CT Reconstruction with Deep Learning: Overview of Approaches and Extensions for DIP

Johannes Leuschner

joint work with
Riccardo Barbano, Javier Antorán, Maximilian Schmidt, Daniel Otero Baguer, Marco Nittscher, Michael Lameter, Andreas Hauptmann, Poulami Somanya Ganguly, Vladyslav Andriiashen, Sophia Bethany Coban, Alexander Denker, Dominik Bauer, Amir Hadjifaradji, Kees Joost Batenburg, Maureen van Eijnatten, Željko Kereta, José Miguel Hernández-Lobato, Bangti Jin and Peter Maaß

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

25 March 2023, 6th Young Scholar Symposium, Chinese University of Hong Kong

Overview

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

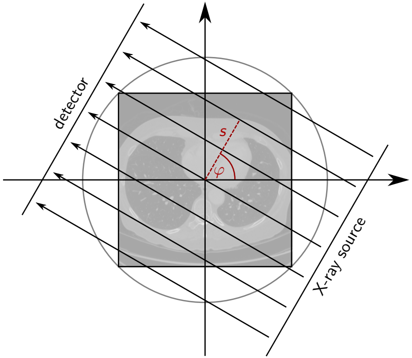


Figure: Parallel beam geometry

- 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:

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

DFG Universität Bremen

Mathematical formulation for parallel beam geometry

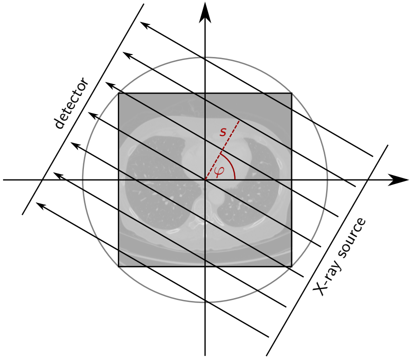


Figure: Parallel beam geometry

- 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:

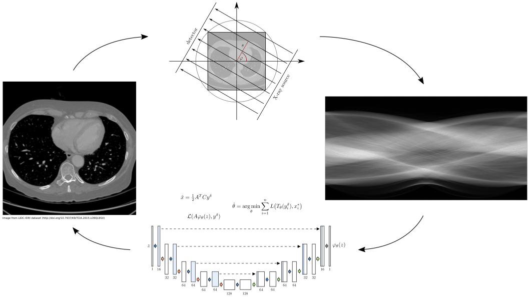
$$\mathcal{A}x(s,\varphi) = -\log\left(\frac{I_1(s,\varphi)}{I_0}\right)$$

discrete linear system $Ax^{\dagger} + \epsilon = y_{\delta}$

$A \in \mathbb{R}_{\geq 0}^{d_\varphi d_s \times d_n^2}, \quad x^{\dagger} \in \mathbb{R}_{\geq 0}^{d_n^2}, \quad y_\delta \in \mathbb{R}^{d_\varphi d_s}$

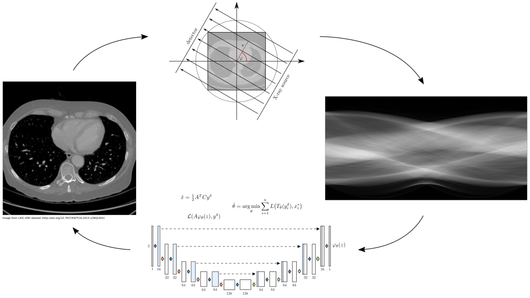
CT reconstruction

- Mildly ill-posed inverse problem: singular values tend to zero (at moderate speed), $\rightsquigarrow$ unstable inversion in the presence of noise
- Reconstruction approaches:
  - Analytical inversion formulas, e.g. filtered back-projection (FBP)
  - Iterative reconstruction
  - Deep-learning-based reconstruction



CT reconstruction

- Mildly ill-posed inverse problem: singular values tend to zero (at moderate speed), $\rightsquigarrow$ unstable inversion in the presence of noise
- Reconstruction approaches:
  - Analytical inversion formulas, e.g. filtered back-projection (FBP)
  - Iterative reconstruction
  - Deep-learning-based reconstruction

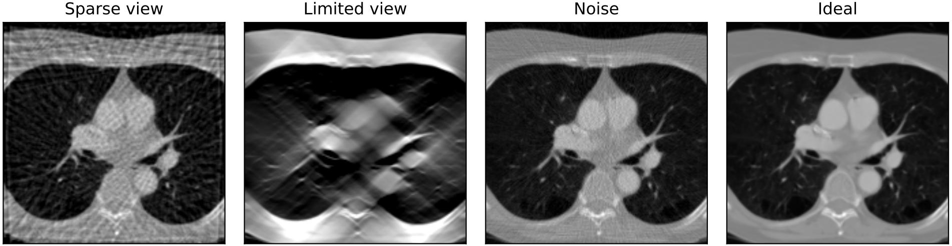


Reconstruction challenges

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

Goals

- Reduce potentially harmful radiation dose
- Reduce scanning time
- Meet technical limitations
- Reduce reconstruction errors (artifacts)



Categories of learned approaches

- Learned pre- and/or post-processing
- Learned iterative reconstruction
- Fully learned reconstruction
- Deep image prior
- etc.

(citations in the following are examples, usually early works of the kind)

Learned post-processing

after classical reconstruction

- dataset of pairs $(\tilde{x}_i, x_i^*)_{i=0,1,\ldots,N-1}$
  - $\tilde{x}_i$: preliminary reconstructions
  - x_i^*: ground truth images
- typical training by minimizing empirical risk [42, 6]:

$$R_{\text{emp}}(\theta) = \mathbb{E}_{(\tilde{x}_i, x_i^*)}\left[L(F_\theta(\tilde{x}_i), x_i^*)\right]$$

L: typically MSE, MAE, SSIM, perceptual loss (or mix of these)

- cGAN-based training with both L and discriminator-based adversarial loss, using a generative adversarial network (GAN) conditioned on the input $\tilde{x}_i$ [38, 39]
- few GAN-based works train on unpaired data $(\tilde{x}_i)_{i=0,1,\ldots,N-1}$, $(x_j^*)_{j=0,1,\ldots,M-1}$, while ensuring correspondence of the output via a loss based on $\tilde{x}_i$ [37, 31]

Learned post-processing

after classical reconstruction

- dataset of pairs $(\tilde{x}_i, x_i^*)_{i=0,1,\dots,N-1}$
  - $\tilde{x}_i$: preliminary reconstructions
  - x_i^*: ground truth images
- typical training by minimizing empirical risk [42, 6]:

$$R_{\text{emp}}(\theta) = \mathbb{E}_{(\tilde{x}_i, x_i^*)}\left[L(F_\theta(\tilde{x}_i), x_i^*)\right]$$

L: typically MSE, MAE, SSIM, perceptual loss (or mix of these)

- cGAN-based training with both L and discriminator-based adversarial loss, using a generative adversarial network (GAN) conditioned on the input $\tilde{x}_i$ [38, 39]
- few GAN-based works train on unpaired data $(\tilde{x}_i)_{i=0,1,\dots,N-1}$, $(x_j^*)_{j=0,1,\dots,M-1}$, while ensuring correspondence of the output via a loss based on $\tilde{x}_i$ [37, 31]

Learned post-processing
after classical reconstruction

- dataset of pairs $(\tilde{x}_i, x_i^*)_{i=0,1,\dots,N-1}$
  - $\tilde{x}_i$: preliminary reconstructions
  - x_i^*: ground truth images
- typical training by minimizing empirical risk [42, 6]:

$$R_{\mathrm{emp}}(\theta) = \mathbb{E}_{(\tilde{x}_i, x_i^*)}\left[L(F_\theta(\tilde{x}_i), x_i^*)\right]$$

L: typically MSE, MAE, SSIM, perceptual loss (or mix of these)

- cGAN-based training with both L and discriminator-based adversarial loss, using a generative adversarial network (GAN) conditioned on the input $\tilde{x}_i$ [38, 39]
- few GAN-based works train on unpaired data $(\tilde{x}_i)_{i=0,1,\dots,N-1}$, $(x_j^*)_{j=0,1,\dots,M-1}$, while ensuring correspondence of the output via a loss based on $\tilde{x}_i$ [37, 31]

Learned post-processing
after classical reconstruction

- dataset of pairs $(\tilde{x}_i, x_i^*)_{i=0,1,\dots,N-1}$
  - $\tilde{x}_i$: preliminary reconstructions
  - x_i^*: ground truth images
- typical training by minimizing empirical risk [42, 6]:

$$R_{\mathrm{emp}}(\theta) = \mathbb{E}_{(\tilde{x}_i, x_i^*)}\left[L(F_\theta(\tilde{x}_i), x_i^*)\right]$$

L: typically MSE, MAE, SSIM, perceptual loss (or mix of these)

- cGAN-based training with both L and discriminator-based adversarial loss, using a generative adversarial network (GAN) conditioned on the input $\tilde{x}_i$ [38, 39]
- few GAN-based works train on unpaired data $(\tilde{x}_i)_{i=0,1,\dots,N-1}$, $(x_j^*)_{j=0,1,\dots,M-1}$, while ensuring correspondence of the output via a loss based on $\tilde{x}_i$ [37, 31]

Learned pre-processing
before classical reconstruction

- dataset of pairs $(y_i^\delta, y_i^*)_{i=0,1,\ldots,N-1}$
  - y_i^δ: degraded measurements
  - y_i^*: reference measurements
- typical training by minimizing empirical risk [19, 14]:

$$R_{\mathrm{emp}}(\theta) = \mathbb{E}_{(y_i^\delta, y_i^*)}\left[L(F_\theta(y_i^\delta), y_i^*)\right]$$

L can include both sinogram- and image-domain losses $\|A^\dagger F_\theta(y_i^\delta) - A^\dagger y_i^*\|$

- cGAN-based training with both L and discriminator-based adversarial loss, using a generative adversarial network (GAN) conditioned on the input $\tilde{x}_i$ [13, 5]

Pre- and post-processing and be combined, e.g. learned end-to-end [41, 25]

Learned pre-processing
before classical reconstruction

- dataset of pairs $(y_i^\delta, y_i^*)_{i=0,1,\dots,N-1}$
  - y_i^δ: degraded measurements
  - y_i^*: reference measurements
- typical training by minimizing empirical risk [19, 14]:

$$R_{\mathrm{emp}}(\theta) = \mathbb{E}_{(y_i^\delta, y_i^*)}\left[L(F_\theta(y_i^\delta), y_i^*)\right]$$

L can include both sinogram- and image-domain losses $\|A^\dagger F_\theta(y_i^\delta) - A^\dagger y_i^*\|$

- cGAN-based training with both L and discriminator-based adversarial loss, using a generative adversarial network (GAN) conditioned on the input $\tilde{x}_i$ [13, 5]

Pre- and post-processing and be combined, e.g. learned end-to-end [41, 25]

Learned pre-processing

before classical reconstruction

- dataset of pairs $(y_i^\delta, y_i^*)_{i=0,1,\ldots,N-1}$
  - y_i^δ: degraded measurements
  - y_i^*: reference measurements
- typical training by minimizing empirical risk [19, 14]:

$$R_{\mathrm{emp}}(\theta) = \mathbb{E}_{(y_i^\delta, y_i^*)}\left[L(F_\theta(y_i^\delta), y_i^*)\right]$$

L can include both sinogram- and image-domain losses $\|A^\dagger F_\theta(y_i^\delta) - A^\dagger y_i^*\|$

- cGAN-based training with both L and discriminator-based adversarial loss, using a generative adversarial network (GAN) conditioned on the input $\tilde{x}_i$ [13, 5]

Pre- and post-processing and be combined, e.g. learned end-to-end [41, 25]

Learned pre-processing

before classical reconstruction

- dataset of pairs $(y_i^\delta, y_i^*)_{i=0,1,\dots,N-1}$
  - y_i^δ: degraded measurements
  - y_i^*: reference measurements
- typical training by minimizing empirical risk [19, 14]:

$$R_{\text{emp}}(\theta) = \mathbb{E}_{(y_i^\delta, y_i^*)}\left[L(F_\theta(y_i^\delta), y_i^*)\right]$$

L can include both sinogram- and image-domain losses $\|A^\dagger F_\theta(y_i^\delta) - A^\dagger y_i^*\|$

- cGAN-based training with both L and discriminator-based adversarial loss, using a generative adversarial network (GAN) conditioned on the input $\tilde{x}_i$ [13, 5]

Pre- and post-processing and be combined, e.g. learned end-to-end [41, 25]

Learned iterative reconstruction

- inspired by iterative reconstruction while introducing learned components
- several approaches:
  - unrolling a finite number of layers
    - end-to-end training [2, 1]
    - greedy iteration-wise training [24, 7]
  - Plug-and-Play priors / Regularization by denoising [40, 36]
  - iterative reconstruction with a learned regularization term [8, 28]
  - alternating a predefined and a relaxed "projection" network step [15, 34]
  - ...

Fully learned reconstruction

- using little operator knowledge
- learning to directly reconstruct an image x from measurements y^δ
- e.g. [23, 16]

Deep Image Prior (DIP)[20]

Forward Model

$$y_\delta = A x + \epsilon$$

DIP Reconstruction Framework

y_δ Measurement

z Network input image, usually i.i.d. noise

θ Network parameters, initialized randomly (default of the DL framework)

$$\theta^\star \in \operatorname*{argmin}_\theta \|A\,\varphi_\theta(z) - y_\delta\|^2 \qquad \textit{unsupervised, only requires } y_\delta$$

$\varphi_{\theta^\star}(z)$ Reconstruction (obtained with early stopping, “regularized” by architecture)

[20] V. Lempitsky et al. “Deep Image Prior”. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 2018, pp. 9446–9454

A benchmark of learned reconstruction methods

Journal paper

Johannes Leuschner, Maximilian Schmidt, Poulami Somanya Ganguly, Vladyslav Andriiashen, Sophia Bethany Coban, Alexander Denker, Dominik Bauer, Amir Hadjifaradji, Kees Joost Batenburg, Peter Maass and Maureen van Eijnatten:
Quantitative Comparison of Deep Learning-Based Image Reconstruction Methods for Low-Dose and Sparse-Angle CT Applications
(2021) Journal of Imaging, vol. 7, no. 3,
doi: 10.3390/jimaging7030044

Code: `https://github.com/jleuschn/dival`
`https://github.com/jleuschn/learned_ct_reco_comparison_paper`
Parameters: `https://github.com/jleuschn/supp.dival`,
`https://zenodo.org/record/4460055`

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 (large-scale simulated datasets with over 30 000 training pairs)

LoDoPaB-CT [21]	Apples-CT [9]
low-dose	sparse-angle

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 (large-scale simulated datasets with over 30 000 training pairs)

LoDoPaB-CT [21]	Apples-CT [9]
low-dose	sparse-angle

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[4] $\hat{\theta} = \min_\theta \|\mathcal{A}F_\theta(z) - y_\delta\|, \quad \hat{x} = F\hat{\theta}(z)$
- **Generative models**: Conditional INN [10] $\hat{x} = \frac{1}{n}\sum_i^n F_\theta\left(z_i, \mathrm{FBP}\left(y_\delta\right)\right), \quad z_i \sim \mathcal{N}(0, I)$
- **Postprocessing**: U-Net[17], U-Net++[43], ISTA U-Net[27], MS-D-CNN[32] $\hat{x} = F_\theta\left(\mathrm{FBP}\left(y_\delta\right)\right)$
- **Fully learned**: iCTU-Net[22] $\hat{x} = F_\theta\left(y_\delta\right)$

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[4] $\hat{\theta} = \min_\theta \|\mathcal{A}F_\theta(z) - y_\delta\|, \quad \hat{x} = F\hat{\theta}(z)$
- **Generative models**: Conditional INN [10] $\hat{x} = \frac{1}{n}\sum_i^n F_\theta\left(z_i, \mathrm{FBP}\left(y_\delta\right)\right), \quad z_i \sim \mathcal{N}(0, I)$
- **Postprocessing**: U-Net[17], U-Net++[43], ISTA U-Net[27], MS-D-CNN[32] $\hat{x} = F_\theta\left(\mathrm{FBP}\left(y_\delta\right)\right)$
- **Fully learned**: iCTU-Net[22] $\hat{x} = F_\theta\left(y_\delta\right)$

[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] $x^{[l+1]} = F_{\theta_l}(x^{[l]}, y_\delta),\ l = 1, \ldots, L,\ \hat{x} = x^{[L]}$
- **Unsupervised**: Deep Image Prior + TV[4] $\hat{\theta} = \min_\theta \|\mathcal{A}F_\theta(z) - y_\delta\|, \quad \hat{x} = F\hat{\theta}(z)$
- **Generative models**: Conditional INN [10] $\hat{x} = \frac{1}{n}\sum_i^n F_\theta\left(z_i, \mathrm{FBP}\left(y_\delta\right)\right), \quad z_i \sim \mathcal{N}(0, I)$
- **Postprocessing**: U-Net[17], U-Net++[43], ISTA U-Net[27], MS-D-CNN[32] $\hat{x} = F_\theta\left(\mathrm{FBP}\left(y_\delta\right)\right)$
- **Fully learned**: iCTU-Net[22] $\hat{x} = F_\theta\left(y_\delta\right)$

[4] 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[4] $\hat{\theta} = \min_\theta \|\mathcal{A}F_\theta(z) - y_\delta\|, \quad \hat{x} = F\hat{\theta}(z)$
- **Generative models**: Conditional INN[10] $\hat{x} = \frac{1}{n}\sum_i^n F_\theta\left(z_i, \mathrm{FBP}\left(y_\delta\right)\right), \quad z_i \sim \mathcal{N}(0, I)$
- **Postprocessing**: U-Net[17], U-Net++[43], ISTA U-Net[27], MS-D-CNN[32] $\hat{x} = F_\theta\left(\mathrm{FBP}\left(y_\delta\right)\right)$
- **Fully learned**: iCTU-Net[22] $\hat{x} = F_\theta\left(y_\delta\right)$

[10] 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] $x^{[l+1]} = F_{\theta_l}(x^{[l]}, y_\delta),\ l = 1, \dots, L,\ \hat{x} = x^{[L]}$
- **Unsupervised**: Deep Image Prior + TV[4] $\hat{\theta} = \min_\theta \|\mathcal{A}F_\theta(z) - y_\delta\|, \quad \hat{x} = F\hat{\theta}(z)$
- **Generative models**: Conditional INN [10] $\hat{x} = \frac{1}{n}\sum_i^n F_\theta\left(z_i, \text{FBP}\left(y_\delta\right)\right), \quad z_i \sim \mathcal{N}(0, I)$
- **Postprocessing**: U-Net[17], U-Net++[43], ISTA U-Net[27], MS-D-CNN[32] $\hat{x} = F_\theta\left(\text{FBP}\left(y_\delta\right)\right)$
- **Fully learned**: iCTU-Net[22] $\hat{x} = F_\theta\left(y_\delta\right)$

[17] Jin et al., 2017, "Deep convolutional neural network for inverse problems in imaging"
[43] Zhou et al., 2018, "Unet++: A nested u-net architecture for medical image segmentation"
[27] Liu et al., 2020, "Interpreting U-Nets via Task-Driven Multiscale Dictionary Learning"
[32] Pelt et al., 2018, "Improving Tomographic Reconstruction from Limited Data Using Mixed-Scale Dense Convolutional Neural Networks"

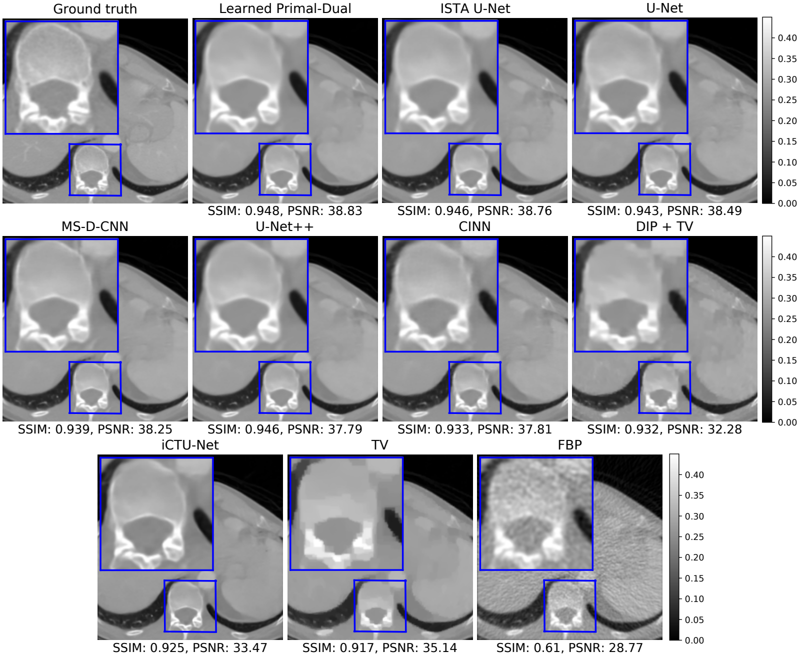
Included Methods - From Modeling to Data-driven

- **Classical reconstruction**: Filtered back-projection (FBP), Total variation (TV), CGLS
- **Learned iterative schemes**: Learned Primal-Dual[1] $x^{[l+1]} = F_{\theta_l}(x^{[l]}, y_\delta),\ l = 1, \dots, L,\ \hat{x} = x^{[L]}$
- **Unsupervised**: Deep Image Prior + TV[4] $\hat{\theta} = \min_\theta \|\mathcal{A}F_\theta(z) - y_\delta\|, \quad \hat{x} = F\hat{\theta}(z)$
- **Generative models**: Conditional INN [10] $\hat{x} = \frac{1}{n}\sum_i^n F_\theta\left(z_i, \mathrm{FBP}\left(y_\delta\right)\right), \quad z_i \sim \mathcal{N}(0, I)$
- **Postprocessing**: U-Net[17], U-Net++[43], ISTA U-Net[27], MS-D-CNN[32] $\hat{x} = F_\theta\left(\mathrm{FBP}\left(y_\delta\right)\right)$
- **Fully learned**: iCTU-Net[22] $\hat{x} = F_\theta\left(y_\delta\right)$

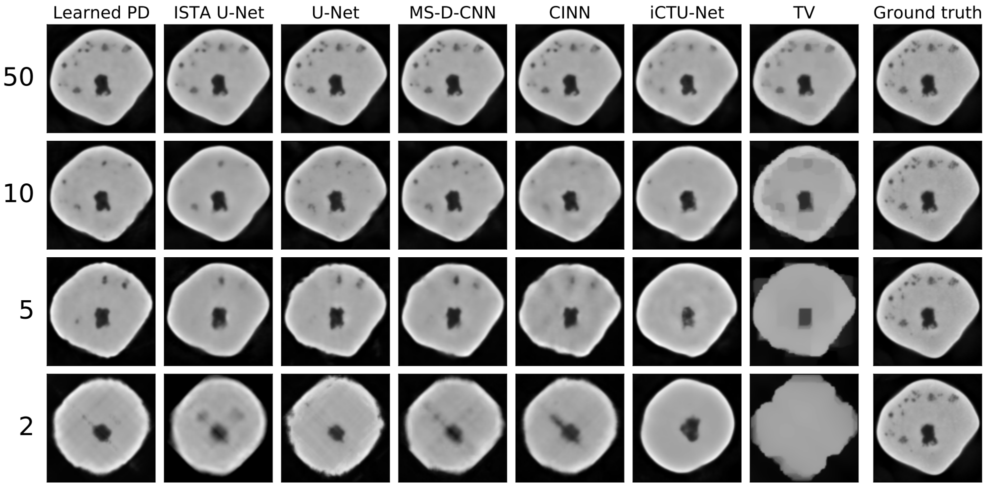
[22] Leuschner et al., 2021, "Quantitative Comparison of Deep Learning-Based Image Reconstruction Methods for Low-Dose and Sparse-Angle CT Applications"

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

Sparse-view reconstruction on Apples-CT with Gaussian noise



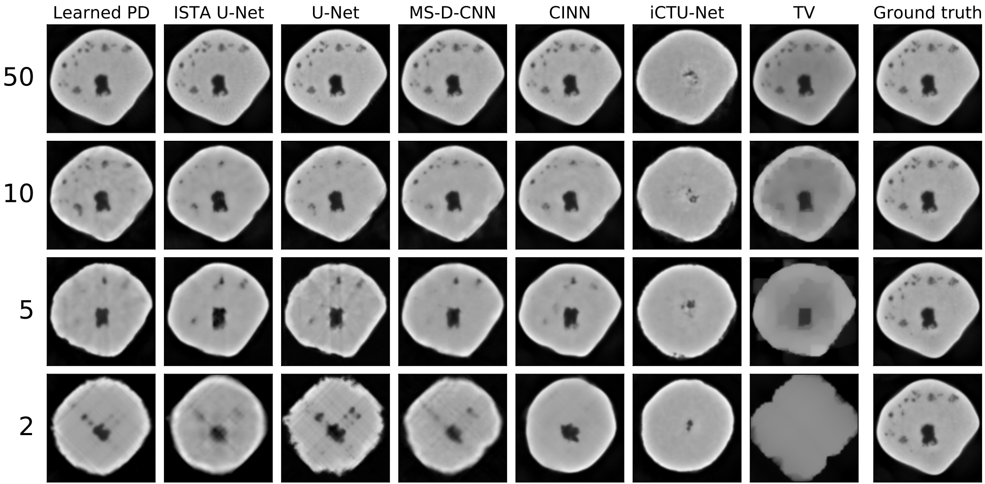
Sparse-view reconstruction on Apples-CT with Gaussian noise

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

Sparse-view reconstruction on Apples-CT with scattering

	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

Sparse-view reconstruction on Apples-CT with scattering

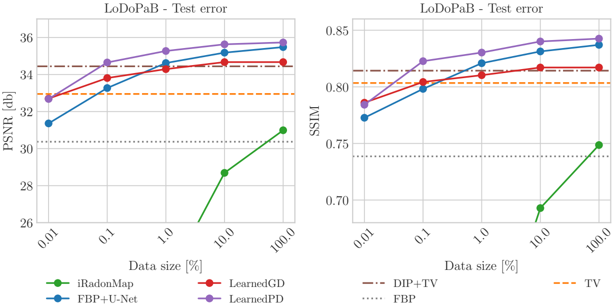


Discussion

- Most learned methods performed similarly well on LoDoPaB-CT (a similar observation has been reported from the fastMRI challenge [18])
  - Learned Primal-Dual (an unrolled iterative method) is among the best-performing methods
- Other important aspects:
  - Data requirements
  - Computational efficiency (and scalability to 3D)
  - Model knowledge (forward operator A, noise model, calibration, ...)
  - Target application

Number of training samples[4]

- 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



[4] 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

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 Avg. improv. over FBP	Apple CT	Rough values for Apple CT Dataset B (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

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
    - ...

Educated Deep Image Prior

Journal paper

Riccardo Barbano, Johannes Leuschner, Maximilian Schmidt, Alexander Denker, Andreas Hauptmann, Peter Maass and Bangti Jin:

An Educated Warm Start For Deep Image Prior-Based Micro CT Reconstruction

(2022) IEEE Trans. Comput. Imaging, vol. 8, pp. 1210-1222, doi: 10.1109/TCI.2022.3233188, arXiv:2111.11926

Code: `https://github.com/educating-dip/educated_deep_image_prior`
Experiments: `https://zenodo.org/record/7234749`

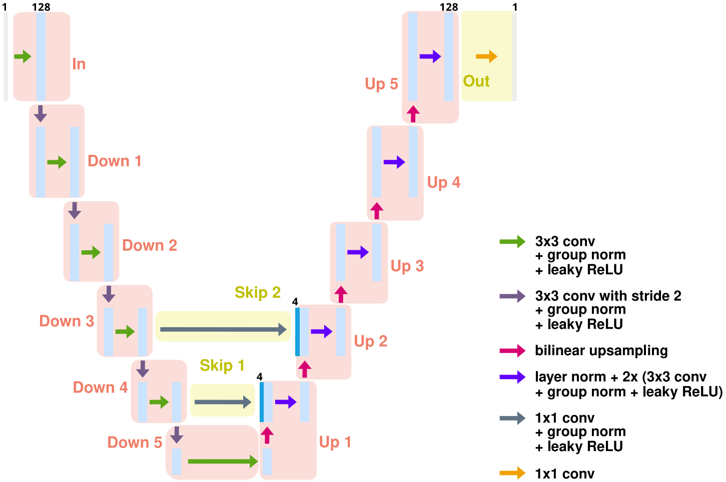
DIP: Features and limitations

$$\theta^\star \in \underset{\theta}{\operatorname{argmin}} \|A\,\varphi_\theta(z) - y_\delta\|^2 + \gamma \mathrm{TV}(\varphi_\theta(z))$$

+ Unsupervised learning: no training data needed, just the measurement y_δ
+ Loss motivated by classical variational formulation
○ Good reconstruction quality, $>$ classical, $<$ learned on in-distribution data
– Computationally expensive: re-“training” for every reconstruction
– Need to identify point for early-stopping

DIP Architecture

φ_θ: typically a CNN,
e.g. U-Net [33]



An Educated Warm Start For DIP-Based µCT Reconstruction

- Can DIP be accelerated by pretraining on synthetic data?
- How does pretraining impact a subsequent unsupervised DIP reconstruction?

Proposed approach (Educated Deep Image Prior, EDIP):

1. **Supervised pretraining (simulated data)**

$$\theta_{\mathrm{s}}^{\star} \in \operatorname*{argmin}_{\theta} \left\{ l_{\mathrm{s}}(\theta) := \frac{1}{N} \sum_{(x^n, y_\delta^n) \in \mathcal{D}} \|\varphi_\theta(A^\dagger y_\delta^n) - x^n\|_2^2 \right\}$$

2. **Unsupervised reconstruction (real-measured data)**

$$\text{Init. } \theta \leftarrow \theta_{\mathrm{s}}^{\star}, \quad \theta_t^{\star} \in \operatorname*{argmin}_{\theta} \left\{ l_{\mathrm{t}}(\theta) := \|A\,\varphi_\theta(z) - y_\delta\|_2^2 + \gamma\,\mathrm{TV}\left(\varphi_\theta(z)\right) \right\},$$

$$x^\star = \varphi_{\theta_{\mathrm{t}}^\star}(z)$$

An Educated Warm Start For DIP-Based µCT Reconstruction

- Can DIP be accelerated by pretraining on synthetic data?
- How does pretraining impact a subsequent unsupervised DIP reconstruction?

Proposed approach (Educated Deep Image Prior, EDIP):

1. **Supervised pretraining (simulated data)**

$$\theta_{\mathrm{s}}^{\star} \in \operatorname*{argmin}_{\theta} \left\{ l_{\mathrm{s}}(\theta) := \frac{1}{N} \sum_{(x^n, y_\delta^n) \in \mathcal{D}} \| \varphi_\theta(A^\dagger y_\delta^n) - x^n \|_2^2 \right\}$$

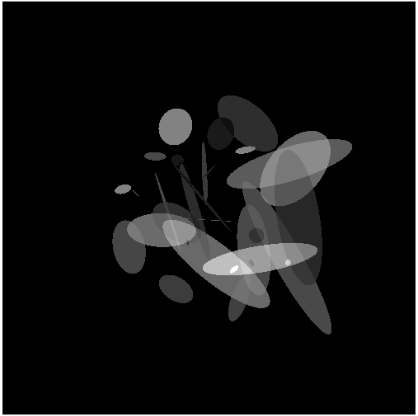
2. **Unsupervised reconstruction (real-measured data)**

$$\text{Init. } \theta \leftarrow \theta_{\mathrm{s}}^{\star}, \quad \theta_t^{\star} \in \operatorname*{argmin}_{\theta} \left\{ l_{\mathrm{t}}(\theta) := \| A \varphi_\theta(A^\dagger y_\delta) - y_\delta \|_2^2 + \gamma \, \mathrm{TV}\left(\varphi_\theta(A^\dagger y_\delta) \right) \right\},$$

$$x^\star = \varphi_{\theta_{\mathrm{t}}^\star}(A^\dagger y_\delta)$$

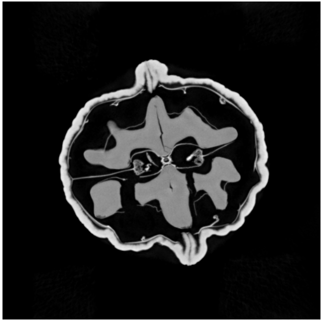
Synthetic Training Dataset of Ellipses/Ellipsoids

- images of ellipses with random position, shape, rotation and intensity values, generated "on-the-fly"
- simulated measurements
  - geometry of target reconstruction task
  - 5 % white noise



X-Ray Walnut Dataset[11]

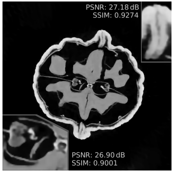
- cone-beam µCT measurements using 3 source positions
- 1200 equidistant angles over $[0, 360^\circ)$
- reduce geometry to 2D volume slice, selecting a subset of measurement pixels
- assemble forward operator as a sparse matrix for image resolution $(501\,\mathrm{px})^2$ from given geometry
- sparse-view task: reconstruct from 120 angles ($10\times$ subsampling)
- ground truth publicly available



[11] H. Der Sarkissian et al. *Cone-Beam X-Ray CT Data Collection Designed for Machine Learning: Samples 1-8*. *Zenodo*. 2019. URL: https://doi.org/10.5281/zenodo.2686726

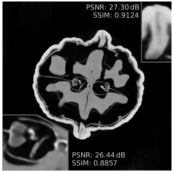
Walnut Reconstruction

EDIP (FBP)



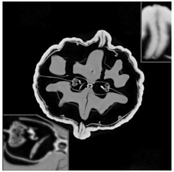
PSNR: 33.31 dB, SSIM: 0.7748

DIP (noise)

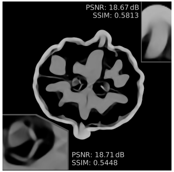


PSNR: 33.92 dB, SSIM: 0.7686

Ground truth

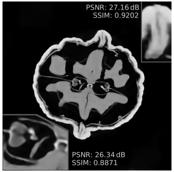


EDIP (FBP) initial



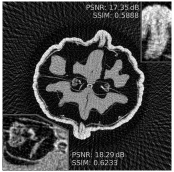
PSNR: 25.67 dB, SSIM: 0.7584

EDIP (FBP) iter. 4500



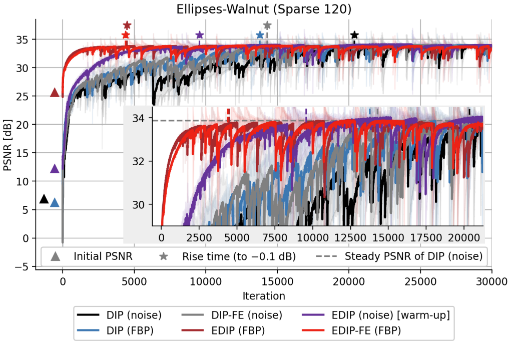
PSNR: 33.68 dB, SSIM: 0.8170

FBP



PSNR: 16.21 dB, SSIM: 0.1398

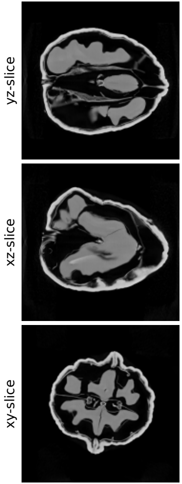
Videos



Walnut in 3D

- 3D geometry:
  - reduce image dimension from $(501\,\mathrm{px})^3$ to $(167\,\mathrm{px})^3$ and sub-sample projections by a factor of 3
  - use approximate adjoint via back-projection (cannot assemble sparse matrix)
  - sparse-view task: reconstruct from 20 angles ($60\times$ subsampling)
- adapted architecture:
  - reduce channels in encoder
  - add conv. layers before output

24 GB VRAM constraint

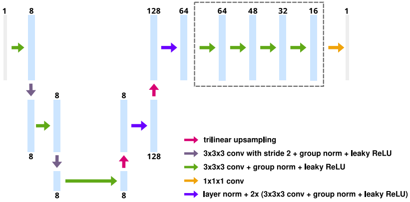


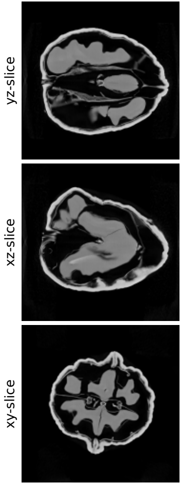
Walnut in 3D

- 3D geometry:
  - reduce image dimension from $(501\,\text{px})^3$ to $(167\,\text{px})^3$ and sub-sample projections by a factor of 3
  - use approximate adjoint via back-projection (cannot assemble sparse matrix)
  - sparse-view task: reconstruct from 20 angles ($60\times$ subsampling)

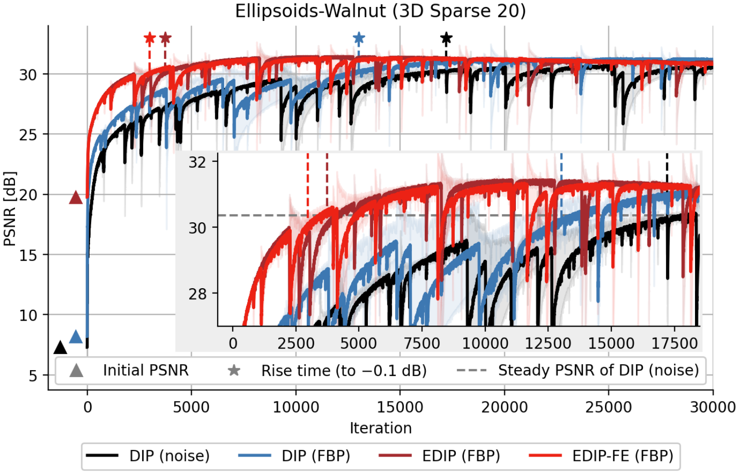
- adapted architecture:
  - reduce channels in encoder
  - add conv. layers before output

24 GB VRAM constraint

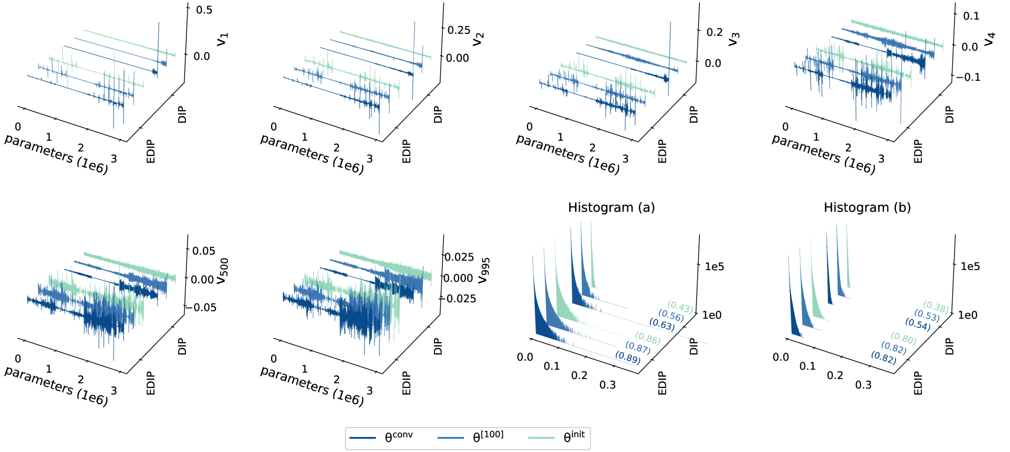




Walnut 3D Reconstruction



Singular Value Analysis



SVD-DIP: optimization by singular value fine-tuning

Conference paper

Marco Nittscher, Michael Falk Lameter, Riccardo Barbano, Johannes Leuschner, Bangti Jin, Peter Maass:
SVD-DIP: Overcoming the Overfitting Problem in DIP-based CT Reconstruction
(2023) will be presented at MIDL 2023 Conference

Code: `https://github.com/anonsvddip/svd_dip`

Singular value fine-tuning for few-shot segmentation [35]

- Adapt pretrained network for segmentation on unseen classes with few samples
- Simply continuing training on few samples risks overfitting
- Classical paradigm: freeze backbone parameters
- *Idea: compute SVD of pretrained backbone parameters and only fine-tune SVs*

Can we adopt this idea for the pretrained EDIP?

[35] Y. Sun et al. "Singular Value Fine-tuning: Few-shot Segmentation requires Few-parameters Fine-tuning". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh et al. 2022. URL: https://openreview.net/forum?id=LEqYZz7cZOI

Singular value fine-tuning for few-shot segmentation [35]

- Adapt pretrained network for segmentation on unseen classes with few samples
- Simply continuing training on few samples risks overfitting
- Classical paradigm: freeze backbone parameters
- *Idea: compute SVD of pretrained backbone parameters and only fine-tune SVs*

Can we adopt this idea for the pretrained EDIP?

[35] Y. Sun et al. "Singular Value Fine-tuning: Few-shot Segmentation requires Few-parameters Fine-tuning". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh et al. 2022. URL: https://openreview.net/forum?id=LEqYZz7cZOI

Singular value fine-tuning

SVD decomposition of 2D convolutional weights following [35]:

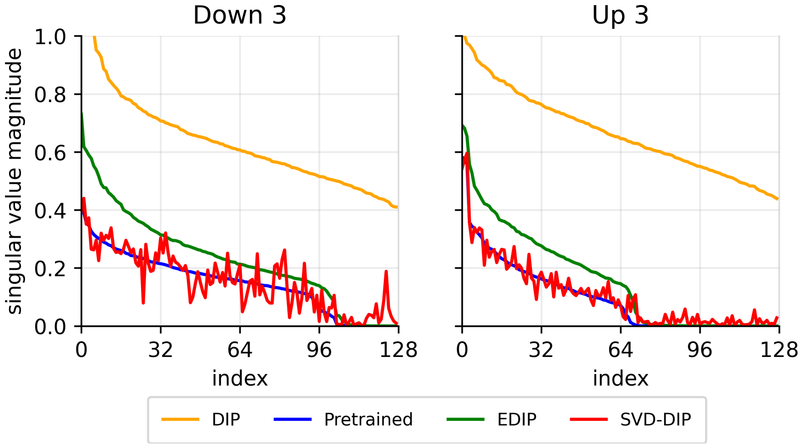
$$W \in \mathbb{R}^{C_{\text{out}} \times C_{\text{in}} \times K \times K} \rightsquigarrow W' \in \mathbb{R}^{C_{\text{out}} \times C_{\text{in}} K^2}$$
$$W' = U'S'V' \quad \text{with} \quad U' \in \mathbb{R}^{C_{\text{out}} \times R}, S' \in \mathbb{R}^{R \times R}, V' \in \mathbb{R}^{R \times C_{\text{in}} K^2}$$

with S' diagonal, $R = \min\{C_{\text{out}}, C_{\text{in}} K^2\}$

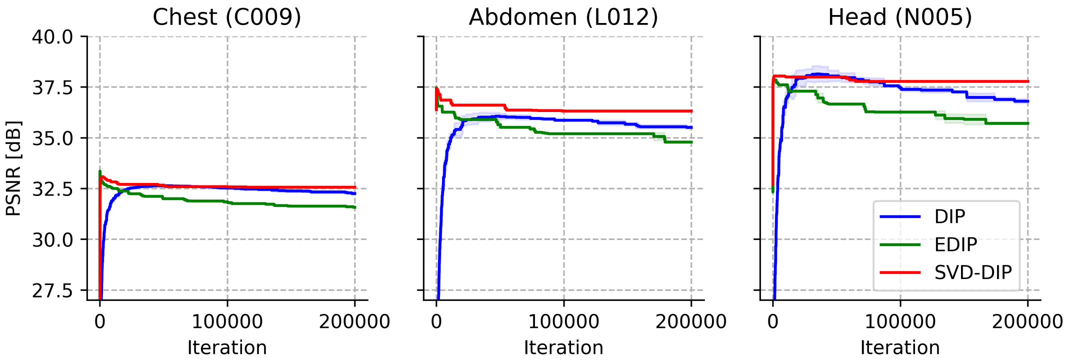
Implementation of 2D convolution with $W' = U'S'V'$:
reshape U' and V' as weight tensors, in between multiply channel-wise with S'

- Compute SVD of pretrained DIP parameters and only fine-tune SVs

SVs on Lotus data of two layers

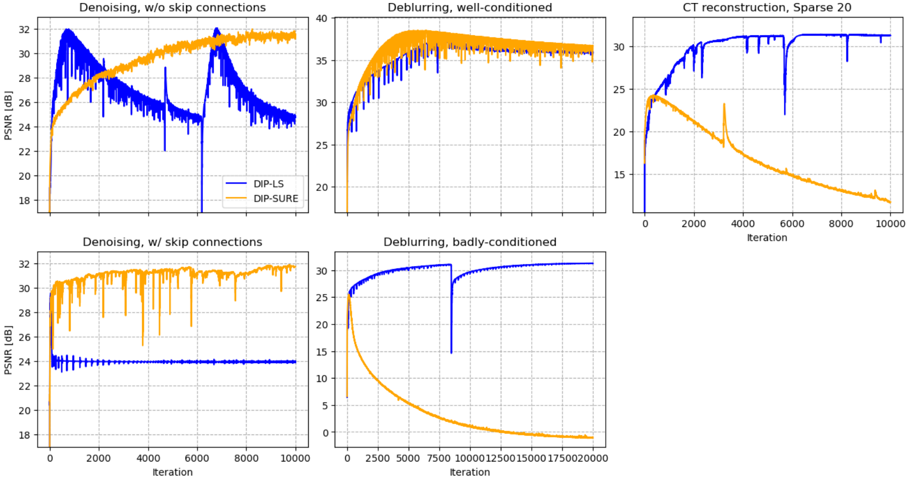


Results on Mayo data



- pretrained on LoDoPaB-200 data (using different FBP filter)

Stein's unbiased risk estimator (SURE) loss to prevent overfitting?



Subspace DIP: optimization on linear parameter manifolds

Preprint

Riccardo Barbano, Javier Antorán, Johannes Leuschner,
José Miguel Hernández-Lobato, Željko Kereta and Bangti Jin:
Fast and Painless Image Reconstruction in Deep Image Prior Subspaces
(2023) arXiv:2302.10279

Code: `https://github.com/anonsubdip/subspace_dip`

Subspace DIP

- DIP parameterizes the image as $x(\theta) = \varphi_\theta(z)$, $\theta \in \mathbb{R}^{d_\theta}$, with large d_θ
- EDIP starts optimization with pretrained parameters θ_{pre}
- Subspace DIP: restrict θ to a low-dimensional affine linear subspace around θ_{pre},

$$\theta(c) = \theta_{\text{pre}} + \sum_{k=1}^{d_{\text{sub}}} B_{:,k}\, c_k = \theta_{\text{pre}} + B\, c,$$

with a basis $B \in \mathbb{R}^{d_\theta \times d_{\text{sub}}}$, coefficients $c \in \mathbb{R}^{d_{\text{sub}}}$, and relatively small d_{sub}

Determining the sparse subspace

- save d_{pre} checkpoints during EDIP pretraining, $\Theta_{\text{pre}} \in \mathbb{R}^{d_\theta \times d_{\text{pre}}}$
- compute truncated top-d_{sub} SVD, $\Theta_{\text{pre}} \approx U\,S\,V^\top$, to obtain $U \in \mathbb{R}^{d_\theta \times d_{\text{sub}}}$
- apply a masking M that sparsifies U in the d_θ dimension by only keeping d_{lev} rows with the highest ℓ_2 norm ("leverage score" [12])

The resulting sparse basis

$$B := M\,U$$

is used to form the subspace

$$\theta(c) = \theta_{\text{pre}} + B\,c.$$

Optimization opportunities

- DIP optimization uses batch size 1, so the gradient descent is not stochastic.
- The low-dimensionality of $c \in \mathbb{R}^{d_{\text{sub}}}$ allows to construct the local curvature needed for second order optimization.

We compare

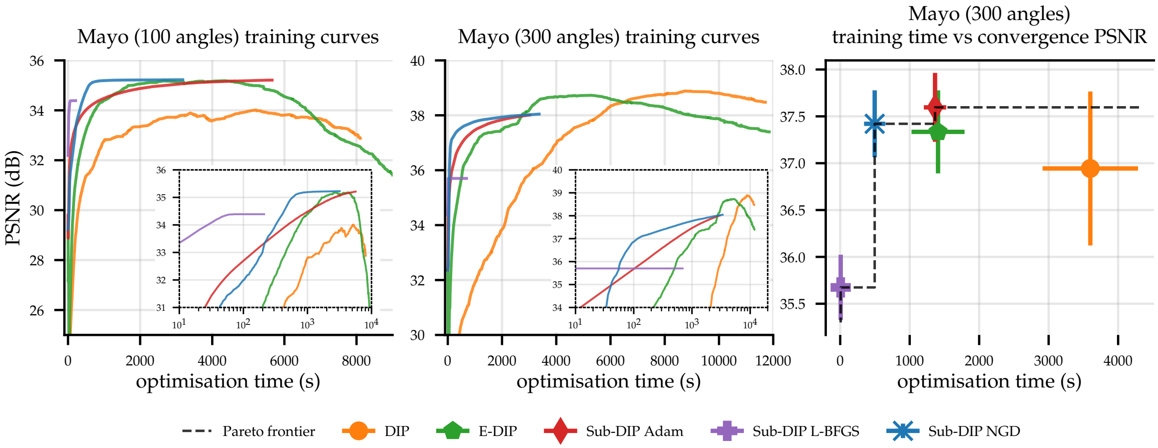
- Adam
- L-BFGS [26] – based on Hessian approximation
- NGD [3, 29] – based on Fisher information approximation

Natural gradient descent (NGD) [3]

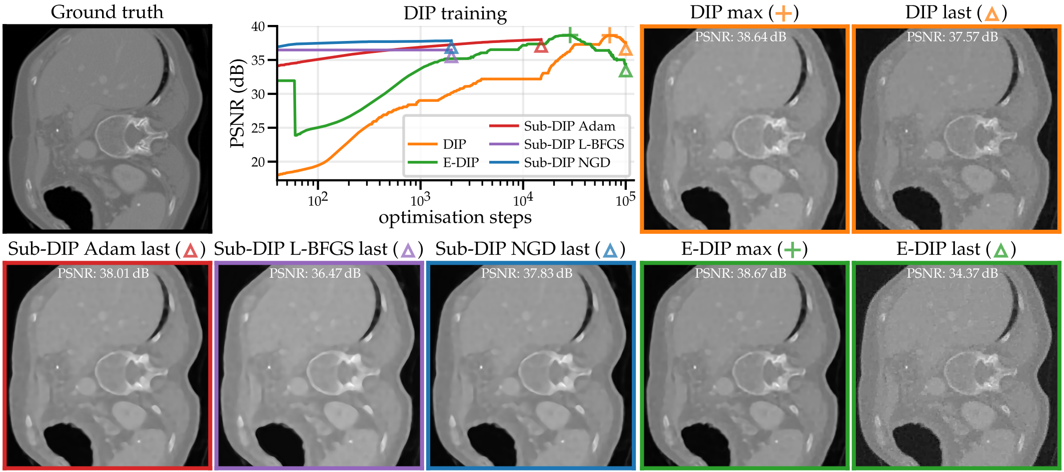
- NGD: $c^{[t+1]} = c^{[t]} - \alpha^{[t]} \tilde{F}^{-1}(c^{[t]}) \nabla L(c^{[t]})$
- $\tilde{F}^{-1}$ inverse Fisher information matrix, L loss function
- excluding the TV regularizer term, which does not depend on the observations, we have $\tilde{F}(c) = (AJ_{\varphi}B)^{\top} AJ_{\varphi}B$
- use online approximation of the Fisher using stochastic estimate
- include damping and scaling parameter for quadratic loss term
- choose momentum and step size via exact Fisher

[3] S.-i. Amari. "Natural Gradient Works Efficiently in Learning". In: *Neural Computation* 10.2 (Feb. 1998), pp. 251–276. ISSN: 0899-7667. eprint: https://direct.mit.edu/neco/article-pdf/10/2/251/813415/089976698300017746.pdf. URL: https://doi.org/10.1162/089976698300017746

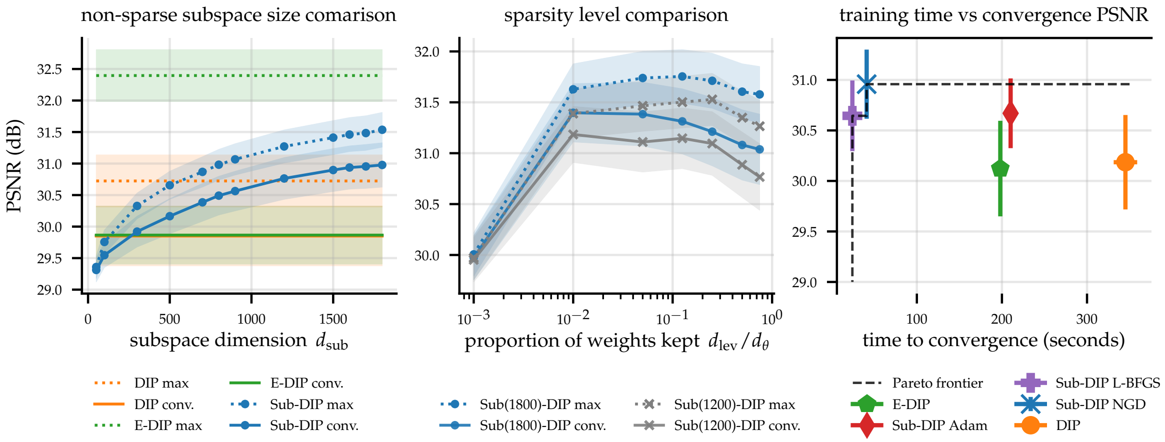
Comparison on Mayo data (using loss-based early stopping, averaged over 10 images)



Comparison on Mayo data (300 angles)



Experiments on CartoonSet



Thank you for your attention!
Comments or questions?

NGD: Quadratic loss

To compute the step $c^{[t+1]} = c^{[t]} + \delta$, consider the quadratic model of the loss

$$M^{[t]}(\delta) = L(c^{[t]}) + \left(\nabla_c L(c^{[t]})\right)^\top \delta + \frac{s}{2}\delta^\top \left(\lambda I_{d_{\text{sub}}} + \tilde{F}(c^{[t]})\right)\delta,$$

which due to $\tilde{F}$ is not a Taylor expansion, but a (maybe poor) convex approximation.

Two parameters introduced:

- $\lambda > 0$: damping, avoids numerical instabilities, adds isotropic curvature
- $s \in (0, 1]$: scaling, can reduce effect of curvature, allows for larger steps (novel)

λ, s updated by Levenberg-Marquardt style rule, clamped with minimum values

NGD: Add momentum [30]

$$\delta = \alpha^{[t]} \Delta^{[t]} + \mu^{[t]} \delta_0$$

- with $\Delta^{[t]} = -\left(F^{[t]}\right)^{-1} \nabla_c L(c^{[t]})$ and previous update direction δ_0
- choose $\alpha^{[t]}, \mu^{[t]}$ to minimize $M^{[t]}(\delta)$ via 2D linear system using the *exact* Fisher information matrix via matrix-vector multiplication

[30] J. Martens et al. "Optimizing Neural Networks with Kronecker-factored Approximate Curvature". In: *Proceedings of the 32nd International Conference on Machine Learning*. Ed. by F. Bach et al. Vol. 37. Proceedings of Machine Learning Research. Lille, France: PMLR, July 2015, pp. 2408–2417. URL: https://proceedings.mlr.press/v37/martens15.html

Bibliography I

[1] J. Adler and O. Öktem. "Learned Primal-Dual Reconstruction". In: *IEEE Transactions on Medical Imaging* 37.6 (2018), pp. 1322–1332.

[2] J. Adler and O. Öktem. "Solving ill-posed inverse problems using iterative deep neural networks". In: *Inverse Problems* 33.12 (Nov. 2017), p. 124007. URL: https://dx.doi.org/10.1088/1361-6420/aa9581.

[3] S.-i. Amari. "Natural Gradient Works Efficiently in Learning". In: *Neural Computation* 10.2 (Feb. 1998), pp. 251–276. ISSN: 0899-7667. eprint: https://direct.mit.edu/neco/article-pdf/10/2/251/813415/089976698300017746.pdf. URL: https://doi.org/10.1162/089976698300017746.

Bibliography II

[4] 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.

[5] J. Bai, X. Dai, Q. Wu, et al. "Limited-view CT Reconstruction Based on Autoencoder-like Generative Adversarial Networks with Joint Loss". In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2018, pp. 5570–5574.

Bibliography III

[6] H. Chen, Y. Zhang, W. Zhang, et al. "Low-dose CT denoising with convolutional neural network". In: *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. 2017, pp. 143–146.

[7] I. Y. Chun, Z. Huang, H. Lim, et al. "Momentum-Net: Fast and convergent iterative neural network for inverse problems". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2020), pp. 1–1.

[8] I. Y. Chun, X. Zheng, Y. Long, et al. "Sparse-view X-ray CT reconstruction using ℓ_1 regularization with learned sparsifying transform". In: *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med.* 2017, pp. 115–9.

Bibliography IV

[9] 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].

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

[11] H. Der Sarkissian, F. Lucka, M. van Eijnatten, et al. *Cone-Beam X-Ray CT Data Collection Designed for Machine Learning: Samples 1-8*. *Zenodo*. 2019. URL: https://doi.org/10.5281/zenodo.2686726.

Bibliography V

[12] P. Drineas, M. Magdon-Ismail, M. W. Mahoney, et al. "Fast Approximation of Matrix Coherence and Statistical Leverage". In: *Journal of Machine Learning Research* 13.111 (2012), pp. 3475–3506. URL: http://jmlr.org/papers/v13/drineas12a.html.

[13] M. U. Ghani and W. C. Karl. "Deep Learning-Based Sinogram Completion for Low-Dose CT". In: *2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*. 2018, pp. 1–5.

Bibliography VI

[14] L. Gjesteby, Q. Yang, Y. Xi, et al. “Deep learning methods to guide CT image reconstruction and reduce metal artifacts”. In: *Medical Imaging 2017: Physics of Medical Imaging*. Ed. by T. G. Flohr, J. Y. Lo, and T. G. Schmidt. Vol. 10132. International Society for Optics and Photonics. SPIE, 2017, 101322W. URL: https://doi.org/10.1117/12.2254091.

[15] H. Gupta, K. H. Jin, H. Q. Nguyen, et al. “CNN-Based Projected Gradient Descent for Consistent CT Image Reconstruction”. In: *IEEE Transactions on Medical Imaging* 37.6 (2018), pp. 1440–1453.

[16] J. He, Y. Wang, and J. Ma. “Radon Inversion via Deep Learning”. In: *IEEE Transactions on Medical Imaging* 39.6 (2020), pp. 2076–2087. ISSN: 1558-254X.

Bibliography VII

[17] 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.

[18] 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).

Bibliography VIII

[19] H. Lee, J. Lee, and S. Cho. "View-interpolation of sparsely sampled sinogram using convolutional neural network". In: *Medical Imaging 2017: Image Processing*. Ed. by M. A. Styner and E. D. Angelini. Vol. 10133. International Society for Optics and Photonics. SPIE, 2017, p. 1013328. URL: https://doi.org/10.1117/12.2254244.

[20] V. Lempitsky, A. Vedaldi, and D. Ulyanov. "Deep Image Prior". In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 2018, pp. 9446–9454.

Bibliography IX

[21] 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). ISSN: 2052-4463. URL: `http://dx.doi.org/10.1038/s41597-021-00893-z`.

[22] 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`.

Bibliography X

[23] Y. Li, K. Li, C. Zhang, et al. "Learning to Reconstruct Computed Tomography Images Directly From Sinogram Data Under A Variety of Data Acquisition Conditions". In: *IEEE Transactions on Medical Imaging* 38.10 (2019), pp. 2469–2481.

[24] Z. Li, S. Ye, Y. Long, et al. "SUPER Learning: A Supervised-Unsupervised Framework for Low-Dose CT Image Reconstruction". In: *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. 2019, pp. 3959–3968.

Bibliography XI

[25] W.-A. Lin, H. Liao, C. Peng, et al. “DuDoNet: Dual Domain Network for CT Metal Artifact Reduction”. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2019, pp. 10504–10513.

[26] D. C. Liu and J. Nocedal. “On the limited memory BFGS method for large scale optimization”. In: *Mathematical Programming* 45.1 (1989), pp. 503–528. ISSN: 1436-4646. URL: https://doi.org/10.1007/BF01589116.

[27] T. Liu, A. Chaman, D. Belius, et al. “Interpreting U-Nets via Task-Driven Multiscale Dictionary Learning”. In: *arXiv preprint* (2020). [arXiv:cs.CV/2011.12815].

Bibliography XII

[28] S. Lunz, O. Öktem, and C.-B. Schönlieb. "Adversarial Regularizers in Inverse Problems". In: *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. NIPS 2018. Montréal, Canada: Curran Associates Inc., 2018, pp. 8516–8525.

[29] J. Martens. "New Insights and Perspectives on the Natural Gradient Method". In: *Journal of Machine Learning Research* 21.146 (2020), pp. 1–76. URL: `http://jmlr.org/papers/v21/17-678.html`.

Bibliography XIII

[30] J. Martens and R. Grosse. "Optimizing Neural Networks with Kronecker-factored Approximate Curvature". In: *Proceedings of the 32nd International Conference on Machine Learning*. Ed. by F. Bach and D. Blei. Vol. 37. Proceedings of Machine Learning Research. Lille, France: PMLR, July 2015, pp. 2408–2417. URL: https://proceedings.mlr.press/v37/martens15.html.

[31] H. S. Park, J. Baek, S. K. You, et al. "Unpaired Image Denoising Using a Generative Adversarial Network in X-Ray CT". In: *IEEE Access* 7 (2019), pp. 110414–110425.

Bibliography XIV

[32] D. M. Pelt, K. J. Batenburg, and J. A. Sethian. “Improving Tomographic Reconstruction from Limited Data Using Mixed-Scale Dense Convolutional Neural Networks”. In: *Journal of Imaging* 4.11 (2018). ISSN: 2313-433X. URL: https://www.mdpi.com/2313-433X/4/11/128.

[33] O. Ronneberger, P. Fischer, and T. Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by N. Navab, J. Hornegger, W. M. Wells, et al. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 978-3-319-24574-4.

Bibliography XV

[34] C. Shen, G. Ma, and X. Jia. "Low-dose CT reconstruction assisted by a global CT image manifold prior". In: *15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*. Ed. by S. Matej and S. D. Metzler. Vol. 11072. International Society for Optics and Photonics. SPIE, 2019, p. 1107205. URL: https://doi.org/10.1117/12.2534959.

[35] Y. Sun, Q. Chen, X. He, et al. "Singular Value Fine-tuning: Few-shot Segmentation requires Few-parameters Fine-tuning". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh, A. Agarwal, D. Belgrave, et al. 2022. URL: https://openreview.net/forum?id=LEqYZz7cZOI.

Bibliography XVI

[36] Y. Sun, J. Liu, and U. Kamilov. “Block Coordinate Regularization by Denoising”. In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, et al. Vol. 32. Curran Associates, Inc., 2019. URL: https://proceedings.neurips.cc/paper/2019/file/9872ed9fc22fc182d371c3e9ed316094-Paper.pdf.

[37] C. Tang, J. Li, L. Wang, et al. “Unpaired Low-Dose CT Denoising Network Based on Cycle-Consistent Generative Adversarial Network with Prior Image Information”. In: *Computational and Mathematical Methods in Medicine* 2019 (2019), p. 8639825. ISSN: 1748-670X. URL: https://doi.org/10.1155/2019/8639825.

Bibliography XVII

[38] J. M. Wolterink, T. Leiner, M. A. Viergever, et al. "Generative Adversarial Networks for Noise Reduction in Low-Dose CT". In: *IEEE Transactions on Medical Imaging* 36.12 (2017), pp. 2536–2545.

[39] Q. Yang, P. Yan, Y. Zhang, et al. "Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss". In: *IEEE Transactions on Medical Imaging* 37.6 (June 2018), pp. 1348–1357. ISSN: 0278-0062.

Bibliography XVIII

[40] D. H. Ye, S. Srivastava, J.-B. Thibault, et al. "Deep Residual Learning for Model-Based Iterative CT Reconstruction Using Plug-and-Play Framework". In: *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2018, pp. 6668–6672.

[41] H. Yuan, J. Jia, and Z. Zhu. "SIPID: A deep learning framework for sinogram interpolation and image denoising in low-dose CT reconstruction". In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. 2018, pp. 1521–1524.

Bibliography XIX

[42] J. Zhao, Z. Chen, L. Zhang, et al. "Few-View CT reconstruction method based on deep learning". In: *2016 IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD)*. 2016, pp. 1–4.

[43] 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.