Deep learning methods for image reconstruction

Samuele Papa

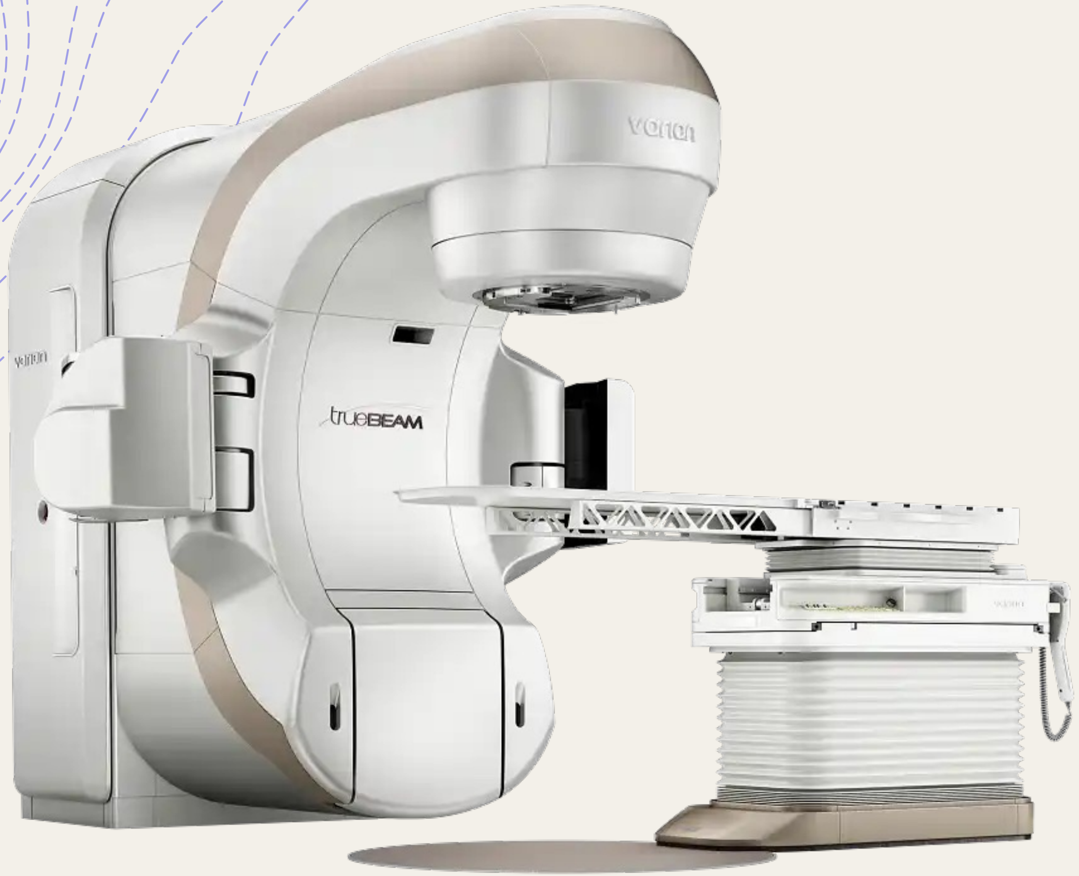
POP ART lab

a collaboration between

Netherlands Cancer Institute – Antoni van Leeuwenhoek · University of Amsterdam · Elekta

DLinRT 2024

Radiotherapy LINACs



Varian CBCT LINAC

Elekta CBCT LINAC

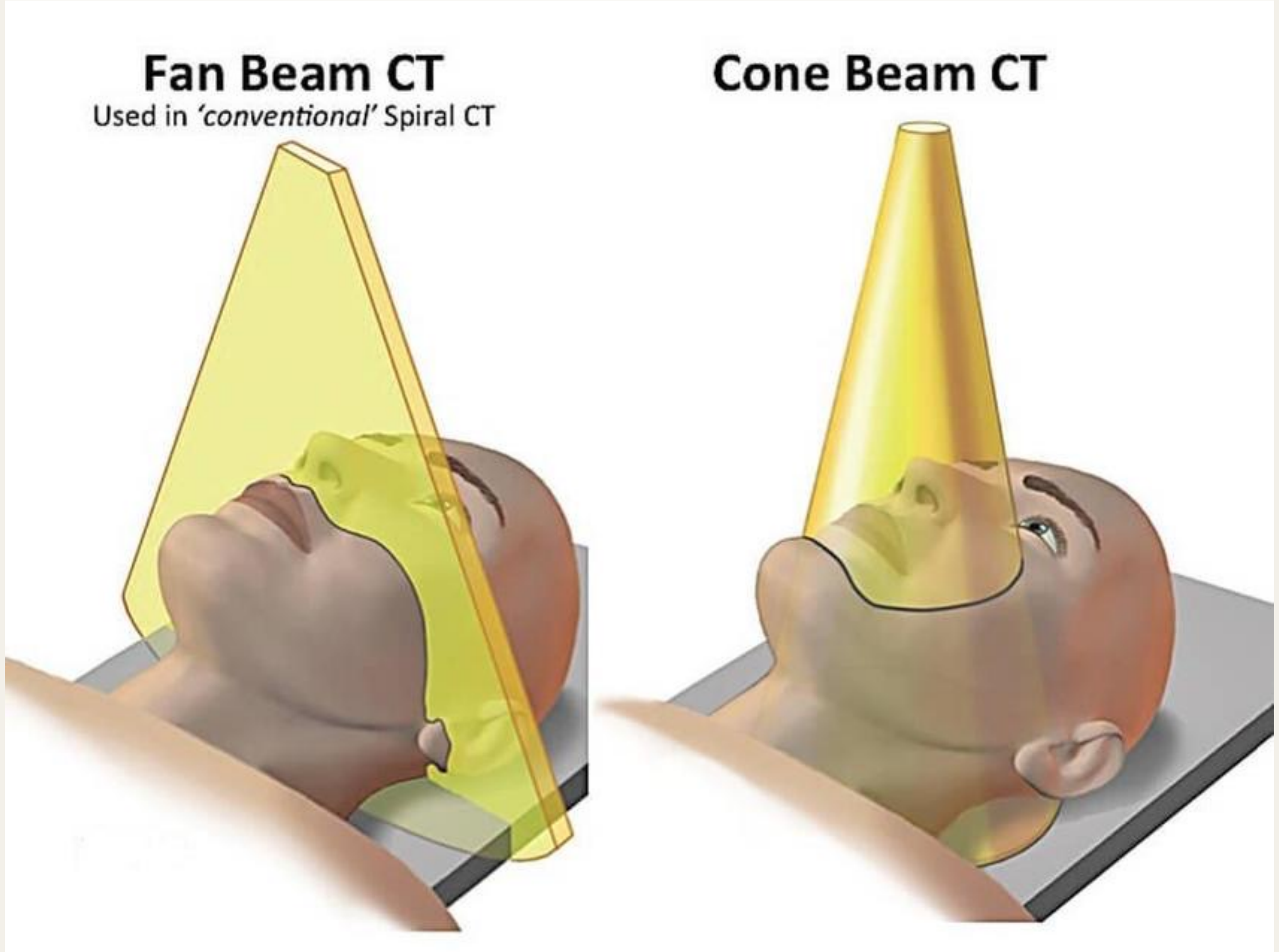
UvA DLinRT 2024

CT vs CBCT

Single slice per rotation.

Low noise.

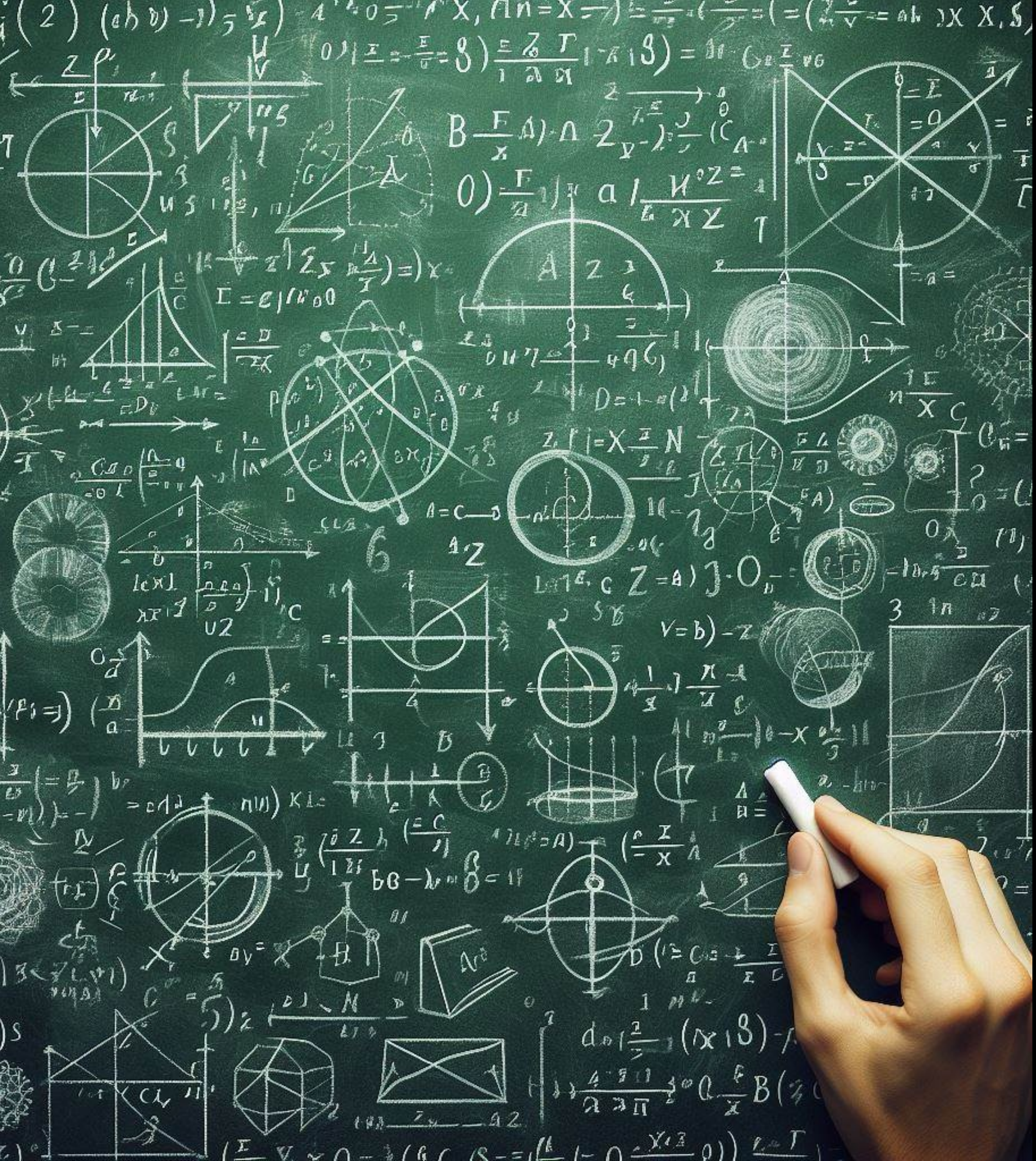
Calibrated HU.



Whole scan per rotation.

High noise.

Un-Calibrated HU.



TRADITIONAL TECHNIQUES

Basis of Computed Tomography

Projection

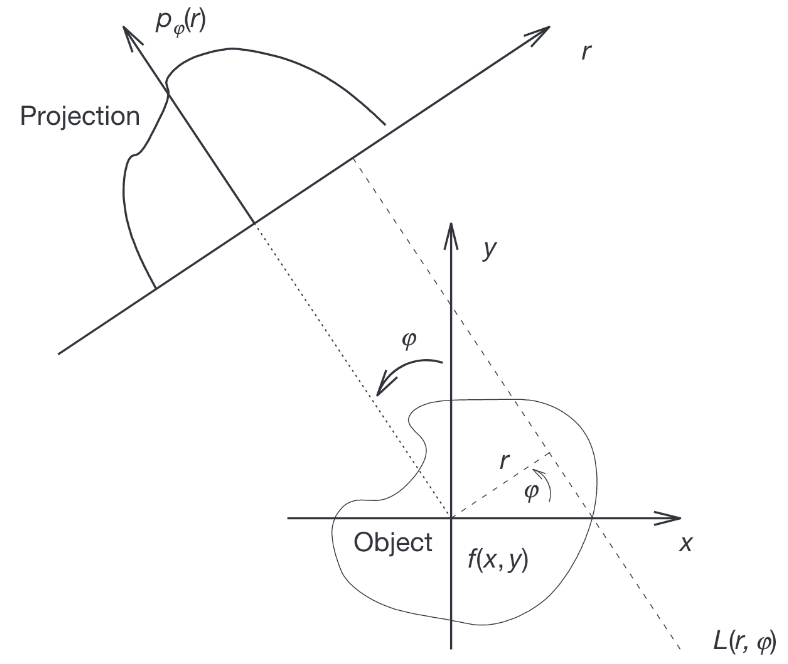


Figure 1 Geometry of the line integrals associated with the Radon transform.

$$p_\varphi(r) = \int_{\mathcal{L}(r,\varphi)} f(x,y)\,\mathrm{d}\ell$$

Back-Projection

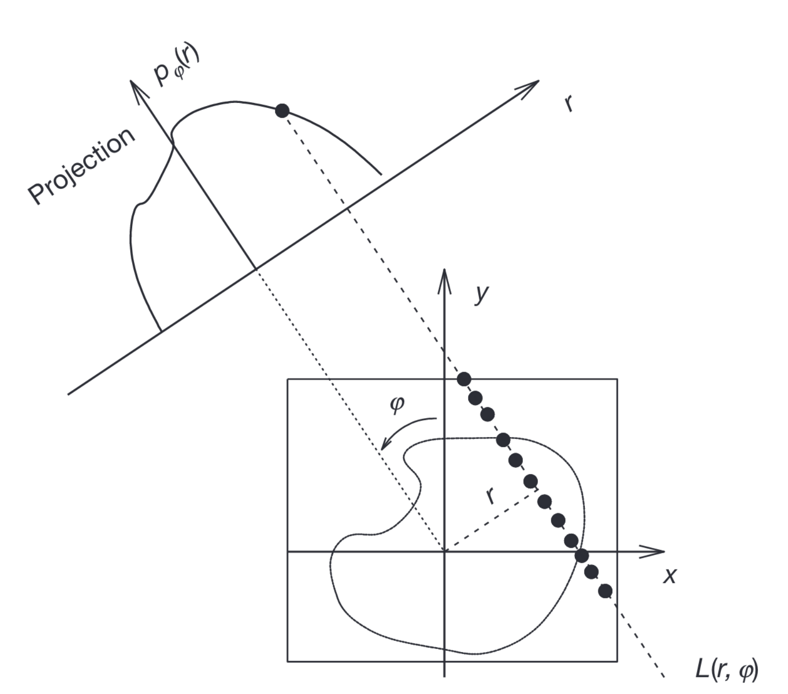


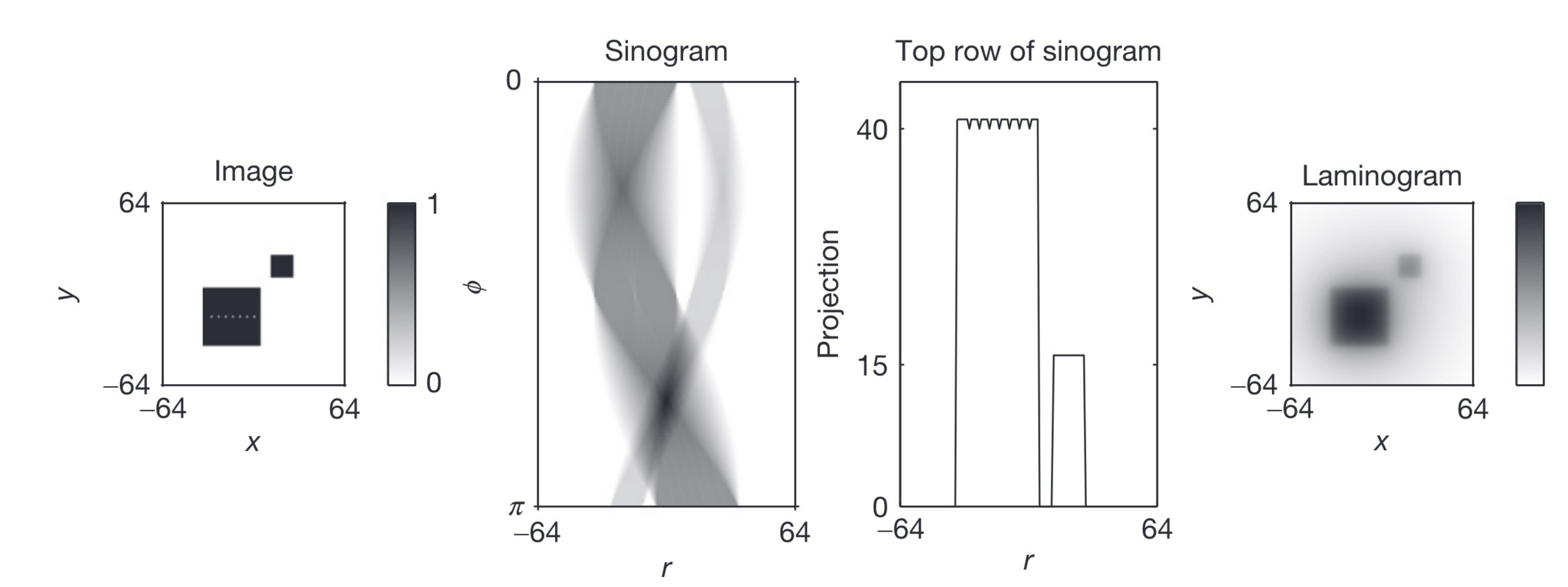
Figure 6 Illustration of back projection operation for a single projection view.

$$f_{\mathrm{b}}(x,y) = \int_0^\pi w(\varphi)\,p_\varphi(x\cos\varphi + y\sin\varphi)\,\mathrm{d}\varphi$$

UvA DLinRT 2024

Fessler, J. A. Fundamentals of CT Reconstruction in 2D and 3D. in *Comprehensive Biomedical Physics* 263–295 (Elsevier, 2014). doi:10.1016/B978-0-444-53632-7.00212-4.

Filtered Back-Projection

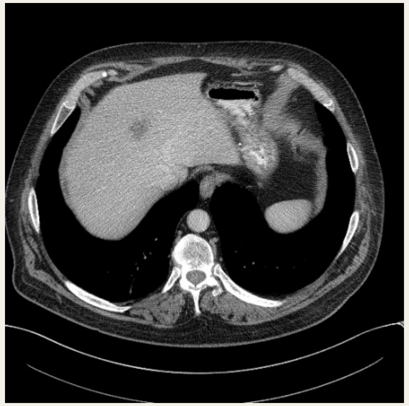


Fessler, J. A. Fundamentals of CT Reconstruction in 2D and 3D. in *Comprehensive Biomedical Physics* 263–295 (Elsevier, 2014). doi:10.1016/B978-0-444-53632-7.00212-4.

Image Reconstruction as Inverse Problem

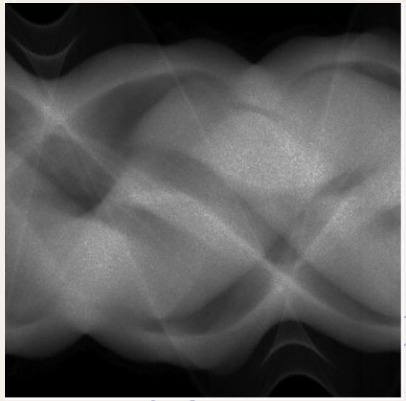
$$y = \mathcal{T}\left(x_{\mathrm{true}}\right) + \delta y$$

$y \in Y$	Data
$x_{\mathrm{true}} \in X$	Image
$\mathcal{T} : X \to Y$	Forward operator
$\delta y \in Y$	Noise



Image

$\xrightarrow{\mathcal{T}}$



Data

Solving the Inverse Problem

$$\min_{x \in X} \mathcal{L}\left(\mathcal{T}(x), y\right)$$

Straightforward approach to inversion

$$\downarrow$$

Find reconstruction that minimizes the negative log-likelihood.

Overfit the measurements.

Noise will affect reconstruction.

Regularization

$$\min_{x \in X} \left[\mathcal{L}\left(\mathcal{T}(x), y\right) + \lambda \mathcal{R}(x)\right]$$

Where:

$\mathcal{R}(x)$ Regularization functional.

λ Regularization parameter.

Add prior information using regularization to reduce effect of noise on the reconstruction.

Iterative Methods for Reconstruction

Total Variation regularization

$$\min_{x \in X} \left[\mathcal{L}\left(\mathcal{T}(x), y\right) + \lambda \|\nabla x\|_1 \right]$$

Spatial gradient as regularization.

Results in smoother reconstructions.

Iterative Methods for Reconstruction

begin

$\sigma :=$ learning rate

$y :=$ projection data

$x^{(0)} :=$ initial guess

for $i := 1, \dots$ **do**

Project the current reconstruction.

$$y^{(i-1)} := \mathcal{T}\left(x^{(i-1)}\right)$$

Compute the loss based on the projection data.

$$L := \|y^{(i-1)} - y\|_2^2 + \lambda \mathcal{R}\left(x^{(i-1)}\right)$$

Update the reconstruction using the gradients.

$$x^{(i)} = x^{(i-1)} - \sigma \nabla_{x^{(i-1)}} L$$

end

end

Iterative Methods for Reconstruction

begin

$\sigma :=$ learning rate

$y :=$ projection data

$x^{(0)} :=$ initial guess

for $i := 1, \ldots$ do

Project the current reconstruction.

$$y^{(i-1)} := \mathcal{T}\left(x^{(i-1)}\right)$$

Compute the loss based on the projection data.

$$L := \|y^{(i-1)} - y\|_2^2 + \lambda \mathcal{R}\left(x^{(i-1)}\right)$$

Update the reconstruction using the gradients.

$$x^{(i)} = x^{(i-1)} - \sigma \nabla_{x^{(i-1)}} L$$

end

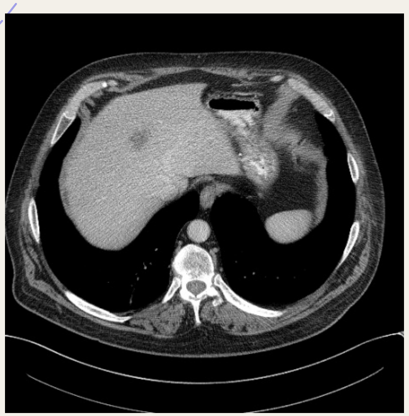
end

Termination condition is an open problem.

DEEP LEARNING BASED

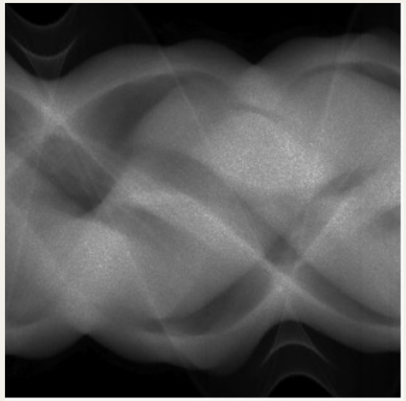
Learned Reconstruction

$$\mathcal{T}_\theta^\dagger : Y \to X$$



Image

$$\xleftarrow{\mathcal{T}_\theta^\dagger}$$



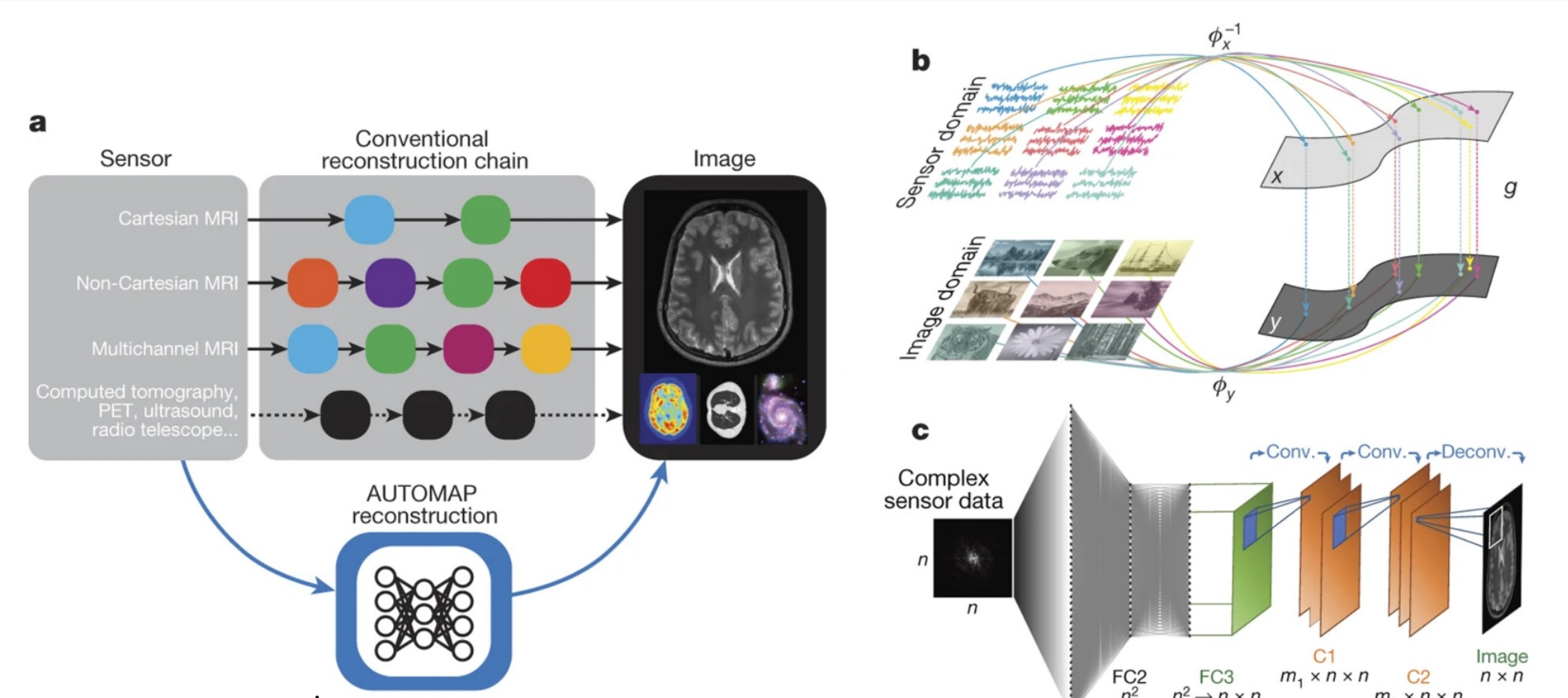
Data

Find *pseudo-inverse* that given the measurements, obtains the clean reconstruction.

Learned refers to finding the best parameters given some training data.

Zhu, B., Liu, J., Cauley, S. et al. Image reconstruction by domain-transform manifold learning. Nature 555, 487–492 (2018). https://doi.org/10.1038/nature25988

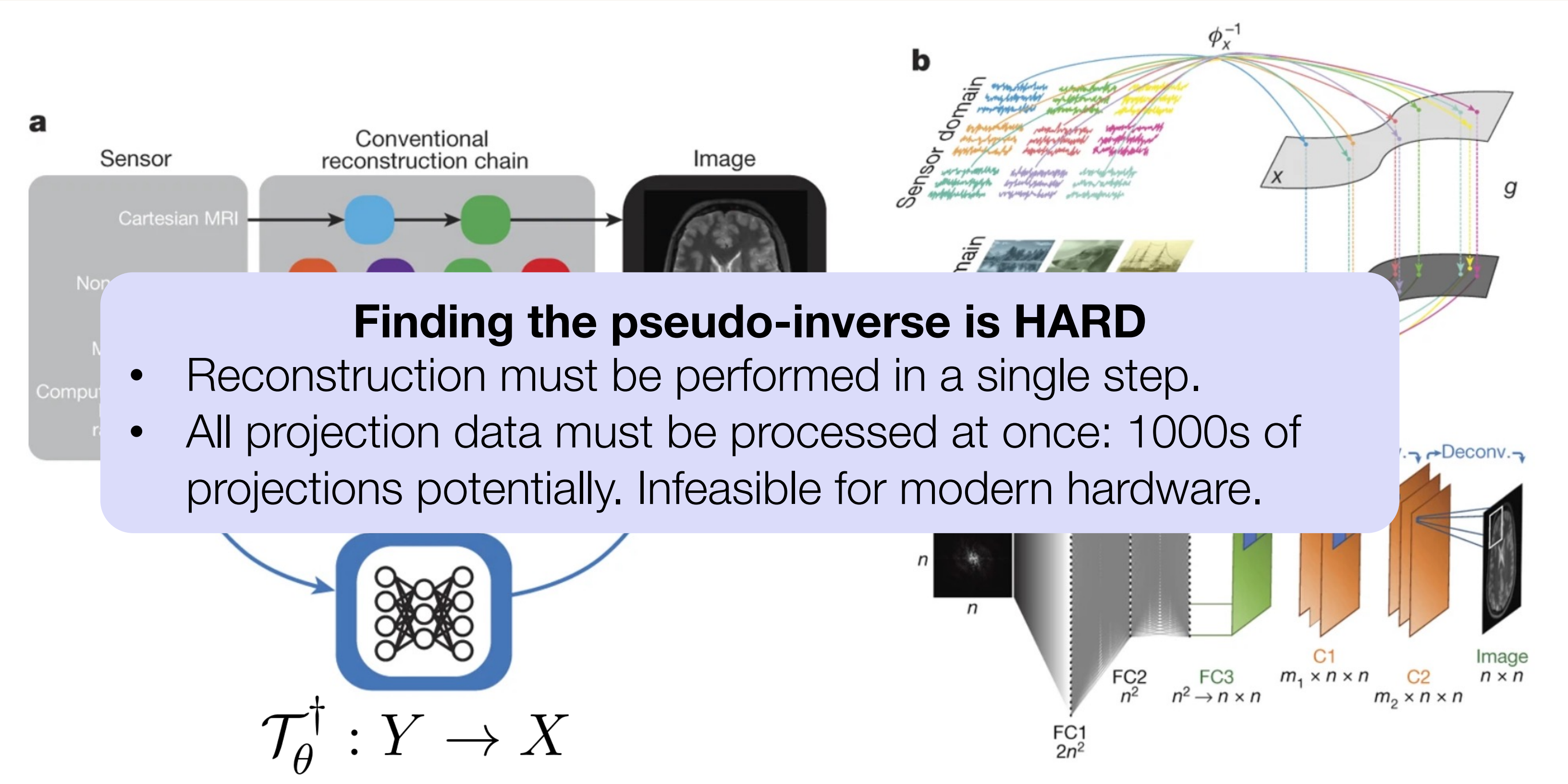
Learned Reconstruction



$$\mathcal{T}_\theta^\dagger : Y \to X$$

Zhu, B., Liu, J., Cauley, S. et al. Image reconstruction by domain-transform manifold learning. Nature 555, 487–492 (2018). https://doi.org/10.1038/nature25988

Learned Reconstruction



Finding the pseudo-inverse is HARD

- Reconstruction must be performed in a single step.
- All projection data must be processed at once: 1000s of projections potentially. Infeasible for modern hardware.

Zhu, B., Liu, J., Cauley, S. et al. Image reconstruction by domain-transform manifold learning. Nature 555, 487–492 (2018). https://doi.org/10.1038/nature25988

Learned Iterative Reconstruction

$$\min_{f \in X} \left[\mathcal{F}\left(\mathcal{K}\left(f\right)\right) + \mathcal{G}\left(f\right)\right]$$

where

$\mathcal{K}$: operator that describes the forward transformation (i.e. the *projection*),

$\mathcal{F}$: objective functional in dual space,

$\mathcal{G}$: objective functional in primal space.

Dual space = projection space

Primal space = image space

Adler, J. & Öktem, O. Learned Primal-dual Reconstruction. IEEE Trans. Med. Imaging 37, 1322–1332 (2018).

Learned Iterative Reconstruction

$$\min_{f \in X} \left[\mathcal{F}\left(\mathcal{K}\left(f\right)\right) + \mathcal{G}\left(f\right)\right]$$

Generalization of the regularized objective:

$$\min_{x \in X} \left[\mathcal{L}\left(\mathcal{T}(x), y\right) + \lambda \mathcal{R}(x)\right]$$

Allows to split the optimization into a primal step and dual step.

Adler, J. & Öktem, O. Learned Primal-dual Reconstruction. IEEE Trans. Med. Imaging 37, 1322–1332 (2018).

Learned Iterative Reconstruction

$$\min_{f \in X} \left[\mathcal{F}\left(\mathcal{K}\left(f \right) \right) + \mathcal{G}\left(f \right) \right]$$

Why use primal-dual?

- Gradient descent on the available projections is noisy.
- Regularization is not enough.

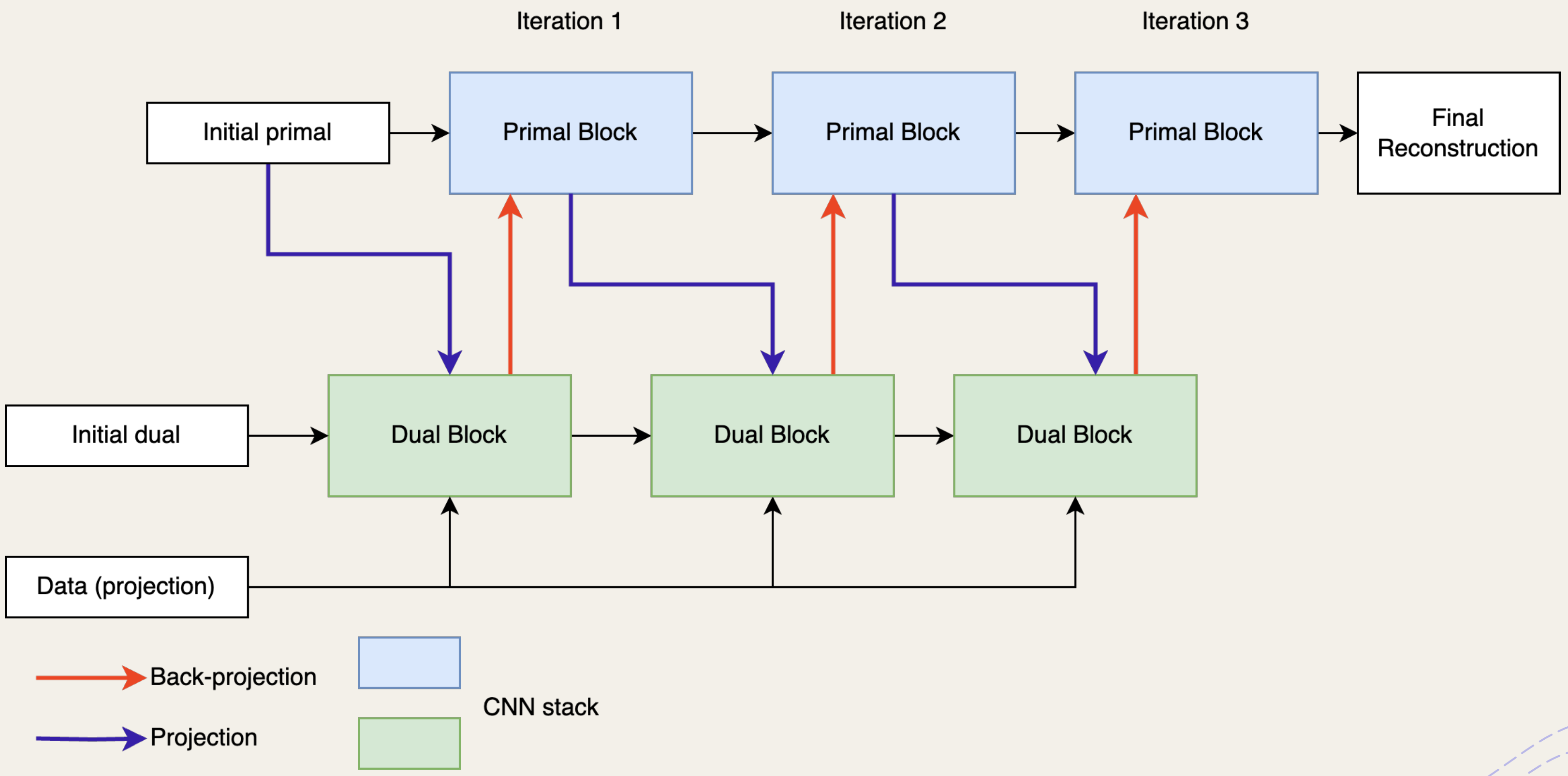
We can solve this problem by unrolling the iterative steps and **learning each step**.

DLinRT 2024

Adler, J. & Öktem, O. Learned Primal-dual Reconstruction. IEEE Trans. Med. Imaging 37, 1322–1332 (2018).

Learned Primal-Dual Model

Inspired by *Primal Dual Hybrid Gradient Method.*



Adler, J. & Öktem, O. Learned Primal-dual Reconstruction. IEEE Trans. Med. Imaging 37, 1322–1332 (2018).

Anatomy of a Block

```
for i in range(n_iter):
    evalop = project(primal)
    update = concat([dual, evalop, projs], axis=-1)

    update = prelu(conv(update))
    update = prelu(conv(update))
    update = conv(update)
    dual = dual + update


    evalop = back_project(dual)
    update = concat([primal, evalop], axis=-1)

    update = prelu(conv(update))
    update = prelu(conv(update))
    update = conv(update)
    primal = primal + update

x_result = primal
```

Projection

Small convolutional stack

Back-Projection

Small convolutional stack

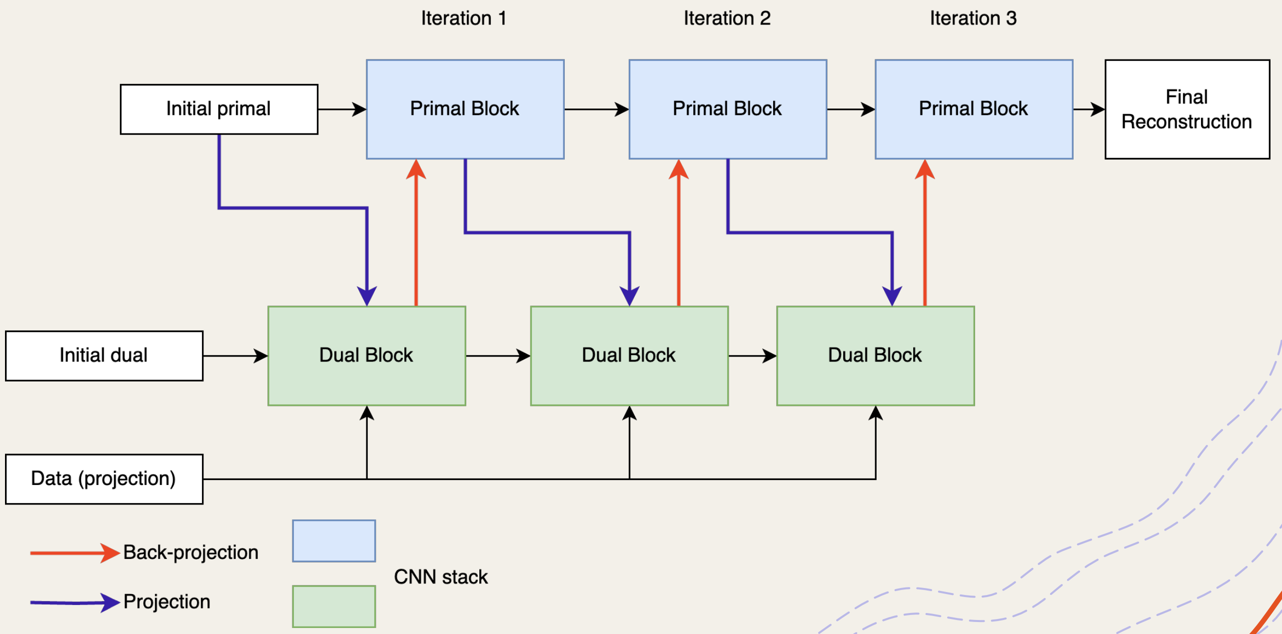
Adler, J. & Öktem, O. Learned Primal-dual Reconstruction. IEEE Trans. Med. Imaging 37, 1322–1332 (2018).

Application: LIRE

Scaling learned primal dual to CBCT

When training large scale models, considerations on **memory** *usage and* **processing** *speed are paramount to making it work.*

1. CBCT projection and back-projection operators require all projections and the entire volume in GPU for fast inference.
2. Several iterations of CNN blocks require storing gradient information for back-propagation (i.e. training).
3. Using external libraries for computing projections is slow.
4. PyTorch does not allow simple compilation of complex combination of operations.



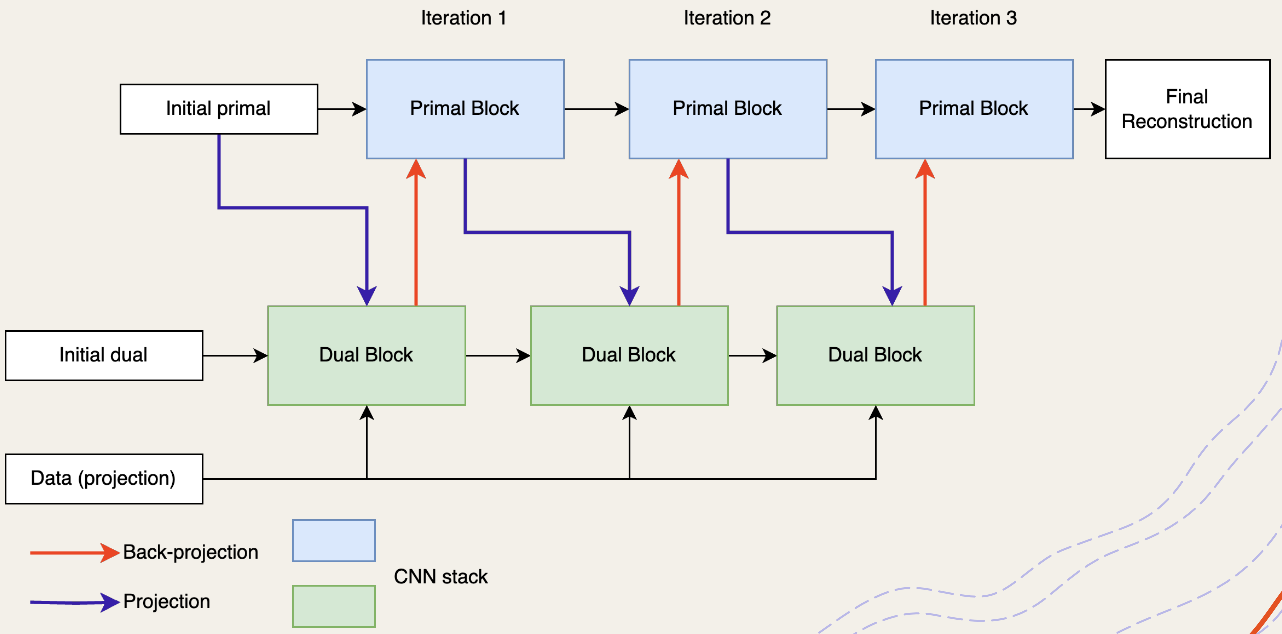
DLinRT 2024

Application: LIRE

Scaling learned primal dual to CBCT

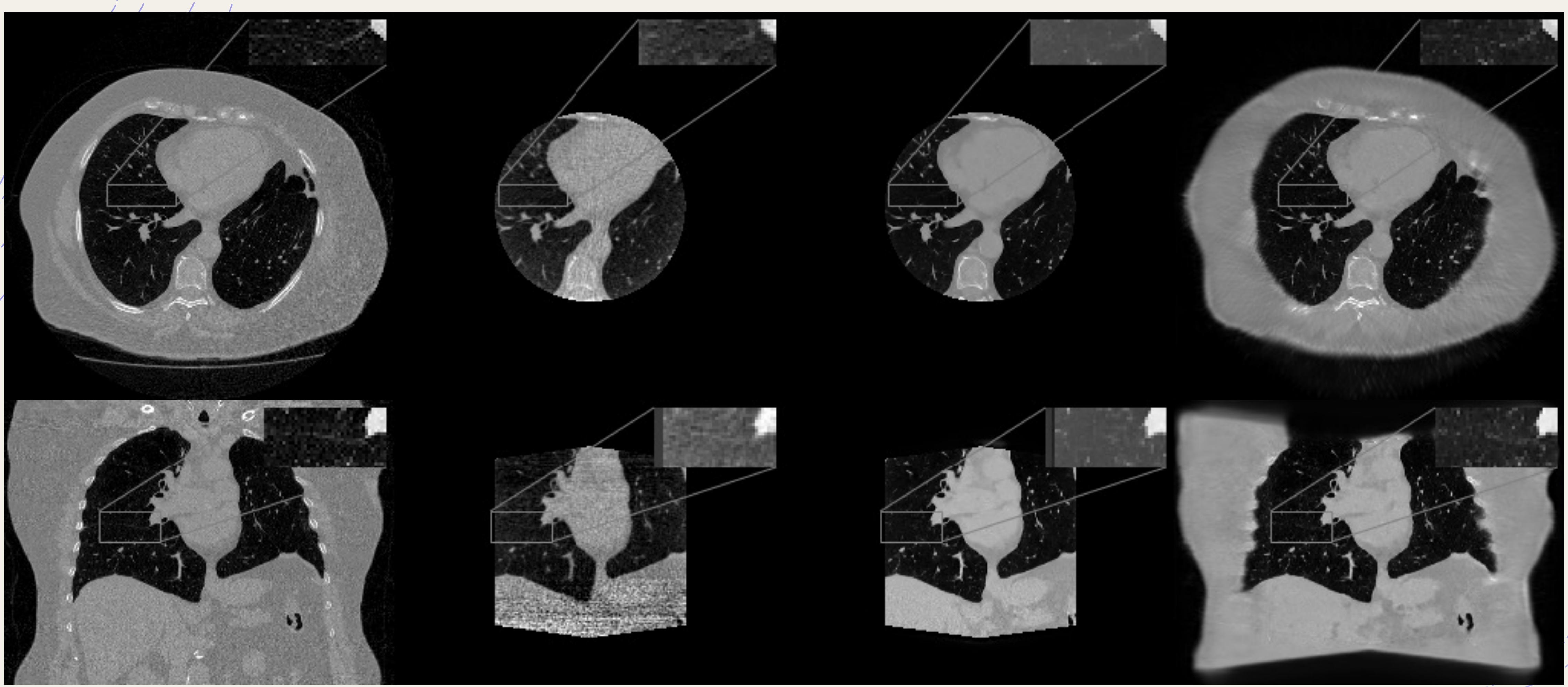
When training large scale models, considerations on **memory** *usage and* **processing** *speed are paramount to making it work.*

1. Mandatory requirement that can only be relaxed if we sacrifice a lot of speed.
2. Use invertible blocks for allowing computation of the gradient from the output to the input. Use tiling mechanism to not store whole feature maps.
3. Write custom CUDA code as a PyTorch extension.
4. Write the whole model as a CUDA kernel. Alternative are possible for CNN-based models.



DLinRT 2024

LIRE Results: Small Field of View



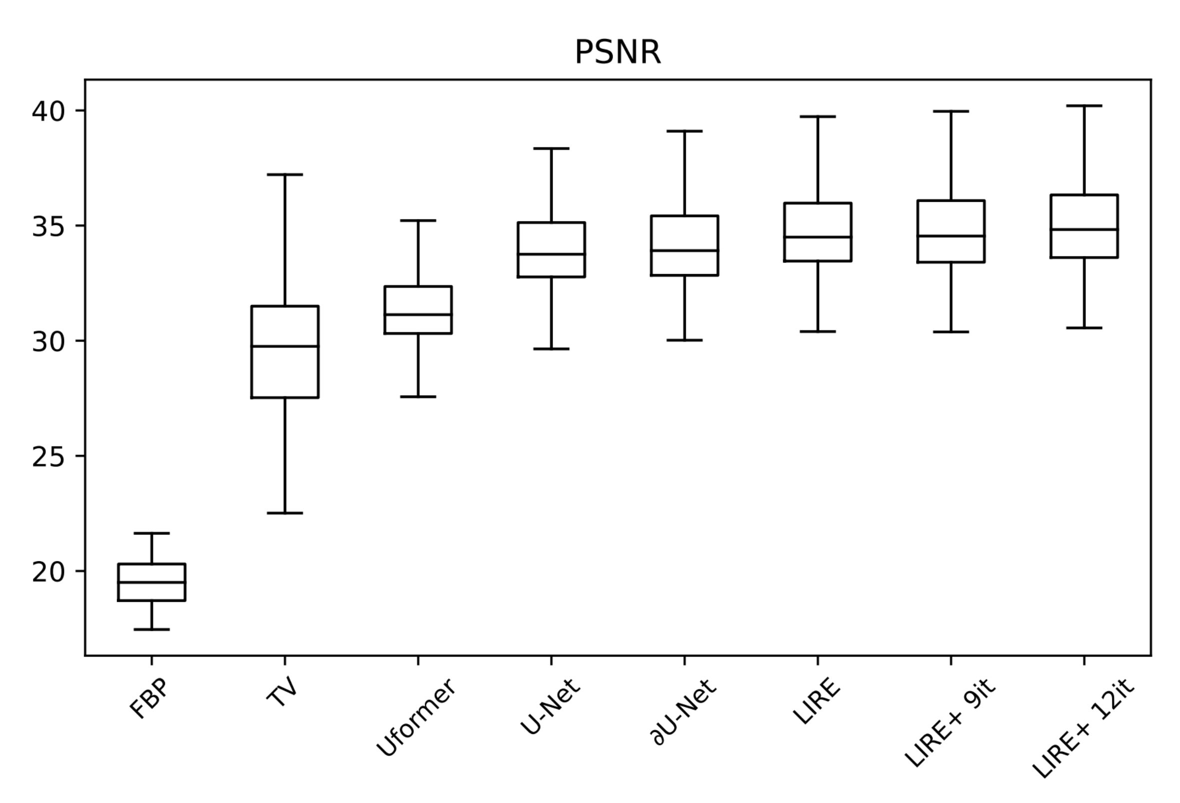
Ground Truth | Iterative: latest commercial CBCT | UNet | LIRE

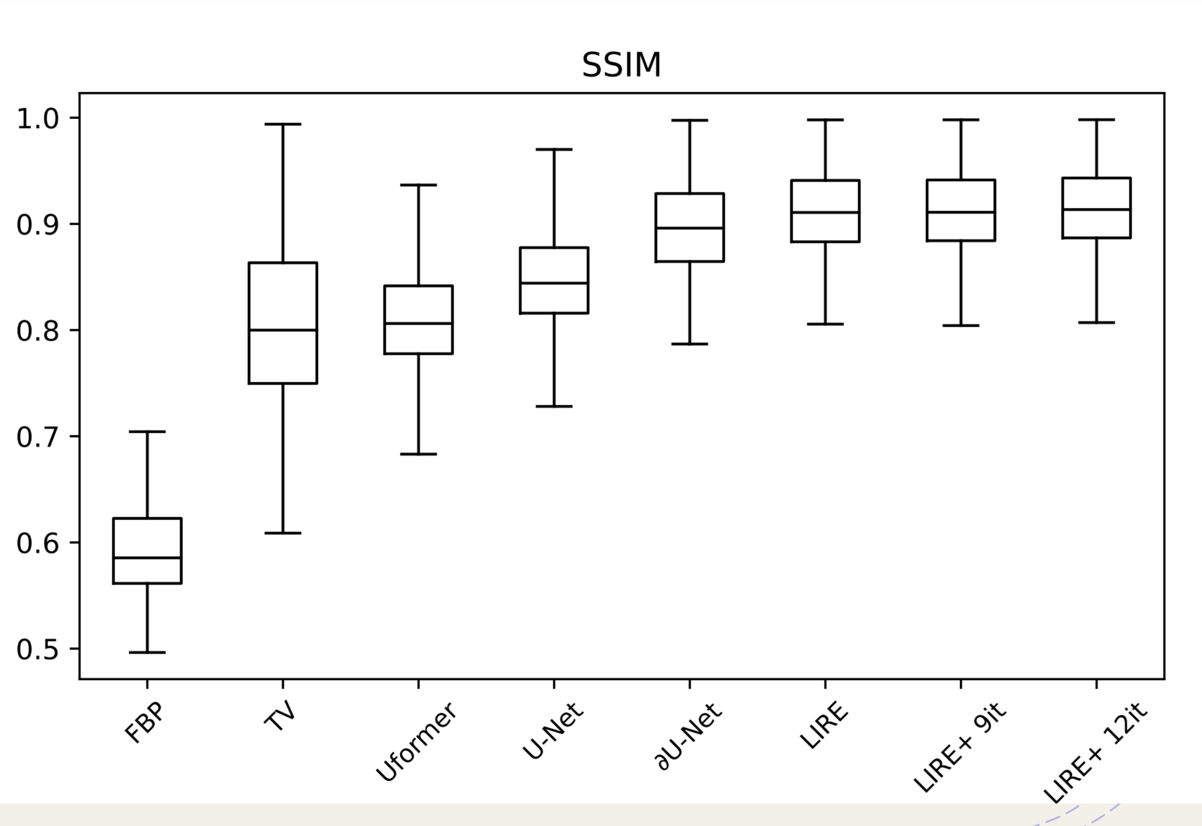
Reconstruction at 1mm resolution with LIRE

GT

LIRE

LIRE Results: Comparison with other methods





Neural fields

Definition 1 A *field* is a quantity defined for all spatial and/or temporal coordinates.

$$f : \mathbb{R}^k \rightarrow \mathbb{R}^n$$

Definition 2 A *neural field* is a field that is parameterized fully or in part by a neural network.

$$f_\theta : \mathbb{R}^k \rightarrow \mathbb{R}^n$$

Examples of fields

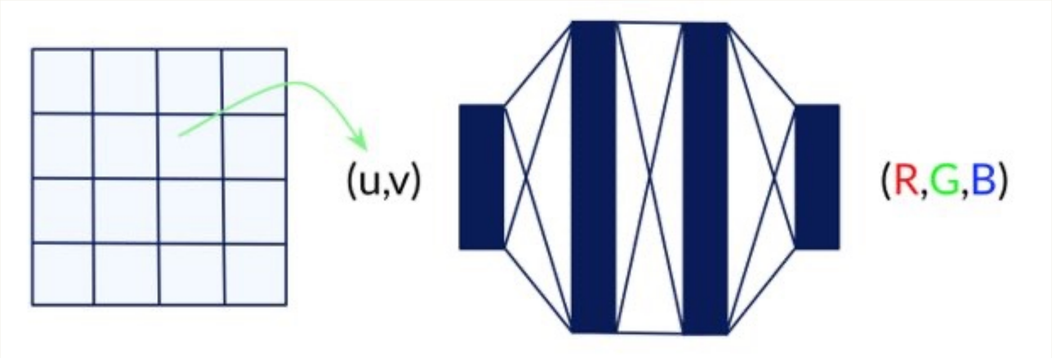
Examples	Field Quantity	Scalar/Vector	Coordinates
Gravitational Field	Force per unit mass (N/kg)	Vector	$\mathbb{R}^n$
3D Paraboloid: $z = x^2 + y^2$	Height z	Scalar	$\mathbb{R}^2$
2D Circle: $r^2 = x^2 + y^2$	Radius r	Scalar	$\mathbb{R}^2$
Signed Distance Field (SDF)	Signed distance	Scalar	$\mathbb{R}^n$
Occupancy Field	Occupancy	Scalar	$\mathbb{R}^n$
Image	RGB intensity	Vector	$\mathbb{Z}^2$ pixel locations x, y
Audio	Amplitude	Scalar	$\mathbb{Z}^1$ time t

UvA

Xie, Yiheng, et al. "Neural fields in visual computing and beyond." Computer Graphics Forum. Vol. 41. No. 2. 2022.

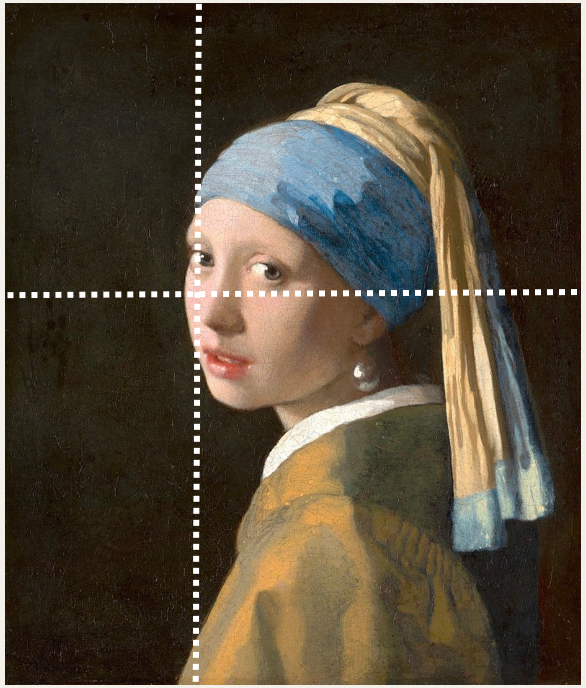
Neural fields

$$f_\theta : \mathbb{R}^k \to \mathbb{R}^n$$



u

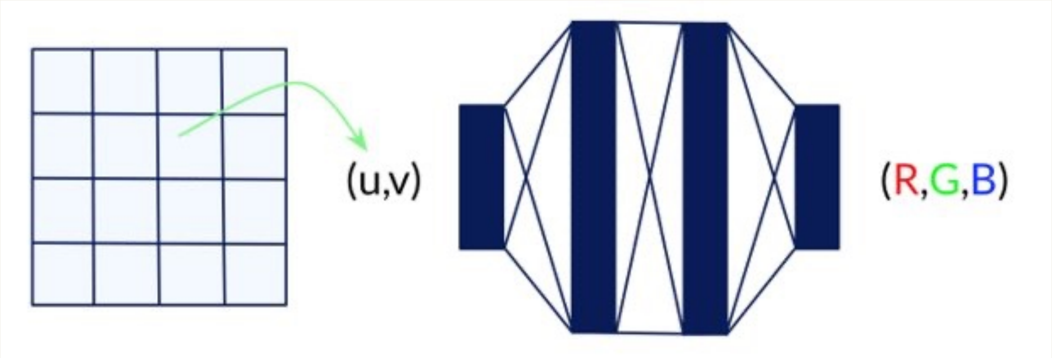
v

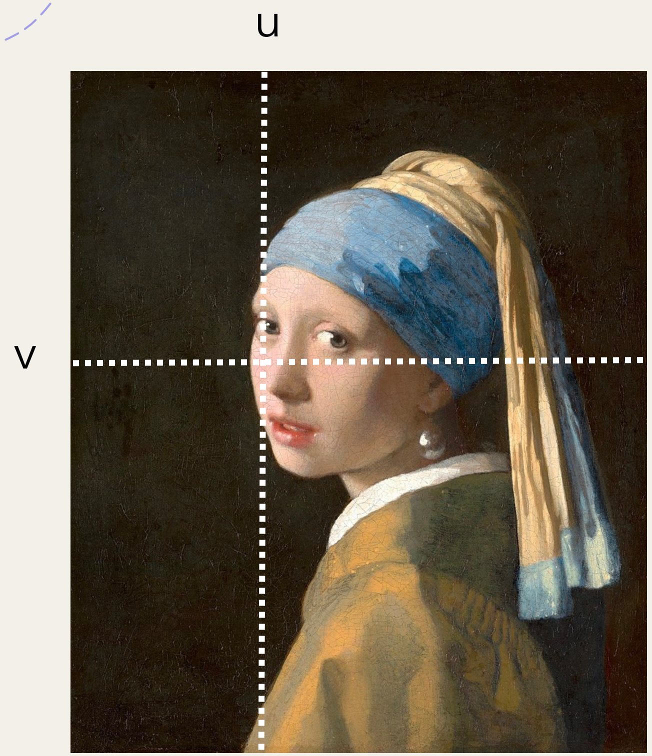


(50, 100) → (242, 211, 160)

Neural fields

$$f_\theta : \mathbb{R}^k \to \mathbb{R}^n$$



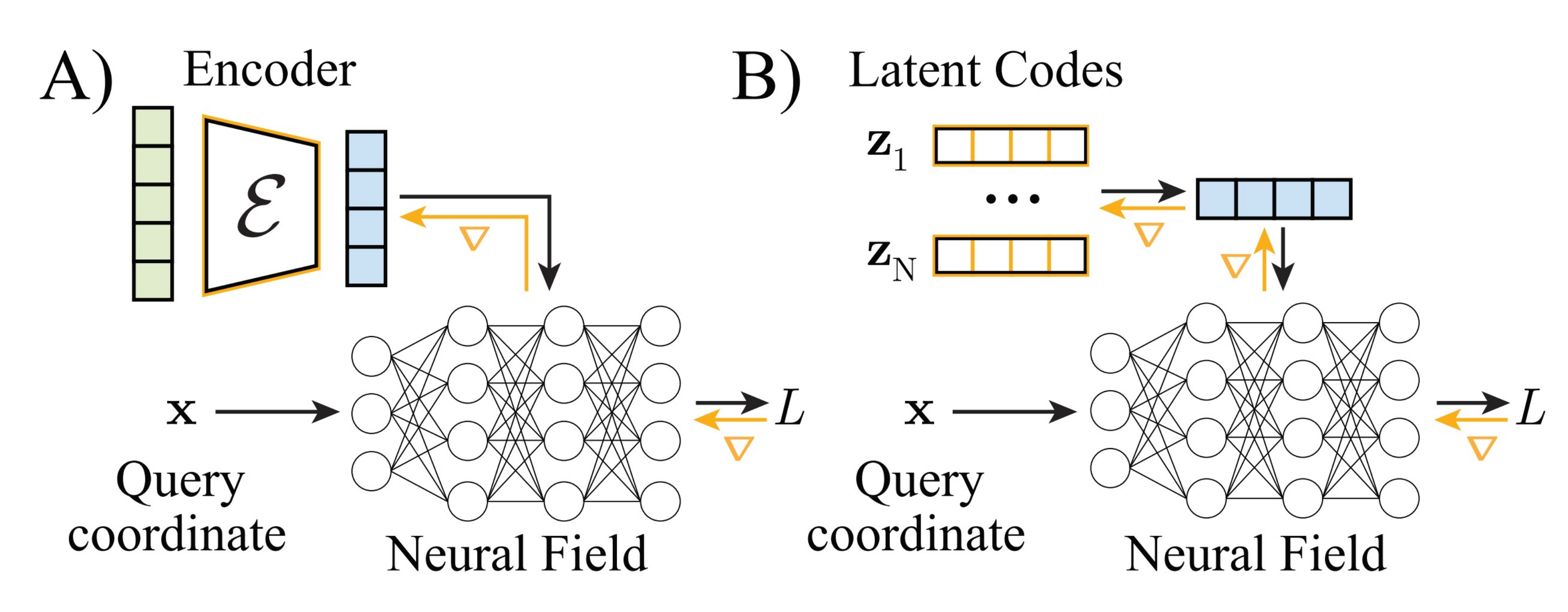


(50, 100) → (242, 211, 160)

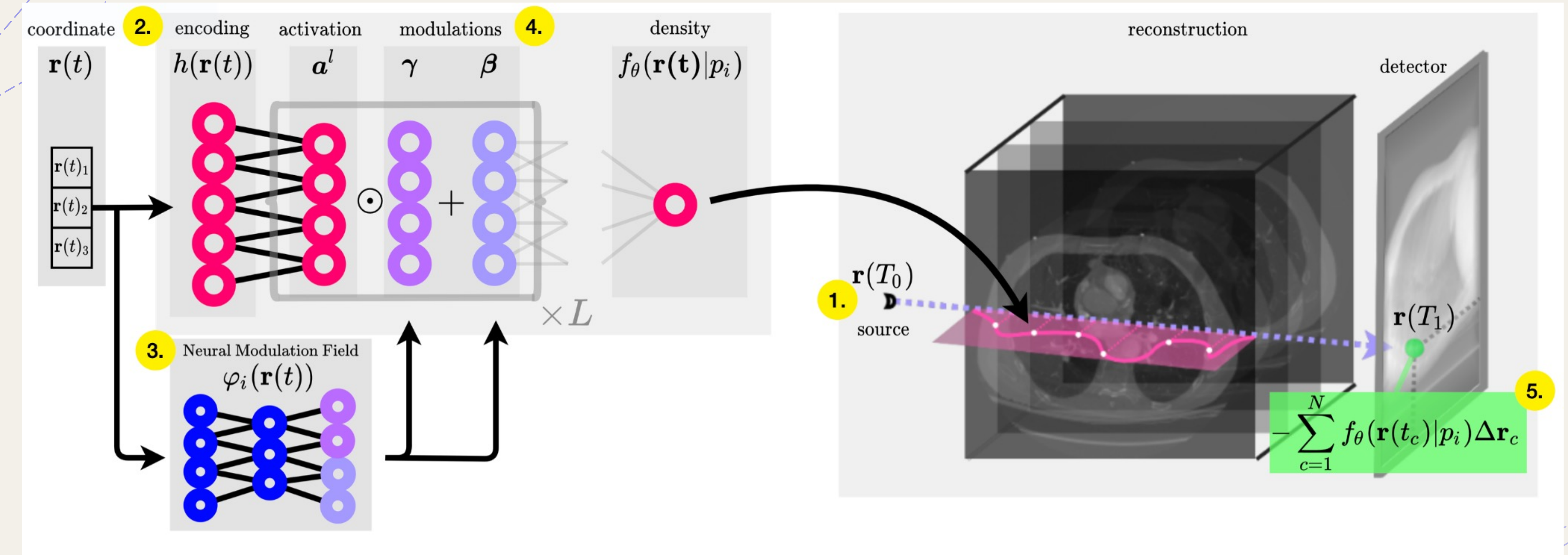
Optimize the network using SGD.
One network per sample

Making Neural Fields Learn from Data

Condition the network on the new measurements.
Train a backbone on all the data available.



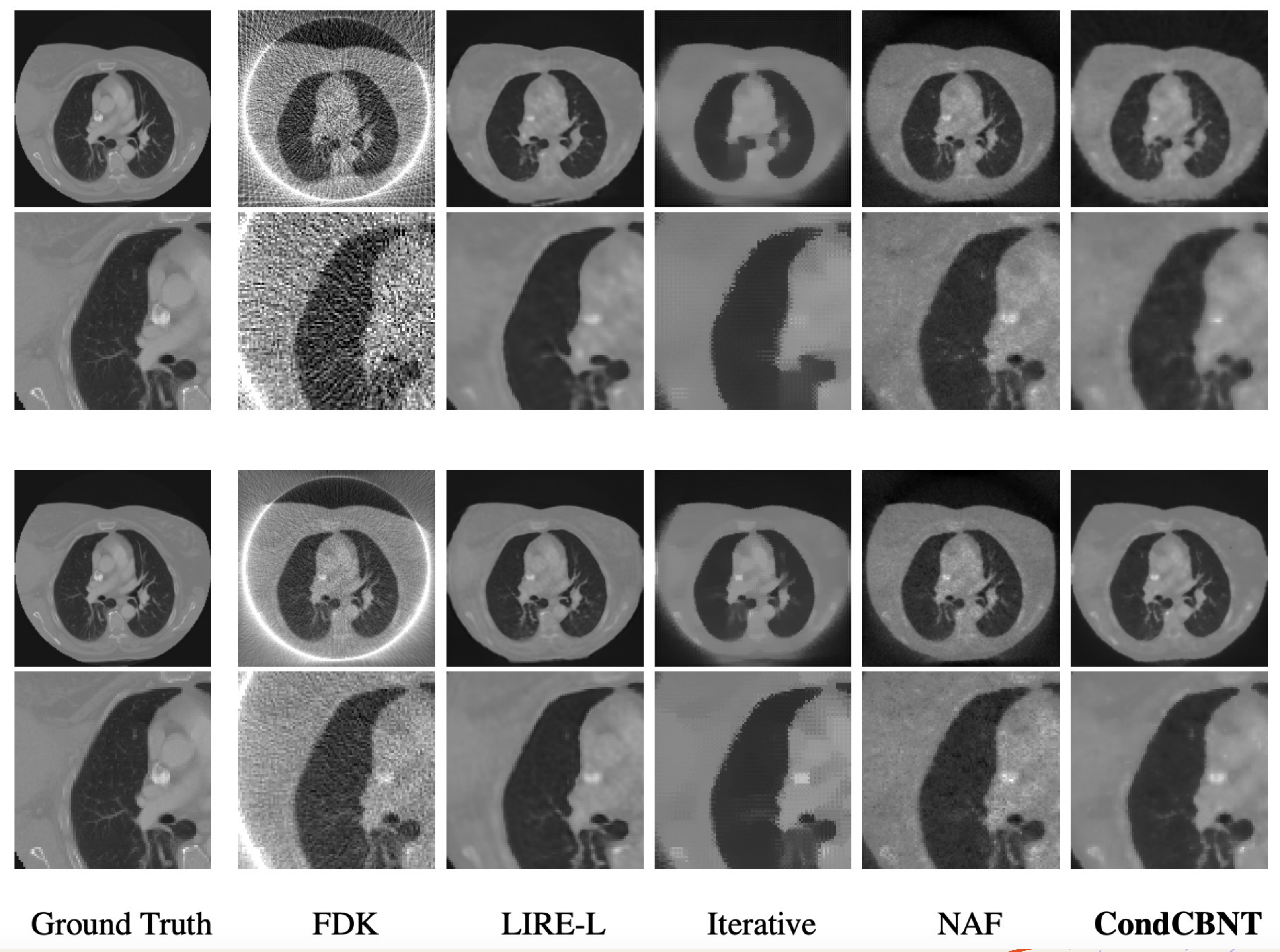
Neural Modulation Fields



Papa, Samuele, et al. "Neural Modulation Fields for Conditional Cone Beam Neural Tomography." arXiv preprint arXiv:2307.08351 (2023).

UvA

Results



Results

		Noisy			Noise-free			
P.	Method	PSNR (↑)	SSIM (↑)	Time (s/vol)	PSNR (↑)	SSIM (↑)	Time (s/vol)	Mem. (MiB)
50	FDK	14.54 ± 2.90	$.20 \pm .07$	0.8	16.09 ± 3.22	$.43 \pm .09$	0.8	100
	Iterative	26.36 ± 2.11	$.70 \pm .08$	7.7	27.13 ± 2.80	$.71 \pm .08$	30.8	300
	LIRE-L	29.48 ± 2.07	$.83 \pm .05$	3.9	-	-	-	2.1k
	NAF	22.83 ± 2.24	$.58 \pm .10$	161	24.26 ± 2.52	$.72 \pm .08$	582	18
	CondCBNT	28.31 ± 1.22	$.80 \pm .05$	124	30.21 ± 1.42	$.86 \pm .05$	647	96
400	FDK	16.43 ± 3.38	$.45 \pm .12$	7	16.71 ± 3.47	$.65 \pm .09$	7	100
	Iterative	28.38 ± 3.27	$.78 \pm .11$	87.4	31.40 ± 6.22	$.91 \pm .07$	174	600
	LIRE-L	30.70 ± 2.25	$.88 \pm .05$	12.8	-	-	-	4k
	NAF	25.93 ± 2.45	$.75 \pm .08$	275	25.04 ± 2.91	$.77 \pm .08$	580	205
	CondCBNT	29.89 ± 1.39	$.86 \pm .05$	763	30.63 ± 1.43	$.88 \pm .04$	595	96

UvA

Summary

Traditional methods cannot learn from data. They leverage the knowledge of the forward operator to optimize a regularized objective.

Learned approach for **direct reconstruction is not feasible**.

Iterative primal-dual method learns to **incrementally reconstruct the image** while also optimizing internal objective in the projection domain.

Outlook and areas of improvements

Learning-based methods should be **physics inspired**.

Computational resources are limited in this domain. Great area for benchmarking new methods.

No real **ground-truth data** is available unless great simulators are developed.

Add **temporal domain** for motion compensation or reconstruction.