to reverse this degraded image chain, an upgrading operator $\mathrm{U}_{\mathrm{sino}}$ is trained to improve $\mathbf{Y'}_S$, the nearest neighbor interpolation of the last-level sinogram $\mathbf{Y}_S$, step by step. To ensure that the upgrading operator is iterable, all the sinograms involved in the reversal process have the same size as $\mathbf{Y}_0$, i.e., $\mathbf{Y'}_s \in \mathbb{R}^{M\times N}$ for all $s = 1, 2, ...S$.

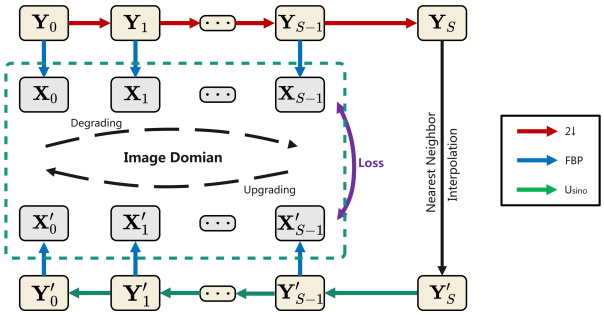

Fig. 2. Degrading and upgrading framework in sinogram domain

We adopt an Attention U-Net [29] as the backbone of the upgrading operator. Figure 3 shows the workflow of the upgrading operator at a timestep s, upgrading $\mathbf{Y'}_s$ to $\mathbf{Y'}_{s-1}$,

$$\mathbf{Y'}_{s-1} = \mathrm{U}_{\mathrm{sino}}(\mathbf{Y'}_s, \mathbf{Y}_S, s) \tag{2}$$

At each timestep, the input $\mathbf{Y'}_s$ is first concatenated channel-wise with the initial input $\mathbf{Y'}_S$, and then fed into the Attention U-Net to predict a residual component, into which the timestep embedding $\mathrm{emb}(s)$ is also encoded to control the degree of upgradation. The residual component is summed with $\mathbf{Y'}_s$ to produce the upgraded sinogram $\mathbf{Y'}_{s-1}$, whose view components corresponding to the original sparse views are finally replaced by $\mathbf{Y}_S$ to ensure that the original projection data are not contaminated during the sinogram upgrading.

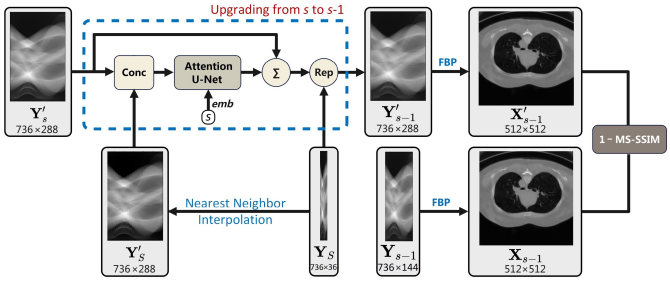

Fig. 3. Workflow of sinogram upgrading module (at timestep $s = 2$ when $S = 3$).

Although $\mathrm{U}_{\mathrm{sino}}$ performs iterative inference in the sinogram domain, the loss function used to train it is defined in the image domain because the ultimate goal is to improve the quality of the reconstructed CT image rather than to complete the sinogram. Thanks to the linear mapping of FBP, the gradient of the loss function can be easily reversed across FBP using the back-propagation algorithm. The multi-scale structural similarity (MS-SSIM) is used to calculate the reconstruction loss, which leads to more accurate textures and avoids the over-

smoothing caused by traditional MSE loss. At the timestep s, the loss is written as

$$l_s = 1 - \text{MS-SSIM}(\mathbf{X'}_{s-1}, \mathbf{X}_{s-1}) \tag{3}$$

where $\mathbf{X'}_{s-1}$ is the FBP reconstruction of the upgraded sinogram $\mathbf{Y'}_{s-1}$, and $\mathbf{X}_{s-1}$ is the label generated in the degradation process (see Fig. 2).

Our SUM can be trained progressively and can infer iteratively in descending order of timesteps, which facilitates it to learn to recover image features at different sparsity levels, alleviating the over-blurring issue often encountered by end-to-end sinogram modules. Compared to the existing iterative pipelines [20], [21], our pipeline is built on an explicit degradation process induced by successive down-sampling of the sinogram, which allows the backbone to be trained progressively and requires very few iterations (S iterations imply an 2^S-fold interpolation of the sinogram) without gradient computation during the inference phase.

Given an original sparse-view sinogram, it can be upgraded to a reasonable rank through few iterations of SUM. Setting more iterations will result in a dramatic increase in computation, as the number of views of all latent sinogram $\{\mathbf{Y'}_s\}$ is 2^S times that of the original sinograms $\mathbf{Y}_S$. Moreover, excessive iterations do not lead to a continuous improvement of the reconstructed image, and the upgradation in the sinogram domain will soon reach saturation, as the original projection data are limited. The following IRM is designed to further refine the reconstructed CT images in the image domain, which employs a noising-based diffusion framework differs from the down-sampling-based framework employed in the sinogram domain.

2) Image refining module: Figure 4 shows the forward workflow of the image-domain refining operator $\mathrm{U}_{\text{image}}$ at a timestep $t = 1, 2, ..., T$, which still uses the Attention U-Net as the backbone, and besides $\mathbf{X'}_0$ (the output of SUM) adds a skip-connected auxiliary channel $\mathbf{X}_S$ from the FBP of the original sinogram to guide the generation of details.

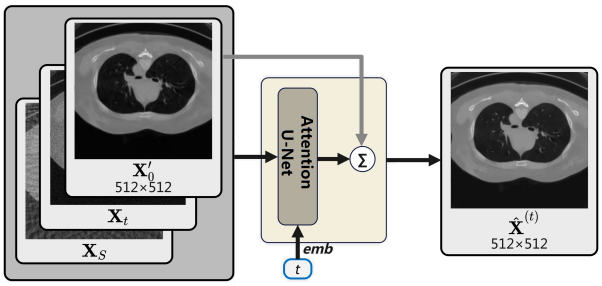

Fig. 4. Forward workflow of image refining module at timestep t.

In the training phase, the noisy sample $\mathbf{X}_t$ is generated by adding Gaussian noise to the initial sample $\mathbf{X}$ (i.e. ground truth) according to a pre-designed single-step noise-adding schedule, as in [24]. At each timestep, the ground truth $\mathbf{X}$, instead of the added noise usually predicted in image generation tasks, is directly predicted after the residual connection with $\mathbf{X'}_0$,

$$\hat{\mathbf{X}}^{(t)} = \mathrm{U}_{\text{image}}(\mathbf{X}_t, \mathbf{X'}_0, \mathbf{X}_S, t), \tag{4}$$

which facilitates subsequent acceleration using a denoising diffusion implicit model (DDIM) [25] in the inference phase. Besides the MS-SSIM, L1 loss is also introduced with a weight λ in the training of IRM for avoiding over smoothing

$$l_t = (1 - \text{MS-SSIM}(\hat{\mathbf{X}}^{(t)}, \mathbf{X})) + \lambda L_1(\hat{\mathbf{X}}^{(t)}, \mathbf{X})). \tag{5}$$

With DDIM, the iterative inference process can be accelerated by a factor of K that is divisible by T, and the noisy sample for the next iteration is calculated by

$$\mathbf{X}_{t-K} = \sqrt{1-\bar{\alpha}_{t-K}}\hat{\mathbf{X}}^{(t)} + \sqrt{1-\bar{\alpha}_{t-K}}\boldsymbol{\epsilon}_t, \tag{6}$$

where $\boldsymbol{\epsilon}_t = (\mathbf{X}_t - \sqrt{\bar{\alpha}_t}\hat{\mathbf{X}}^{(t)})/\sqrt{1-\bar{\alpha}_t}$, and $\bar{\alpha}_t$ is the coefficient from the noise schedule.

III. Experiments and Results

A. Data and settings

The proposed method is validated on NBIA database [30], in which 6500 chest CT slices from 20 patients and 7400 abdomen CT slices from 18 patients serve as the training data, 600 chest slices from another 2 patients serve as validation data, and 800 abdomen slices from 2 patients serve as testing data. In addition, the performance of the proposed method on a untrained organ is also tested with 1000 head slices from 5 patients, which is obviously more challenging than the abdominal testing as the training set does not contain any head slice.

For simplifying computation, the raw data of helix scanning are transformed to parallel-beam with Helix2Fan [31] and MATLAB script, yielding the full-view sinograms of size 736×1152 with angles covering $[0, 2\pi)$. Note that the proposed pipeline can also be used for the fan-beam scanning by simply modifying the FBP layers accordingly. The original sparse-view sinograms, down-sampled from the 1152-view sinograms, preserves 36 views covering $[0, \pi)$. SUM is first trained to upgrade the 36-view sinograms to 288 views through $S = 3$ iterations. IRM is then trained to refine the FBP reconstructed images of the upgraded 288-view sinograms by using the FBP of the 1152-view sinograms as the ground truth. More sophisticated training strategies, such as separate training followed by joint fine-tuning, may also be employed with adequate hardware resources. All the CT images reconstructed in our work have a size of 512×512. Our DDDM is trained on Pytorch 1.11 @ NVIDIA A100 40G with AdamW [33] optimizer, 1e-4 learning rate, and 0.3 dropout rate. FBPConvNet [4], DuDoTrans and MIST-net are trained on the same database for comparison. Three similarity metrics, PSNR, SSIM, and the learned perceptual image patch similarity (LPIPS) [34], are used to quantitatively evaluate the reconstructed images.

B. Results

Figures 5 and 6 show two sets of reconstructed CT images, from the abdominal and untrained head testing data, respectively. It is clear that all four deep-learning methods can significantly improve the reconstruction of 36-view sinograms, compared to the classical FBP method. However, on the

abdomen data, DuDoTrans and MIST-net result in noticeable clumpy artifacts and edge blurring (see red circles in Fig. 5), and are inferior to FBPConvNet and DDDM in terms of overall visual perception. Our DDDM reconstructs the edge details more accurately than FBPConvNet. As marked in yellow in Fig. 5, edge sticking, dissolution of weak features, and even spurious black spots appear in the reconstructed image of FBPConvNet.

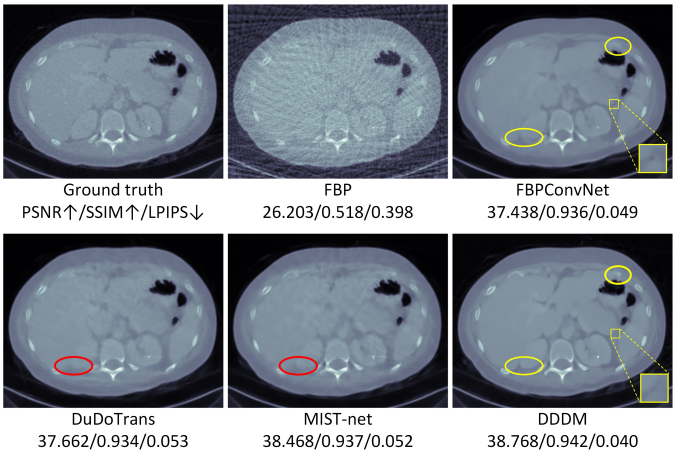

Fig. 5. Reconstructed CT images for an abdomen test sample.

On the head data, the advantages of DDDM over other methods are more obvious (see the enlarged yellow boxes in Fig. 6), which shows that the proposed model generalizes well and performs well even on untrained organs.

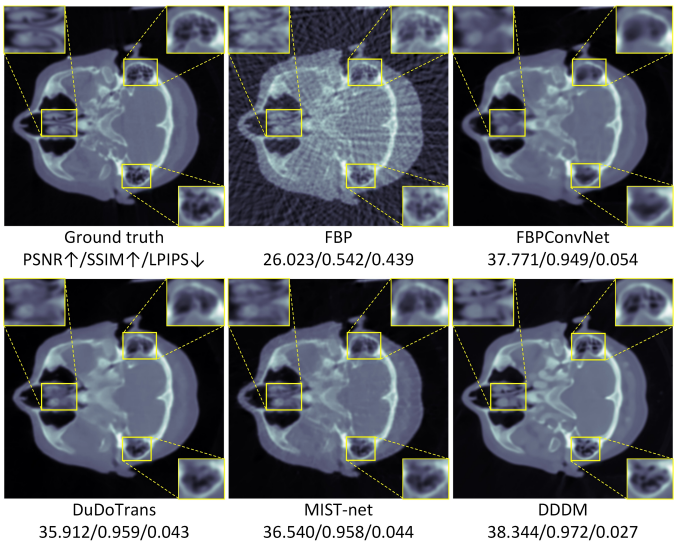

Fig. 6. Reconstructed CT images for a head test sample.

Tables I and II give the average performance metrics of different methods in the abdomen and head testing sets, respectively. Consistent with the visual inspection, our method performs significantly better than the baseline methods on both testing sets, especially in the LPIPS metric, which indicates that the reconstructed images by our DDDM are closer to the ground truth in terms of human perception. In addition, the performance of DDDM on the untrained head data is surprising, showing that the proposed dual-domain framework has very good generalization across organs and can be widely applied to low-dose reconstruction of different organs.

TABLE I
QUANTITATIVE RESULTS OBTAINED BY DIFFERENT METHODS IN THE ABDOMEN TESTING SET

Methods	PSNR↑	SSIM↑	LPIPS↓
FBP	26.529	0.517	0.376
FBPConvNet	36.873	0,913	0.0596
DuDoTrans	36.901	0.912	0.0637
MIST-net	37.103	0.911	0.0633
DDDM	**38.132**	**0.921**	**0.0486**

TABLE II
QUANTITATIVE RESULTS OBTAINED BY DIFFERENT METHODS IN THE HEAD TESTING SET

Methods	PSNR↑	SSIM↑	LPIPS↓
FBP	28.182	0.639	0.429
FBPConvNet	39.372	0.973	0.0640
DuDoTrans	39.322	0.975	0.0585
MIST-net	38.470	0.970	0.0730
DDDM	**42.066**	**0.984**	**0.0420**

IV. CONCLUSION

In this letter, a new sparse-view CT reconstruction pipeline, DDDM, is proposed, which introduces the idea of diffusion in the dual domains to improve the reconstruction process step by step. In the sinogram domain, the sparse-view reconstruction is defined for the first time as an inverse process of image degradation induced by successive down-sampling of the sinogram, for which a diffusion-like model is designed and trained to upgrade the sparse-view sinogram step-by-step. In the image domain, an improved denoising diffusion implicit model is used to further refine the reconstructed images, which introduces auxiliary channels from the output of the sinogram-domain module and the original sparse-view sinogram at each timestep to guide the *generation* of image details. Experiments based on publicly available database show that the proposed method has obvious advantages over existing deep-learning baseline methods, especially in the recovery of image details. It is also surprising that the proposed model has good generalization and its performance on an untrained organ is excellent. In this work, the effectiveness of applying diffusion models in dual domains to enhance sparse-view CT reconstruction is demonstrated with a simple parallel-beam scanning setting, a classical U-Net backbone, and a serial dual-domain architecture. In the future, more advanced parallel dual-domain architectures, better backbone networks other than U-Net, and closer-to-clinical scanning settings can be tried to improve the dual-domain diffusion models.

REFERENCES

[1] W. Du, H. Chen, P. Liao, H. Yang, G. Wang and Y. Zhang, "Visual Attention Network for Low-Dose CT," in *IEEE Signal Processing Letters*, vol. 26, no. 8, pp. 1152-1156, Aug. 2019, doi: 10.1109/LSP.2019.2922851.

[2] P. Bao et al., "Convolutional Sparse Coding for Compressed Sensing CT Reconstruction," in *IEEE Transactions on Medical Imaging*, vol. 38, no. 11, pp. 2607-2619, Nov. 2019, doi: 10.1109/TMI.2019.2906853.

[3] Y. Han and J. C. Ye, "Framing U-Net via Deep Convolutional Framelets: Application to Sparse-View CT," in *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1418-1429, June 2018, doi: 10.1109/TMI.2018.2823768.

[4] K. H. Jin, M. T. McCann, E. Froustey and M. Unser, "Deep Convolutional Neural Network for Inverse Problems in Imaging," in *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509-4522, Sept. 2017, doi: 10.1109/TIP.2017.2713099.

[5] C. Wang, K. Shang, H. Zhang, Q. Li, and S. K. Zhou, "DuDoTrans: Dual-domain transformer for sparse-view CT reconstruction," *Machine Learning for Medical Image Reconstruction*, pp. 84–94, 2022. doi:10.1007/978-3-031-17247-2-9.

[6] J. Pan, H. Zhang, W. Wu, Z. Gao, and W. Wu, "Multi-domain integrative Swin Transformer Network for sparse-view tomographic reconstruction," *Patterns*, vol. 3, no. 6, p. 100498, 2022. doi:10.1016/j.patter.2022.100498.

[7] X. Gao et al., "Attention-based dual-branch deep network for sparse-view computed tomography image reconstruction," *Quantitative Imaging in Medicine and Surgery*, vol. 13, no. 3, pp. 1360–1374, 2023. doi:10.21037/qims-22–609.

[8] H. Lee, J. Lee, H. Kim, B. Cho and S. Cho, "Deep-Neural-Network-Based Sinogram Synthesis for Sparse-View CT Image Reconstruction," in *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 109-119, March 2019, doi: 10.1109/TRPMS.2018.2867611.

[9] Y. Li, K. Li, C. Zhang, J. Montoya and G. -H. Chen, "Learning to Reconstruct Computed Tomography Images Directly From Sinogram Data Under A Variety of Data Acquisition Conditions," in *IIEEE Transactions on Medical Imaging*, vol. 38, no. 10, pp. 2469-2481, Oct. 2019, doi: 10.1109/TMI.2019.2910760.

[10] Z. Zhang, X. Liang, X. Dong, Y. Xie and G. Cao, "A Sparse-View CT Reconstruction Method Based on Combination of DenseNet and Deconvolution," in *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1407-1417, June 2018, doi: 10.1109/TMI.2018.2823338.

[11] R. Ge, Y. He, and C. Xia, "DDPNet: A Novel Dual-Domain Parallel Network for Low-Dose CT Reconstruction," in *the 25th International Conference on Medical Image Computing and Computer Assisted Intervention*.

[12] X. Ye, Z. Sun, R. Xu, Z. Wang, and H. Li, "Low-Dose CT Reconstruction via Dual-Domain Learning and Controllable Modulation," in *the 25th International Conference on Medical Image Computing and Computer Assisted Intervention*.

[13] Z. Zhang, L. Yu, X. Liang, W. Zhao, and L. Xing, "TransCT: Dual-path transformer for low dose computed tomography," in *MICCAI* 2021, pp. 55–64, 2021. doi:10.1007/978-3-030-87231-1-6.

[14] C. Sun, K. Deng, Y. Liu, and H. Yang, "A lightweight dual-domain attention framework for sparse-view CT reconstruction," *2022 IEEE 8th International Conference on Computer and Communications (ICCC)*, 2022. doi:10.1109/iccc56324.2022.10065958.

[15] L. Yang, Z. Li, R. Ge, J. Zhao, H. Si and D. Zhang, "Low-Dose CT Denoising via Sinogram Inner-Structure Transformer," in *IEEE Transactions on Medical Imaging*, vol. 42, no. 4, pp. 910-921, April 2023, doi: 10.1109/TMI.2022.3219856.

[16] W. Wang et al., "An End-to-End Deep Network for Reconstructing CT Images Directly From Sparse Sinograms," in *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1548-1560, 2020, doi: 10.1109/TCI.2020.3039385.

[17] M. Patwari, R. Gutjahr, R. Raupach, and A. Maier, "JBFnet - low dose CT denoising by trainable joint bilateral filtering," in *MICCAI* 2020, pp. 506–515, 2020. doi:10.1007/978-3-030-59713-9-49.

[18] D. Wu, K. Kim, G. El Fakhri and Q. Li, "Iterative Low-Dose CT Reconstruction With Priors Trained by Artificial Neural Network," in *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2479-2486, Dec. 2017, doi: 10.1109/TMI.2017.2753138.

[19] W. Xia et al., "A transformer-based iterative reconstruction model for sparse-view CT reconstruction," in *MICCAI* 2022. doi:10.1007/978-3-031-16446-0-75.

[20] Y. Zhang et al., "LEARN++: Recurrent Dual-Domain Reconstruction Network for Compressed Sensing CT," in *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 7, no. 2, pp. 132-142, Feb. 2023, doi: 10.1109/TRPMS.2022.3222213.

[21] W. Wu, D. Hu, C. Niu, H. Yu, V. Vardhanabhuti and G. Wang, "DRONE: Dual-Domain Residual-based Optimization NEtwork for Sparse-View CT Reconstruction," in *IEEE Transactions on Medical Imaging*, vol. 40, no. 11, pp. 3002-3014, Nov. 2021, doi: 10.1109/TMI.2021.3078067.

[22] H. Chen et al., "LEARN: Learned Experts' Assessment-Based Reconstruction Network for Sparse-Data CT," in *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1333-1347, June 2018, doi: 10.1109/TMI.2018.2805692.

[23] J. Ho, P. Abbeel, and A. Jain, "Denoising Diffusion Probabilistic Models," in *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*.

[24] A. Nichol and P. Dhariwal, "Improved Denoising Diffusion Probabilistic Models," in *Proceedings of the 38th International Conference on Machine Learning*.

[25] J. Song, C. Meng, and S. Ermon, "Denoising Diffusion Implicit Models," in *International Conference on Learning Representations (2021)*.

[26] A. Bansal, E. Borgnia, H. Chu, J. S. Li, H. Kazemi, F. Huang, M. Goldblum, J. Geiping, and T. Goldstein, "Cold Diffusion: Inverting Arbitrary Image Transforms Without Noise," 2022. [Online]. Available: arXiv:2208.09392.

[27] C. Peng, P. Guo, S. K. Zhou, V. Patel, and R. Chellappa, "Towards performant and reliable undersampled MR reconstruction via diffusion model sampling," in *the 25th International Conference on Medical Image Computing and Computer Assisted Intervention*.

[28] Y. Xie and Q. Li, "Measurement-conditioned Denoising Diffusion Probabilistic Model for Under-sampled Medical Image Reconstruction," in *the 25th International Conference on Medical Image Computing and Computer Assisted Intervention*.

[29] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention U-Net: Learning Where to Look for the Pancreas," 2018. [Online]. Available: arXiv:1804.03999.

[30] T. R. Moen et al., "Low-dose CT image and projection dataset," *Medical Physics*, vol. 48, no. 2, pp. 902–911, 2020. doi:10.1002/mp.14594.

[31] F. Wagner, M. Thies, L. Pfaff, O. Aust, S. Pechmann, D. Weidner, N. Maul, M. Rohleder, M. Gu, J. Utz, F. Denzinger, and A. Maier, "On the Benefit of Dual-domain Denoising in a Self-supervised Low-dose CT Setting," 2022. [Online]. Available: arXiv:2211.01111.

[32] A. Paszke et al., "Automatic differentiation in PyTorch," in *Proc. Neural Inf. Process. Syst.*, 2017.

[33] I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization," 2019. [Online]. Available: arXiv:1711.05101.

[34] R. Zhang, P. Isola, A. A. Efros, E. Shechtman and O. Wang, "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, pp. 586-595, doi: 10.1109/CVPR.2018.00068.