Quantitative Comparison of Deep Learning-Based Image Reconstruction Methods for Low-Dose and Sparse-Angle CT Applications

Johannes Leuschner, Maximilian Schmidt, Poulami S. Ganguly, Vladyslav Andriiashen, Sophia B. Coban, Alexander Denker, Daniel O. Baguer, Dominik Bauer, Amir Hadjifaradji, Kees J. Batenburg, Peter Maass, Maureen van Eijnatten

Center for Industrial Mathematics
University of Bremen
jleuschn@uni-bremen.de

Table of Contents

Computed Tomography (CT)

A typical clinical CT scanner
Photo by daveynin / CC BY 2.0

An industrial nano CT system
©Fraunhofer IIS, Image from
https://www.iis.fraunhofer.de/en/pr/2020/20200604_ntct.html

Mathematical formulation for parallel beam geometry

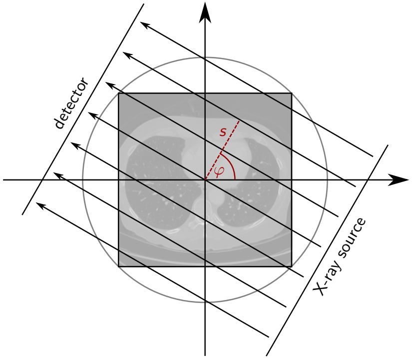


Figure: Parallel beam geometry

- Radon transform $\mathcal{A}x(s,\varphi)$ simulates the attenuation of a single beam

$$\mathcal{A}x(s,\varphi) = \int_{\mathbb{R}} x\left(s\theta + t\theta^{\perp}\right)\,\mathrm{d}t$$

$$\theta = (\cos\varphi, \sin\varphi)^T,\ \varphi \in [0,\pi)$$

- Beer-Lambert's law states:

$$\mathcal{A}x(s,\varphi) = -\log\left(\frac{I_1(s,\varphi)}{I_0}\right)$$

The linear system of equations (for 2D image reconstruction)

$$\mathcal{A}x(s,\varphi) = -\log\left(\frac{I_1(s,\varphi)}{I_0}\right) = y$$

- Number of angles n_a, number of detector pixels n_d
  $\rightsquigarrow$ $m = n_a \cdot n_d$ measured intensities
- Discretize image domain with $n_{\mathrm{im}} \times n_{\mathrm{im}}$ uniform pixels
  $\rightsquigarrow$ $n = n_{\mathrm{im}}^2$ attenuation values to be reconstructed

$$Ax^{\dagger} + \epsilon = y_{\delta}, \qquad A \in \mathbb{R}_{\geq 0}^{m\times n}, \quad x^{\dagger} \in \mathbb{R}_{\geq 0}^{n}, \quad y_{\delta} \in \mathbb{R}^{m}$$

Reconstruction task and learning

- Mildly ill-posed inverse problem
- Classical methods:
  - Filtered back-projection
  - Iterative reconstruction
- Deep learning methods show convincing results for reconstruction and segmentation

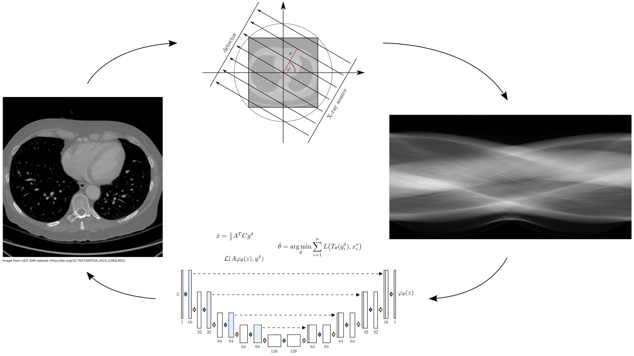


Image from LIDC-IDRI dataset (http://doi.org/10.7937/K9/TCIA.2015.LO9QL95X)

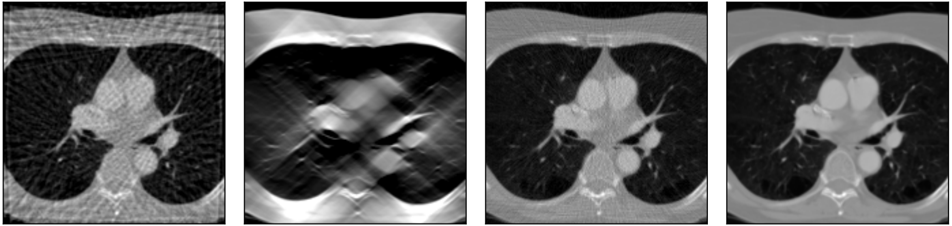
Reconstruction challenges

- Few angles (sparse view)
- Limited angle range (limited view)
- Low intensity (noise)

Goals in applications

- Reduce potentially harmful radiation dose
- Reduce scanning time
- Meet technical limitations

Sparse view | Limited view | Noise | Ideal



Applications

Clinical CT

- Diagnostics
- Screening
- Virtual treatment planning
- . . .

Industrial CT

- Non-destructive testing (NDT)
- Assembly analysis
- . . .

Scientific CT

- Micro CT / Nano CT
  - Material science
  - Biomedical research
- . . .

Benchmark data

LoDoPaB-CT data [9]

Apples-CT data [4]

Applications

Clinical CT

- Diagnostics
- Screening
- Virtual treatment planning
- . . .

Industrial CT

- Non-destructive testing (NDT)
- Assembly analysis
- . . .

Scientific CT

- Micro CT / Nano CT
  - Material science
  - Biomedical research
- . . .

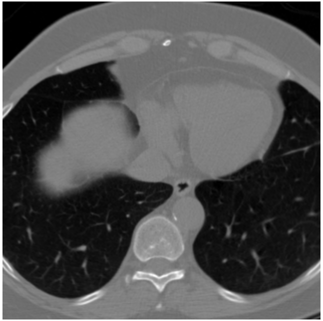
Benchmark data

LoDoPaB-CT data [9]

Apples-CT data [4]

Low-Dose Parallel Beam (LoDoPaB)-CT Dataset [9]

- Ground truth: 40 000 normal-dose thoracic CT slices from the LIDC/IDRI Database [2]
- Measurements: 1000 angles and 513 detector pixels
- Image size: 362 px $\times$ 362 px
- Low-dose simulation with Poisson noise
- Sampling ratio: ≈ 3.9 (oversampling)
- Public challenge at `lodopab.grand-challenge.org`
- Easy access with our Python library DIV$\alpha\ell$ [10]

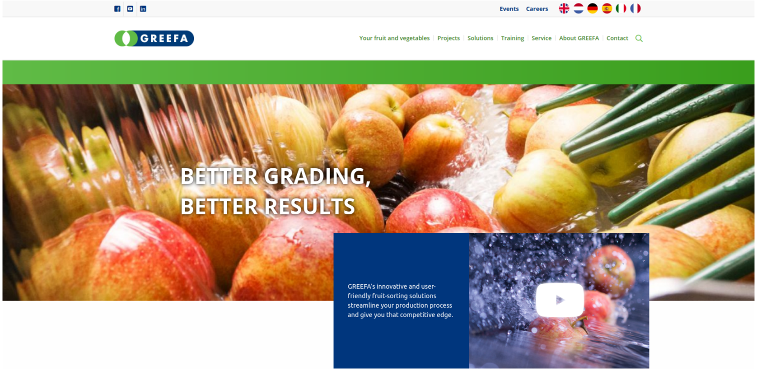


[2] Armato III et al., 2011, "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans"

[9] Leuschner et al., 2021, "LoDoPaB-CT, a benchmark dataset for low-dose computed tomography reconstruction"

Apples-CT Datasets [4]: Application of fruit sorting

- Data provided by company GREEFA, based in Tricht, NL
- Defect detection / segmentation task
- High speed requirements
- Very few angles (for viability)

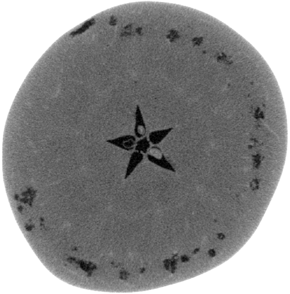


https://www.greefa.com/

[4] S. B. Coban et al. *Parallel-beam X-ray CT datasets of apples with internal defects and label balancing for machine learning*. 2020. arXiv: 2012.13346 [cs.LG]

Apples-CT Datasets [4]

- Ground truth: 70 000 high quality CT slices of apples with and without internal defects
- Simulated parallel beam measurements: 50, 10, 5, 2 angles and 1377 detector pixels
- Image size: 972 px $\times$ 972 px
- 3 noise settings: Noise-free, Gaussian, scattering
- Sampling ratio: $\approx$ 0.07-0.003
- Public challenge at `apples-ct.grand-challenge.org`



Bitterpit defect within the slice

[4] S. B. Coban et al. *Parallel-beam X-ray CT datasets of apples with internal defects and label balancing for machine learning*. 2020. arXiv: 2012.13346 [cs.LG]

Kick-off event

- Participants from many institutions
- Speakers from Industry and Academia
- Two challenges: LoDoPaB-CT and Apples-CT

Code Sprint 2020
June 15-24th

http://dival.math.uni-bremen.de/code_sprint_2020/

Included Methods - From Modeling to Data-driven

- **Classical reconstruction**: Filtered back-projection (FBP), Total variation (TV), CGLS
- **Learned iterative schemes**: Learned Primal-Dual [1] $x^{[l+1]} = F_{\theta_l}(x^{[l]}, y_\delta),\ l = 1, \dots, L,\ \hat{x} = x^{[L]}$
- **Unsupervised**: Deep Image Prior + TV [3] $\hat{\theta} = \min_\theta \|\mathcal{A}F_\theta(z) - y_\delta\|, \quad \hat{x} = F\hat{\theta}(z)$
- **Generative models**: Conditional INN [5] $\hat{x} = \frac{1}{n}\sum_i^n F_\theta\left(z_i, \mathrm{FBP}\left(y_\delta\right)\right), \quad z_i \sim \mathcal{N}(0, I)$
- **Postprocessing**: U-Net [7], U-Net++ [15], ISTA U-Net [12], MS-D-CNN [13] $\hat{x} = F_\theta\left(\mathrm{FBP}\left(y_\delta\right)\right)$
- **Fully learned**: iCTU-Net [11] $\hat{x} = F_\theta\left(y_\delta\right)$

Included Methods - From Modeling to Data-driven

- **Classical reconstruction**: Filtered back-projection (FBP), Total variation (TV), CGLS
- **Learned iterative schemes**: Learned Primal-Dual [1] $x^{[l+1]} = F_{\theta_l}(x^{[l]}, y_\delta),\ l = 1, \dots, L,\ \hat{x} = x^{[L]}$
- **Unsupervised**: Deep Image Prior + TV [3] $\hat{\theta} = \min_\theta \|\mathcal{A}F_\theta(z) - y_\delta\|, \quad \hat{x} = F\hat{\theta}(z)$
- **Generative models**: Conditional INN [5] $\hat{x} = \frac{1}{n}\sum_i^n F_\theta\left(z_i, \mathrm{FBP}\left(y_\delta\right)\right), \quad z_i \sim \mathcal{N}(0, I)$
- **Postprocessing**: U-Net [7], U-Net++ [15], ISTA U-Net [12], MS-D-CNN [13] $\hat{x} = F_\theta\left(\mathrm{FBP}\left(y_\delta\right)\right)$
- **Fully learned**: iCTU-Net [11] $\hat{x} = F_\theta\left(y_\delta\right)$

[1] Adler et al., 2018, "Learned Primal-Dual Reconstruction"

Included Methods - From Modeling to Data-driven

- **Classical reconstruction**: Filtered back-projection (FBP), Total variation (TV), CGLS
- **Learned iterative schemes**: Learned Primal-Dual [1] $x^{[l+1]} = F_{\theta_l}(x^{[l]}, y_\delta),\ l = 1, ..., L,\ \hat{x} = x^{[L]}$
- **Unsupervised**: Deep Image Prior + TV [3] $\hat{\theta} = \min_\theta \|\mathcal{A}F_\theta(z) - y_\delta\|, \quad \hat{x} = F\hat{\theta}(z)$
- **Generative models**: Conditional INN [5] $\hat{x} = \frac{1}{n}\sum_i^n F_\theta\left(z_i, \mathrm{FBP}\left(y_\delta\right)\right), \quad z_i \sim \mathcal{N}(0, I)$
- **Postprocessing**: U-Net [7], U-Net++ [15], ISTA U-Net [12], MS-D-CNN [13] $\hat{x} = F_\theta\left(\mathrm{FBP}\left(y_\delta\right)\right)$
- **Fully learned**: iCTU-Net [11] $\hat{x} = F_\theta\left(y_\delta\right)$

[3] Baguer et al., 2020, "Computed tomography reconstruction using deep image prior and learned reconstruction methods"

Included Methods - From Modeling to Data-driven

- **Classical reconstruction**: Filtered back-projection (FBP), Total variation (TV), CGLS
- **Learned iterative schemes**: Learned Primal-Dual [1] $x^{[l+1]} = F_{\theta_l}(x^{[l]}, y_\delta),\ l = 1, \dots, L,\ \hat{x} = x^{[L]}$
- **Unsupervised**: Deep Image Prior + TV [3] $\hat{\theta} = \min_\theta \|\mathcal{A}F_\theta(z) - y_\delta\|, \quad \hat{x} = F\hat{\theta}(z)$
- **Generative models**: Conditional INN [5] $\hat{x} = \frac{1}{n}\sum_i^n F_\theta\left(z_i, \mathrm{FBP}\left(y_\delta\right)\right), \quad z_i \sim \mathcal{N}(0, I)$
- **Postprocessing**: U-Net [7], U-Net++ [15], ISTA U-Net [12], MS-D-CNN [13] $\hat{x} = F_\theta\left(\mathrm{FBP}\left(y_\delta\right)\right)$
- **Fully learned**: iCTU-Net [11] $\hat{x} = F_\theta\left(y_\delta\right)$

[5] Denker et al., 2020, *Conditional Normalizing Flows for Low-Dose Computed Tomography Image Reconstruction*

Included Methods - From Modeling to Data-driven

- **Classical reconstruction**: Filtered back-projection (FBP), Total variation (TV), CGLS
- **Learned iterative schemes**: Learned Primal-Dual [1] $x^{[l+1]} = F_{\theta_l}(x^{[l]}, y_\delta),\ l = 1, ..., L,\ \hat{x} = x^{[L]}$
- **Unsupervised**: Deep Image Prior + TV [3] $\hat{\theta} = \min_\theta \|\mathcal{A}F_\theta(z) - y_\delta\|, \quad \hat{x} = F\hat{\theta}(z)$
- **Generative models**: Conditional INN [5] $\hat{x} = \frac{1}{n}\sum_i^n F_\theta\left(z_i, \mathrm{FBP}\left(y_\delta\right)\right), \quad z_i \sim \mathcal{N}(0, I)$
- **Postprocessing**: U-Net [7], U-Net++ [15], ISTA U-Net [12], MS-D-CNN [13] $\hat{x} = F_\theta\left(\mathrm{FBP}\left(y_\delta\right)\right)$
- **Fully learned**: iCTU-Net [11] $\hat{x} = F_\theta\left(y_\delta\right)$

[7] Jin et al., 2017, "Deep convolutional neural network for inverse problems in imaging"
[15] Zhou et al., 2018, "Unet++: A nested u-net architecture for medical image segmentation"
[12] Liu et al., 2020, *Interpreting U-Nets via Task-Driven Multiscale Dictionary Learning*
[13] Pelt et al., 2018, "Improving tomographic reconstruction from limited data using Mixed-Scale Dense convolutional neural networks"

Included Methods - From Modeling to Data-driven

- **Classical reconstruction**: Filtered back-projection (FBP), Total variation (TV), CGLS
- **Learned iterative schemes**: Learned Primal-Dual [1] $x^{[l+1]} = F_{\theta_l}(x^{[l]}, y_\delta),\ l = 1, ..., L,\ \hat{x} = x^{[L]}$
- **Unsupervised**: Deep Image Prior + TV [3] $\hat{\theta} = \min_\theta \|\mathcal{A}F_\theta(z) - y_\delta\|, \quad \hat{x} = F\hat{\theta}(z)$
- **Generative models**: Conditional INN [5] $\hat{x} = \frac{1}{n}\sum_i^n F_\theta\left(z_i, \mathrm{FBP}\left(y_\delta\right)\right), \quad z_i \sim \mathcal{N}(0, I)$
- **Postprocessing**: U-Net [7], U-Net++ [15], ISTA U-Net [12], MS-D-CNN [13] $\hat{x} = F_\theta\left(\mathrm{FBP}\left(y_\delta\right)\right)$
- **Fully learned**: iCTU-Net [11] $\hat{x} = F_\theta\left(y_\delta\right)$

[11] Leuschner et al., 2021, "Quantitative Comparison of Deep Learning-Based Image Reconstruction Methods for Low-Dose and Sparse-Angle CT Applications"

Performance measures

Standard metrics (based on reference ground truth image):

- Peak signal-to-noise ratio: pixel-wise comparison

$$\mathrm{PSNR}\left(\hat{x}, x^\dagger\right) := 10\log_{10}\left(\frac{L^2}{\mathrm{MSE}\left(\hat{x}, x^\dagger\right)}\right)$$

- Structural similarity [14]: compares overall image structure

$$\mathrm{SSIM}\left(\hat{x}, x^\dagger\right) := \frac{1}{M}\sum_{j=1}^{M}\frac{\left(2\hat{\mu}_j\mu_j + C_1\right)\left(2\Sigma_j + C_2\right)}{\left(\hat{\mu}_j^2 + \mu_j^2 + C_1\right)\left(\hat{\sigma}_j^2 + \sigma_j^2 + C_2\right)}.$$

where index j selects sliding window (e.g. of size 7×7)

Data discrepancy $\mathcal{D}_Y\left(\mathcal{A}\hat{x}, y_\delta\right)$

Compare forward projection $y = A\hat{x}$ to observation y_δ

Poisson regression loss on LoDoPaB-CT:

$$-\ell_{\mathrm{Pois}}(y \mid y_\delta) = -\sum_{j=1}^{m} N_0 \exp(-y_{\delta,j}\mu_{\max})(-y_j\mu_{\max} + \ln(N_0)) - N_0 \exp(-y_j\mu_{\max})$$

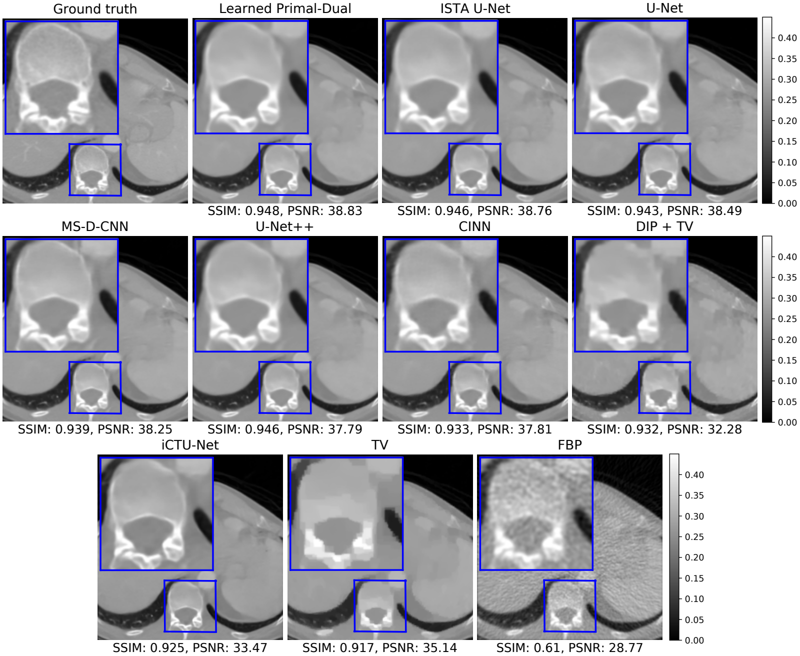
Mean squared error on Apples-CT:

$$\mathrm{MSE}_Y(y, y_\delta) = \frac{1}{m}\|y - y_\delta\|_2^2$$

Relates directly to the likelihood under the assumed noise model for LoDoPaB-CT and for the Gaussian noise setting of Apples-CT

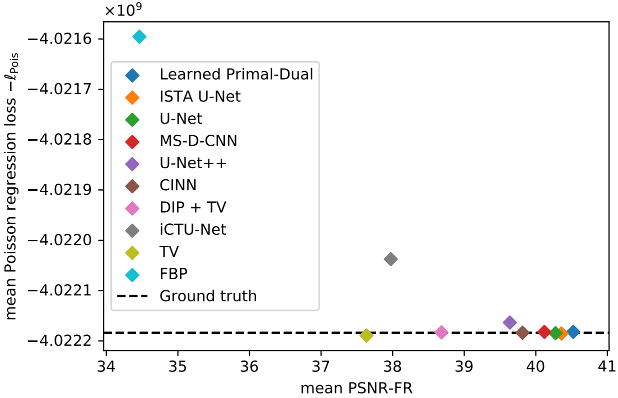
Reconstruction performance on LoDoPaB-CT

Method	PSNR	SSIM	#Params
Learned Primal-Dual	36.25 ± 3.70	0.866 ± 0.115	874,980
ISTA U-Net	36.09 ± 3.69	0.862 ± 0.120	83,396,865
U-Net	36.00 ± 3.63	0.862 ± 0.119	613,322
MS-D-CNN	35.85 ± 3.60	0.858 ± 0.122	181,306
U-Net++	35.37 ± 3.36	0.861 ± 0.119	9,170,079
CINN	35.54 ± 3.51	0.854 ± 0.122	6,438,332
DIP + TV	34.41 ± 3.29	0.845 ± 0.121	hyperp.
iCTU-Net	33.70 ± 2.82	0.844 ± 0.120	147,116,792
TV	33.36 ± 2.74	0.830 ± 0.121	(hyperp.)
FBP	30.19 ± 2.55	0.727 ± 0.127	(hyperp.)


Ground truth
Learned Primal-Dual
SSIM: 0.948, PSNR: 38.83
ISTA U-Net
SSIM: 0.946, PSNR: 38.76
U-Net
SSIM: 0.943, PSNR: 38.49
MS-D-CNN
SSIM: 0.939, PSNR: 38.25
U-Net++
SSIM: 0.946, PSNR: 37.79
CINN
SSIM: 0.933, PSNR: 37.81
DIP + TV
SSIM: 0.932, PSNR: 32.28
iCTU-Net
SSIM: 0.925, PSNR: 33.47
TV
SSIM: 0.917, PSNR: 35.14
FBP
SSIM: 0.61, PSNR: 28.77
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00

Data consistency on LoDoPaB-CT

- Discrepancy term is only explicitly used in TV and DIP + TV
- Still, the data discrepancy of most methods is close to the empirical mean noise level

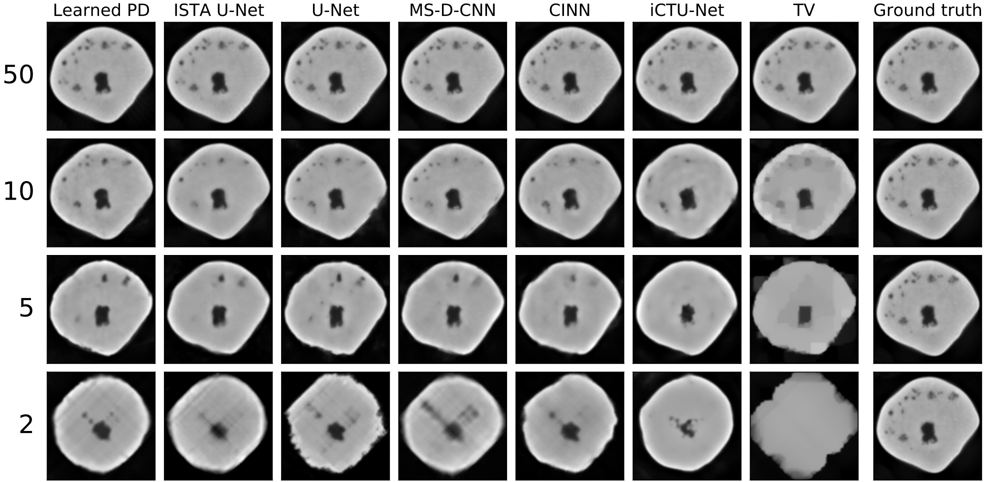


Note: lower data discrepancy translates to highest likelihood for each individual reconstruction, but discrepancies lower than the mean discrepancy w.r.t. the ground truth usually indicate overfitting to the noise.

Reconstruction on Apples-CT Dataset A: Noise-Free

Noise-Free	**PSNR**				**SSIM**			
Number of Angles	50	10	5	2	50	10	5	2
Learned Primal-Dual	38.72	35.85	30.79	22.00	0.901	0.870	0.827	0.740
ISTA U-Net	38.86	34.54	28.31	20.48	0.897	0.854	0.797	0.686
U-Net	39.62	33.51	27.77	19.78	0.913	0.803	0.803	0.676
MS-D-CNN	39.85	34.38	28.45	20.55	0.913	0.837	0.776	0.646
CINN	39.59	34.84	27.81	19.46	0.913	0.871	0.762	0.674
iCTU-Net	36.07	29.95	25.63	19.28	0.878	0.847	0.824	0.741
TV	39.27	29.00	22.04	15.95	0.915	0.783	0.607	0.661
CGLS	33.05	21.81	12.60	15.25	0.780	0.619	0.537	0.615
FBP	30.39	17.09	15.51	13.97	0.714	0.584	0.480	0.438

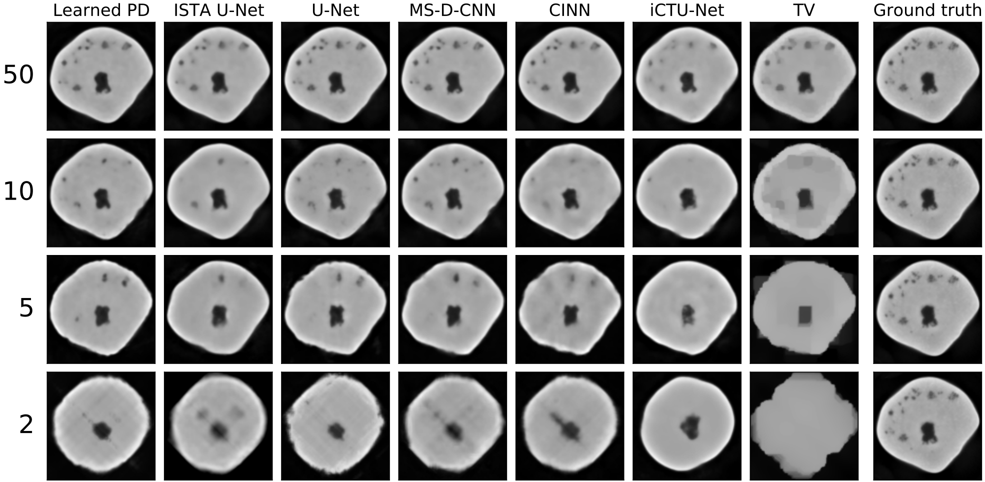
Reconstruction on Apples-CT Dataset A: Noise-free



Reconstruction on Apples-CT Dataset B: Gaussian noise

Gaussian Noise	PSNR				SSIM			
Number of Angles	50	10	5	2	50	10	5	2
Learned Primal-Dual	36.62	33.76	29.92	21.41	0.878	0.850	0.821	0.674
ISTA U-Net	36.04	33.55	28.48	20.71	0.871	0.851	0.811	0.690
U-Net	36.48	32.83	27.80	19.86	0.882	0.818	0.789	0.706
MS-D-CNN	36.67	33.20	27.98	19.88	0.883	0.831	0.748	0.633
CINN	36.77	31.88	26.57	19.99	0.888	0.771	0.722	0.637
iCTU-Net	32.90	29.76	24.67	19.44	0.848	0.837	0.801	0.747
TV	32.36	27.12	21.83	16.08	0.833	0.752	0.622	0.637
CGLS	27.36	21.09	14.90	15.11	0.767	0.624	0.553	0.616
FBP	27.88	17.09	15.51	13.97	0.695	0.583	0.480	0.438

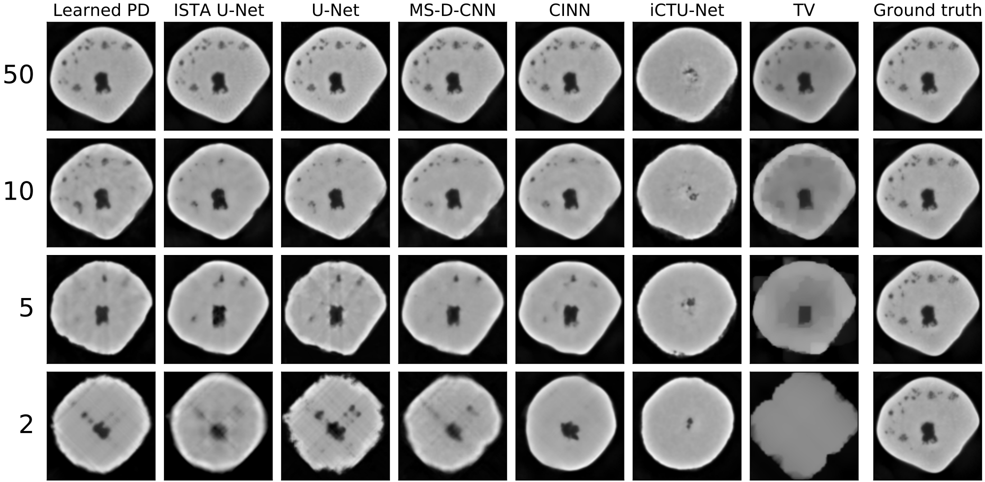
Reconstruction on Apples-CT Dataset B: Gaussian noise



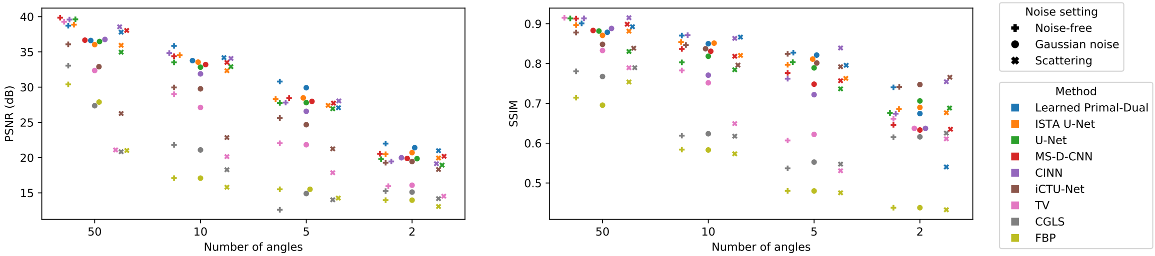
Reconstruction on Apples-CT Dataset C: Scattering

Scattering Noise	PSNR				SSIM			
Number of Angles	50	10	5	2	50	10	5	2
Learned Primal-Dual	37.80	34.19	27.08	20.98	0.892	0.866	0.796	0.540
ISTA U-Net	35.94	32.33	27.41	19.95	0.881	0.820	0.763	0.676
U-Net	34.96	32.91	26.93	18.94	0.830	0.784	0.736	0.688
MS-D-CNN	38.04	33.51	27.73	20.19	0.899	0.818	0.757	0.635
CINN	38.56	34.08	28.04	19.14	0.915	0.863	0.839	0.754
iCTU-Net	26.26	22.85	21.25	18.32	0.838	0.796	0.792	0.765
TV	21.09	20.14	17.86	14.53	0.789	0.649	0.531	0.611
CGLS	20.84	18.28	14.02	14.18	0.789	0.618	0.547	0.625
FBP	21.01	15.80	14.26	13.06	0.754	0.573	0.475	0.433

Reconstruction on Apples-CT Dataset C: Scattering

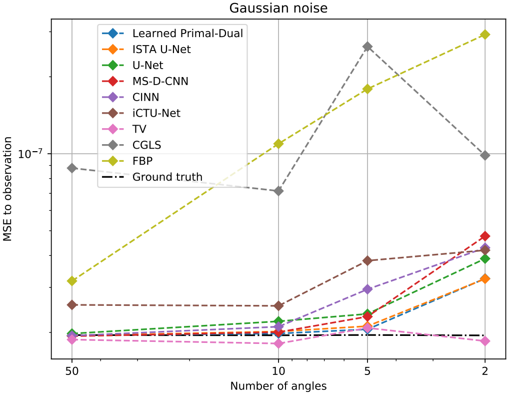


Reconstruction performance on Apples-CT Datasets



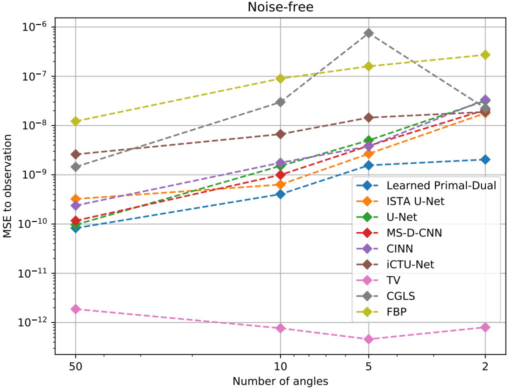
Data consistency on Apples-CT

- Undersampling: Reconstruction problem is ambiguous (can add any element from null-space of A)
- Discrepancy is suboptimally high for learned methods, increasing with fewer angles
- TV maintains suitable data discrepancy, but does not perform better in terms of PSNR and SSIM



Data consistency on Apples-CT

- Undersampling: Reconstruction problem is ambiguous (can add any element from null-space of A)
- Discrepancy is suboptimally high for learned methods, increasing with fewer angles
- TV maintains suitable data discrepancy, but does not perform better in terms of PSNR and SSIM



Discussion

- Most learned methods performed similarly well on LoDoPaB-CT (a similar observation has been reported from the fastMRI challenge [8])
  - Learned Primal-Dual (an unrolled iterative method) is among the best-performing methods
- Other important aspects:
  - Computational efficiency
  - Data requirements
  - Model knowledge (forward operator A, noise model, calibration, ...)
  - Target application

Computational requirements and reconstruction speed

- Learned methods: Ressource intensive training, but fast inference (reconstruction) (special case CINN: relies on sampling)
- DIP + TV: no standard training, but very slow due to “retraining” for each reconstruction
- Classical methods: iterative reconstruction can be relatively time-consuming (TV)

Transfer to 3D reconstruction

- Current benchmark only covers simple 2D parallel beam geometry
- 3D requires much more computational ressources
- Slice-wise application of 2D neural networks is probably suboptimal compared to methods directly targeting the 3D setting
- To transfer the compared methods to 3D, adaptations would be required to overcome ressource limitations (currently)
- Learned iterative schemes require efficient evaluation of A, post-processing slightly more flexible; iterative multi-scale approach [6] addresses scalability

[6] A. Hauptmann et al. "Multi-Scale Learned Iterative Reconstruction". In: *IEEE Transactions on Computational Imaging* 6 (2020), pp. 843–856

Feature summary

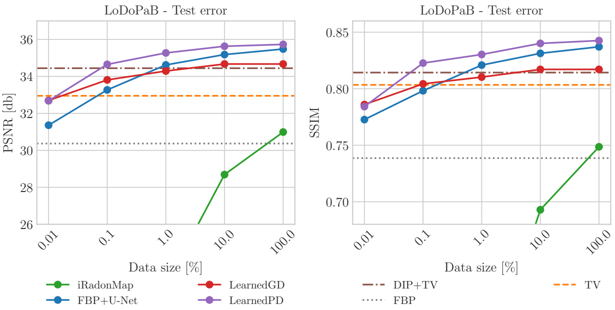
Model	Reconstruction Error (Image Metrics)		Training Time	Recon-Struction Time	GPU Memory	Learned Para-Meters	Uses $\mathcal{D}_Y$ Discre-Pancy	Operator Required
Learned P.-D.	⋆⋆	⋆	⋆⋆⋆⋆	⋆⋆	⋆⋆⋆⋆	⋆⋆	no	⋆⋆⋆
ISTA U-Net	⋆⋆	⋆	⋆⋆⋆	⋆⋆	⋆⋆⋆	⋆⋆⋆	no	⋆⋆
U-Net	⋆⋆	⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆	no	⋆⋆
MS-D-CNN	⋆⋆	⋆	⋆⋆⋆⋆	⋆⋆	⋆⋆	⋆	no	⋆⋆
U-Net++	⋆⋆	-	⋆⋆	⋆⋆	⋆⋆⋆	⋆⋆⋆	no	⋆⋆
CINN	⋆⋆	⋆	⋆⋆	⋆⋆⋆	⋆⋆⋆	⋆⋆⋆	no	⋆⋆
DIP + TV	⋆⋆⋆	-	-	⋆⋆⋆⋆	⋆⋆	3+	yes	⋆⋆⋆⋆
iCTU-Net	⋆⋆⋆	⋆⋆	⋆⋆	⋆⋆	⋆⋆⋆	⋆⋆⋆⋆	no	⋆
TV	⋆⋆⋆	⋆⋆⋆	-	⋆⋆⋆	⋆	3	yes	⋆⋆⋆⋆
CGLS	-	⋆⋆⋆⋆	-	⋆	⋆	1	yes	⋆⋆⋆⋆
FBP	⋆⋆⋆⋆	⋆⋆⋆⋆	-	⋆	⋆	2	no	⋆⋆⋆⋆
Legend	LoDoPaB	Apple CT	Rough values for Apple CT Dataset B					
	Avg. improv. over FBP		(varying for different setups and datasets)					
⋆⋆⋆⋆	0%	0–15%	>2 weeks	>10 min	>10 GiB	$>10^8$		Direct
⋆⋆⋆	12–16%	25–30%	>5 days	>30 s	>3 GiB	$>10^6$		In network
⋆⋆	17–20%	40–45%	>1 day	>0.1 s	>1.5 GiB	$>10^5$		For input
⋆		50–60%		$\leq$0.02 s	$\leq$1 GiB	$\leq 10^5$		Only concept

Generalization to other CT setups

- Current benchmark only covers simple 2D parallel beam geometry
- Standard geometries: helical fan-beam, cone-beam
- For changes in the scanning setup, retraining or transfer learning is required
- Learned methods can partially compensate for some model deviations when trained with suitable data
  - Example Apples-CT with scattering: classical methods would require manual model adaptation (or including an alternative way to learn the model deviation)

Number of training samples [3]

- Learned methods usually rely on large datasets
- Fully learned approaches require much more data, Learned Primal-Dual also works well with few training samples
- DIP+TV performs well in the low-data regime



[3] D. O. Baguer et al. "Computed tomography reconstruction using deep image prior and learned reconstruction methods". In: *Inverse Problems* 36.9 (Sept. 2020), p. 094004. URL: https://doi.org/10.1088%2F1361-6420%2Faba415

Requirements in target applications

- PSNR and SSIM do not fully represent reconstruction quality
- Different target applications require different reconstruction features, e.g.
  - Medical imaging:
    - TV-smoothed reconstructions to see overall organ shape
    - Detail-preserving reconstruction to see texture inside organs
  - Industrial CT:
    - Indicative reconstructions for a subsequent defect detection task
    - ...

Conclusions

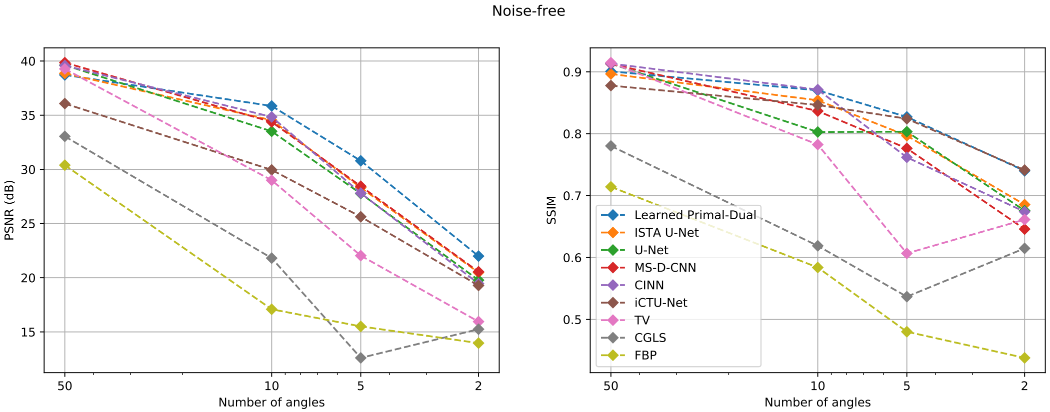
- Data-driven methods can improve reconstruction quality over classical methods
- Learned Primal-Dual and post-processing methods perform similarly well in a variety of settings
- Choice of reconstruction method depends on
  - Training data availability
  - Model knowledge
  - Target application
  - ...
- More methods to compare (e.g. learned regularization)!

Source code, network parameters and reconstructions are publicly available

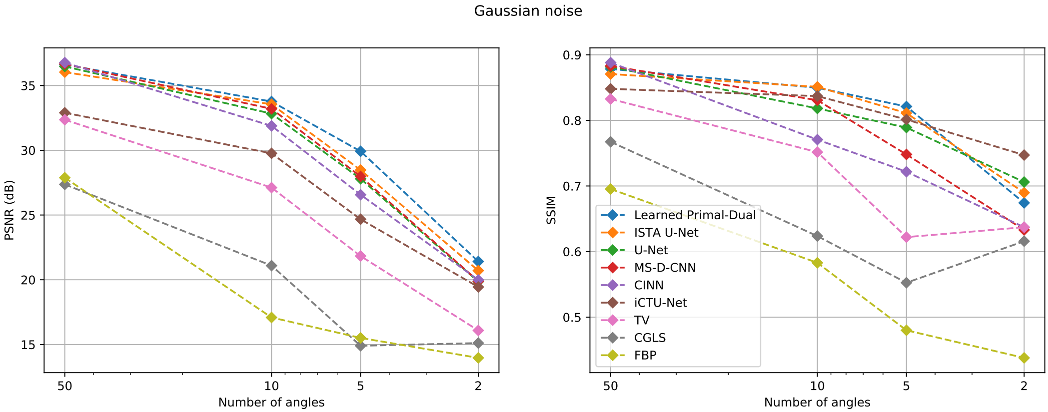
https://zenodo.org/record/4479816
https://zenodo.org/record/4460055
https://zenodo.org/record/4459962
https://zenodo.org/record/4459250

Thank you for your attention!

Reconstruction performance on Apples-CT Dataset A

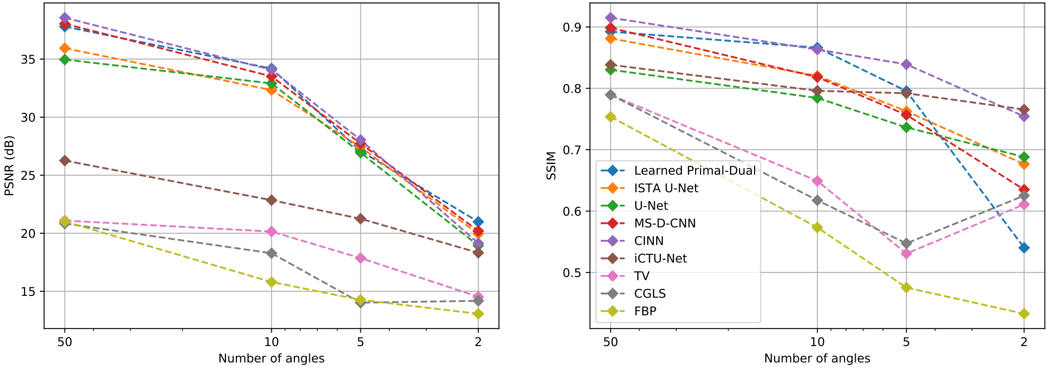


Reconstruction performance on Apples-CT Dataset B



Reconstruction performance on Apples-CT Dataset C

Scattering



References I

[1] J. Adler and O. Öktem. "Learned Primal-Dual Reconstruction". In: *IEEE Transactions on Medical Imaging* 37.6 (June 2018), pp. 1322–1332. URL: https://doi.org/10.1109/TMI.2018.2799231.

[2] S. G. Armato III, G. McLennan, L. Bidaut, et al. "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans". In: *Med. Phys.* 38.2 (Feb. 2011), pp. 915–931. ISSN: 0094-2405. URL: https://doi.org/10.1118/1.3528204.

References II

[3] D. O. Baguer, J. Leuschner, and M. Schmidt. “Computed tomography reconstruction using deep image prior and learned reconstruction methods”. In: *Inverse Problems* 36.9 (Sept. 2020), p. 094004. URL: https://doi.org/10.1088%2F1361-6420%2Faba415.

[4] S. B. Coban, V. Andriiashen, P. S. Ganguly, et al. *Parallel-beam X-ray CT datasets of apples with internal defects and label balancing for machine learning*. 2020. arXiv: 2012.13346 [cs.LG].

[5] A. Denker, M. Schmidt, J. Leuschner, et al. *Conditional Normalizing Flows for Low-Dose Computed Tomography Image Reconstruction*. 2020. arXiv: 2006.06270 [eess.IV].

References III

[6] A. Hauptmann, J. Adler, S. Arridge, et al. "Multi-Scale Learned Iterative Reconstruction". In: *IEEE Transactions on Computational Imaging* 6 (2020), pp. 843–856.

[7] K. H. Jin, M. T. McCann, E. Froustey, et al. "Deep convolutional neural network for inverse problems in imaging". In: *IEEE Transactions on Image Processing* 26.9 (2017), pp. 4509–4522. URL: https://doi.org/10.1109/TIP.2017.2713099.

[8] F. Knoll, T. Murrell, A. Sriram, et al. "Advancing machine learning for MR image reconstruction with an open competition: Overview of the 2019 fastMRI challenge". In: *Magnetic Resonance in Medicine* (2020).

References IV

[9] J. Leuschner, M. Schmidt, D. O. Baguer, et al. "LoDoPaB-CT, a benchmark dataset for low-dose computed tomography reconstruction". In: *Scientific Data* 8.1 (Apr. 2021), p. 109. ISSN: 2052-4463. URL: https://doi.org/10.1038/s41597-021-00893-z.

[10] J. Leuschner, M. Schmidt, and D. Erzmann. *Deep Inversion Validation Library*. https://github.com/jleuschn/dival. 2019.

[11] J. Leuschner, M. Schmidt, P. S. Ganguly, et al. "Quantitative Comparison of Deep Learning-Based Image Reconstruction Methods for Low-Dose and Sparse-Angle CT Applications". In: *Journal of Imaging* 7.3 (2021). ISSN: 2313-433X. URL: https://www.mdpi.com/2313-433X/7/3/44.

References V

[12] T. Liu, A. Chaman, D. Belius, et al. *Interpreting U-Nets via Task-Driven Multiscale Dictionary Learning*. 2020. arXiv: 2011.12815 [cs.CV].

[13] D. Pelt, J. Batenburg, and J. Sethian. "Improving tomographic reconstruction from limited data using Mixed-Scale Dense convolutional neural networks". In: *Journal of Imaging* 4.11 (Oct. 2018), pp. 128–128. URL: https://doi.org/10.3390/jimaging4110128.

[14] Z. Wang, A. C. Bovik, H. R. Sheikh, et al. "Image quality assessment: from error visibility to structural similarity". In: *IEEE Transactions on Image Processing* 13.4 (Apr. 2004), pp. 600–612. ISSN: 1057-7149.

References VI

[15] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, et al. "Unet++: A nested u-net architecture for medical image segmentation". In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, 2018, pp. 3–11. URL: https://doi.org/10.1007/978-3-030-00889-5_1.