Dual Domain Closed-loop Learning for Sparse-view CT Reconstruction

Yi Guo[a, b], Yongbo Wang[a, b], Manman Zhu[a, b], Dong Zeng[a, b], Zhaoying Bian[a, b], Xi Tao[a, b], and Jianhua Ma[a, b]

[a]School of Biomedical Engineering, Southern Medical University, Guangdong, China
[b]Guangdong Artificial Intelligence and Digital Economy Laboratory (Guangzhou), Guangdong, China

ABSTRACT

Sparse view sampling is one of the effective ways to reduce radiation dose in CT imaging. However, artifacts and noise in sparse-view filtered back projection reconstructed CT images are obvious that should be removed effectively to maintain diagnostic accuracy. In this paper, we propose a novel sparse-view CT reconstruction framework, which integrates the projection-to-image and image-to-projection mappings to build a dual domain closed-loop learning network. For simplicity, the proposed framework is termed a closed-loop learning reconstruction network (CLRrcon). Specifically, the primal mapping (i.e., projection-to-image mapping) contains a projection domain network, a backward projection module, and an image domain network. The dual mapping (i.e., image-to-projection mapping) contains an image domain network and a forward projection module. All modules are trained simultaneously during the network training stage, and only the first mapping is used in the network inference stage. It should be noted that both the inference time and hardware requirements do not increase compared with traditional hybrid domain networks. Experiments on low-dose CT data demonstrate the proposed CLRecon model can obtain promising reconstruction results in terms of edge preservation, texture recovery, and reconstruction accuracy in the sparse-view CT reconstruction task.

Keywords: Computed tomography, closed-loop learning, image reconstruction, sparse view

1. INTRODUCTION

Computed Tomography (CT) has been widely used in modern medical diagnosis and treatment due to its fast imaging speed, high resolution, etc. However, patients will receive lots of radiation doses during the CT examination, which is becoming a concern. Sparse-view scans can effectively reduce the radiation dose, but with the decrease of the scan views, the image quality will degrade when using the traditional filtered back projection (FBP) algorithm.

Numerous model-based iterative algorithms (MBIR) have been proposed for sparse-view CT reconstruction in the past decade. With correct prior assumptions,[1, 2] the iterative algorithm can obtain high-quality images. However, prior information is often manually selected, which cannot achieve desired results when it is not completely consistent with the actual collected projection data. In addition, the iterative algorithm requires repeated forward and backward projection until the desired image is obtained, which is time-consuming and requires a lot of computing resources.

In recent years, deep learning has achieved great success in the CT reconstruction task. The deep learning based reconstruction methods can be divided into the following three categories. The first category trains a network mapping from low-dose data to normal-dose data in the image domain or projection domain.[3–5] Some scholars attempt to combine the image domain network with the projection domain network to form a hybrid model.[6, 7] The second category expands the iterative reconstruction algorithms into a network.[8–10] The third category builds a projection-to-image reconstruction network.[11–13] Through sufficient training, the network can directly reconstruct the image without artifacts from sparse-view projection data. In addition, Tao et al.[14]
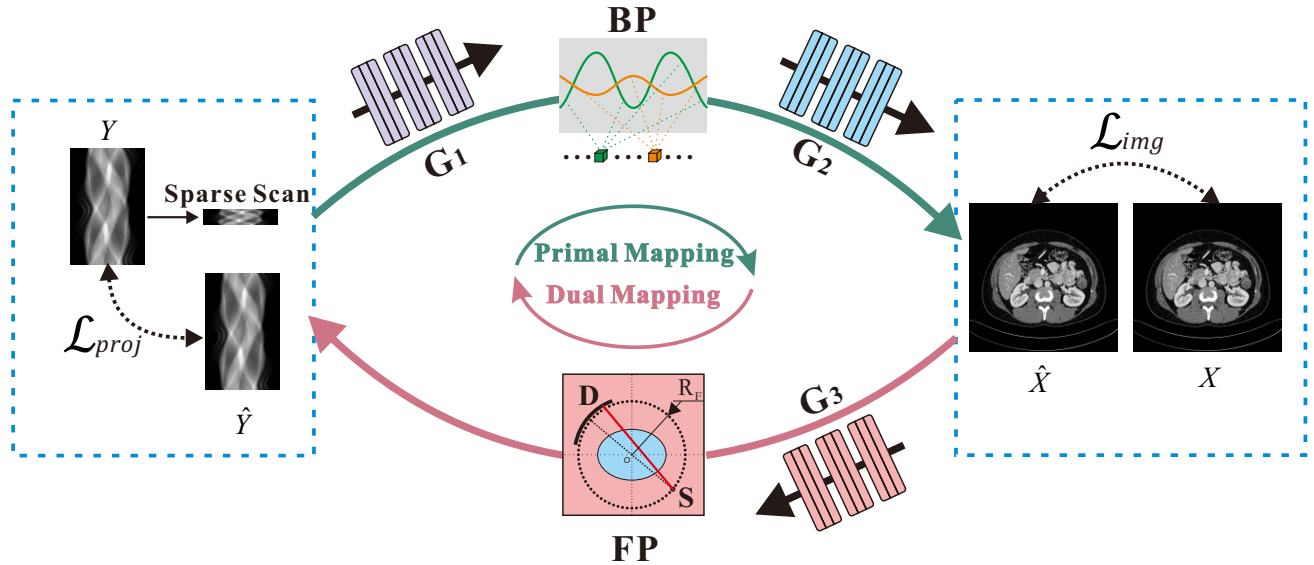

Figure 1. The proposed closed-loop learning reconstruction framework.

proposed to learn in the view-by-view backprojection tensor (VVBP-Tensor), and the experiment found that the results of this learning framework were significantly improved. At present, the application of deep learning in CT reconstruction mainly completes the learning from projection to image. However, this process lacks the constraint of comparing the difference between the calculated projection by projecting the reconstructed image and the measured projection like iterative algorithms.

Different from traditional supervised learning or semi-supervised learning, dual learning forms a closed-loop system by creating a dual problem of the primary problem. The primary problem and dual problem can mutually promote learning each other through this closed-loop system, so as to obtain better learning. Dual learning has been a great success in natural image processing such as image super-resolution[15] and raindrop removal.[16] For CT reconstruction, the advantages of traditional iterative reconstruction algorithms in low-dose CT reconstruction are mainly reflected in the use of forward and backward projection operators to update the target image by error feedback. This process implies a constraint that the estimated projection obtained by projecting the reconstructed image should be consistent with the measured projection. Inspired by these works, we propose a closed-loop learning reconstruction model (CLRecon) for sparse-view CT reconstruction. In the proposed CLRecon, the original problem is learning the mapping from measured projection to image, and the dual problem is learning the mapping from image to measured projection. The mapping from image to projection can effectively constrain the mapping from projection to image to learn in the right direction, thus helping to improve the quality of reconstruction.

2. METHODS

2.1 Overview

Fig. 1 depicts the overview of our proposed closed-loop learning reconstruction framework. It consists of two learning mappings. The primal mapping is used to learn the transformation from sparse-view projection to reconstructed image, and the second one learns the dual mapping from reconstructed image to projection. The primal mapping includes a projection domain network, a gradient returnable backward projection module which is the implementation of FBP and an image domain network. The dual learning mapping consists of an image domain network and a gradient returnable forward projection module.

2.2 Network Architecture

We build the backward projection module by implementing the FBP algorithm based on PyTorch deep learning library,[17] which can retain gradient during forward propagation. Considering the sparsity of the system
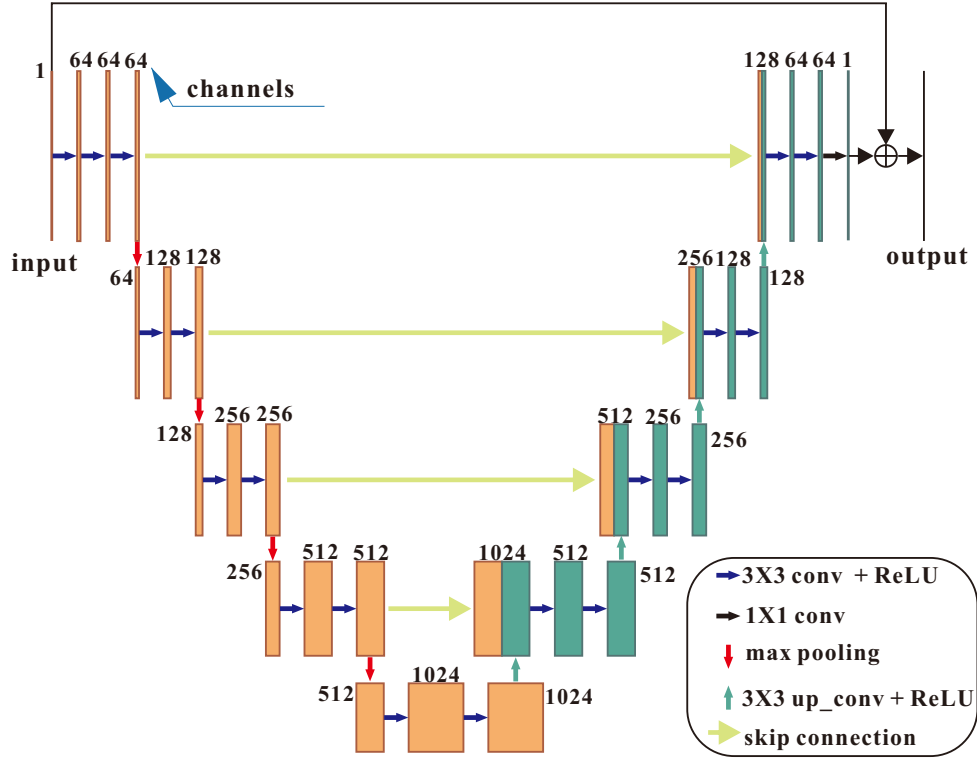
Figure 2. The architecture of the base network.

matrix, we construct it by sparse matrix. This greatly reduces memory requirements and makes it possible to project the reconstructed image when training the network. In our proposed closed-loop learning sparse-view CT reconstruction framework, sub-network G_1, G_2, G_3 can be any network. In our next experiment, we chose FBPConvNet[18] as the base network and we remove the batch normalization layer as shown in Fig. 2.

2.3 Loss Functions

We adopt the simple mean square error (MSE) loss to train the network. To make network training constrained by both label image and measured projection, we add the loss in both the image domain and projection domain. For image domain loss, we can formulate it as follows:

$$\mathcal{L}_{img} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \|\hat{X}_{ij} - X_{ij}\|_2^2, \tag{1}$$

where $\hat{X}$ is the output of the network and X is the label image. M is the number of samples in a batch and N is the number of pixels in a sample. The projection domain loss can be formulated as follows:

$$\mathcal{L}_{proj} = \frac{1}{MP} \sum_{i=1}^{M} \sum_{j=1}^{P} \|\hat{Y}_{ij} - Y_{ij}\|_2^2, \tag{2}$$

where $\hat{Y}$ is the output of the network and Y is the measured projection. P is the number of pixels in a sample. If we have a full-view projection, Y can be a full-view projection for better performance. The full objective function contains the image domain loss and the projection domain loss as follows:

$$\mathcal{L}_{mse} = \lambda \mathcal{L}_{img} + (1 - \lambda)\mathcal{L}_{proj}, \tag{3}$$

where λ is a parameter that balances image loss and the projection loss, and in this work, it is set to be 0.8.
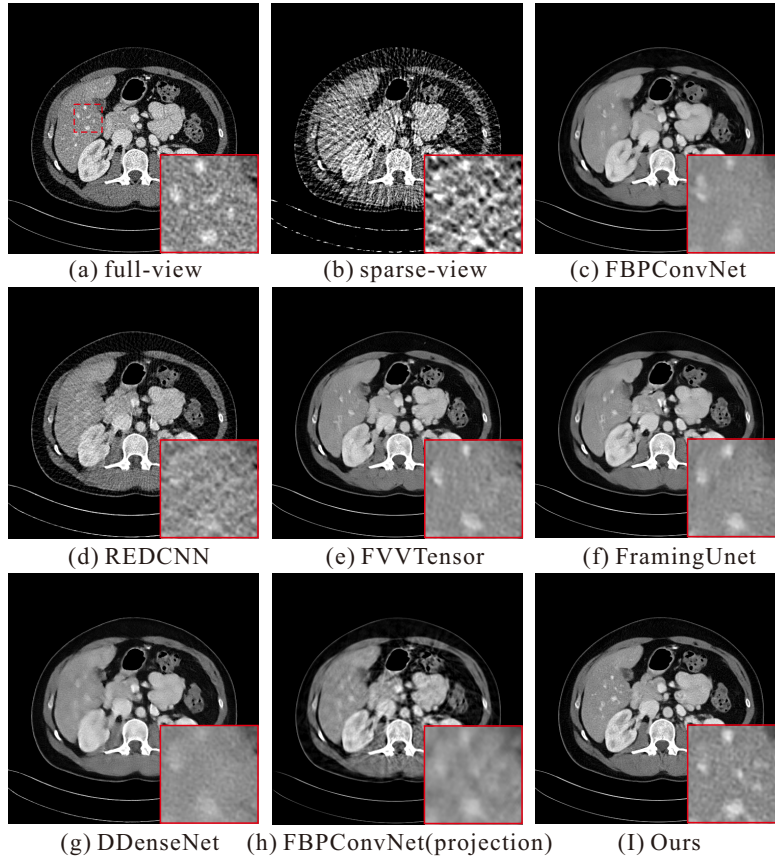
Figure 3. Reconstruction results of different networks with a view number of 72. Display window: [-160, 240] HU.

3. EXPERIMENTS

3.1 Data

We used AAPM Low Dose CT Grand Challenge datasets for evaluation which consists of routine dose CT and the corresponding simulated LDCT data.[19] We used 10976 slices from 34 patients for training the network and 1095 slices from 6 patients for validation and testing. We projected the images to obtain the simulation fan-beam projection data. The geometry parameters projection were set as projection view number of 1152, detector bin number of 736, image pixel space of $0.6934mm \times 0.6934mm$, and detector bin width of $1.2858\ mm$. In our experiment, we extracted 72 views with equal angle distribution to simulate the sparse-view scan.

3.2 Implementation Details

The framework was implemented in Python based on PyTorch deep learning library.[17] All reconstruction images have a size of 512×512 and the sinograms are with a size of $n \times 1152$, where n is the projection views. The Adam optimizer[20] was used to optimize the whole framework with the parameters $(\beta_1, \beta_2) = (0.9, 0.999)$. The learning rate drop linearly from 10^{-3} to 10^{-5}. We trained the network on NVIDIA GeForce RTX 3090 GPUs.

4. RESULTS

4.1 Experimental Results on Mayo Data

4.1.1 Qualitative analysis

We compared our method with the recent deep-learning-based methods, including FBPConvNet,[18] FramingUNet,[21] REDCNN,[22] DDenseNet,[3] FVVtensor.[14] FBPConvNet is closer to the network we used, but we did

(a) full-view   (b) sparse-view   (c) G_2

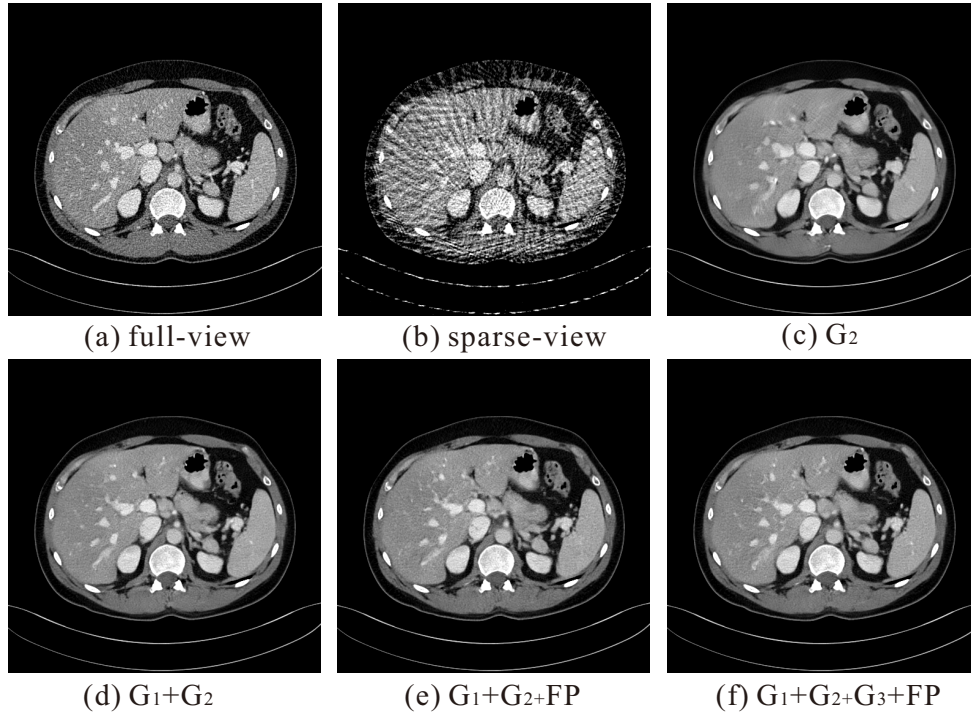(d) G_1+G_2 (e) G_1+G_2+FP (f) $G_1+G_2+G_3+FP$

Figure 4. Reconstruction results of different combinations of modules. Display window: [-160, 240]HU.

Table 1. Quantitative comparison of different models.

Model	RMSE	SSIM
FBPConvNet	4.6328	0.8710
REDCNN	4.9033	0.8417
FVVTensor	4.2655	0.8827
FramingUnet	4.4261	0.8782
DDenseNet	4.4699	0.8770
FBPConvNet (projection)	5.0693	0.8603
Ours	**3.7063**	**0.9004**

not use the batch normalization layer as it did. FramingUNet, REDCNN, and DDenseNet optimized the network structure for better performance. FVVtensor learns in VVBP-Tensor and obtains good results. We also applied the FBPConvNet in the projection domain to repair sparse views projection directly.

Fig. 3 shows the visual comparisons of our method and other methods on the reconstructed images with 72 views. We show the full-view image, sparse-view image, and results of different models. To better compare the results of different networks, the ROI in the image is enlarged at the bottom of the image. It can be observed that our network can achieve excellent results in edge preservation and artifact removal.

4.1.2 Quantitative comparisons with state-of-the-art methods

Table 1 shows the quantitative comparison results of our method and other methods. It's the average of all the test slices. We can observe that our model achieves lower root mean square error (RMSE) and higher structured similarity index (SSIM) than other methods, which can prove that our network can obtain better reconstruction quality under these sparse-view degradation levels.

Table 2. Quantitative comparison of different combinations of modules.

Model	RMSE	SSIM
G_2	4.1230	0.8858
$G_1 + G_2$	3.9299	0.8937
$G_1 + G_2 + \text{FP}$	3.7685	0.8984
$G_1 + G_2 + G_3 + \text{FP}$	**3.7063**	**0.9004**

4.2 Ablation Study

In our proposed CLRecon, we add image-to-projection mapping to the primal mapping (i.e., projection-to-image mapping) and form a closed-loop learning system. To show the effectiveness of this procedure, we compared the reconstruction results of different combinations of G_1, G_2, G_3 and forward projection module (FP). Fig. 4 shows the visual comparisons of different combinations of modules. We can observe that the image quality can be improved after adding G_3 and FP. Table 2 shows the quantitative comparison results of different experiments. We can see that it can achieve lower RMSE and SSIM with forward projection module and G_3. The network complexity did not increase because image-to-projection mapping was not used during the test stage. The improvement of reconstructed image quality lies in the improvement of network learning strategy rather than the increase of network depth. It proves that the proposed closed-loop learning can improve the reconstruction quality in sparse-view CT reconstruction.

5. CONCLUSION

We have presented a closed-loop learning reconstruction model (CLRecon) for sparse-view CT reconstruction in this paper. Specifically, the primal mapping is used to learn the transformation from sparse-view projection to reconstructed image, and the dual mapping learns from reconstructed image to projection. Our experiment shows that the addition of these two modules can improve the quality of the reconstructed image. Since this mapping is only used in network training, the network parameters are not increased. The improvement of network performance lies in the change of learning strategy.

ACKNOWLEDGMENTS

REFERENCES

[1] Sidky, E. Y. and Pan, X., "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Physics in Medicine & Biology* **53**(17), 4777 (2008).

[2] Kim, K., El Fakhri, G., and Li, Q., "Low-dose ct reconstruction using spatially encoded nonlocal penalty," *Medical physics* **44**(10), e376–e390 (2017).

[3] Zhang, Z., Liang, X., Dong, X., Xie, Y., and Cao, G., "A sparse-view ct reconstruction method based on combination of densenet and deconvolution," *IEEE transactions on medical imaging* **37**(6), 1407–1417 (2018).

[4] Lee, H., Lee, J., Kim, H., Cho, B., and Cho, S., "Deep-neural-network-based sinogram synthesis for sparse-view ct image reconstruction," *IEEE Transactions on Radiation and Plasma Medical Sciences* **3**(2), 109–119 (2018).

[5] Shan, H., Padole, A., Homayounieh, F., Kruger, U., Khera, R. D., Nitiwarangkul, C., Kalra, M. K., and Wang, G., "Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose ct image reconstruction," *Nature Machine Intelligence* **1**(6), 269–276 (2019).

[6] Zheng, A., Gao, H., Zhang, L., and Xing, Y., "A dual-domain deep learning-based reconstruction method for fully 3d sparse data helical ct," *Physics in Medicine & Biology* **65**(24), 245030 (2020).

[7] Zhang, Q., Hu, Z., Jiang, C., Zheng, H., Ge, Y., and Liang, D., "Artifact removal using a hybrid-domain convolutional neural network for limited-angle computed tomography imaging," *Physics in Medicine & Biology* **65**(15), 155010 (2020).

[8] Chen, H., Zhang, Y., Zhang, W., Sun, H., Liao, P., He, K., Zhou, J., and Wang, G., "Learned experts' assessment-based reconstruction network (" learn") for sparse-data ct," *arXiv preprint arXiv:1707.09636* (2017).

[9] Zhang, H., Liu, B., Yu, H., and Dong, B., "Metainv-net: meta inversion network for sparse view ct image reconstruction," *IEEE Transactions on Medical Imaging* **40**(2), 621–634 (2020).

[10] Hauptmann, A., Adler, J., Arridge, S., and Öktem, O., "Multi-scale learned iterative reconstruction," *IEEE transactions on computational imaging* **6**, 843–856 (2020).

[11] Zhu, B., Liu, J. Z., Cauley, S. F., Rosen, B. R., and Rosen, M. S., "Image reconstruction by domain-transform manifold learning," *Nature* **555**(7697), 487–492 (2018).

[12] He, J., Wang, Y., and Ma, J., "Radon inversion via deep learning," *IEEE transactions on medical imaging* **39**(6), 2076–2087 (2020).

[13] Shen, L., Zhao, W., and Xing, L., "Patient-specific reconstruction of volumetric computed tomography images from a single projection view via deep learning," *Nature biomedical engineering* **3**(11), 880–888 (2019).

[14] Tao, X., Wang, Y., Lin, L., Hong, Z., and Ma, J., "Learning to reconstruct ct images from the vvbp-tensor," *IEEE Transactions on Medical Imaging* (2021).

[15] Guo, Y., Chen, J., Wang, J., Chen, Q., Cao, J., Deng, Z., Xu, Y., and Tan, M., "Closed-loop matters: Dual regression networks for single image super-resolution," in [*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*], 5407–5416 (2020).

[16] Ye, Y., Chang, Y., Zhou, H., and Yan, L., "Closing the loop: Joint rain generation and removal via disentangled image translation," in [*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*], 2053–2062 (2021).

[17] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al., "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems* **32** (2019).

[18] Ongie, G., Jalal, A., Metzler, C. A., Baraniuk, R. G., Dimakis, A. G., and Willett, R., "Deep learning techniques for inverse problems in imaging," *IEEE Journal on Selected Areas in Information Theory* **1**(1), 39–56 (2020).

[19] "Low dose ct grand challenge."

[20] Kingma, D. P. and Ba, J., "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980* (2014).

[21] Han, Y. and Ye, J. C., "Framing u-net via deep convolutional framelets: Application to sparse-view ct," *IEEE transactions on medical imaging* **37**(6), 1418–1429 (2018).

[22] Chen, H., Zhang, Y., Kalra, M. K., Lin, F., Chen, Y., Liao, P., Zhou, J., and Wang, G., "Low-dose ct with a residual encoder-decoder convolutional neural network," *IEEE transactions on medical imaging* **36**(12), 2524–2535 (2017).