Computational Imaging:
Reconciling **Physical** and **Learned** Models

Ulugbek S. Kamilov
Computational Imaging Group (CIG)

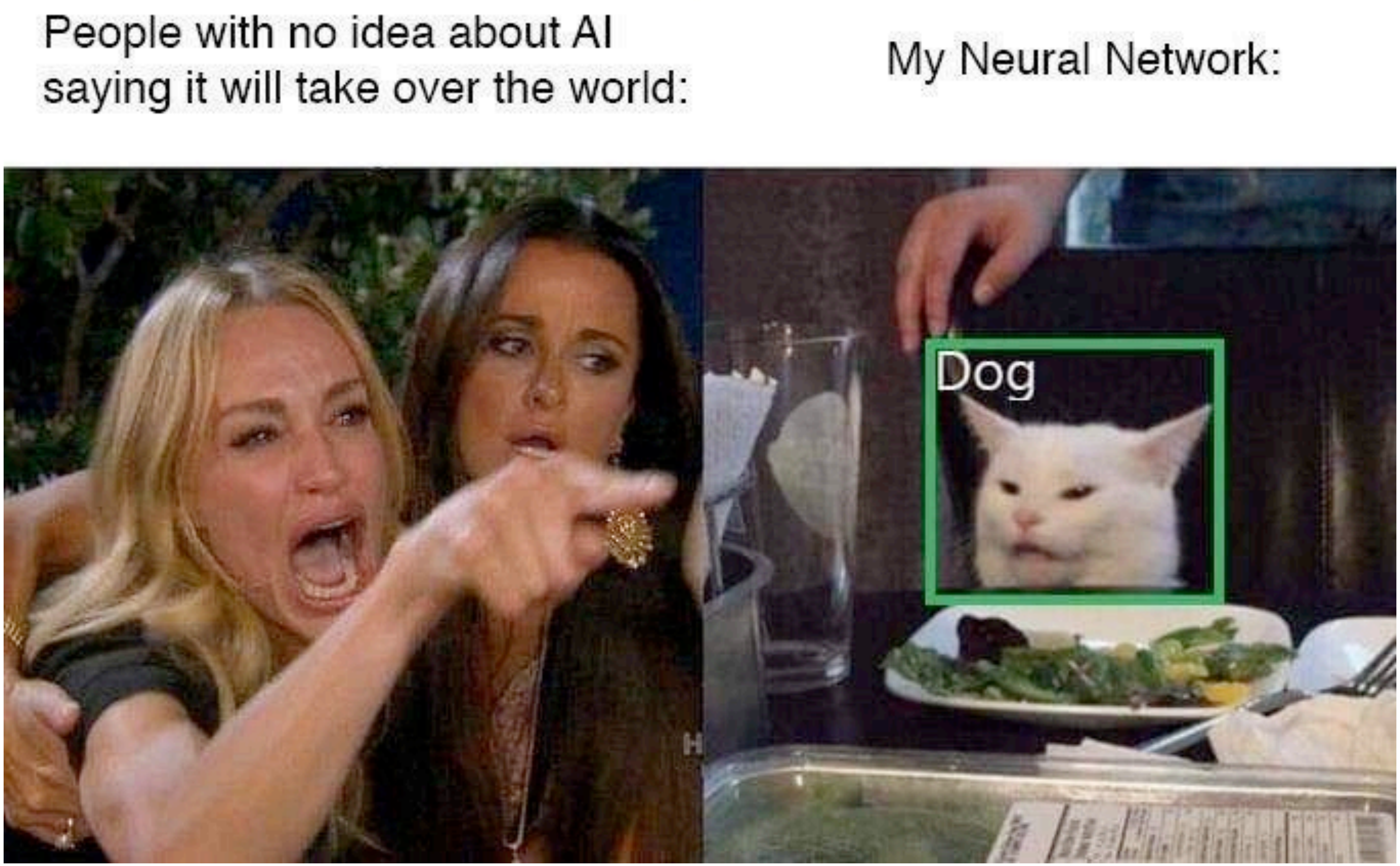
Connect with me:
Twitter: @ukmlv
Web: cigroup.wustl.edu
Email: kamilov@wustl.edu

Washington University in St. Louis

Machine learning will change **imaging technology**, but the current technology is still in its infancy

Source: Twitter

Machine learning will change imaging technology, but the current technology is still in its infancy



Source: Twitter

Machine learning will change imaging technology, but the current technology is still in its infancy

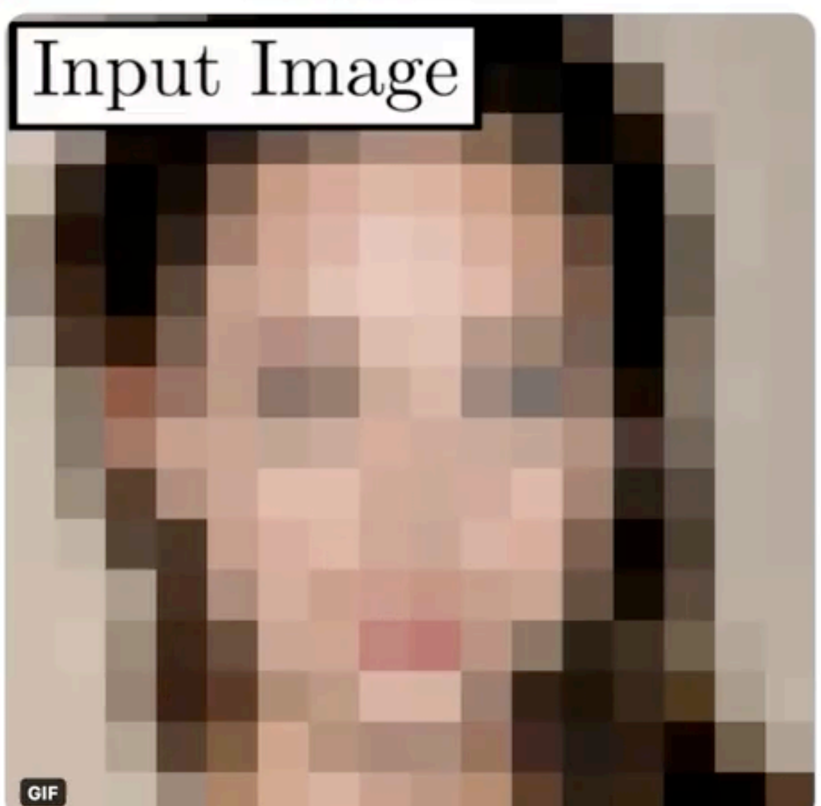
Bomze
@tg_bomze

Face Depixelizer

Given a low-resolution input image, model generates high-resolution images that are perceptually realistic and downscale correctly.

😺GitHub: github.com/tg-bomze/Face-...
📙Colab: colab.research.google.com/github/tg-bomz...

P.S. Colab is based on the
github.com/adamian98/pulse



4:56 PM · Jun 19, 2020 · Twitter Web App

Source: Twitter

Machine learning will change imaging technology, but the current technology is still in its infancy

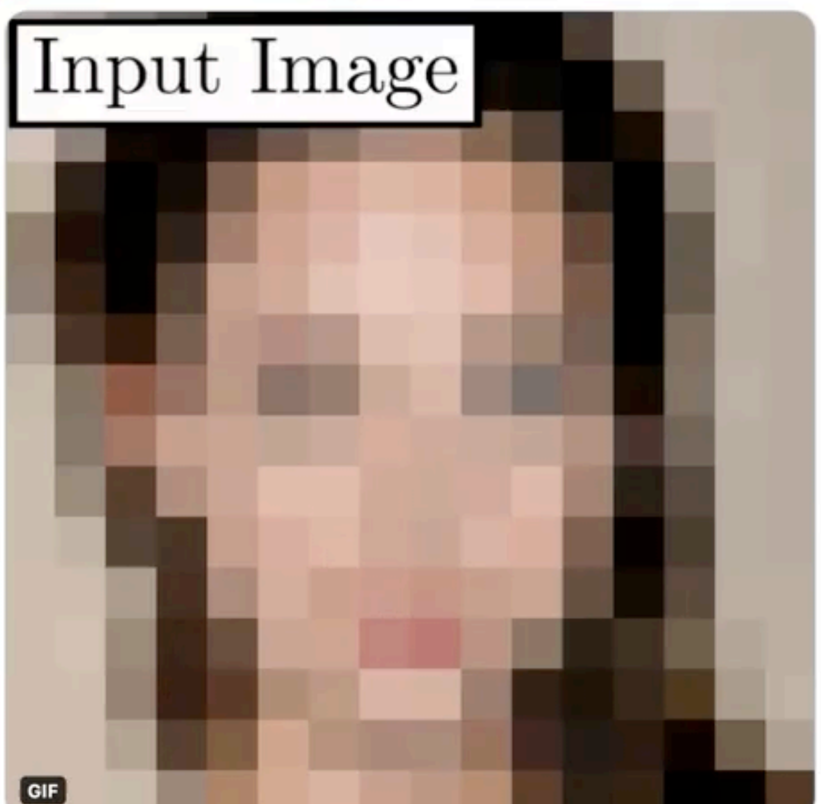
Bomze
@tg_bomze

Face Depixelizer

Given a low-resolution input image, model generates high-resolution images that are perceptually realistic and downscale correctly.

😺GitHub: github.com/tg-bomze/Face-...
📙Colab: colab.research.google.com/github/tg-bomz...

P.S. Colab is based on the github.com/adamian98/pulse



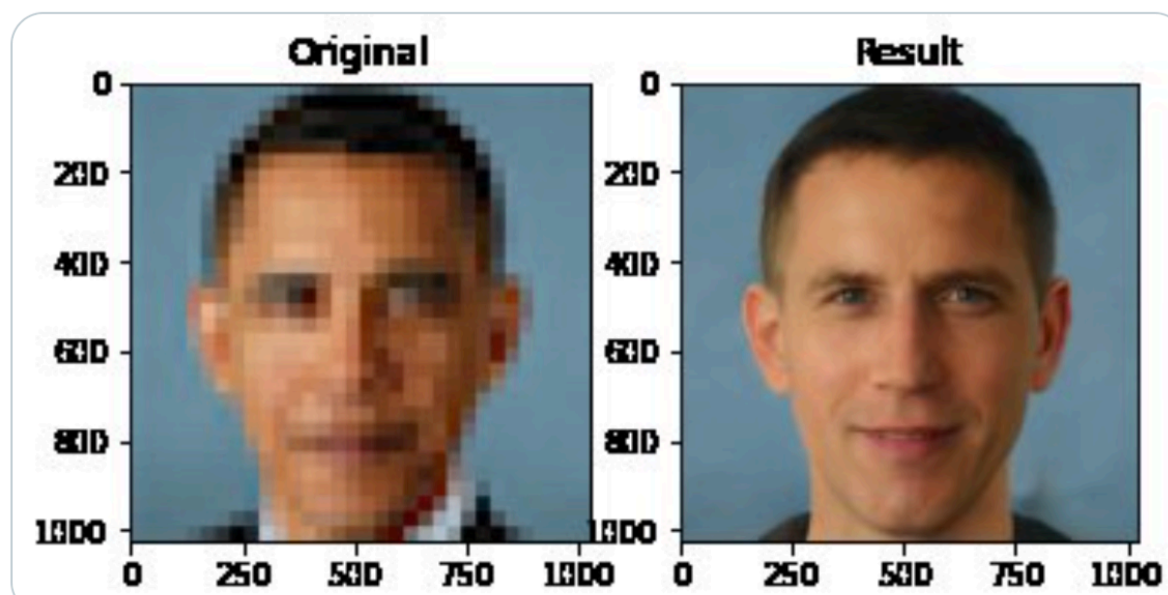
4:56 PM · Jun 19, 2020 · Twitter Web App

Counter example:

Chicken3gg
@Chicken3gg

Replying to @tg_bomze

🤔🤔🤔



7:14 AM · Jun 20, 2020 · Twitter for Android

Source: Twitter

Biomedical imaging is going through a paradigm shift driven by machine learning

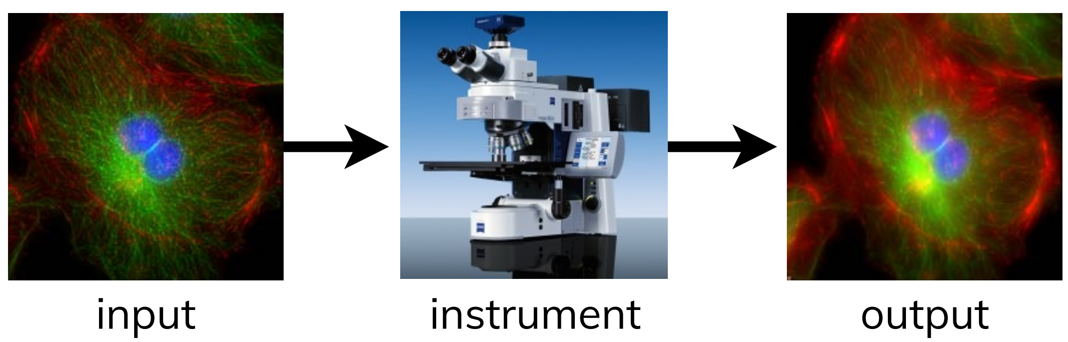
Biomedical imaging is going through a paradigm shift driven by machine learning

Past: Focus on hardware for image formation

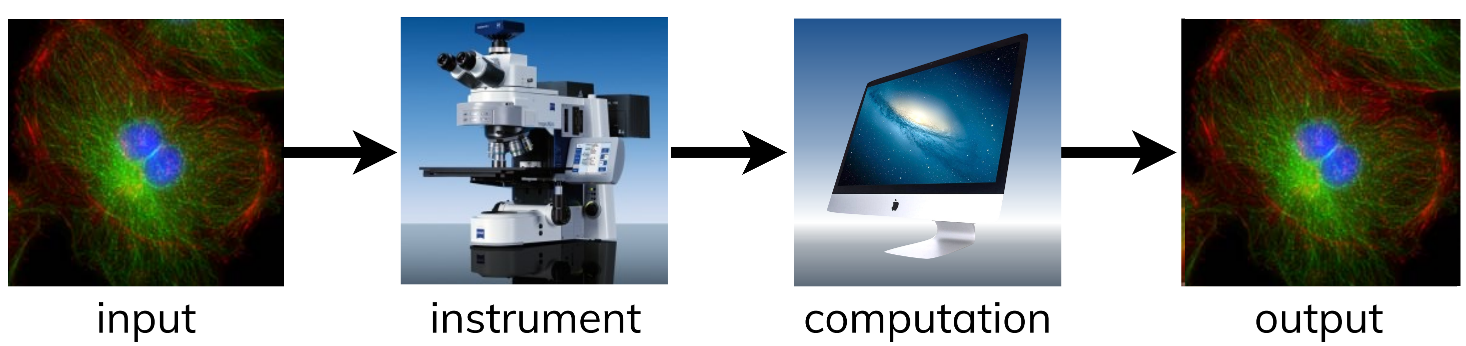
input instrument output

Biomedical imaging is going through a paradigm shift driven by machine learning

Past: Focus on hardware for image formation



Present: Use digital image processing for better performance



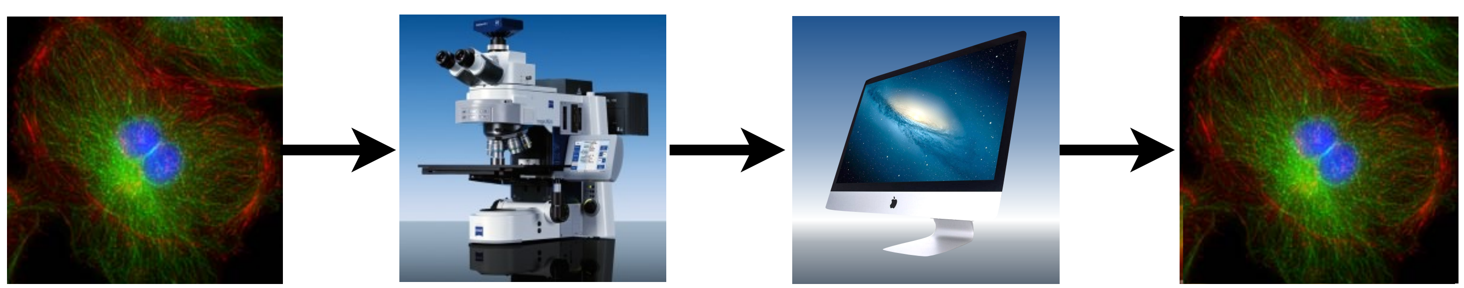
Washington University in St.Louis

Biomedical imaging is going through a paradigm shift driven by machine learning

Past: Focus on hardware for image formation

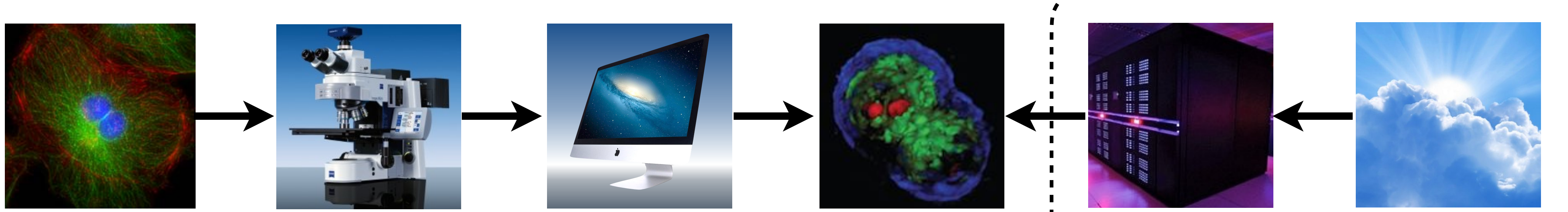
input instrument output

Present: Use digital image processing for better performance



input instrument computation output

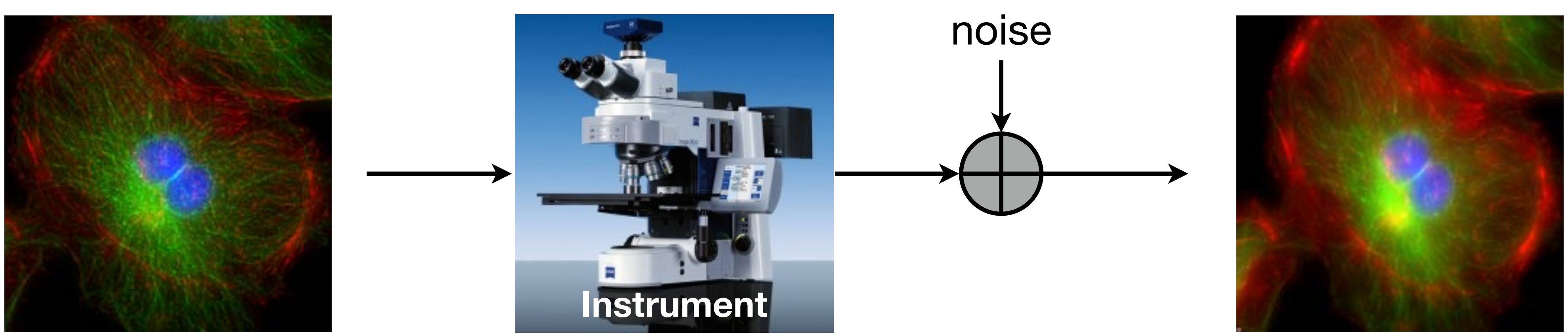
Future: AI for retrieving ***hidden*** information



input instrument computation output learning external data

The vast majority of imaging problems can be formulated as **inverse problems**

The vast majority of imaging problems can be formulated as inverse problems

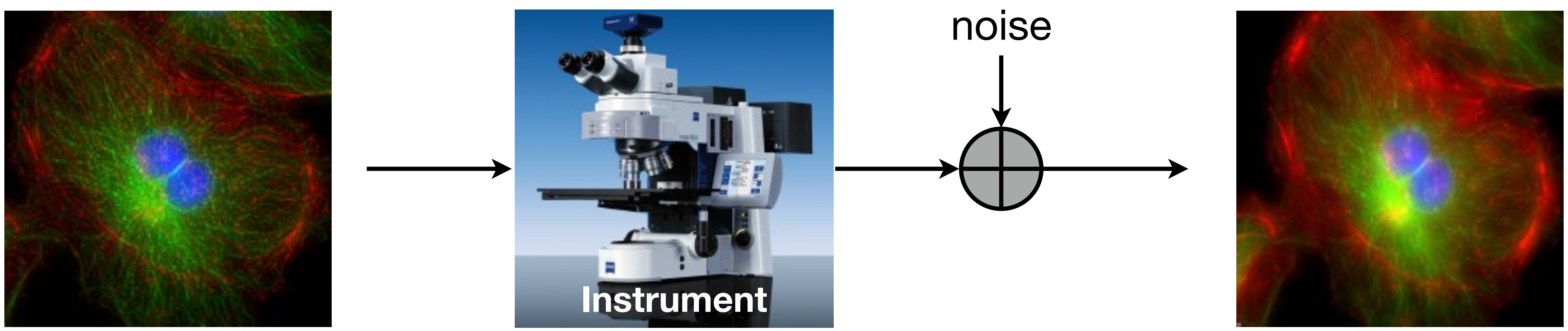


$\boldsymbol{x}$ $\qquad$ $\mathbf{H}$ $\qquad$ $\boldsymbol{y} = \mathbf{H}(\boldsymbol{x}) + \boldsymbol{e}$

Source: Michael Unser

The vast majority of imaging problems can be formulated as inverse problems

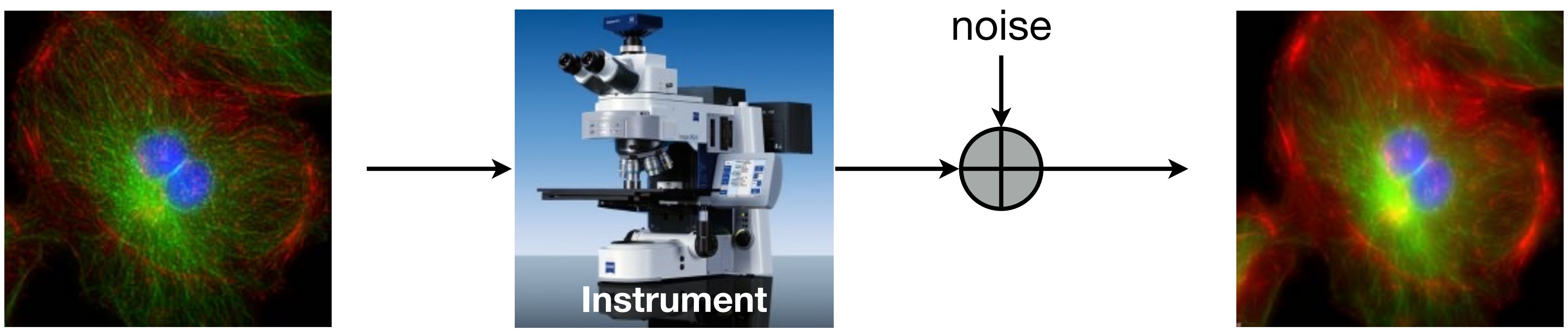
Forward problem: generate $\boldsymbol{y}$ from $\boldsymbol{x}$



$\boldsymbol{x}$ $\quad$ $\mathbf{H}$ $\quad$ $\boldsymbol{y} = \mathbf{H}(\boldsymbol{x}) + \boldsymbol{e}$

The vast majority of imaging problems can be formulated as inverse problems

Forward problem: generate $\boldsymbol{y}$ from $\boldsymbol{x}$



$\boldsymbol{x}$ $\qquad$ $\mathbf{H}$ $\qquad$ $\boldsymbol{y} = \mathbf{H}(\boldsymbol{x}) + \boldsymbol{e}$

Inverse problem: recover $\boldsymbol{x}$ from $\boldsymbol{y}$

Source: Michael Unser

The vast majority of imaging problems can be formulated as inverse problems

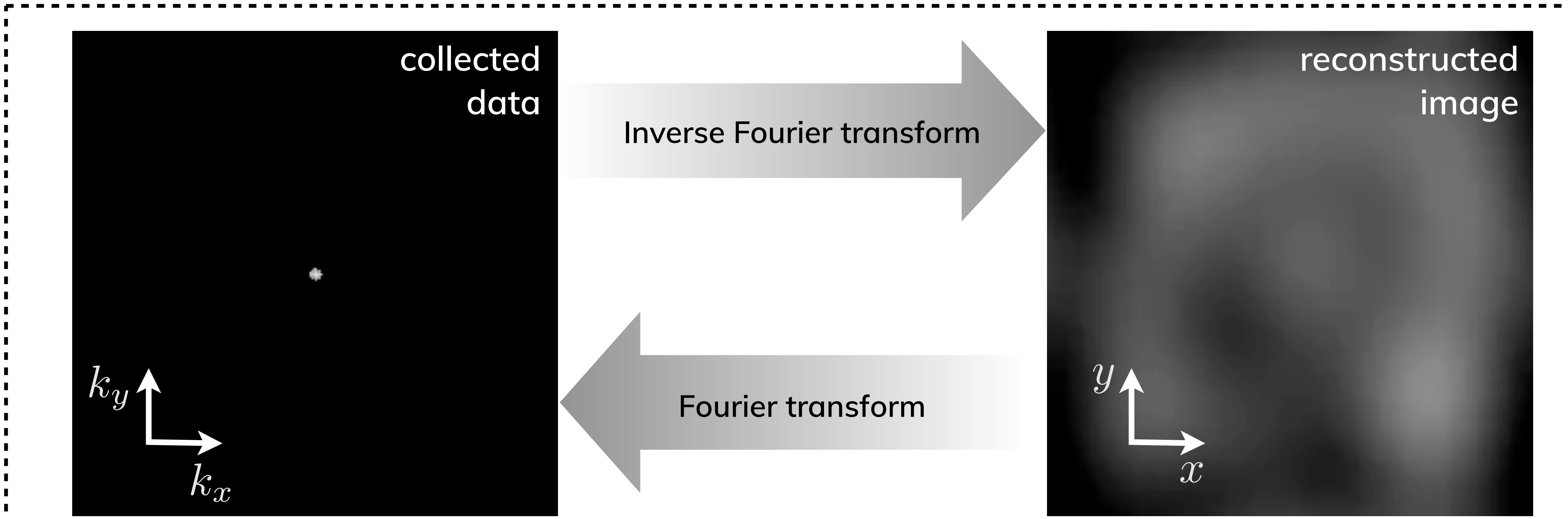
Imaging Problem	Light Source	Forward Model	Variations
2D or 3D tomography	coherent x-ray	$\boldsymbol{y}_i = \mathbf{R}_{\boldsymbol{\theta}_i}\boldsymbol{x}$	parallel, cone beam
3D deconvolution microscopy	fluorescence	$\boldsymbol{y} = \mathbf{H}\boldsymbol{x}$	brightfield, confocal, light sheet
structured illumination microscopy (SIM)	fluorescence	$\boldsymbol{y}_i = \mathbf{H}\mathbf{W}_i\boldsymbol{x}$	full 3D reconstruction, non-sinusoidal patters
positron emission tomography (PET)	gamma rays	$\boldsymbol{y}_i = \mathbf{H}_{\boldsymbol{\theta}_i}\boldsymbol{x}$	list mode with time-of-flight
magnetic resonance imaging (MRI)	radio frequency	$\boldsymbol{y} = \mathbf{S}\mathbf{F}\boldsymbol{x}$	uniform or nonuniform sampling in k-space
Cardiac MRI (parallel, nonuniform)	radio frequency	$\boldsymbol{y}_{t,i} = \mathbf{S}_t\mathbf{F}\mathbf{W}_i\boldsymbol{x}$	gated or nongated, retrospective registration
optical diffraction tomography (ODT)	coherent light	$\boldsymbol{y}_i = \mathbf{W}_i\mathbf{F}\boldsymbol{x}$	with holography or gating interferometry

Source: Michael Unser

Example: **Magnetic resonance imaging (MRI)** collects data in the spatial-frequency domain

Example: Magnetic resonance imaging (MRI) collects data in the spatial-frequency domain

Example: Magnetic resonance imaging (MRI) collects data in the spatial-frequency domain

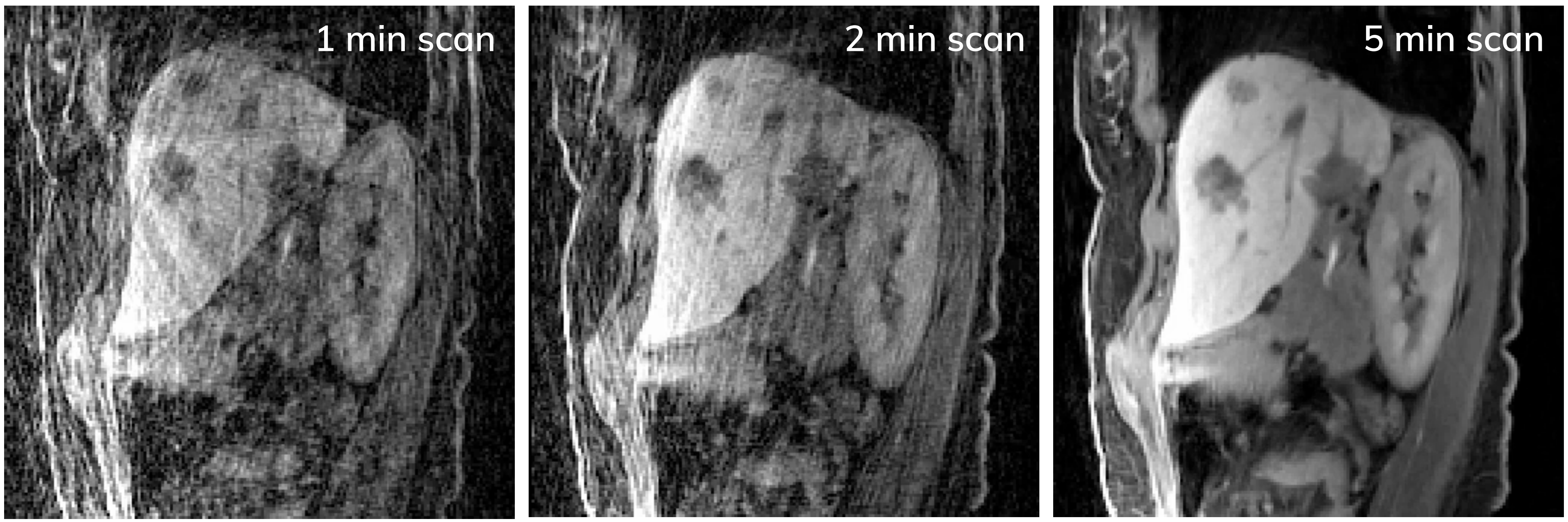


Source: Brian Hargreaves

Longer data collection leads to **better images**,
but it implies more **discomfort for patients**

Washington University in St. Louis

Example: Magnetic resonance imaging (MRI) collects data in the spatial-frequency domain

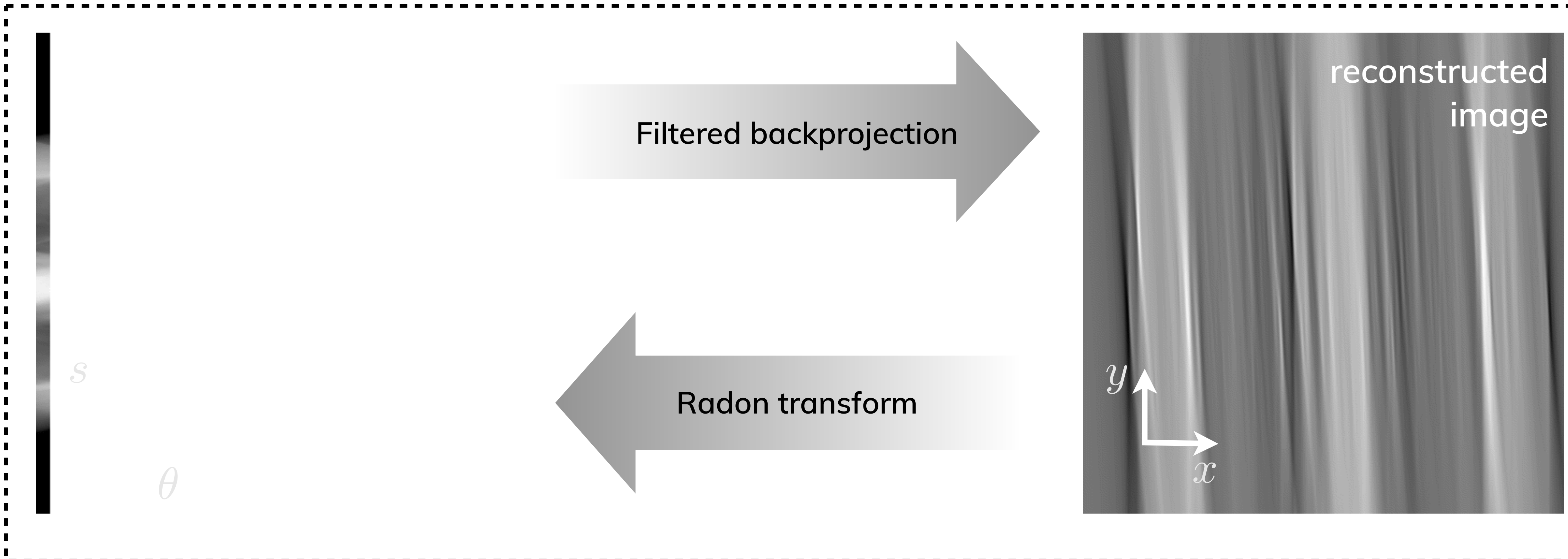


Longer data collection leads to **better images**,
but it implies more **discomfort for patients**

Liu *et al.*, “RARE: Image Reconstruction using Deep Priors Learned without Ground Truth,” *IEEE J. Sel. Topics Signal Process.*, October 2020

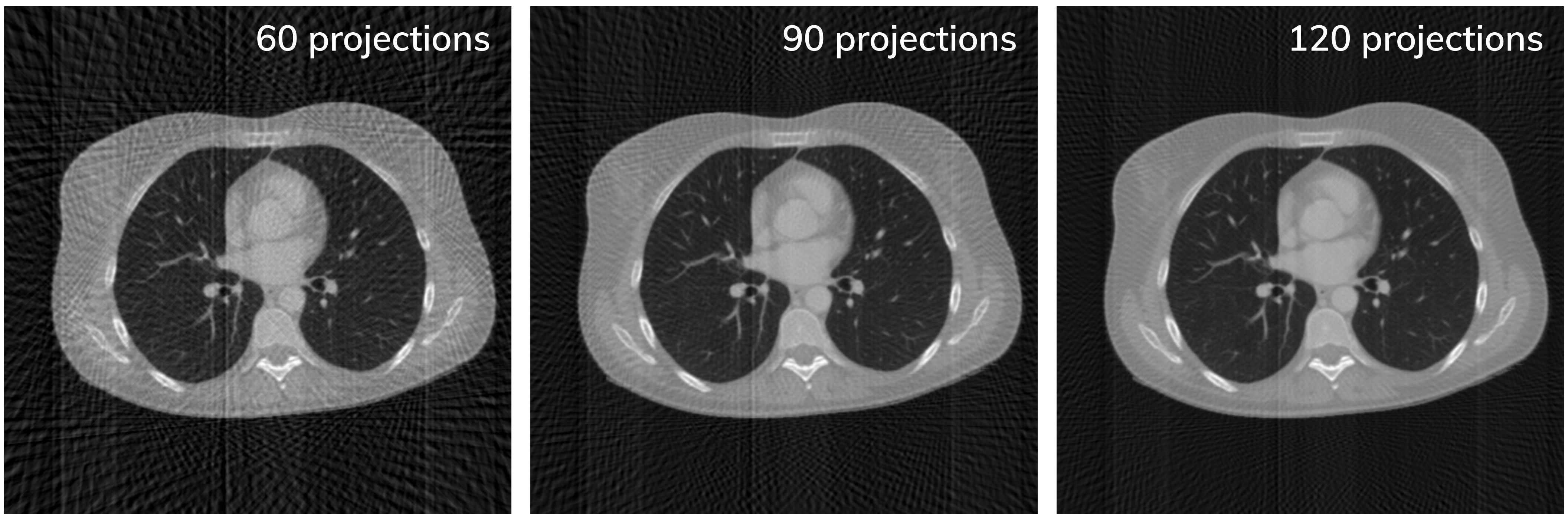
Example: **Tomographic imaging** forms a single volumetric image from multiple projections

Example: Tomographic imaging forms a single volumetric image from multiple projections



More projections lead to **better images**, but also **higher x-ray radiation dose**

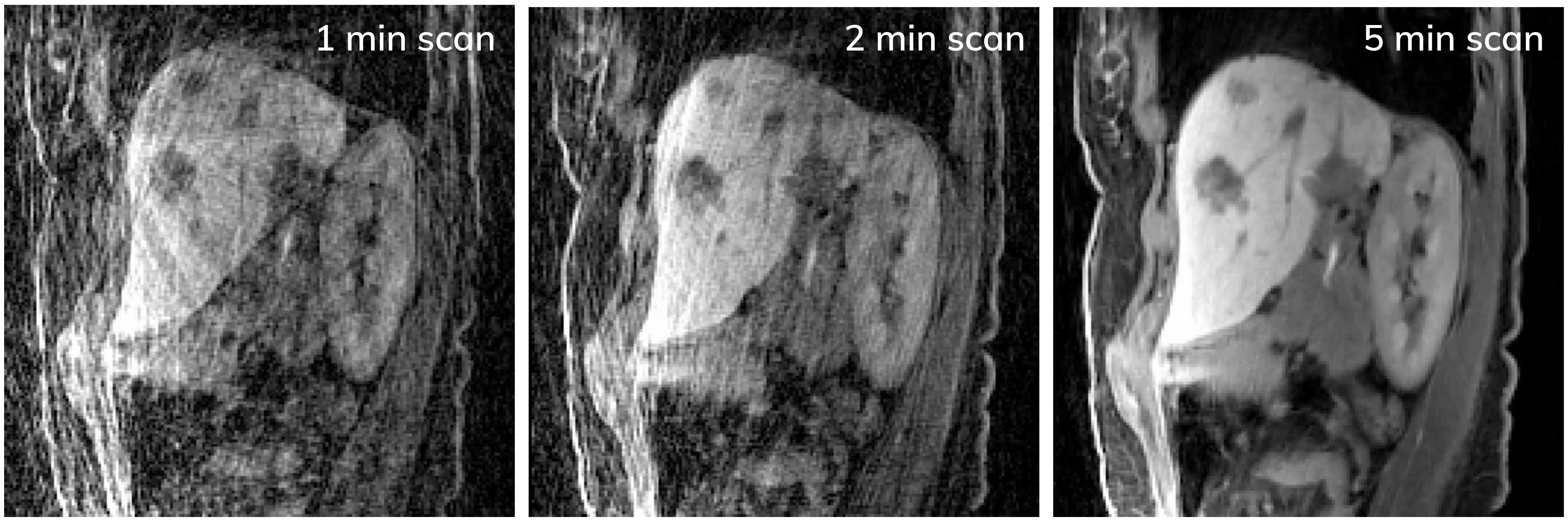
Example: Tomographic imaging forms a single volumetric image from multiple projections



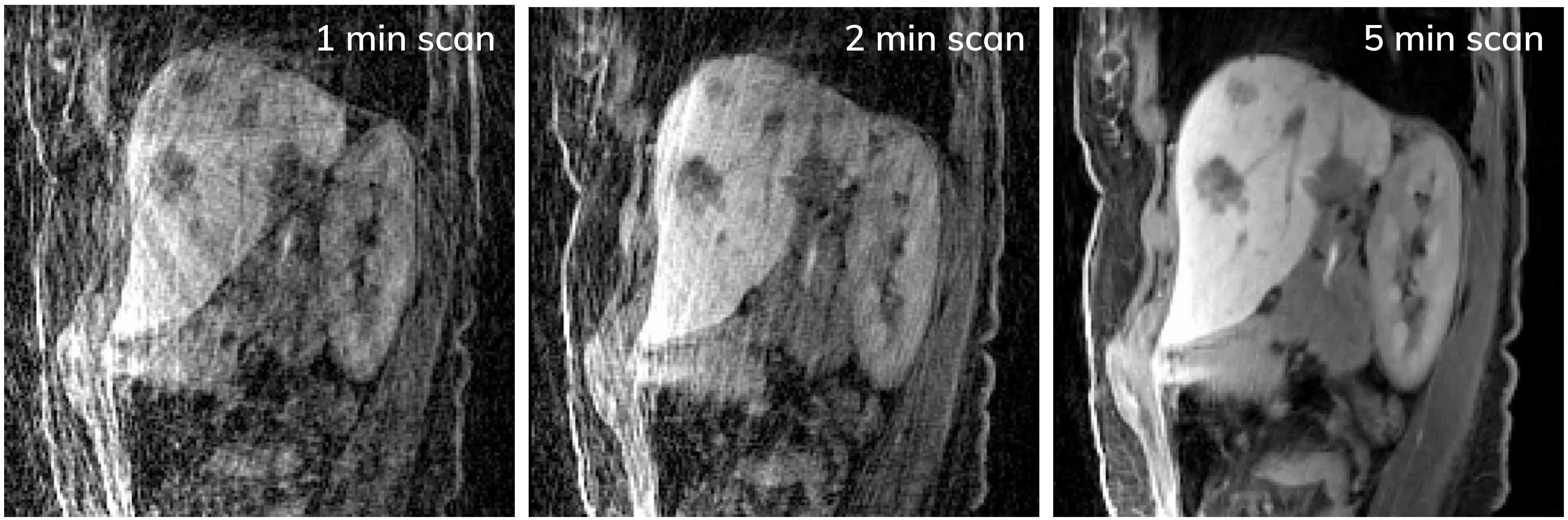
More projections lead to **better images**, but also **higher x-ray radiation dose**

Three challenges in biomedical imaging:
slow acquisition, imaging artifacts, and big data

Three challenges in biomedical imaging: slow acquisition, imaging artifacts, and big data

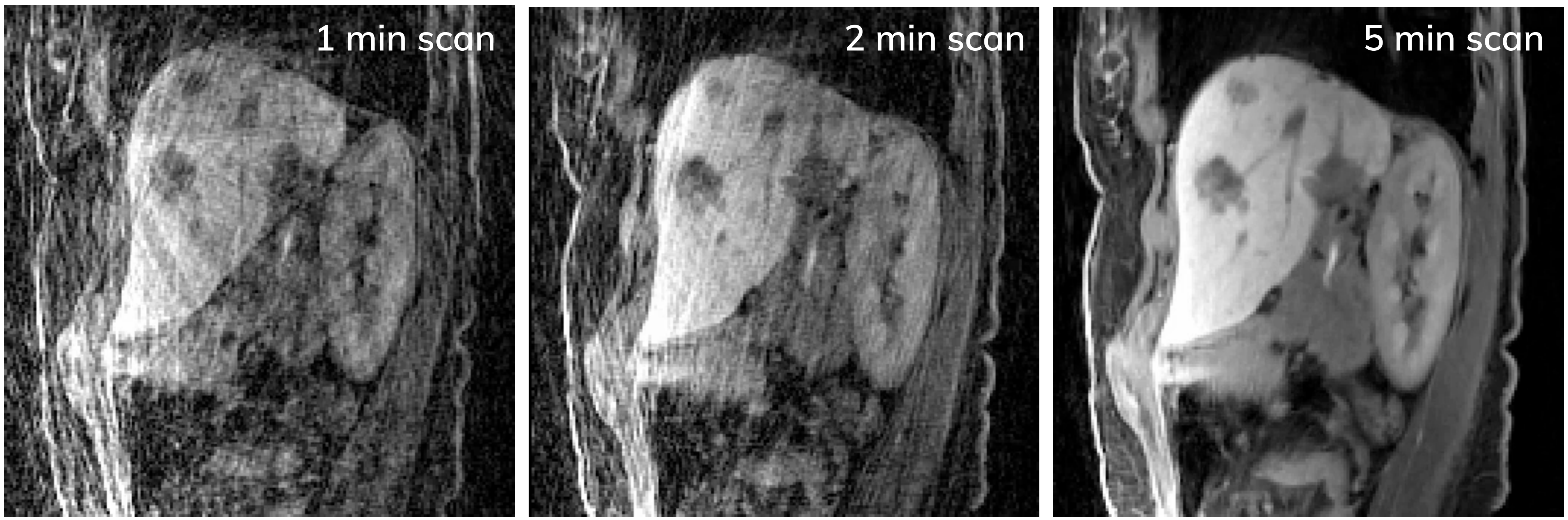


Three challenges in biomedical imaging: slow acquisition, imaging artifacts, and big data



Challenge #1: Acquisition is too slow for some applications
Due to sequential and indirect acquisition of data

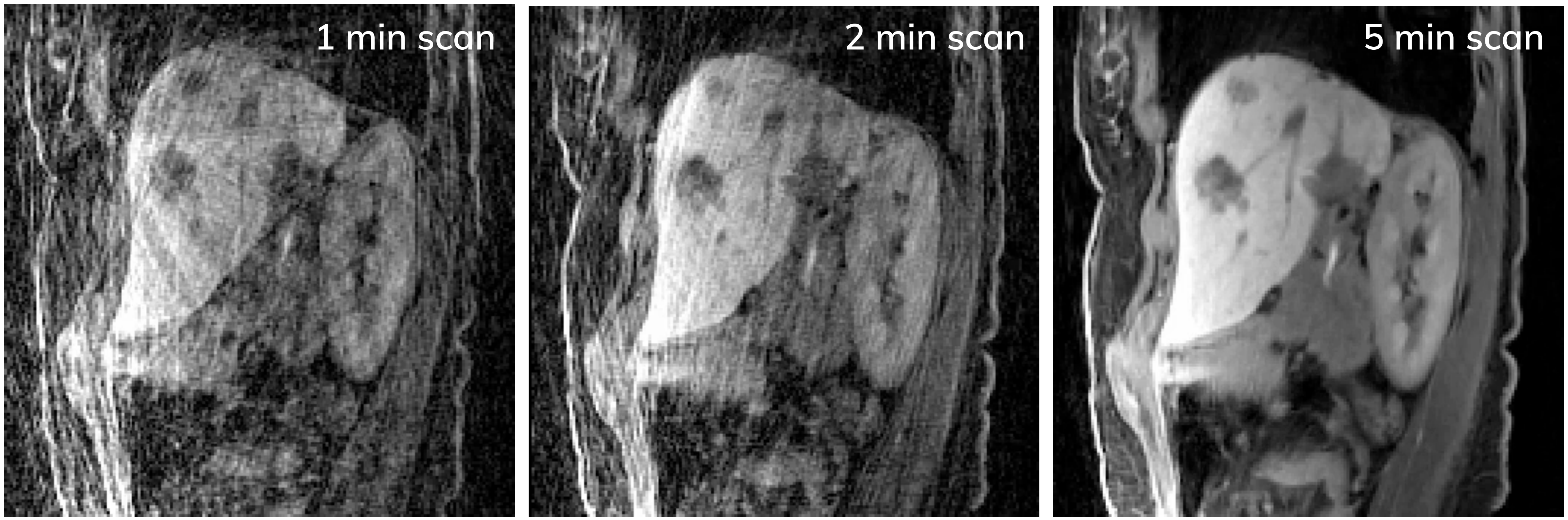
Three challenges in biomedical imaging: slow acquisition, imaging artifacts, and big data



Challenge #1: Acquisition is too slow for some applications
Due to sequential and indirect acquisition of data

Challenge #2: Reconstructed images contain artifacts:
Due to undersampling, model mismatch, and noise

Three challenges in biomedical imaging: slow acquisition, imaging artifacts, and big data



Challenge #1: Acquisition is too slow for some applications
Due to sequential and indirect acquisition of data

Challenge #2: Reconstructed images contain artifacts:
Due to undersampling, model mismatch, and noise

Challenge #3: High computational/memory requirements:
Due to large volumes of data to process

Outline for the rest of the talk

- Regularization by Artifact Removal (**RARE**)
  Integrating physical models and learned deep priors
- Efficient model-based deep learning (**SGD-Net**)
  Approximating physical layers for complexity gains

Outline for the rest of the talk

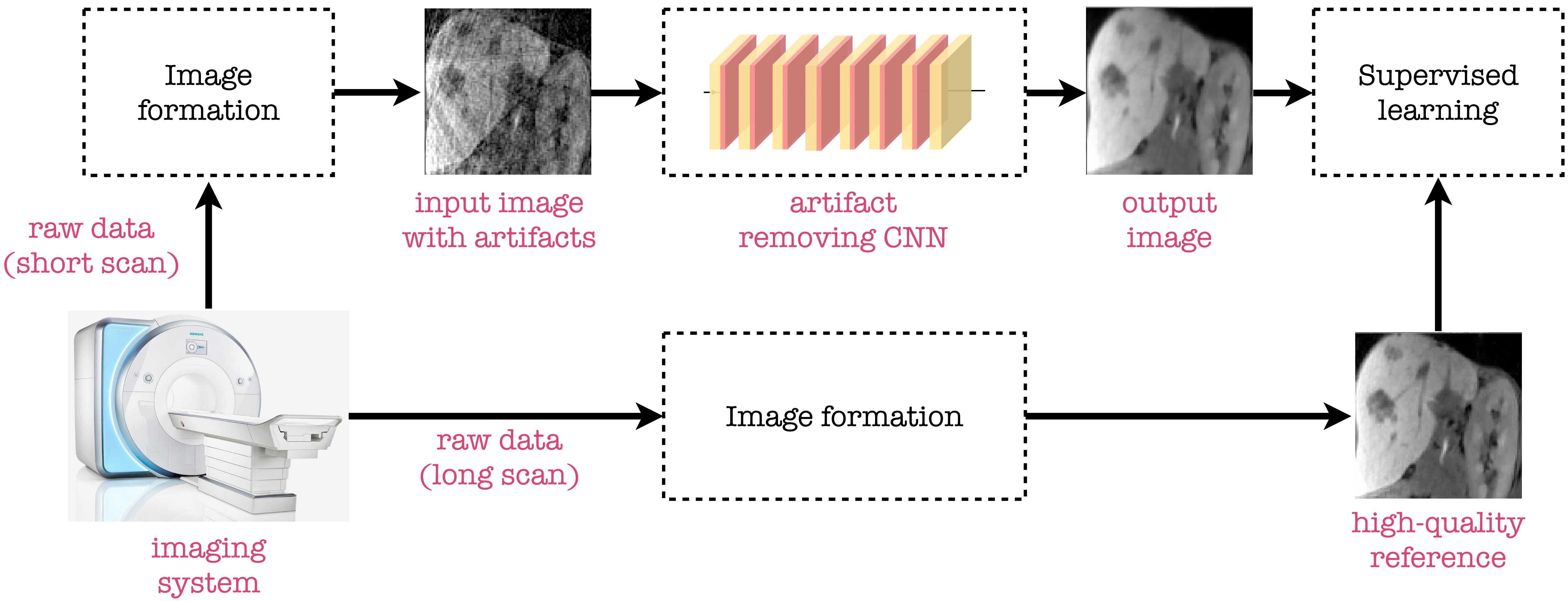
- Regularization by Artifact Removal (**RARE**)
  Integrating physical models and learned deep priors

- Efficient model-based deep learning (**SGD-Net**)
  Approximating physical layers for complexity gains

Simple recipe for DL-based image formation: Train a **supervised artifact-removing CNN**

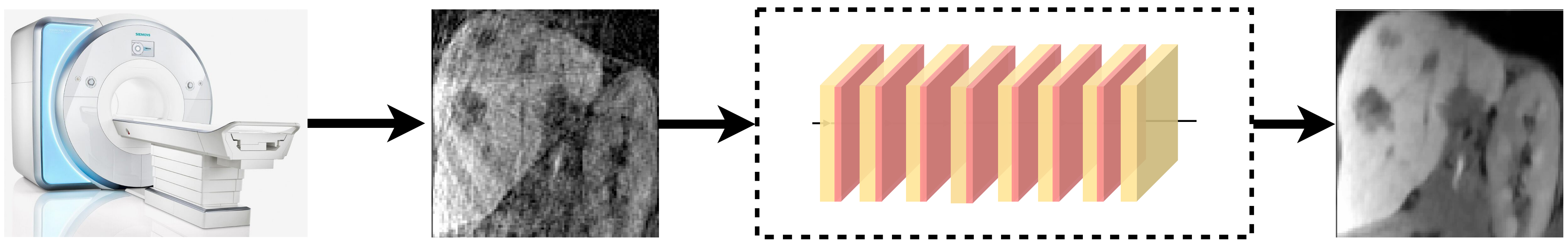
Simple recipe for DL-based image formation: Train a supervised artifact-removing CNN

Example: Train a deep neural net to remove artifacts from an image



Simple recipe for DL-based image formation: Train a supervised artifact-removing CNN

Example: Train a deep neural net to remove artifacts from an image



imaging system

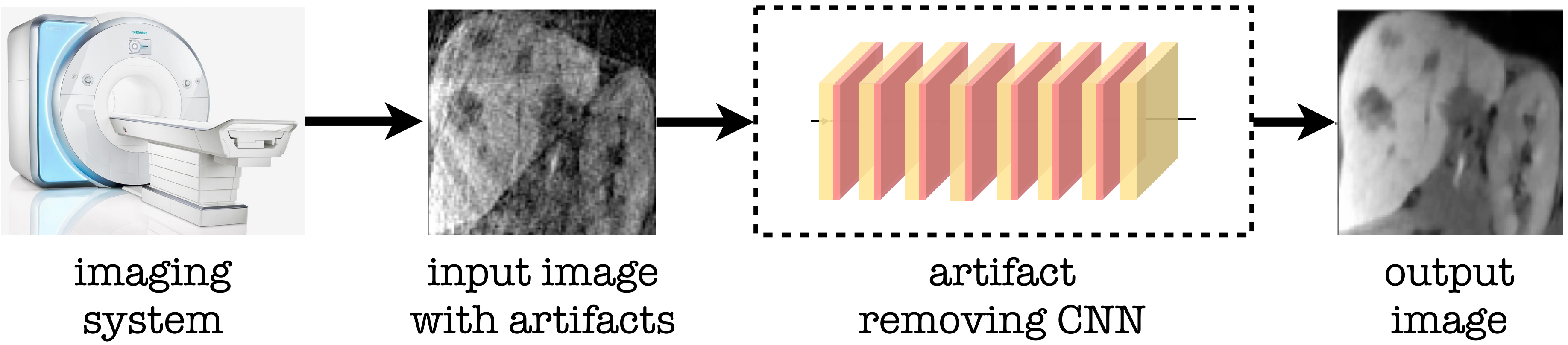
input image with artifacts

artifact removing CNN

output image

Simple recipe for DL-based image formation: Train a supervised artifact-removing CNN

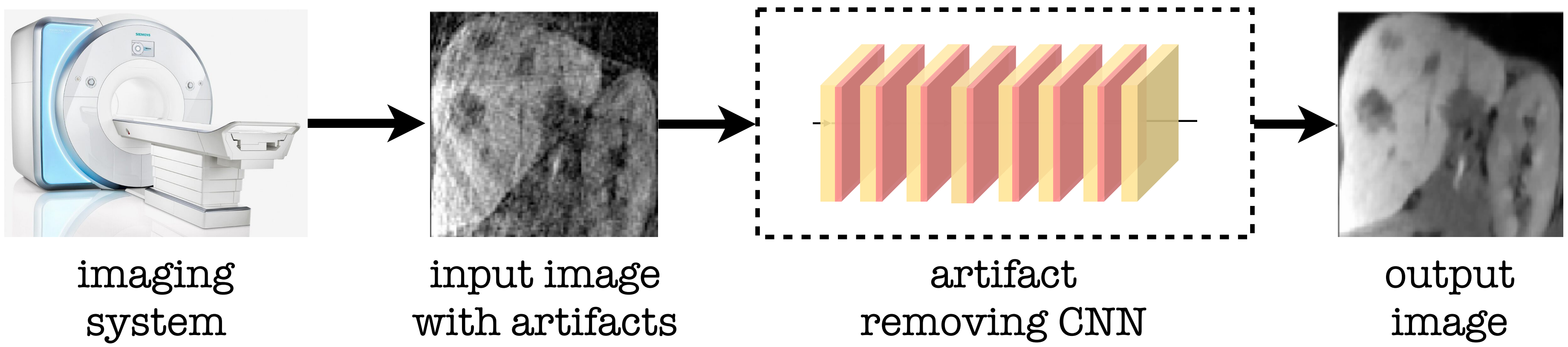
Example: Train a deep neural net to remove artifacts from an image



Question: What are some **advantages** of this approach?

Simple recipe for DL-based image formation: Train a supervised artifact-removing CNN

Example: Train a deep neural net to remove artifacts from an image



Question: What are some advantages of this approach?

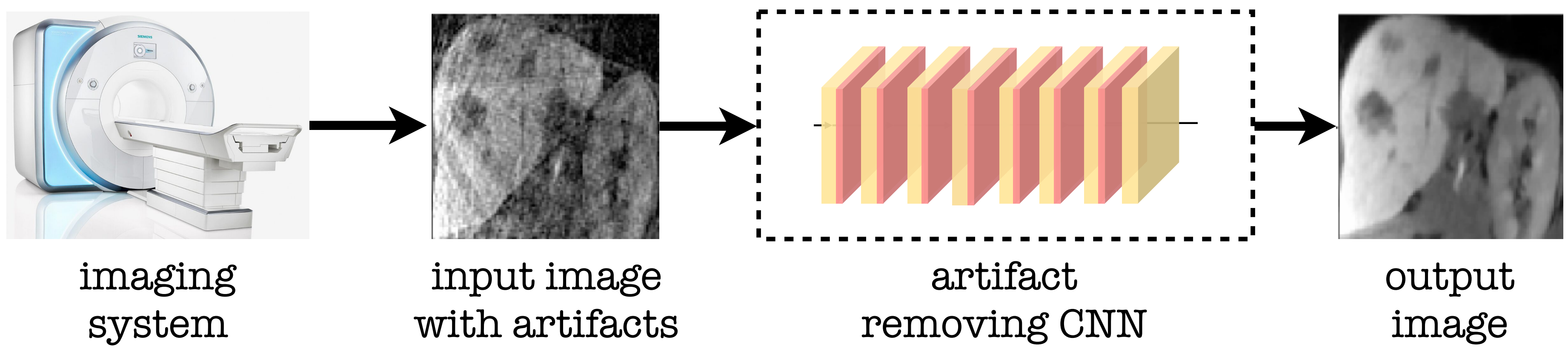
Advantage #1: Very easy to implement and deploy

Use existing deep learning frameworks and architectures

Washington University in St. Louis

Simple recipe for DL-based image formation: Train a supervised artifact-removing CNN

Example: Train a deep neural net to remove artifacts from an image



Question: What are some advantages of this approach?

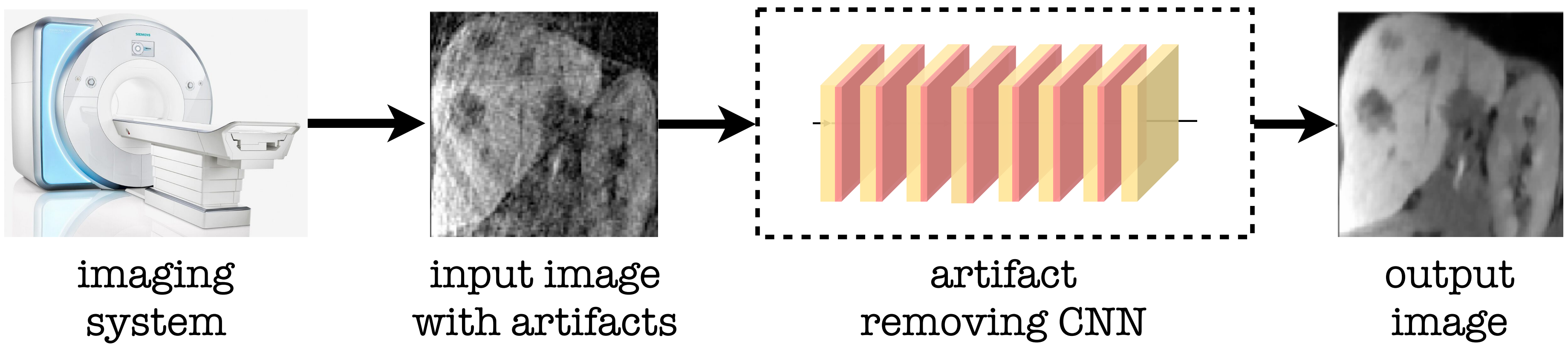
Advantage #1: Very easy to implement and deploy

Advantage #2: Very fast at test time

Simple pass through CNN (seconds) vs. optimization (hours)

Simple recipe for DL-based image formation: Train a supervised artifact-removing CNN

Example: Train a deep neural net to remove artifacts from an image



Question: What are some advantages of this approach?

Advantage #1: Very easy to implement and deploy

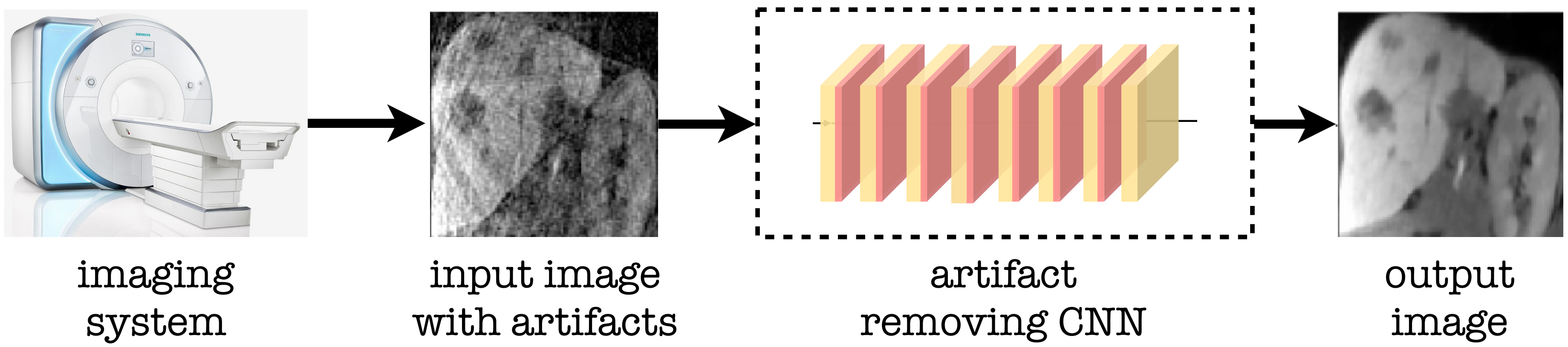
Advantage #2: Very fast at test time

Advantage #3: No need to explicitly model anything

Everything is learned automatically from data

Simple recipe for DL-based image formation: Train a supervised artifact-removing CNN

Example: Train a deep neural net to remove artifacts from an image



Question: What are some advantages of this approach?

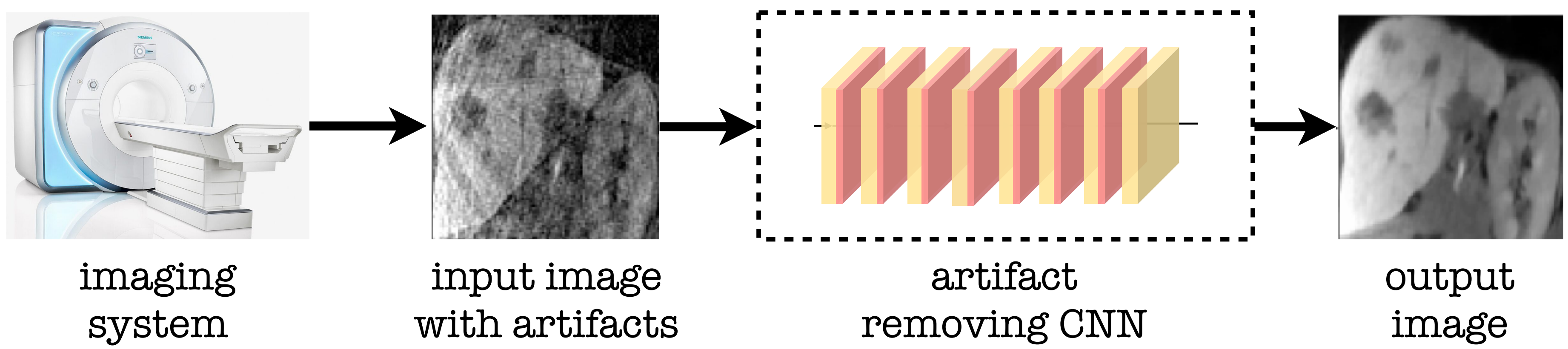
Advantage #1: Very easy to implement and deploy

Advantage #2: Very fast at test time

Advantage #3: No need to explicitly model anything

Simple recipe for DL-based image formation: Train a supervised artifact-removing CNN

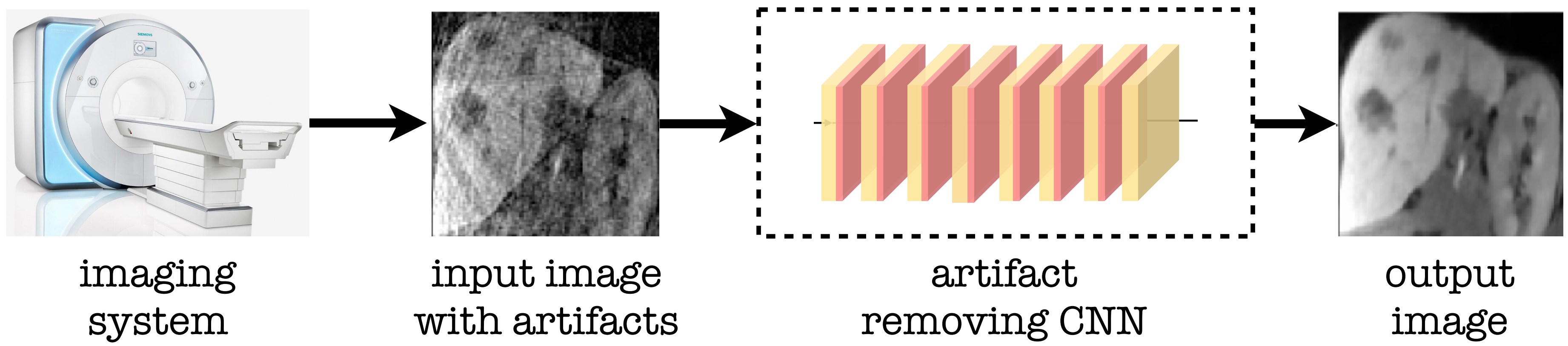
Example: Train a deep neural net to remove artifacts from an image



Question: What are some **limitations** of this approach?

Simple recipe for DL-based image formation: Train a supervised artifact-removing CNN

Example: Train a deep neural net to remove artifacts from an image



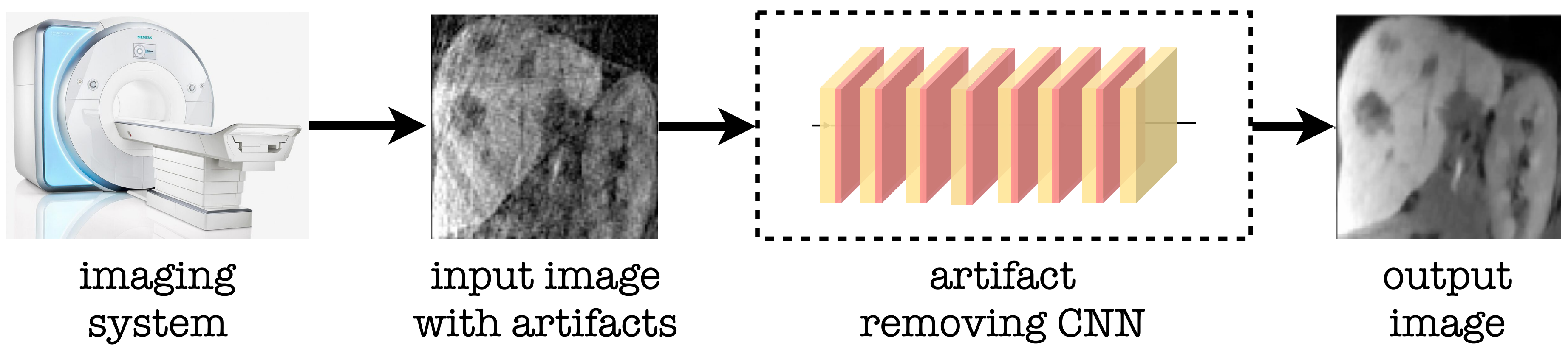
Question: What are some limitations of this approach?

Limitation #1: Does not allow for efficient **model adaptation**

One must **retrain** the model, which is computationally expensive!

Simple recipe for DL-based image formation: Train a supervised artifact-removing CNN

Example: Train a deep neural net to remove artifacts from an image



Question: What are some limitations of this approach?

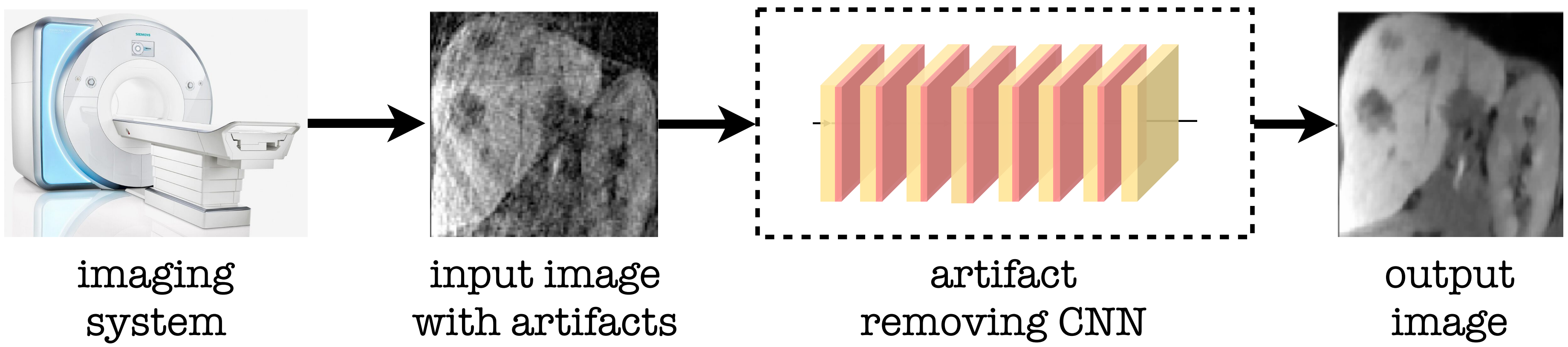
Limitation #1: Does not allow for efficient model adaptation

Limitation #2: Needs **ground truth** for training the CNN

Limits applicability to some important imaging problems

Simple recipe for DL-based image formation: Train a supervised artifact-removing CNN

Example: Train a deep neural net to remove artifacts from an image



Question: What are some limitations of this approach?

Limitation #1: Does not allow for efficient model adaptation

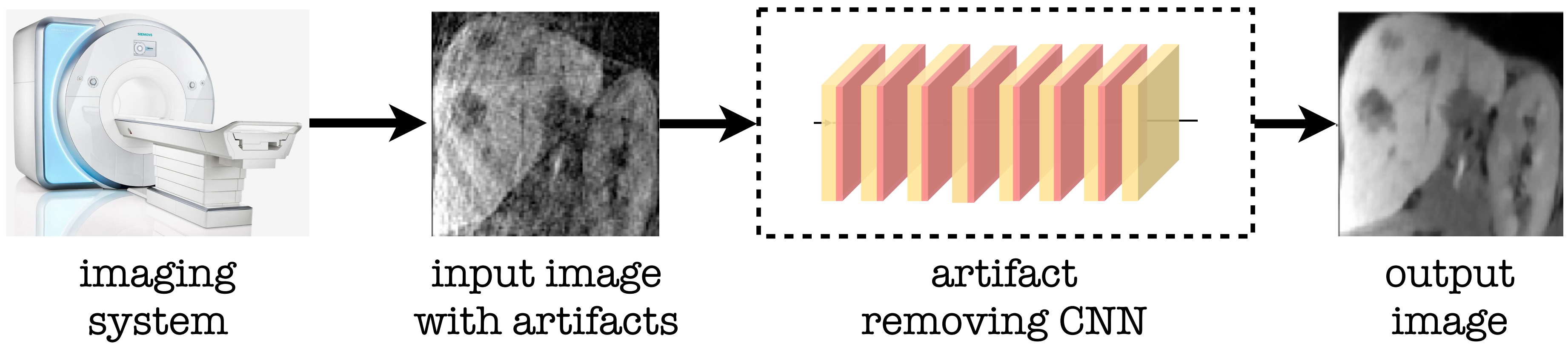
Limitation #2: Needs ground truth for training the CNN

Limitation #3: Does not exploit known **physical models**

Why re-learn something we know?

Simple recipe for DL-based image formation: Train a supervised artifact-removing CNN

Example: Train a deep neural net to remove artifacts from an image



Question: What are some limitations of this approach?

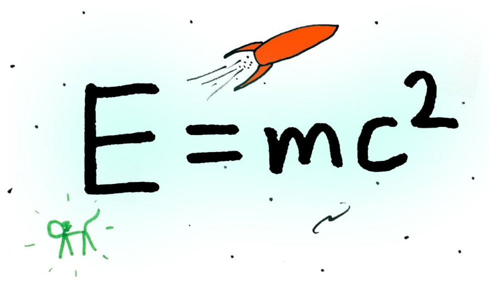
Limitation #1: Does not allow for efficient model adaptation

Limitation #2: Needs ground truth for training the CNN

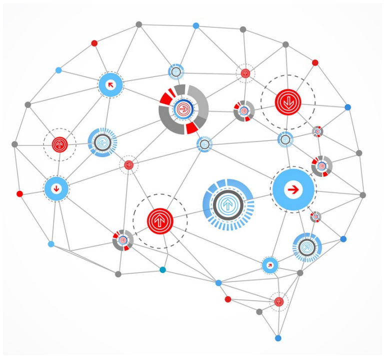
Limitation #3: Does not exploit known physical models

Idea: Address limitations by combining **model-based optimization** and **deep learning**

Idea: Address limitations by combining model-based optimization and deep learning



Models



Learning

Idea: Address limitations by combining model-based optimization and deep learning

Regularization by denoising [Romano'17]

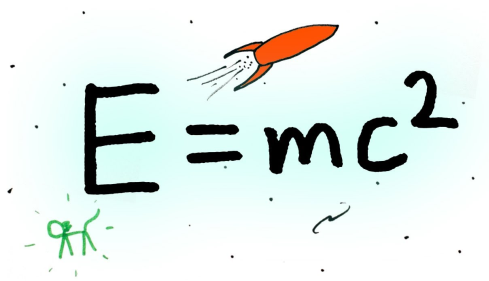
Regularization by artifact removal [Liu'20]

Consensus Equilibrium [Buzzard'19]

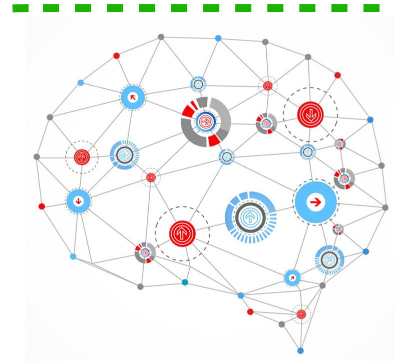
Plug-and-play Priors [Venkat'13]

SIMBA [Wu'20]

PnP-FISTA [Kamilov'17]



Models



Learning

Idea: Address limitations by combining model-based optimization and deep learning

Regularization by denoising
[Romano'17]

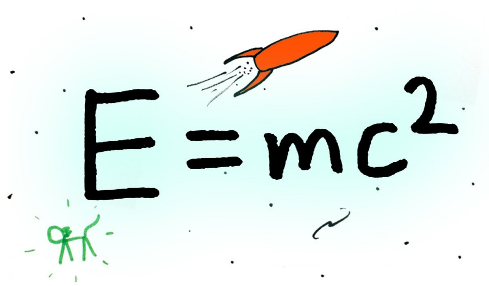
Regularization by artifact removal
[Liu'20]

Consensus Equilibrium
[Buzzard'19]

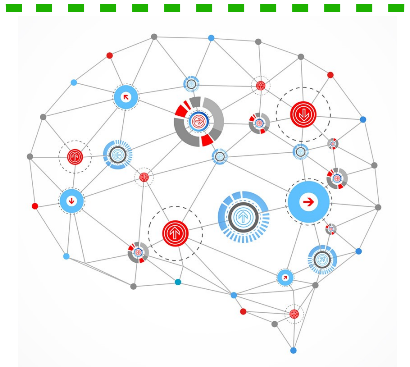
Plug-and-play Priors
[Venkat'13]

SIMBA
[Wu'20]

PnP-FISTA
[Kamilov'17]



Models



Learning

Neural Sensors
[Martel'20]

TNRD
[Chen'15]

Learning Tomo
[Kamilov'15]

LISTA
[Gregor'10]

SGD-Net
[Liu'21]

MoDL
[Aggarwal'19]

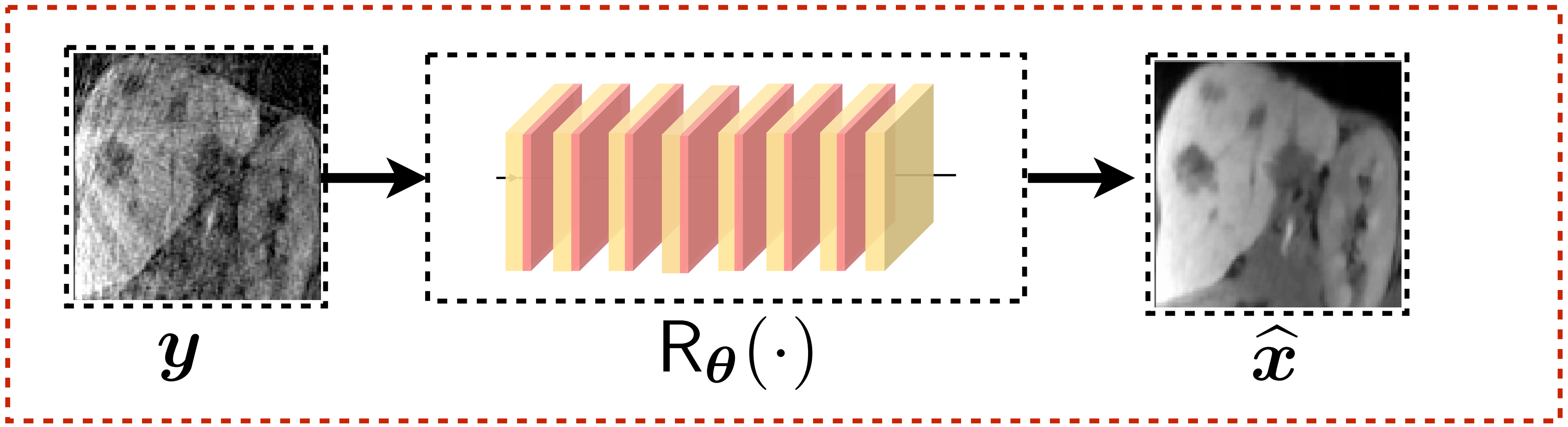
Regularization by Artifact Removal (RARE) uses artifact removing deep networks as image priors

Regularization by Artifact Removal (RARE) uses artifact removing deep networks as image priors

Liu *et al.*, “RARE: Image Reconstruction using Deep Priors Learned without Ground Truth,” *IEEE J. Sel. Topics Signal Process.*, October 2020

Regularization by Artifact Removal (RARE) uses artifact removing deep networks as image priors

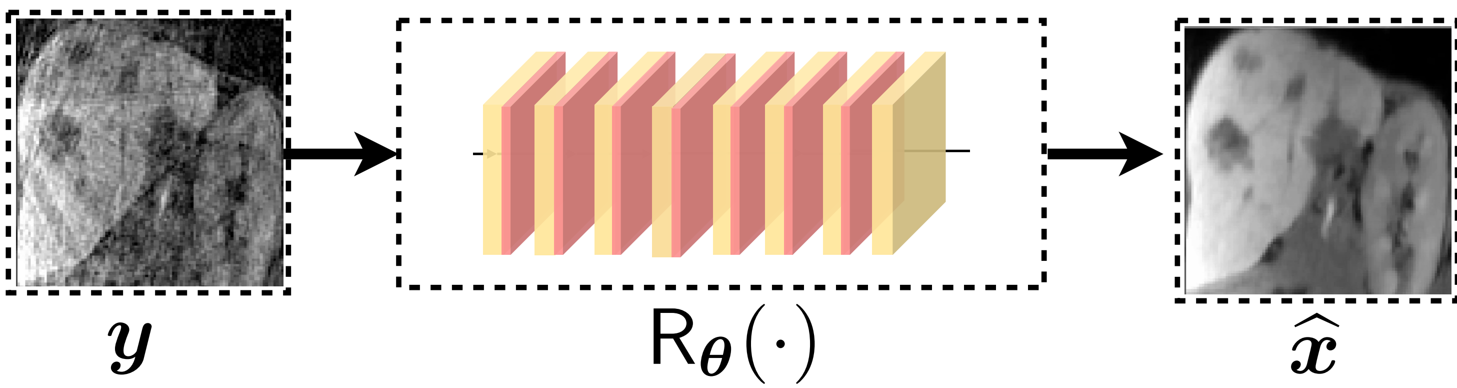
Idea: Pre-train an artifact removing CNN on a dataset of images



Liu *et al.*, "RARE: Image Reconstruction using Deep Priors Learned without Ground Truth," *IEEE J. Sel. Topics Signal Process.*, October 2020

Regularization by Artifact Removal (RARE) uses artifact removing deep networks as image priors

Idea: Pre-train an artifact removing CNN on a dataset of images



Combine the CNN with the **physical-model of the instrument**

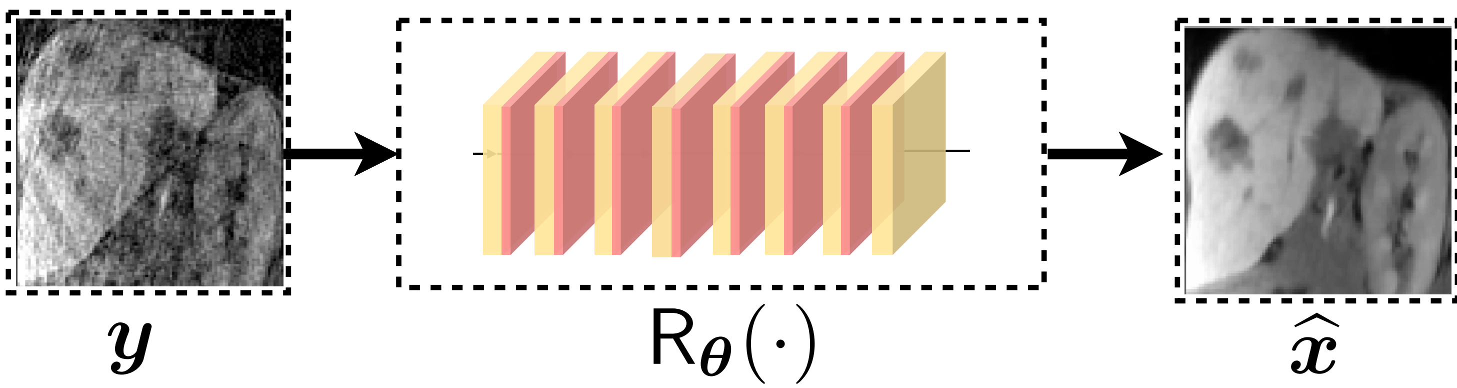
Example data fit: $g_{\boldsymbol{\phi}}(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{H}_{\boldsymbol{\phi}}\boldsymbol{x}\|_2^2$

Includes the physical model of the imaging instrument

The **data-fidelity term** measures the distance between the actual measurements and predicted ones!

Regularization by Artifact Removal (RARE) uses artifact removing deep networks as image priors

Idea: Pre-train an artifact removing CNN on a dataset of images



Combine the CNN with the **physical-model of the instrument**

Question: How an we use information from both the data-fidelity term and the CNN?

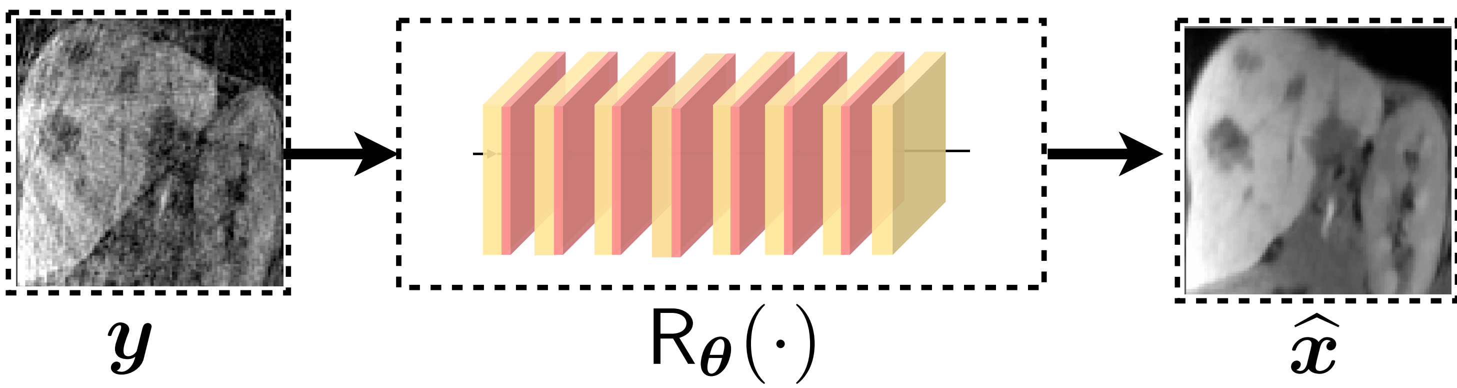
Example data fit: $g_{\boldsymbol{\phi}}(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{H}_{\boldsymbol{\phi}}\boldsymbol{x}\|_2^2$

$\boldsymbol{\phi}$ = parameters of the measurement model

$\boldsymbol{\theta}$ = parameters of the CNN

Regularization by Artifact Removal (RARE) uses artifact removing deep networks as image priors

Idea: Pre-train an artifact removing CNN on a dataset of images



RARE #1: Based on **Regularization by Denoising (RED)**

$$\boldsymbol{x}^t \leftarrow \boldsymbol{x}^{t-1} - \gamma \mathsf{G}(\boldsymbol{x}^{t-1})$$

"gradient" descent

$$\mathsf{G}(\boldsymbol{x}) := \nabla g_{\boldsymbol{\phi}}(\boldsymbol{x}) + \tau(\boldsymbol{x} - \mathsf{R}_{\boldsymbol{\theta}}(\boldsymbol{x}))$$

improve data fit

reduce image artifacts

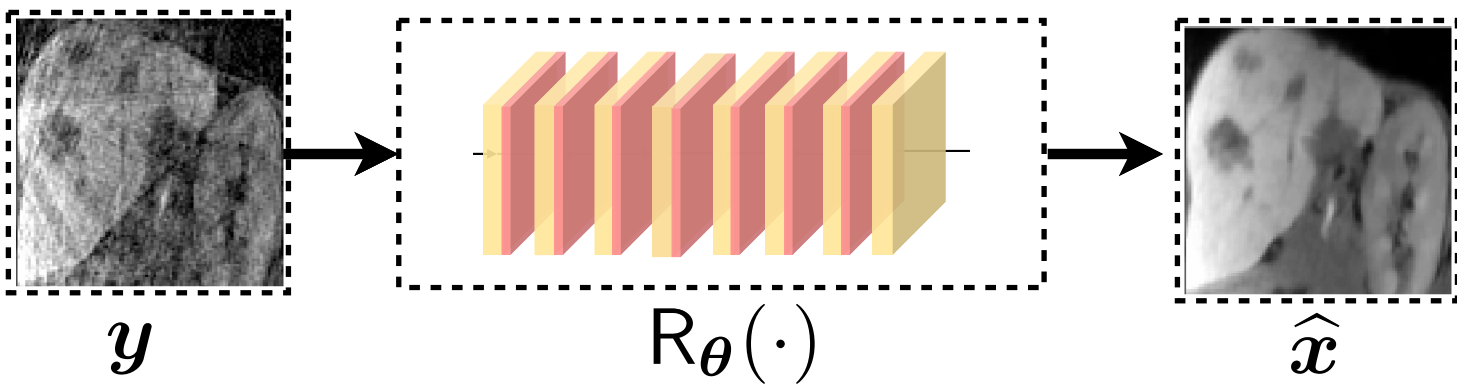
Example data fit: $g_{\boldsymbol{\phi}}(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{H}_{\boldsymbol{\phi}}\boldsymbol{x}\|_2^2$

Romano *et al.*, "The Little Engine That Could: Regularization by Denoising," *SIAM J. Imaging Sci.*, vol. 10, no. 4, 2017

Washington University in St. Louis

Regularization by Artifact Removal (RARE) uses artifact removing deep networks as image priors

Idea: Pre-train an artifact removing CNN on a dataset of images



RARE #1: Based on Regularization by Denoising (RED)

$$\boldsymbol{x}^t \leftarrow \boldsymbol{x}^{t-1} - \gamma \mathsf{G}(\boldsymbol{x}^{t-1}) \qquad \mathsf{G}(\boldsymbol{x}) := \nabla g_{\boldsymbol{\phi}}(\boldsymbol{x}) + \tau(\boldsymbol{x} - \mathsf{R}_{\boldsymbol{\theta}}(\boldsymbol{x}))$$

RARE #2: Based on **Plug-and-Play Priors (PnP)**

$$\boldsymbol{x}^t \leftarrow \mathsf{R}_{\boldsymbol{\theta}}(\boldsymbol{z}^t)$$ reduce image artifacts

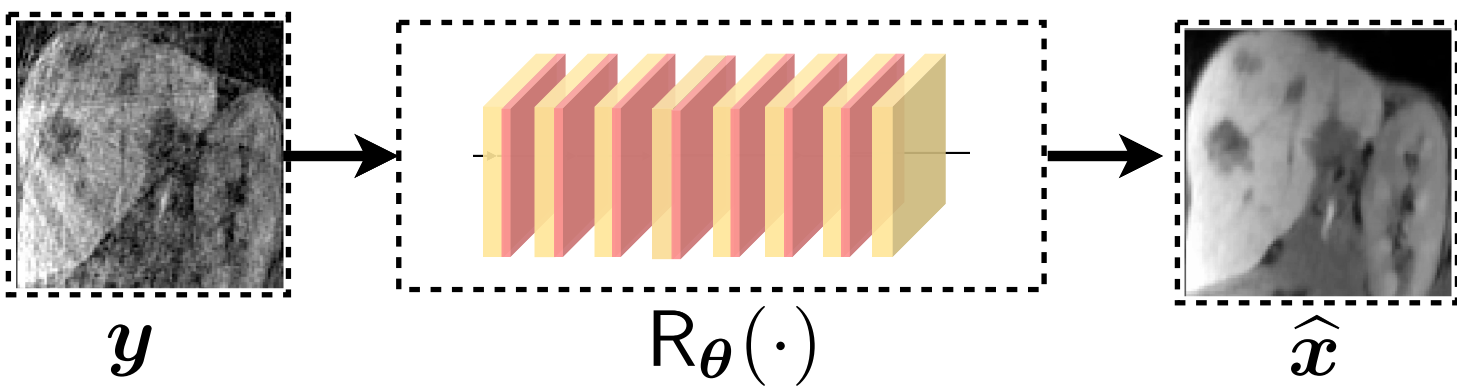
$$\boldsymbol{z}^t \leftarrow \boldsymbol{x}^{t-1} - \gamma \nabla g_{\boldsymbol{\phi}}(\boldsymbol{x}^{t-1})$$ improve data fit

Venkatakrishnan *et al.*, "Plug-and-Play Priors for Model Based Reconstruction," *Proc. IEEE GlobalSIP*, pp. 945-948, December 2013.

Kamilov *et al.*, "A Plug-and-Play Priors Approach for Solving Nonlinear Imaging Inverse Problems," *IEEE Signal Process. Lett.*, 2017

Regularization by Artifact Removal (RARE) uses artifact removing deep networks as image priors

Idea: Pre-train an artifact removing CNN on a dataset of images



RARE #1: Based on Regularization by Denoising (RED)

$$\boldsymbol{x}^t \leftarrow \boldsymbol{x}^{t-1} - \gamma \mathsf{G}(\boldsymbol{x}^{t-1}) \qquad \mathsf{G}(\boldsymbol{x}) := \nabla g_{\boldsymbol{\phi}}(\boldsymbol{x}) + \tau(\boldsymbol{x} - \mathsf{R}_{\boldsymbol{\theta}}(\boldsymbol{x}))$$

RARE #2: Based on Plug-and-Play Priors (PnP)

$$\boldsymbol{x}^t \leftarrow \mathsf{R}_{\boldsymbol{\theta}}(\boldsymbol{z}^t) \qquad \boldsymbol{z}^t \leftarrow \boldsymbol{x}^{t-1} - \gamma \nabla g_{\boldsymbol{\phi}}(\boldsymbol{x}^{t-1})$$

RARE leverages **DL-priors** while controlling **fidelity to data!**

RARE might look heuristic, but it has a rigorous foundation in **monotone operator theory**

RARE might look heuristic, but it has a rigorous foundation in monotone operator theory

Suppose there exists a vector that satisfies

$$\boxed{\boldsymbol{x}^* \in \mathsf{Zer}(\nabla g) \cap \mathsf{Fix}(\mathsf{R})}$$

$$\mathsf{Zer}(\nabla g) = \{\boldsymbol{x} \in \mathbb{R}^n : \nabla g(\boldsymbol{x}) = \boldsymbol{0}\}$$

$$\mathsf{Fix}(\mathsf{R}) = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} = \mathsf{R}(\boldsymbol{x})\}$$

J. Liu, S. Asif, B. Wohlberg, and U. S. Kamilov, “Recovery Analysis for Plug-and-Play Priors using the Restricted Eigenvalue Condition”, arXiv:2106.03668, 2021

RARE might look heuristic, but it has a rigorous foundation in monotone operator theory

Suppose there exists a vector that satisfies

$$\boxed{\boldsymbol{x}^* \in \mathsf{Zer}(\nabla g) \cap \mathsf{Fix}(\mathsf{R})}$$

$$\mathsf{Zer}(\nabla g) = \{\boldsymbol{x} \in \mathbb{R}^n : \nabla g(\boldsymbol{x}) = \mathbf{0}\}$$

$$\mathsf{Fix}(\mathsf{R}) = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} = \mathsf{R}(\boldsymbol{x})\}$$

Question: How can we interpret such a vector?

J. Liu, S. Asif, B. Wohlberg, and U. S. Kamilov, "Recovery Analysis for Plug-and-Play Priors using the Restricted Eigenvalue Condition", arXiv:2106.03668, 2021

RARE might look heuristic, but it has a rigorous foundation in monotone operator theory

Suppose there exists a vector that satisfies

$$\boldsymbol{x}^* \in \mathsf{Zer}(\nabla g) \cap \mathsf{Fix}(\mathsf{R})$$

$$\mathsf{Zer}(\nabla g) = \{\boldsymbol{x} \in \mathbb{R}^n : \nabla g(\boldsymbol{x}) = \mathbf{0}\}$$

Consistent with the measurements

$$\mathsf{Fix}(\mathsf{R}) = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} = \mathsf{R}(\boldsymbol{x})\}$$

Artifact-free according to the prior

J. Liu, S. Asif, B. Wohlberg, and U. S. Kamilov, "Recovery Analysis for Plug-and-Play Priors using the Restricted Eigenvalue Condition", arXiv:2106.03668, 2021

RARE might look heuristic, but it has a rigorous foundation in monotone operator theory

Suppose there exists a vector that satisfies

$$\boldsymbol{x}^* \in \mathsf{Zer}(\nabla g) \cap \mathsf{Fix}(\mathsf{R})$$

Consider the following assumptions

Assumption 1. The function g is convex and L-Lipschitz continuous.

Assumption 2. The AR operator R is a contraction.

Definition 1. We say that an operator S is λ-Lipschitz continuous if $\|\mathsf{S}(\boldsymbol{x}) - \mathsf{S}(\boldsymbol{y})\|_2 \leq \lambda \|\boldsymbol{x} - \boldsymbol{y}\|_2$, for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$.

RARE might look heuristic, but it has a rigorous foundation in monotone operator theory

Suppose there exists a vector that satisfies

$$\boldsymbol{x}^* \in \mathsf{Zer}(\nabla g) \cap \mathsf{Fix}(\mathsf{R})$$

Consider the following assumptions

Assumption 1. The function g is convex and L-Lipschitz continuous.

Assumption 2. The AR operator R is a contraction.

Theorem 1. Both variants of RARE (based on PnP and RED) converge to a vector in $\mathsf{Zer}(\nabla g) \cap \mathsf{Fix}(\mathsf{R}) = \mathsf{Fix}(\mathsf{T}) = \mathsf{Zer}(\mathsf{G})$, where $\mathsf{T} = \mathsf{R}(\mathsf{I} - \gamma \nabla g)$.

PnP and RED are equivalent under the assumptions above!

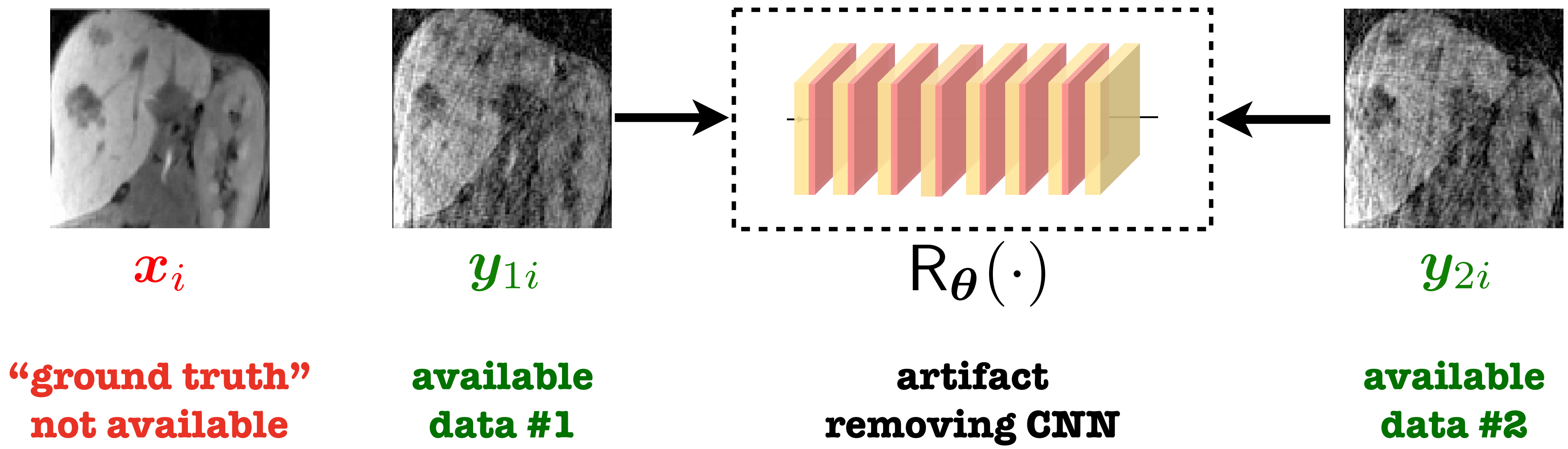
J. Liu, S. Asif, B. Wohlberg, and U. S. Kamilov, "Recovery Analysis for Plug-and-Play Priors using the Restricted Eigenvalue Condition", arXiv:2106.03668, 2021

Washington University in St. Louis

Artifact2Artifact (A2A) is a technique for training CNN priors for RARE **without ground truth**

Artifact2Artifact (A2A) is a technique for training CNN priors for RARE without ground truth

Consider multiple independent views for each object
(**examples**: radial lines in MRI, projections in CT)

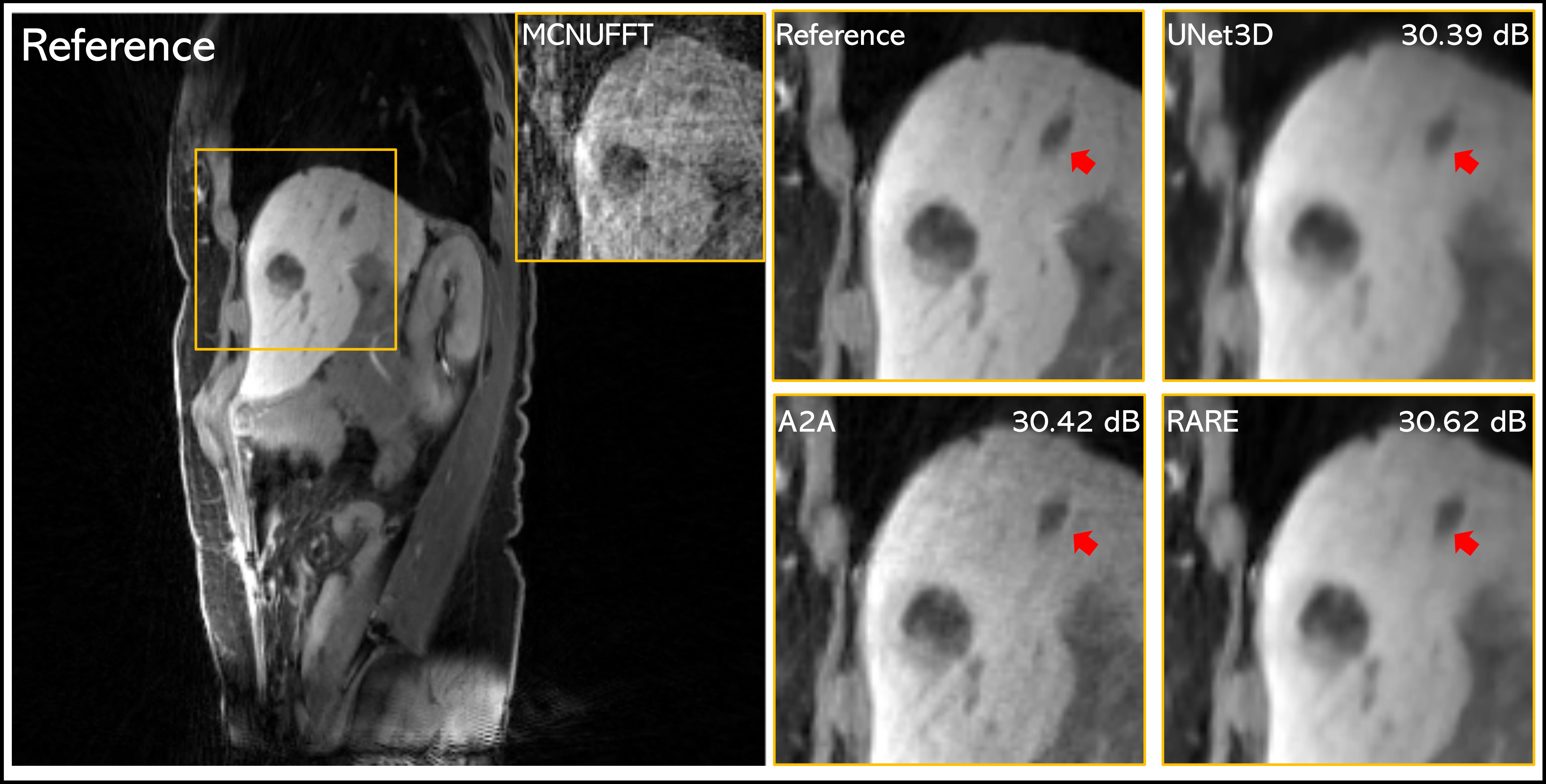


Artifact2Artifact (A2A) uses two independent acquisitions of the same object as training labels for the AR prior!

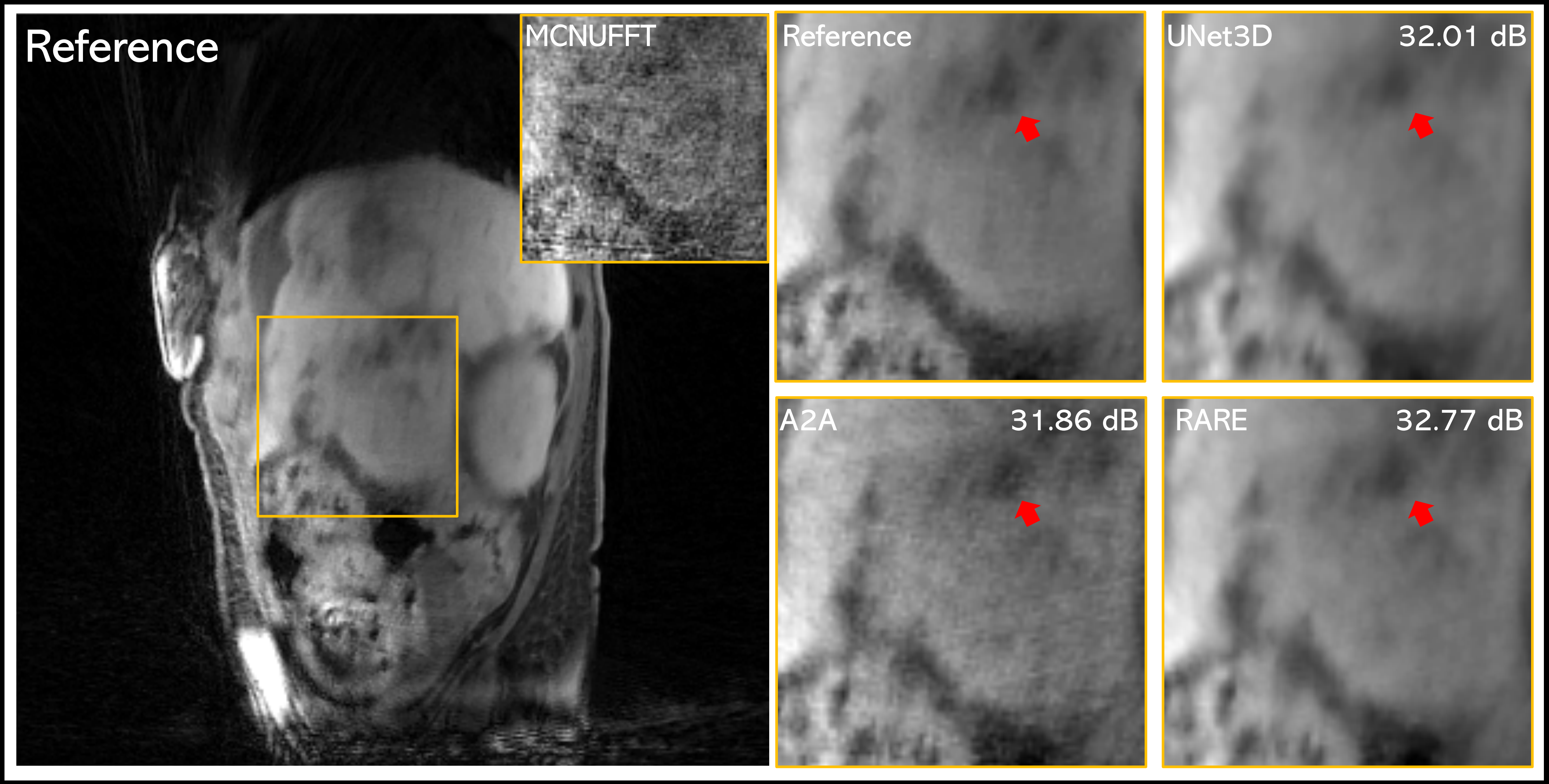
Lehtinen *et al.*, "Noise2Noise: Learning Image Restoration without Clean Data," *Proc. ICML*, 2018

Liu *et al.*, "RARE: Image Reconstruction using Deep Priors Learned without Ground Truth," *IEEE J. Sel. Topics Signal Process.*, October 2020

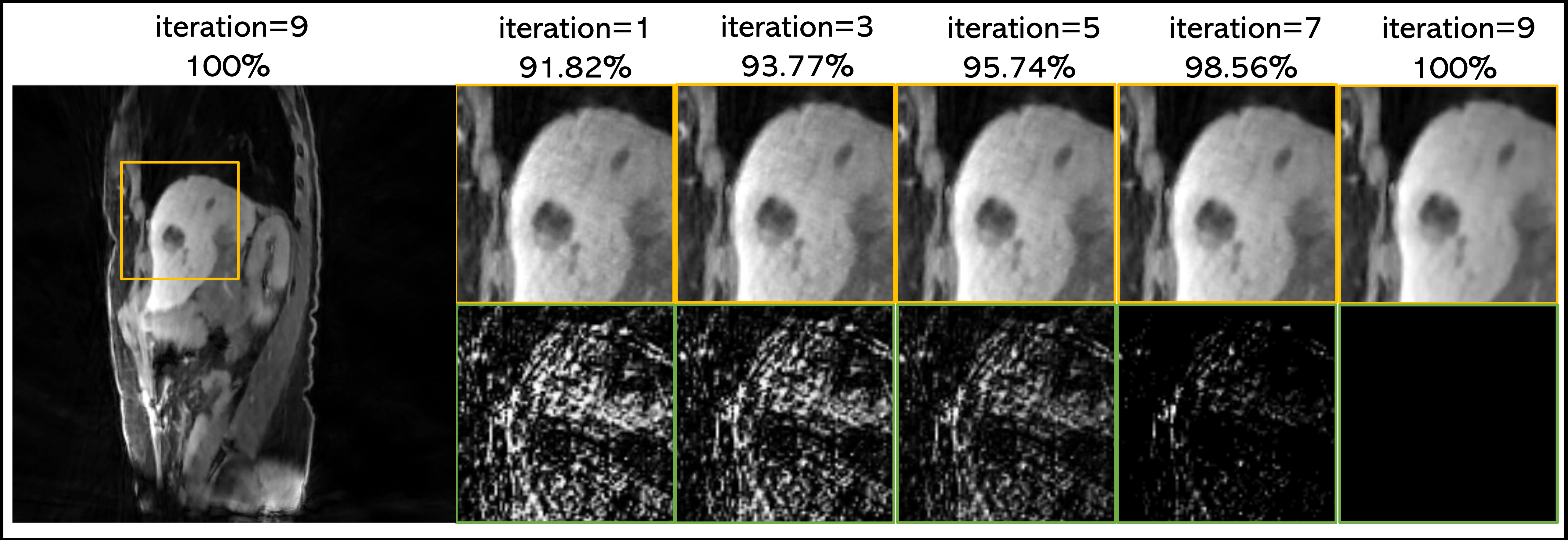
Application: Reconstructing 10 motion phases from a 1 minute free-breathing MRI scan



Application: Reconstructing 10 motion phases from a 1 minute free-breathing MRI scan



Application: Reconstructing 10 motion phases from a 1 minute free-breathing MRI scan



Outline for the rest of the talk

- Regularization by Artifact Removal (**RARE**)
  Integrating physical models and learned deep priors

- Efficient model-based deep learning (**SGD-Net**)
  Approximating physical layers for complexity gains

SIMBA decomposes a large-scale imaging problem into a sequence of partial updates

SIMBA decomposes a large-scale imaging problem into a sequence of partial updates

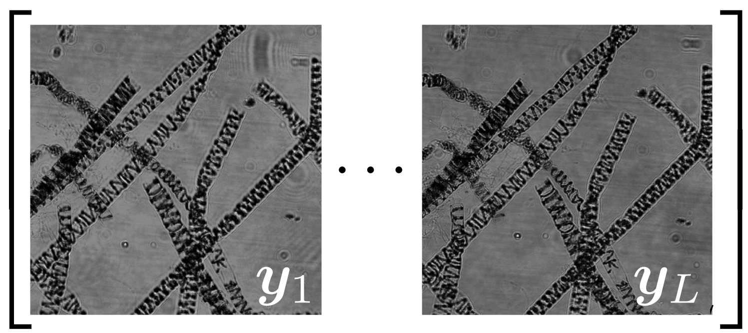
Wu *et al.*, "SIMBA: Scalable Inversion in Optical Tomography using Deep Denoising Priors," *IEEE J. Sel. Topics Signal Process.*, October 2020

Sun *et al.*, "A Provably Convergent Asynchronous Block Parallel Stochastic Method using Deep Denoising Priors," *Proc. ICLR*, May 2021

SIMBA decomposes a large-scale imaging problem into a sequence of partial updates

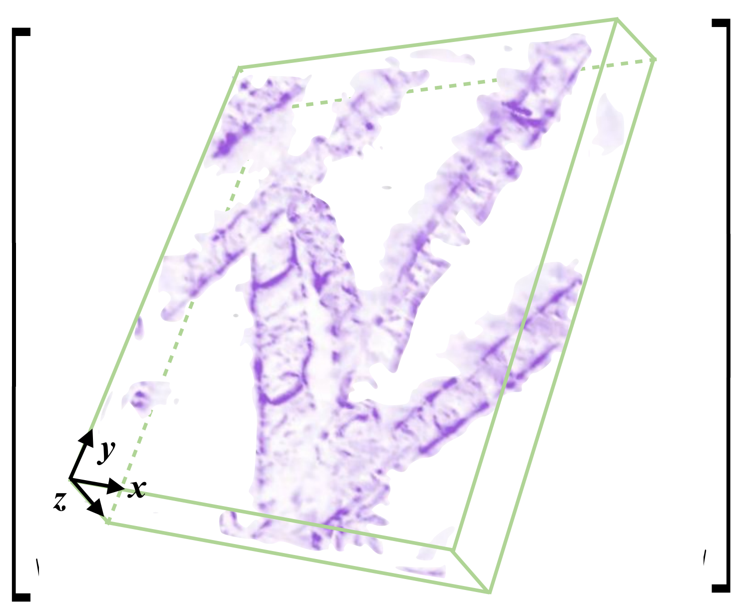
Traditional PnP/RED algorithms are use all the data at every iteration, which **significantly limits their scalability**

$$\boldsymbol{y} = \left[\ \boldsymbol{y}_1 \cdots \boldsymbol{y}_L\ \right]$$



Large measurement challenge:
10^2–10^6 measurements,
each with 10^6 pixels

$$\boldsymbol{x} = \left[\quad\right]$$



Large image challenge:
3D (space), 4D (space, time),
5D (space, time, spectrum)
images with 10^6–10^{12} voxels

SIMBA decomposes a large-scale imaging problem into a sequence of partial updates

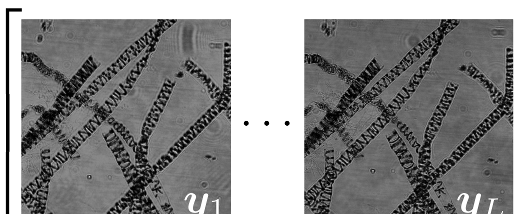
Traditional PnP/RED algorithms are use all the data at every iteration, which significantly limits their scalability

Batch data-fidelity: $g(\boldsymbol{x}) = \dfrac{1}{L}\displaystyle\sum_{\ell=1}^{L} g_\ell(\boldsymbol{x})$

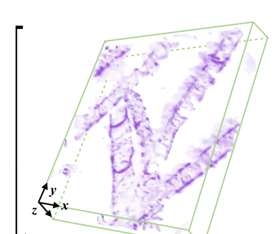
Complexity grows with the # of measurements

Example term: $g_\ell(\boldsymbol{x}) = \dfrac{1}{2}\|\boldsymbol{y}_\ell - \boldsymbol{H}_\ell\boldsymbol{x}\|_2^2$

$\boldsymbol{y} =$



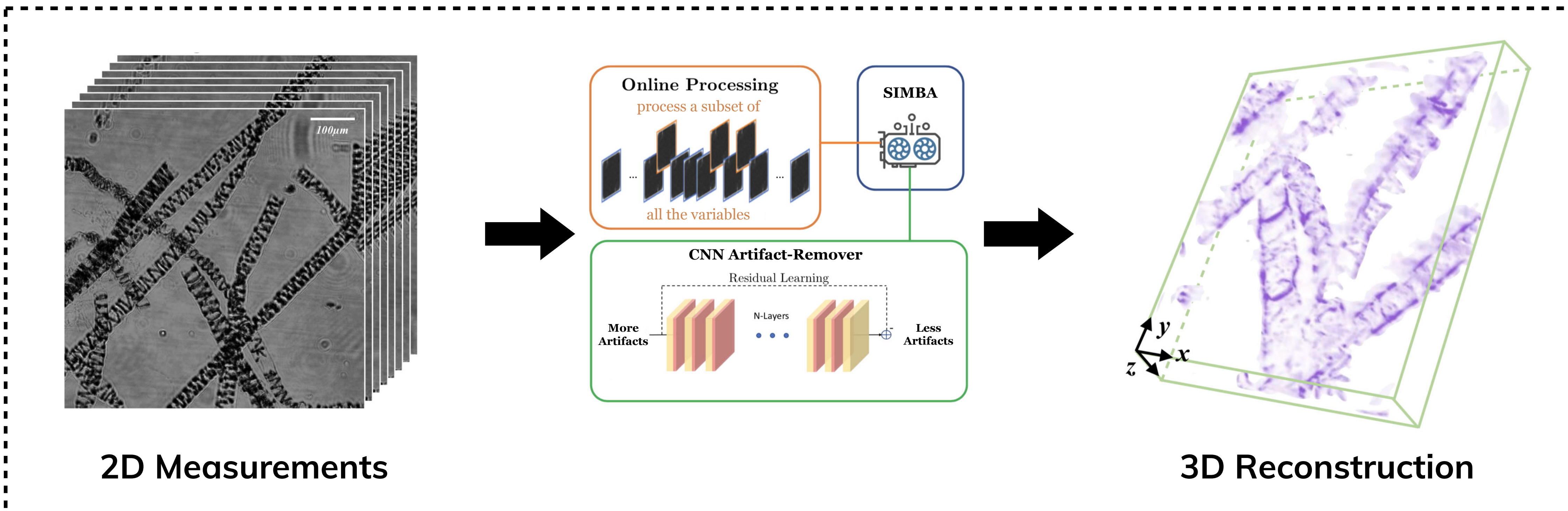
$\boldsymbol{x} =$



SIMBA decomposes a large-scale imaging problem into a sequence of partial updates

Traditional PnP/RED algorithms are use all the data at every iteration, which significantly limits their scalability

SIMBA uses only a subset of variables at a time which makes it scalable to datasets that are too large for batch processing



SIMBA decomposes a large-scale imaging problem into a sequence of partial updates

Traditional PnP/RED algorithms are use all the data at every iteration, which significantly limits their scalability

SIMBA uses only a subset of variables at a time which makes it scalable to datasets that are too large for batch processing

Batch data-fidelity: $g(\boldsymbol{x}) = \dfrac{1}{L}\sum_{\ell=1}^{L} g_\ell(\boldsymbol{x})$

Complexity grows with the # of measurements

Online data-fidelity: $\widehat{g}(\boldsymbol{x}) = \dfrac{1}{B}\sum_{b=1}^{B} g_{\ell_b}(\boldsymbol{x})$

Complexity independent from the # of measurements

Example term: $g_\ell(\boldsymbol{x}) = \dfrac{1}{2}\|\boldsymbol{y}_\ell - \boldsymbol{H}_\ell \boldsymbol{x}\|_2^2$

SIMBA decomposes a large-scale imaging problem into a sequence of partial updates

Traditional PnP/RED algorithms are use all the data at every iteration, which significantly limits their scalability

SIMBA uses only a subset of variables at a time which makes it scalable to datasets that are too large for batch processing

SIMBA uses only B << L measurements per iteration

RARE

$$\boxed{\nabla g(\boldsymbol{x}^{k-1}) \leftarrow \mathsf{FullGradient}(\boldsymbol{x}^{k-1})}$$
$$\mathsf{G}(\boldsymbol{x}^{k-1}) \leftarrow \nabla g(\boldsymbol{x}^{k-1}) + \tau(\boldsymbol{x}^{k-1} - \mathsf{R}_{\boldsymbol{\theta}}(\boldsymbol{x}^{k-1}))$$
$$\boldsymbol{x}^{k} \leftarrow \boldsymbol{x}^{k-1} - \gamma\mathsf{G}(\boldsymbol{x}^{k-1})$$

SIMBA

$$\boxed{\widehat{\nabla} g(\boldsymbol{x}^{k-1}) \leftarrow \mathsf{MinibatchGradient}(\boldsymbol{x}^{k-1})}$$
$$\mathsf{G}(\boldsymbol{x}^{k-1}) \leftarrow \widehat{\nabla} g(\boldsymbol{x}^{k-1}) + \tau(\boldsymbol{x}^{k-1} - \mathsf{R}_{\boldsymbol{\theta}}(\boldsymbol{x}^{k-1}))$$
$$\boldsymbol{x}^{k} \leftarrow \boldsymbol{x}^{k-1} - \gamma\mathsf{G}(\boldsymbol{x}^{k-1})$$

Example term: $g_\ell(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{y}_\ell - \boldsymbol{H}_\ell\boldsymbol{x}\|_2^2$

Washington University in St.Louis

SIMBA decomposes a large-scale imaging problem into a sequence of partial updates

Traditional PnP/RED algorithms are use all the data at every iteration, which significantly limits their scalability

SIMBA uses only a subset of variables at a time which makes it scalable to datasets that are too large for batch processing

Theorem 1. Run SIMBA for $t \geq 1$ iterations under Assumptions 1-3 using a fixed step-size $0 < \gamma \leq 1/(L + 2\tau)$ and a fixed minibatch size $B = t$. Then, we have

$$\mathbb{E}\left[\frac{1}{t}\sum_{k=1}^{t} \|\mathsf{G}(\boldsymbol{x}^{k-1})\|_2^2\right] \leq \frac{C}{\sqrt{t}},$$

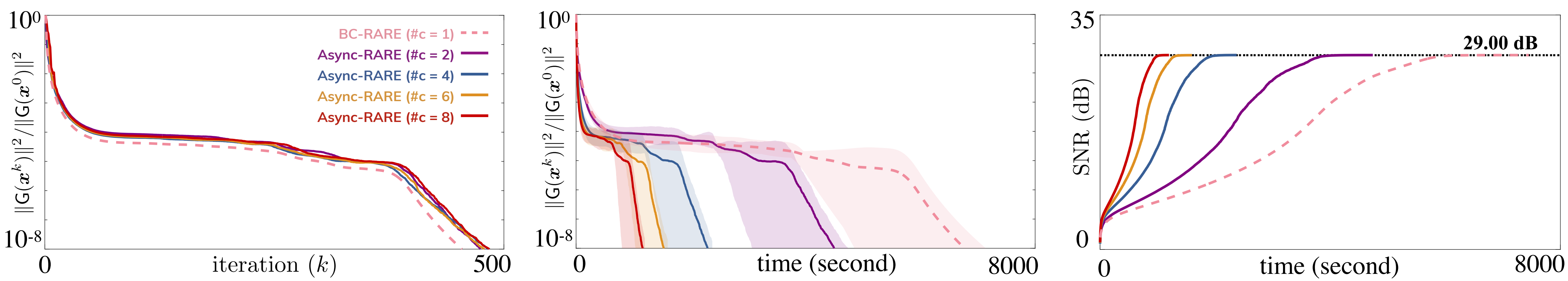
where $C > 0$ is a constant.

This convergence behavior is similar to that of SGD

Wu *et al.*, "SIMBA: Scalable Inversion in Optical Tomography using Deep Denoising Priors," *IEEE J. Sel. Topics Signal Process.*, October 2020

SIMBA leads to faster image reconstruction when using several processing cores

SIMBA leads to faster image reconstruction when using several processing cores

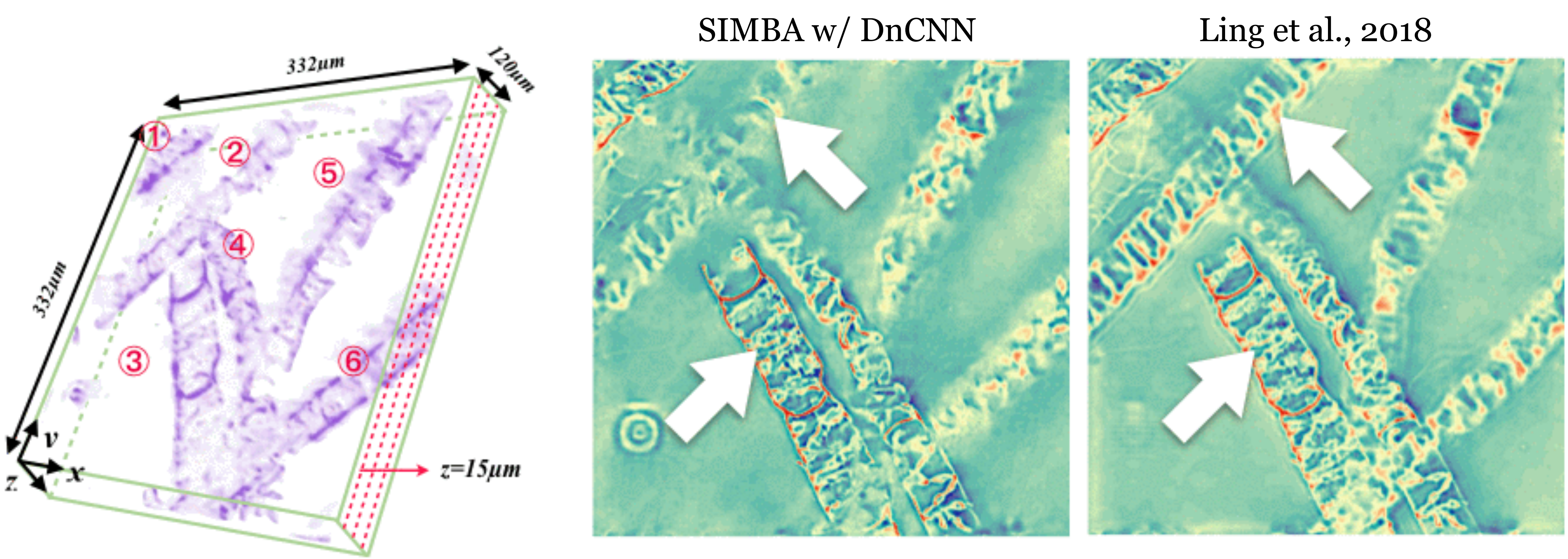


Method	SNR	Time	Speed-Up
RARE (1-core)	29.01 dB	1.8 hrs	-
Sync-SIMBA (8-core)	29.00 dB	38.9 mins	2.8x
Async-SIMBA-BG (8-core)	29.01 dB	17.9 mins	6.1x
Async-SIMBA-SG (8-core)	28.08 dB	13.0 mins	8.4x

Accelerations of up to 8.4x for image reconstruction in compressive sensing

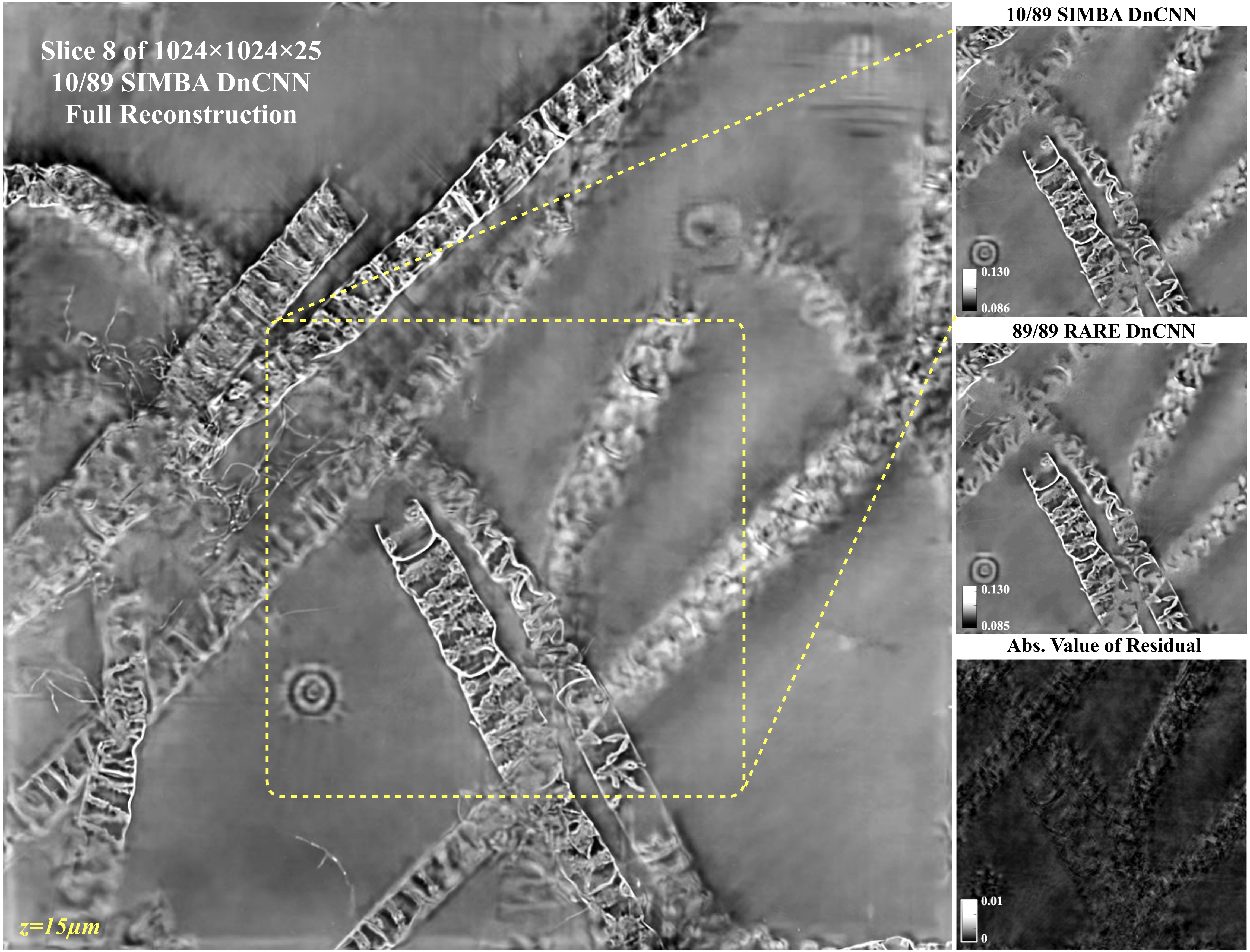
Sun *et al.*, "A Provably Convergent Asynchronous Block Parallel Stochastic Method using Deep Denoising Priors," *Proc. ICLR*, May 2021

SIMBA leads to better image quality when combined with deep priors



Significant improvements in sectioning capability in
Intensity Diffraction Tomography (IDT)

Wu *et al.*, "SIMBA: Scalable Inversion in Optical Tomography using Deep Denoising Priors," *IEEE J. Sel. Topics Signal Process.*, October 2020



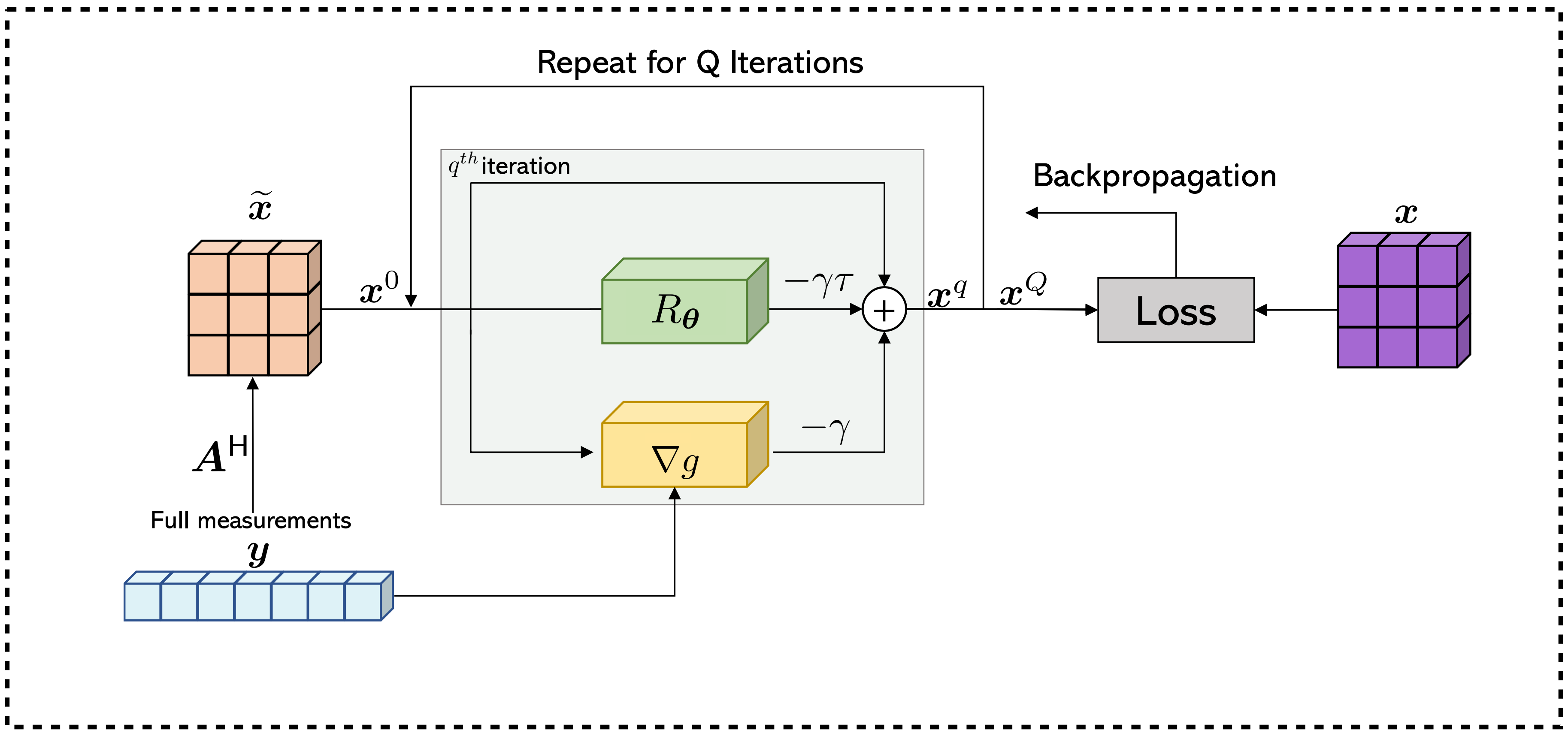
SIMBA is as good as full RARE

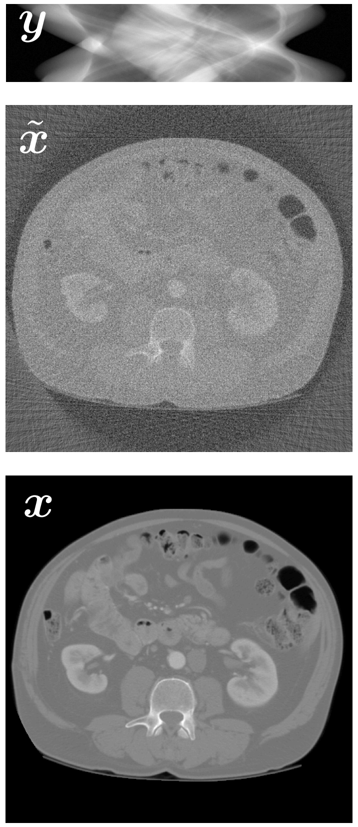
SGD-Net is a model-based deep network obtained by **“unfolding”** iterations of SIMBA

Washington University in St. Louis

SGD-Net is a model-based deep network obtained by "unfolding" iterations of SIMBA

An **"optimal"** artifact-removal CNN can be designed by unfolding truncated RARE and training it end-to-end





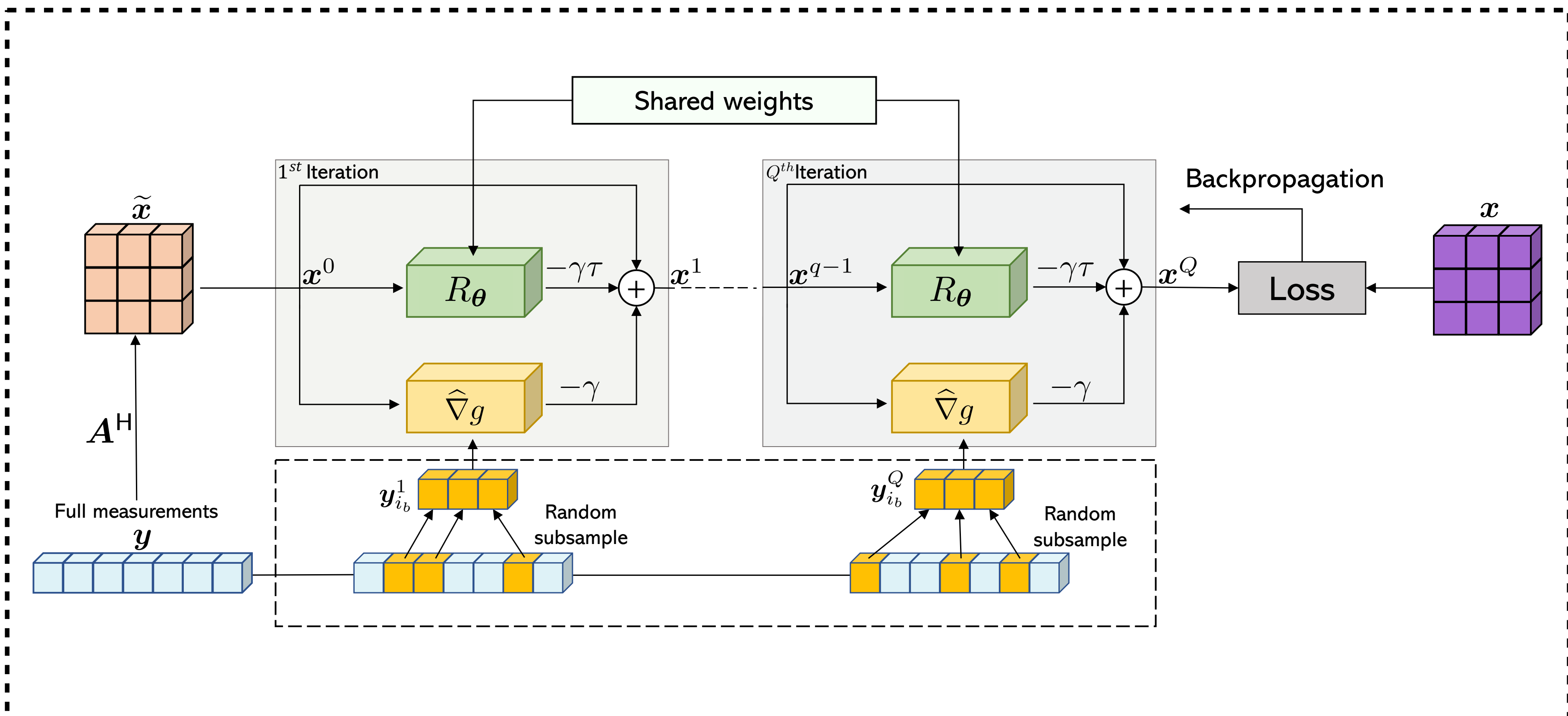
Limitation: Computational and GPU-memory complexities of data-consistency layers scale with the number of projections!

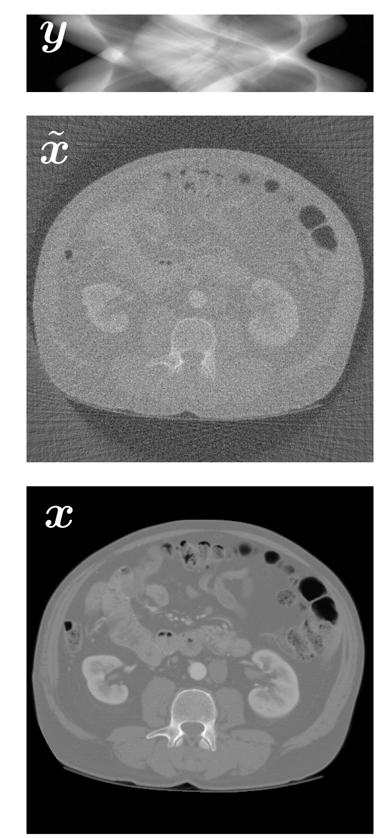
Liu *et al.*, "SGD-Net: Efficient Model-Based Deep Learning with Theoretical Guarantees," *IEEE Trans. Comput. Imag.*, in press.

SGD-Net is a model-based deep network obtained by “unfolding” iterations of SIMBA

An “optimal” artifact-removal CNN can be designed by unfolding truncated RARE and training it end-to-end

SGD-Net improves scalability of training and testing by directly unfolding SIMBA that processes data in minibatches





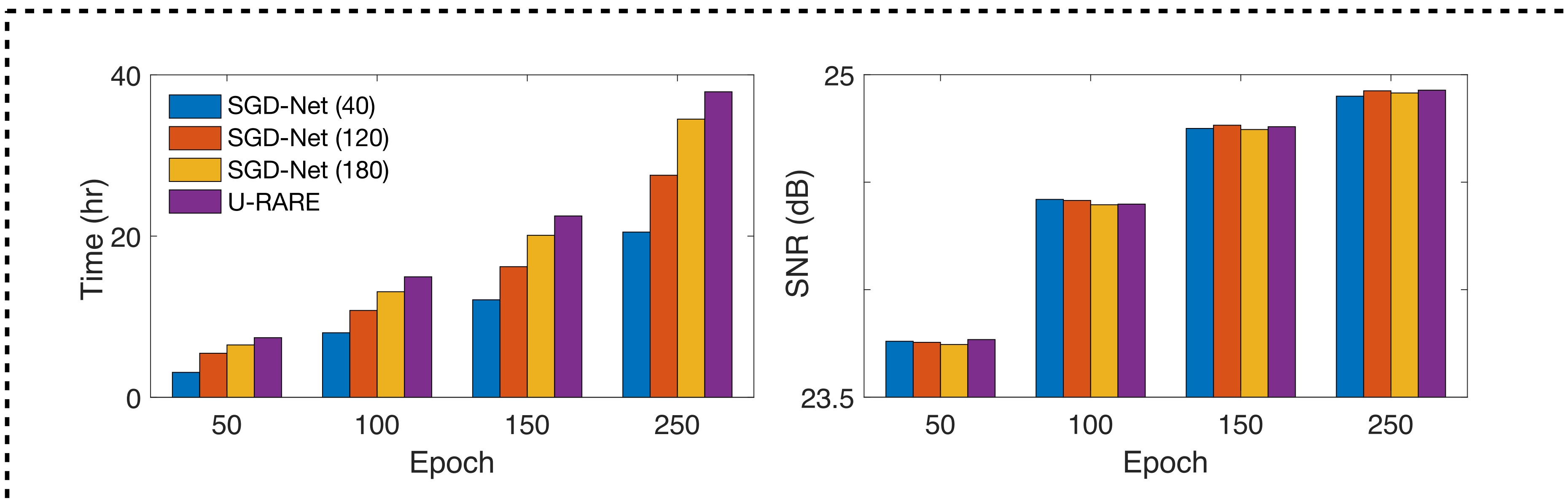
Key idea: Use minibatches in data-consistency layers!

SGD-Net is a model-based deep network obtained by “unfolding” iterations of SIMBA

An “optimal” artifact-removal CNN can be designed by unfolding truncated RARE and training it end-to-end

SGD-Net improves scalability of training and testing by directly unfolding SIMBA that processes data in minibatches

We **theoretically show** that SGD-Net can be trained to approximate the full unfolded RARE to any desired precision (see the paper)



Training accelerations of up to 2x for the same image quality in IDT and CT!

SGD-Net can significantly **reduce** the training time and the usage of GPU-memory

SGD-Net is competitive in terms of image quality with some of the best deep learning methods

Metric / Method \ Input-SNR (dB)	SNR 20-5	SNR 20	SNR 20+5	SSIM 20-5	SSIM 20	SSIM 20+5	**#Iterations**	**Size** Model/Measurement	**Time** CPU/GPU
TV	24.26	24.31	24.39	0.887	0.890	0.891	250	———/1.01 GB	87.58s/10.66s
U-Net	24.27	24.33	24.35	0.887	0.889	0.889	–	118.2 MB/———	0.925s/0.012s
ISTA-Net^{+}	24.39	24.41	24.47	0.889	0.890	0.890	12	6.90 MB/1.01 GB	18.36s/0.402s
RED-DnCNN	24.54	24.61	24.67	0.890	0.892	0.893	220	2.29 MB/1.01 GB	197.5s/4.144s
SGD-Net (40)	24.84	24.94	24.96	0.896	0.899	0.901	8	29.6 MB/0.17 GB	7.443s/0.322s
SGD-Net (120)	24.87	24.93	24.94	0.898	0.899	0.900	8	29.6 MB/0.51 GB	16.51s/0.617s
U-RED	24.89	24.93	24.94	0.898	0.899	0.900	8	29.6 MB/1.01 GB	31.23s/0.943s

Intensity Diffraction Tomography (IDT) with 240 measurements!

SGD-Net reduces complexity of model-based deep learning, while offering comparable or better imaging quality

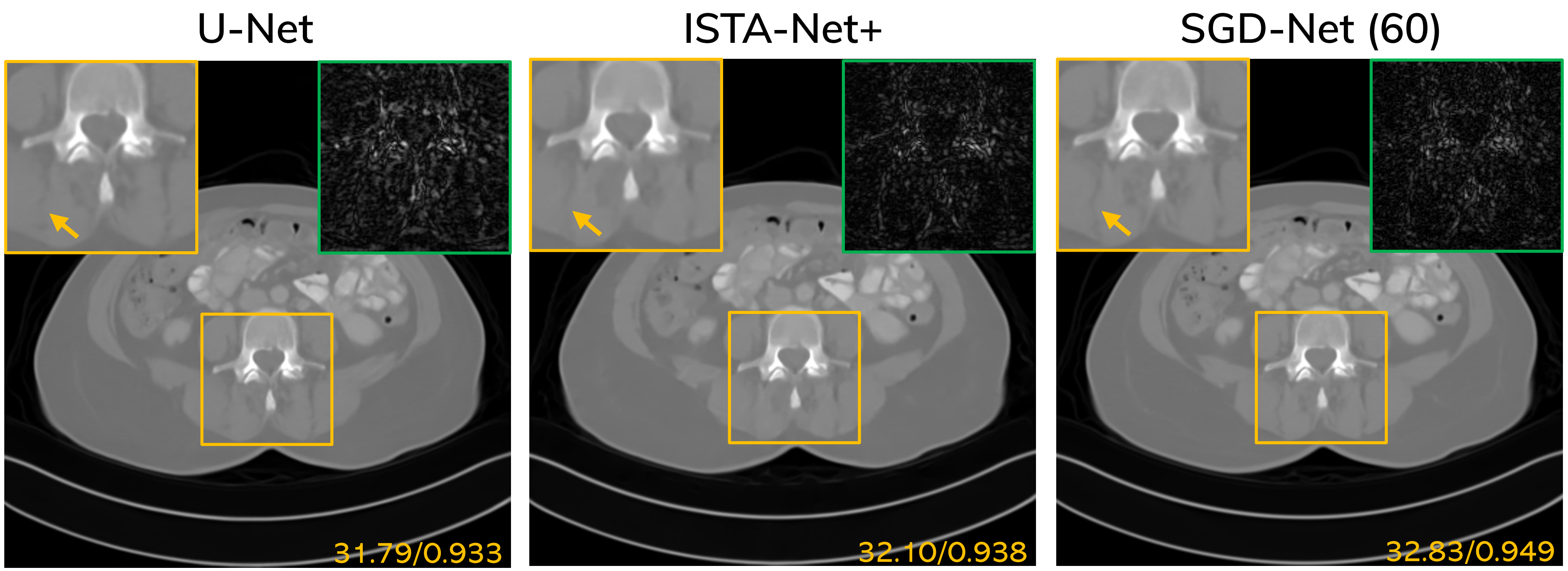
SGD-Net is competitive in terms of image quality with some of the best deep learning methods

Views	Metric	Method							
		FBP	TV	U-Net	RED-DnCNN	ISTA-Net^{+}	SGD-Net (30)	SGD-Net (60)	U-RED
90	SNR	17.56	30.09	31.17	31.93	32.01	32.76	**32.88**	32.87
	SSIM	0.362	0.924	0.930	0.935	0.934	0.942	**0.943**	**0.943**
120	SNR	20.03	31.23	32.54	33.13	33.17	33.91	33.95	**34.01**
	SSIM	0.449	0.929	0.936	0.941	0.940	0.948	0.949	**0.950**
180	SNR	23.19	32.97	34.04	34.49	34.61	35.44	**35.46**	**35.46**
	SSIM	0.582	0.940	0.948	0.950	0.951	0.957	**0.958**	**0.958**
Time (views=180)	CPU	0.859s	304.1s	2.061s	460.3s	15.95s	11.56s	13.31s	20.72s
	GPU	0.147s	13.58s	0.217s	5.177s	0.331s	0.269s	0.278s	0.325s

Sparse-view Computerized Tomography (CT)

SGD-Net reduces complexity of model-based deep learning, while offering comparable or better imaging quality

SGD-Net is competitive in terms of image quality with some of the best deep learning methods



SGD-Net significantly improves over several existing methods in terms of imaging quality!

Outline for the rest of the talk

- Regularization by Artifact Removal (**RARE**)
  Integrating physical models and learned deep priors

- Efficient model-based deep learning (**SGD-Net**)
  Approximating physical layers for complexity gains

One more thing: How does RARE compare with the methods based on using **generative models** as priors?

Artifact-removal (AR) priors improve over AWGN denoising priors when used in PnP/RED

Artifact-removal (AR) priors improve over AWGN denoising priors when used in PnP/RED

Recovery Analysis for Plug-and-Play Priors using the Restricted Eigenvalue Condition

Jiaming Liu
Washington University in St. Louis
jiaming.liu@wustl.edu

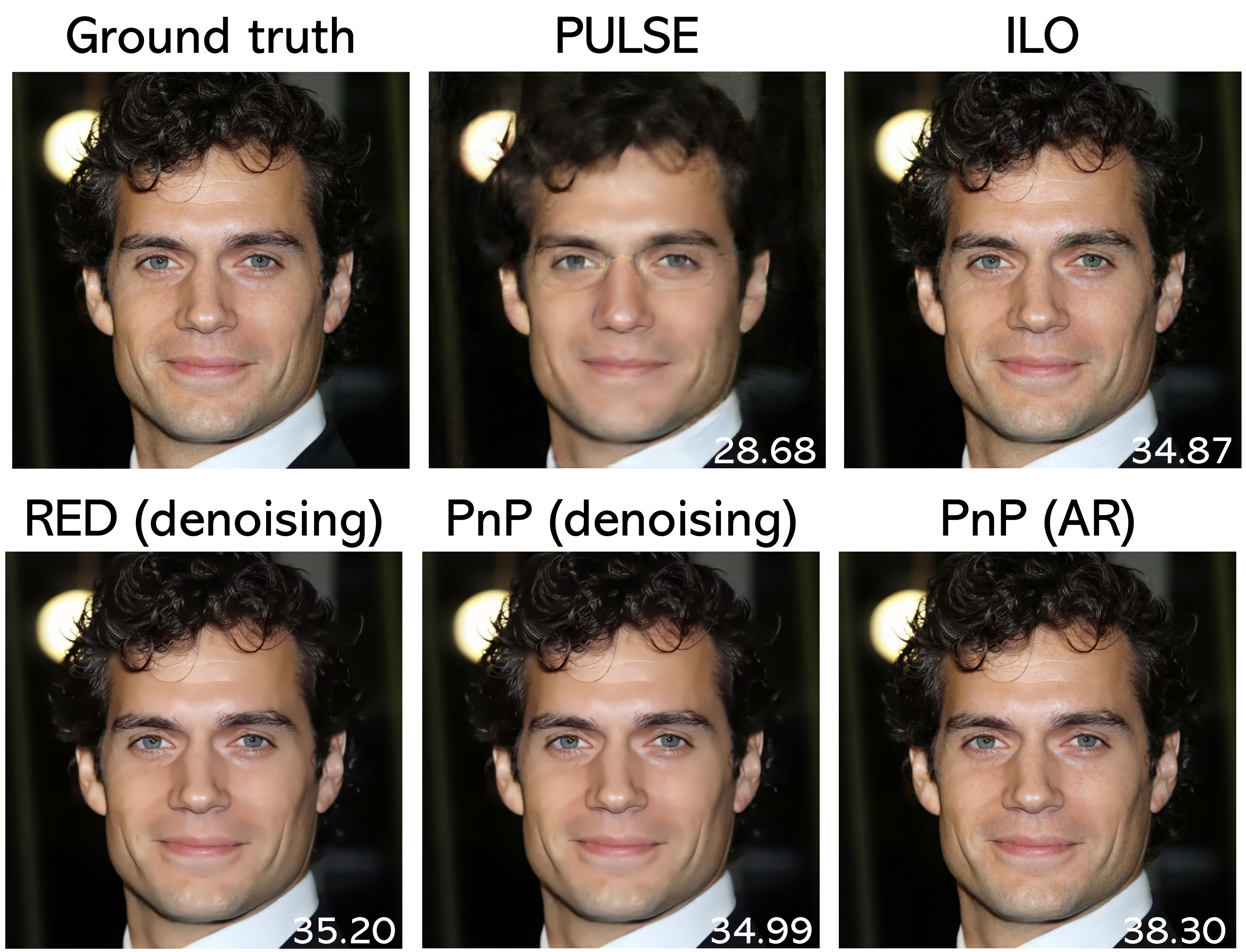
M. Salman Asif
University of California, Riverside
sasif@ece.ucr.edu

Brendt Wohlberg
Los Alamos National Laboratory
brendt@ieee.org

Ulugbek S. Kamilov
Washington University in St. Louis
kamilov@wustl.edu

J. Liu, S. Asif, B. Wohlberg, and U. S. Kamilov, “Recovery Analysis for Plug-and-Play Priors using the Restricted Eigenvalue Condition”, arXiv:2106.03668, 2021

Artifact-removal (AR) priors improve over AWGN denoising priors when used in PnP/RED



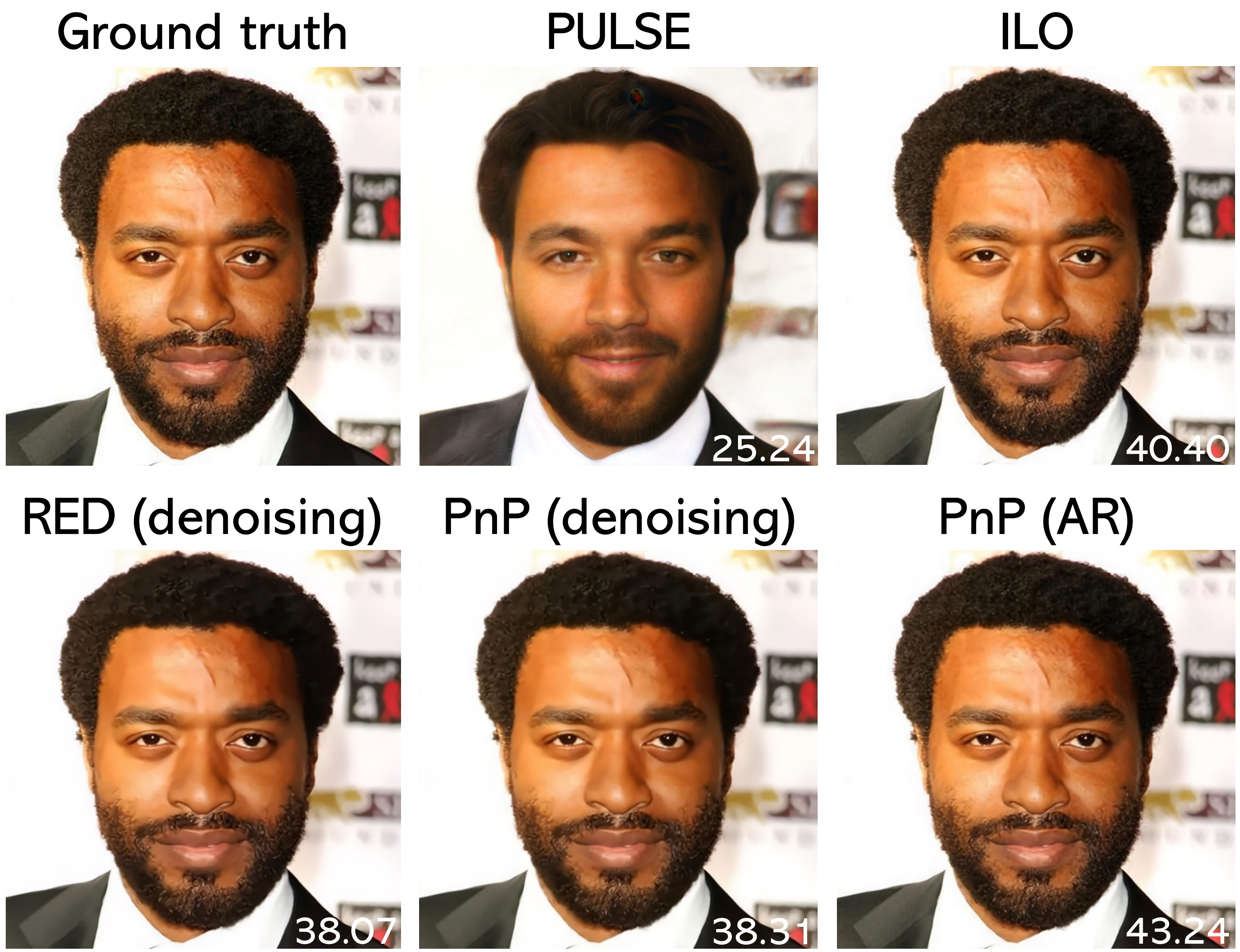
1. AR is better than AWGN (**expected!**)
2. RED is nearly equivalent to PnP (**somehow surprising!**)
3. PnP (AR) is competitive with PULSE and ILO (**surprising!**)

Table 3: Average PSNR (dB) values for several algorithms on test images from CelebA HQ.

Method \ CS Ratio	10%	20%	30%	40%	50%
TV	32.13	35.24	37.41	39.35	41.29
PULSE [34]	27.45	29.98	33.06	34.25	34.77
ILO [35]	36.15	40.98	43.46	47.89	48.21
RED (denoising)	35.46	41.59	45.65	48.13	52.17
PnP (denoising)	35.61	41.51	45.71	48.05	52.24
PnP (AR)	**39.19**	**44.20**	**48.66**	**51.32**	**53.89**

Washington University in St. Louis

Artifact-removal (AR) priors improve over AWGN denoising priors when used in PnP/RED



1. AR is better than AWGN (**expected!**)
2. RED is nearly equivalent to PnP (**somehow surprising!**)
3. PnP (AR) is competitive with PULSE and ILO (**surprising!**)

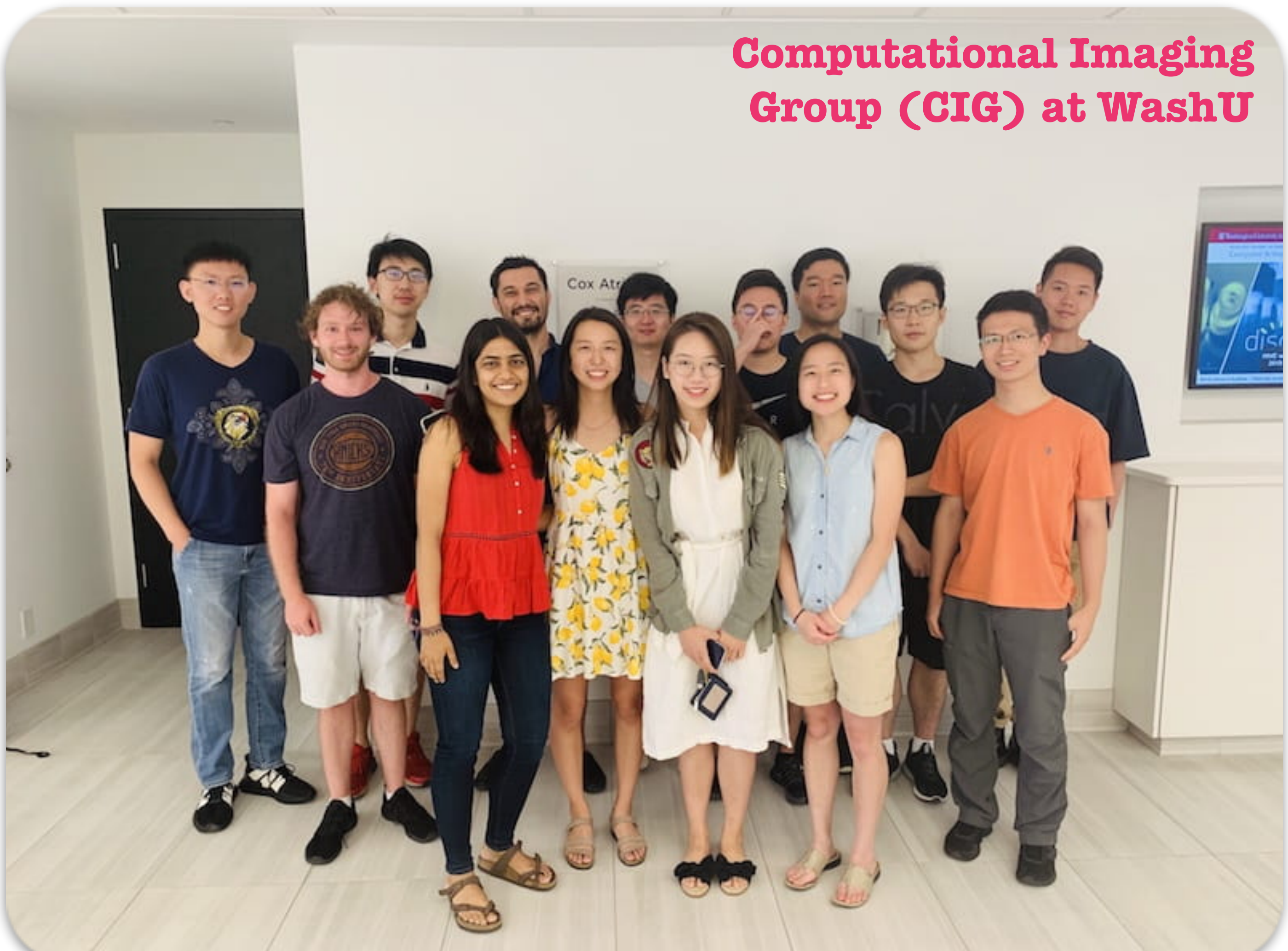
Table 3: Average PSNR (dB) values for several algorithms on test images from CelebA HQ.

Method \ CS Ratio	10%	20%	30%	40%	50%
TV	32.13	35.24	37.41	39.35	41.29
PULSE [34]	27.45	29.98	33.06	34.25	34.77
ILO [35]	36.15	40.98	43.46	47.89	48.21
RED (denoising)	35.46	41.59	45.65	48.13	52.17
PnP (denoising)	35.61	41.51	45.71	48.05	52.24
PnP (AR)	**39.19**	**44.20**	**48.66**	**51.32**	**53.89**

To conclude

- We increasingly rely on **deep learning** for characterizing complex high-dimensional statistical distributions
- **RARE** is a theoretically sound algorithm that combines an artifact removing CNN with data consistency
- **SGD-Net** is a model-based network that uses minibatches to reduce complexity of data-consistency layers

Computational Imaging Group (CIG) at WashU focuses on algorithms and math for imaging



Learn more about what we do at: **cigroup.wustl.edu**

Washington University in St. Louis

Computational Imaging Group (CIG) at WashU focuses on algorithms and math for imaging

CONTACT INFO

Prof. Ulugbek Kamilov
Email: kamilov@wustl.edu
Twitter: @ukmlv

Web: http://cigroup.wustl.edu
Group Twitter: @wustlcig

ACKNOWLEDGEMENTS

Yu Sun
Jiaming Liu
Xiaojian Xu
Weijie Gan
Ray Wu

Dr. Brendt Wohlberg
Prof. Hongyu An
Prof. Lei Tian
Dr. Cihat Eldeniz
Dr. Hassan Mansour

Support:

CCF-181391

Connect with me