Patterns, Volume *3*

Supplemental information

Multi-domain integrative Swin transformer network for sparse-view tomographic reconstruction

Jiayi Pan, Heye Zhang, Weifei Wu, Zhifan Gao, and Weiwen Wu

Supplemental Experimental Procedures

Appendix A

More simulated data reconstructions

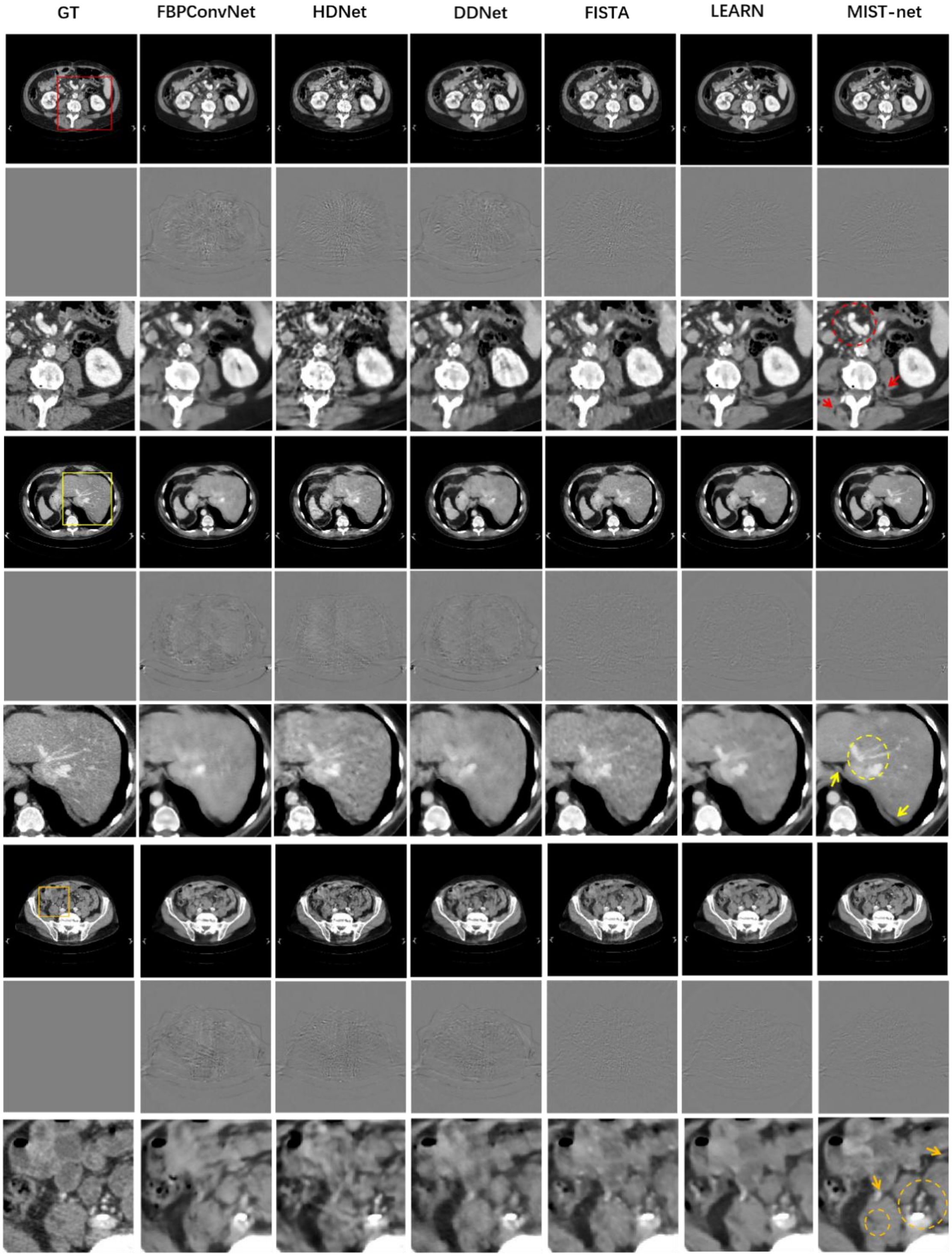

Figure S1. More simulated data reconstructions from sparse-view data by using different networks. The 1st-7th columns stand for the FBP reconstruction from full-view data, FBPconvNet, HDNet, DDNet, FISTA, LEARN and MIST-net counterparts from 48 views. The display windows for the reconstructed images are [-160 240] HU.

We also do the experiments on numerical datasets with 64 views. The results can be found in Figure S2.

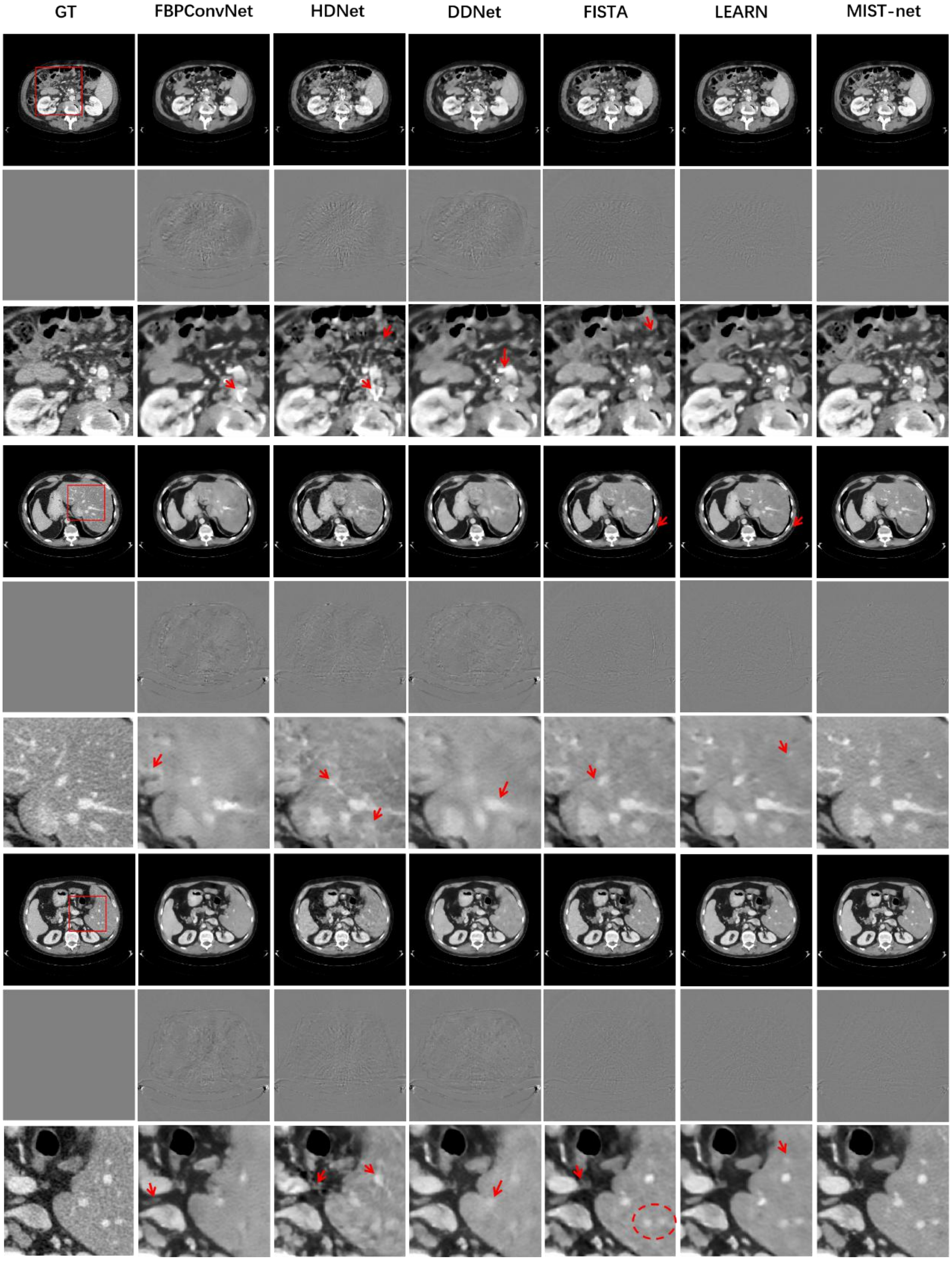

Figure S2. The simulated data reconstructions from 64 views by using different networks. The 1st-7th columns stand for the FBP reconstruction from full-view data, FBPconvNet, HDNet, DDNet, FISTA, LEARN and MIST-net. The display windows for the reconstructed images are [-160 240] HU.

Appendix B

The statistical quantitative evaluations

The statistical quantitative evaluations result from testing datasets were computed in terms of RMSE, PSNR and SSIM, and their results were summarized in Table S1. It can be seen that our MIST-net can obtain the best quantitative statistical results in terms of mean and standard deviation than other competitors.

Table S1

Quantitative evaluation of 64 projections reconstruction results from simulated testing datasets

Views	Methods	RMSE	PSNR	SSIM
64	FBPconvNet	25.9083 ± 4.5941	38.6168 ± 1.6026	0.9634 ± 0.0174
	HDNet	22.8740 ± 4.2488	39.2675 ± 1.2331	0.9670 ± 0.0144
	DDNet	24.1746 ± 4.7592	39.7018 ± 1.3685	0.9655 ± 0.0163
	FISTA	17.4884 ± 3.1788	42.0121 ± 1.2369	0.9769 ± 0.0100
	LEARN	15.9171 ± 3.6576	42.8622 ± 1.3997	0.9814 ± 0.0123
	MIST-net	$\mathbf{14.5492 \pm 3.2807}$	$\mathbf{43.6418 \pm 1.4070}$	$\mathbf{0.9844 \pm 0.00101}$

In addition to the performance metrics, we also compare the complexity and runtime of all competitors. The test time represents the time taken to predict a total of 391 test datasets. Table S2 shows that our MIST-net train and test much faster than LEARN and FISTA.

Table S2

Time-consuming comparison of different methods (48 projections)

Views		FBPconvNet	HDNet	DDNet	FISTA	LEARN	MIST-net
48	Epochs	150	150(×2)	100	40	40	40
	Input Size	512×512	48×880	512×512	48×880	48×880	48×880
	Output Size	512×512	512×512	512×512	512×512	512×512	512×512
	#Param.	9.8M	9.8M(×2)	6.4M	1.3M	3.0M	12.0M
	Train Time	21.3h	37.5h	41.2h	30.1h	83.1h	38.7h
	Test Time	40s	63s	66s	105s	316s	98s

Appendix C

More noise experiment results

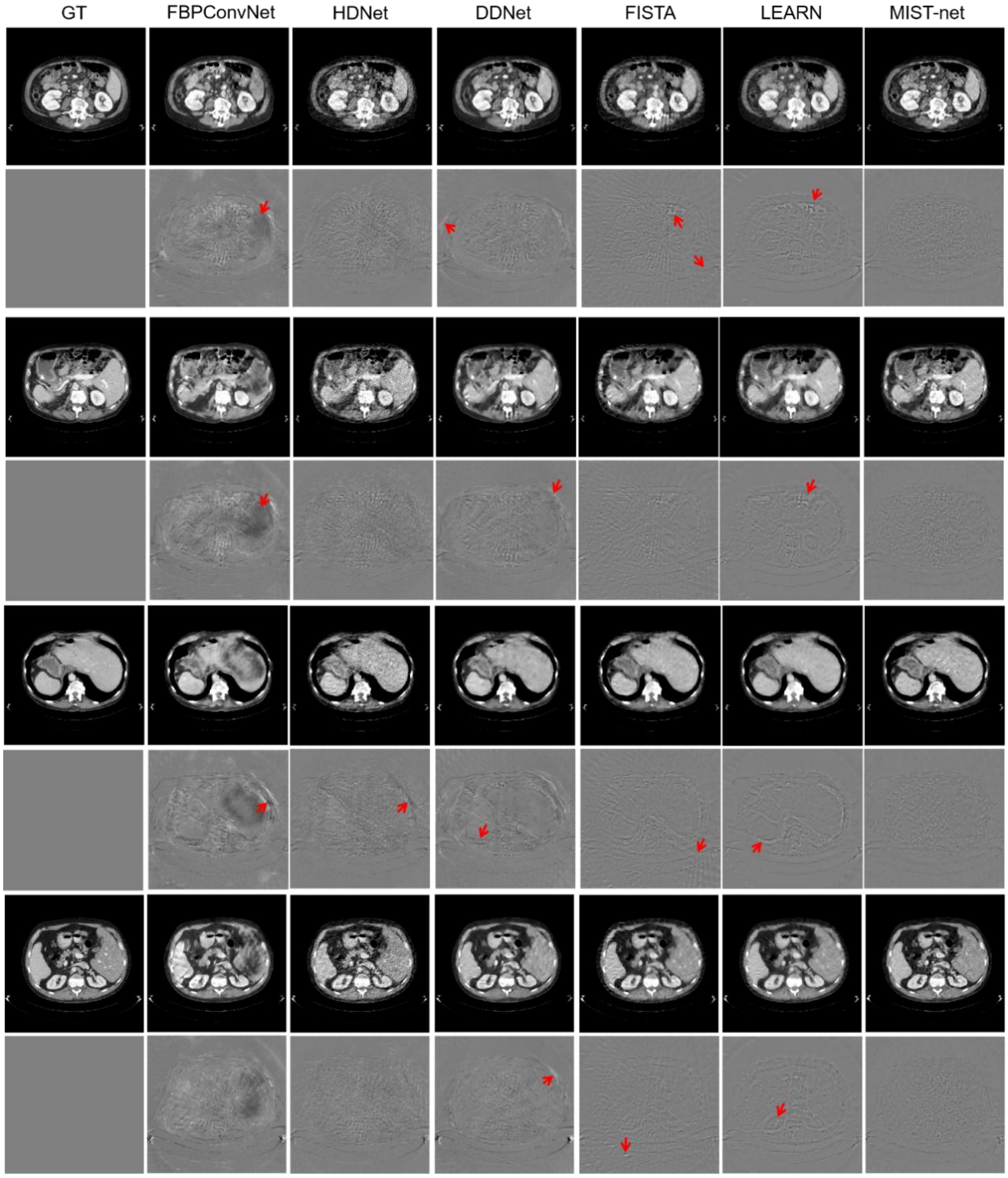

Figure S3. The generalization of different deep reconstruction networks against noise on simulation datasets. The 1st-7th columns stand for the ground truth, FBPconvNet, HDNet, DDNet, FISTA, LEARN and MIST-net from 48 views. The display window of reconstructed images is [-160 240] HU.

Appendix D

More network details

Table S3

Parametric structure for all the layers in the encoder-decoder block

Layers	Parameters	Input Channel	Output Channel
A0	3×3Conv+BN+ReLU	1	32
	3×3Conv+BN+ReLU		
A1	2×2 MaxPooling	32	64
	3×3Conv+BN+ReLU		
	3×3Conv+BN+ReLU		
A2	2×2 MaxPooling	64	128
	3×3Conv+BN+ReLU		
	3×3Conv+BN+ReLU		
A3	2×2 MaxPooling	128	256
	3×3Conv+BN+ReLU		
	3×3Conv+BN+ReLU		
A4	2×2 MaxPooling	256	512
	3×3Conv+BN+ReLU		
	3×3Conv+BN+ReLU		
B0	Upsample+3×3Conv+BN+ReLU	512	256
	Concatenation		
	3×3Conv+BN+ReLU		
	3×3Conv+BN+ReLU		
B1	Upsample+3×3Conv+BN+ReLU	256	128
	Concatenation		
	3×3Conv+BN+ReLU		
	3×3Conv+BN+ReLU		
B2	Upsample+3×3Conv+BN+ReLU	128	64
	Concatenation		
	3×3Conv+BN+ReLU		
	3×3Conv+BN+ReLU		
B3	Upsample+3×3Conv+BN+ReLU	64	32
	Concatenation		
	3×3Conv+BN+ReLU		
	3×3Conv+BN+ReLU		
B4	1×1 Convolution	32	1
B5	Residual Connection	1	1

Table S4

Parametric structure for all the layers in the edge enhancement Rec-network

Layers	**Parameters**	**Input Channel**	**Output Channel**
C0	Sobel Convolution	1	32
	Concatenation		33
C1	3×3Conv+ReLU	33	32
	3×3Conv+ReLU	32	32
	Concatenation		65
C2	3×3Conv+ReLU	65	32
	3×3Conv+ReLU	32	32
	Concatenation		65
C3	3×3Conv+ReLU	65	32
	3×3Conv+ReLU	32	32
	Concatenation		65
C4	3×3Conv+ReLU	65	32
	3×3Conv+ReLU	32	32
	Concatenation		65
C5	3×3Conv+ReLU	65	32
	3×3Conv+ReLU	32	32
	Concatenation		65
C6	3×3Conv+ReLU	65	32
	3×3Conv+ReLU	32	32
	Concatenation		65
C7	3×3Conv+ReLU	65	32
	3×3Conv+ReLU	32	32
	Concatenation		65
C8	3×3Conv+ReLU	65	32
	3×3Conv	32	1
	Residual Connection	1	1

Table S5

Parametric structure for all the layers in the Swin Rec-former

Layers	Parameters	Window Size	Head Numbers	Head Numbers
D0	3×3convolution			
D1	Patch-Embed Block	8×8	6	1×1
	Swin Transformer Block (×6)			
	Patch-UnEmbed Block			
	3×3convolution+ Residual Connection			
D2	Patch-Embed Block	8×8	6	1×1
	Swin Transformer Block (×6)			
	Patch-UnEmbed Block			
	3×3convolution+ Residual Connection			
D3	Patch-Embed Block	8×8	6	1×1
	Swin Transformer Block (×6)			
	Patch-UnEmbed Block			
	3×3convolution+ Residual Connection			
D4	Patch-Embed Block	8×8	6	1×1
	Swin Transformer Block (×6)			
	Patch-UnEmbed Block			
	3×3convolution+ Residual Connection			
D5	3×3convolution+ Residual Connection			