Dual-domain Sparse View CT Reconstruction using Denoising Diffusion Probabilistic Model

Xin Tie (xtie@wisc.edu, 9081372113)
Minyi Dai (mdai26@wisc.edu, 9079876687)
Hao Zhang (hzhang978@wisc.edu, 9084624668)

Introduction and Motivation

In medical imaging, reconstruction from acquired data typically presumes that the image object is stationary during the acquisition process. However, in dynamic imaging scenarios, this assumption may not hold true. Failing to account for the object motion leads to the emergence of motion artifacts, which can significantly affect radiologists' diagnosis. To deal with the dynamic acquisitions, two potential solutions have been proposed: upgrading the hardware system to shorten the scan time and developing image reconstruction algorithms to generate artifact-free images from an undersampled data set [1].

Computed tomography (CT) is a non-invasive diagnostic imaging tools, known for fast acquisitions, high spatial resolution, and broad availability. Sparse-view CT reconstruction in the dynamic acquisitions is an intriguing yet challenging problem, which involves acquiring only a portion of the full projection data. This low-cost and efficient technique also presents an opportunity to reduce radiation dose while maintaining the desired diagnostic performance. The main challenge in sparse-view reconstruction is angular undersampling that violates Nyquist's criterion, causing aliasing artifacts [2]. These artifacts manifest as streaks in the CT images and can obscure low-contrast objects of interest, such as lesions. As the number of acquired view projection data decreases, the resulting artifacts become more pronounced and severe. Figure 1 depicts the impact of reducing the sampled data to one-eighth of the full data set on the image quality.

Method and Related Works

Sparse-view CT reconstruction is an ill-posed inverse problem and there does not exist an analytical solution. In the past, two paradigms have been proposed to recover high-quality images from severely undersampled data. The first paradigm, compressed sensing (CS) [3-5], reformulates an image reconstruction problem to a convex optimization problem with two terms: the data fidelity term, which enforces consistency with the acquired data, and the regularization term, which transforms the image to a new space and promotes the sparsity in that space [3]. This optimization problem can be solved by a range of gradient-descent based methods, such as conjugate gradient and ADMM. In the end, the reconstructed image balances data fidelity and artifact removal [1]. The second paradigm, deep-learning-based methods [1-2,6-10], enables fast and higher-quality image reconstruction compared to CS-based methods. This class of methods can be categorized into three groups: (1) converting artifact-contaminated images to artifact-free images (2) inpainting undersampled projection data to generate a full data set, then applying filter backprojection (FBP), and (3) directly reconstructing images from undersampled projection data without explicit use of the classic reconstruction algorithms (e.g., FBP). In most cases, deep neural networks trained purely on the image domain performed reasonably well, like RED-CNN [11]. However, they did not fully leverage the information from the projection domain. Lack of consistency check with the acquired data may cause false negative lesions and false positive lesion-like structures [1] in the reconstructed images. Similarly, only using the projection data to train the model loses information of spatial correlations in the image domain. In this work, we propose a dual-domain sparse view CT reconstruction pipeline that leverages both sinograms and images. Specfically, we first use a deep neural network to convert the artifact-contaminated CT image to an artifact-reduced image, followed by a Radon transform to convert the image to the sinogram. In the sinogram domain, we develop another model to correct the inpainted projection data to obtain a more accurate estimate of the full data set. Finally, we reconstruct the image using FBP.

In recent years, diffusion models [12-21] hold a great promise in image generation. It overcomes the problems encountered in its counterpart generative adversarial model (GAN), including mode collapse and unstable training. Additionally, diffusion models have shown potential for solving inverse problems [15] in image reconstruction, by gradually adding Gaussian noise and using a network to learn the reverse process, i.e., denoising. In this study, we aim to utilize the diffusion model to address artifact removal in the image domain and sinogram correction in the projection domain.

Dataset

This study employs a retrospective dual-energy dataset, consisting of 44 clinical pulmonary CT angiography (CTA) exams. The dataset encompasses both sinograms and images obtained at 80 kVp and 140 kVp, with only the latter being included in this work due to its superior image quality and fewer beam hardening artifacts. Among 44 human subjects, 36 (4531 slices) will be randomly selected for model development and the remaining 8 cases (938 slices) will be used for testing.

Evaluation

Two evaluation metrics will be used to quantify the performance of our models. The first one is relative root mean squared error (rRMSE). It reflects the accuracy of HU (or CT numbers) in the reconstructed CT images with respect to the ground truth, i.e., fully sampled images. The second one is structure similarity index metric (SSIM), which measures the similarity between two images in terms of perceived quality, taking into account the structural information present in the images, such as pixel correlations, brightness, and image contrast.

Milestones

03/15	Finish the training of the image-domain diffusion model
03/31	Finish the training of the projection-domain diffusion model
04/04	Submit the midterm report
04/15	Finish the model evaluation and comparison
04/20	Finish the slides for project presentation
05/05	Finish the project website

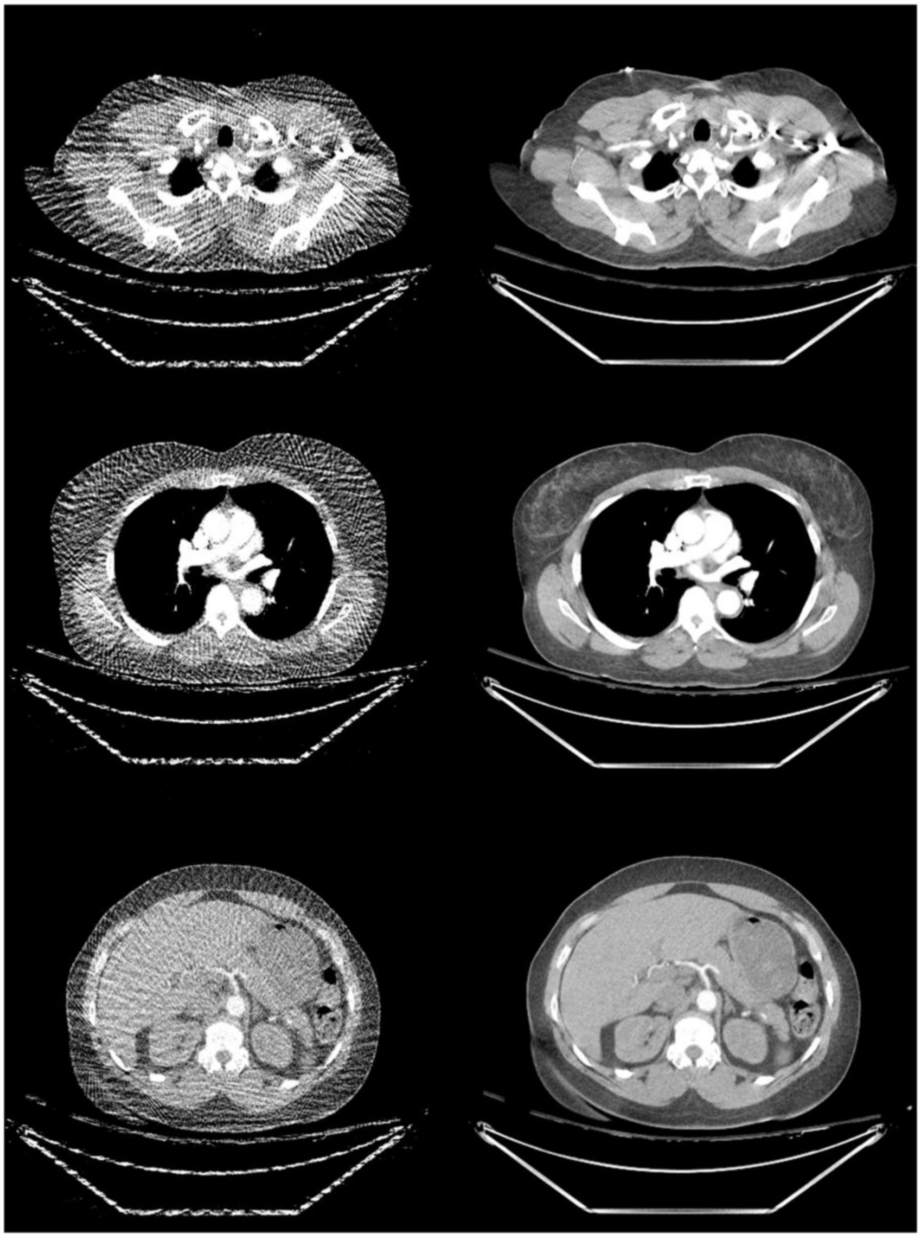

Figure 1: Impact of undersampling on CT image quality. Left: sparse-view CT with 123 views; Right: dense-view CT (or fully sampled CT) with 984 views.

References:

[1] C. Zhang, Y. Li, and G. Chen, “Accurate and robust sparse-view angle CT image reconstruction using deep learning and prior image constrained compressed sensing (DL-PICCS),” *Medical Physics*, vol. 48, no. 10, pp. 5765–5781, Oct. 2021, doi: 10.1002/mp.15183.

[2] W. Xia, W. Cong, and G. Wang, “Patch-Based Denoising Diffusion Probabilistic Model for Sparse-View CT Reconstruction.” arXiv, Nov. 18, 2022. Accessed: Feb. 17, 2023. [Online]. Available: http://arxiv.org/abs/2211.10388

[3] H. Kudo, T. Suzuki, and E. A. Rashed, "Image reconstruction for sparse-view CT and interior CT— introduction to compressed sensing and differentiated backprojection," *Quantitative Imaging in Medicine and Surgery*, vol. 3, no. 3, 2013.
[4] J. Bai, Y. Liu, and H. Yang, "Sparse-View CT Reconstruction Based on a Hybrid Domain Model with Multi-Level Wavelet Transform," *Sensors*, vol. 22, no. 9, p. 3228, Apr. 2022, doi: 10.3390/s22093228.
[5] Z. Qu, X. Yan, J. Pan, and P. Chen, "Sparse View CT Image Reconstruction Based on Total Variation and Wavelet Frame Regularization," *IEEE Access*, vol. 8, pp. 57400–57413, 2020, doi: 10.1109/ACCESS.2020.2982229.
[6] Z. Fu, H. W. Tseng, S. Vedantham, A. Karellas, and A. Bilgin, "A residual dense network assisted sparse view reconstruction for breast computed tomography," *Sci Rep*, vol. 10, no. 1, p. 21111, Dec. 2020, doi: 10.1038/s41598-020-77923-0.
[7] Z. Zhang, X. Liang, X. Dong, Y. Xie, and G. Cao, "A Sparse-View CT Reconstruction Method Based on Combination of DenseNet and Deconvolution," *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1407–1417, Jun. 2018, doi: 10.1109/TMI.2018.2823338.
[8] W. Xia, Z. Yang, Q. Zhou, Z. Lu, Z. Wang, and Y. Zhang, "A Transformer-Based Iterative Reconstruction Model for Sparse-View CT Reconstruction," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*, vol. 13436, L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, and S. Li, Eds. Cham: Springer Nature Switzerland, 2022, pp. 790–800. doi: 10.1007/978-3-031-16446-0_75.
[9] B. Zhou, X. Chen, S. K. Zhou, J. S. Duncan, and C. Liu, "DuDoDR-Net: Dual-domain data consistent recurrent network for simultaneous sparse view and metal artifact reduction in computed tomography," *Medical Image Analysis*, vol. 75, p. 102289, Jan. 2022, doi: 10.1016/j.media.2021.102289.
[10] Y. Han and J. C. Ye, "Framing U-Net via Deep Convolutional Framelets: Application to Sparse-view CT." arXiv, Mar. 28, 2018. Accessed: Feb. 17, 2023. [Online]. Available: http://arxiv.org/abs/1708.08333
[11] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, and P. Liao, "Low-Dose CT with a Residual Encoder-Decoder Convolutional Neural Network (RED-CNN)".
[12] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models".
[13] Y. Song and S. Ermon, "Generative Modeling by Estimating Gradients of the Data Distribution".
[14] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-Based Generative Modeling through Stochastic Differential Equations." arXiv, Feb. 10, 2021. Accessed: Feb. 21, 2023. [Online]. Available: http://arxiv.org/abs/2011.13456
[15] Y. Song, L. Shen, L. Xing, and S. Ermon, "Solving Inverse Problems in Medical Imaging with Score-Based Generative Models." arXiv, Jun. 15, 2022. Accessed: Feb. 21, 2023. [Online]. Available: http://arxiv.org/abs/2111.08005
[16] L. Yang *et al.*, "Diffusion Models: A Comprehensive Survey of Methods and Applications." arXiv, Oct. 23, 2022. Accessed: Feb. 16, 2023. [Online]. Available: http://arxiv.org/abs/2209.00796
[17] F.-A. Croitoru, V. Hondru, R. T. Ionescu, and M. Shah, "Diffusion Models in Vision: A Survey." arXiv, Dec. 20, 2022. Accessed: Feb. 21, 2023. [Online]. Available: http://arxiv.org/abs/2209.04747
[18] C. Luo, "Understanding Diffusion Models: A Unified Perspective." arXiv, Aug. 25, 2022. Accessed: Feb. 17, 2023. [Online]. Available: http://arxiv.org/abs/2208.11970

[19] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, “High-Resolution Image Synthesis with Latent Diffusion Models.” arXiv, Apr. 13, 2022. Accessed: Feb. 21, 2023. [Online]. Available: http://arxiv.org/abs/2112.10752
[20] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi, “Image Super-Resolution Via Iterative Refinement,” *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1–14, 2022, doi: 10.1109/TPAMI.2022.3204461.
[21] A. Lugmayr, M. Danelljan, A. Romero, F. Yu, R. Timofte, and L. Van Gool, “RePaint: Inpainting using Denoising Diffusion Probabilistic Models.” arXiv, Aug. 31, 2022. Accessed: Feb. 21, 2023. [Online]. Available: http://arxiv.org/abs/2201.09865