SAGIVTECH

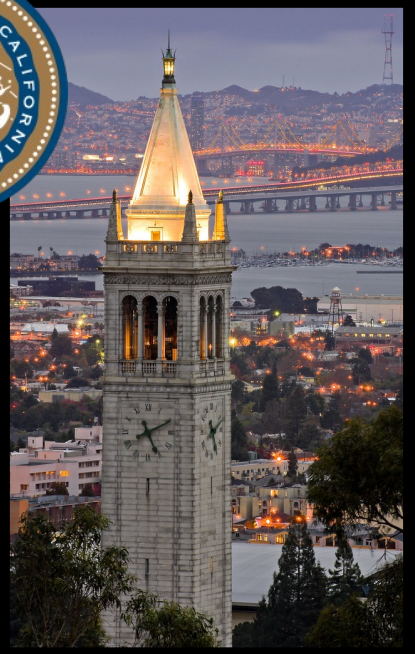
IMVC 2024

Data crimes: The Risk in Naïve Training of Medical AI Algorithms

Efrat Shimron

Technion

Berkeley
UNIVERSITY OF CALIFORNIA



Postdoc (2020-2023)

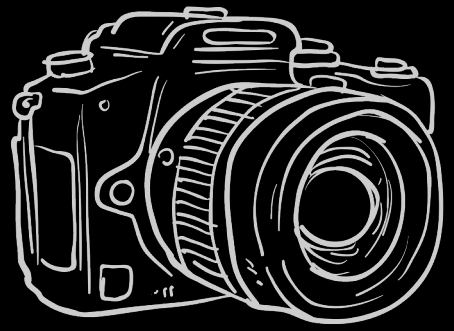
Stanford University

Technion
Israel Institute of Technology

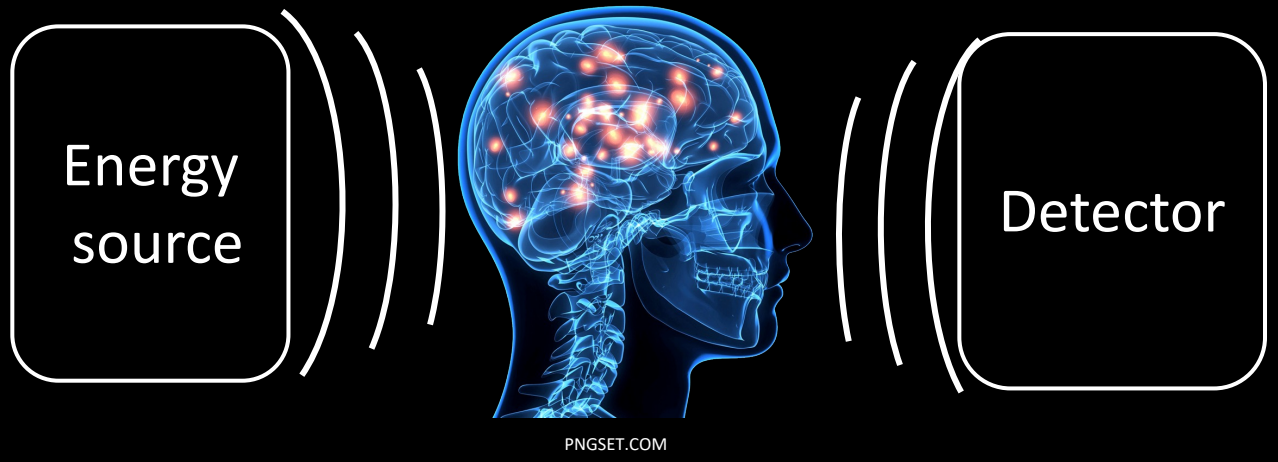
Medical AI & MRI lab

Introduction to Medical Imaging

Photography



Medical imaging



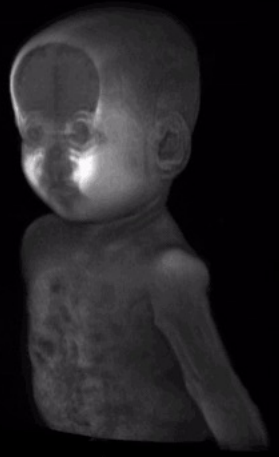
Introduction to MRI

MRI

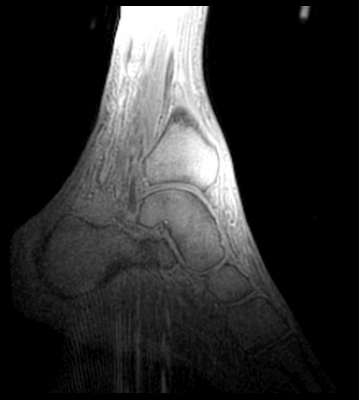
Introduction to MRI

MRI

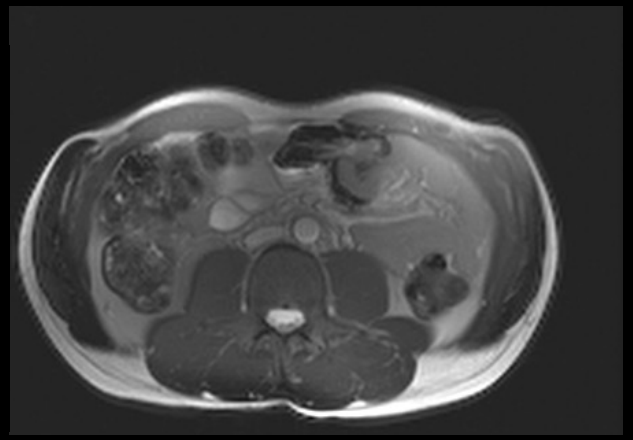
- ✓ Safe – no ionizing radiation
- ✓ Rich visual information
- ✓ Captures static & dynamic data



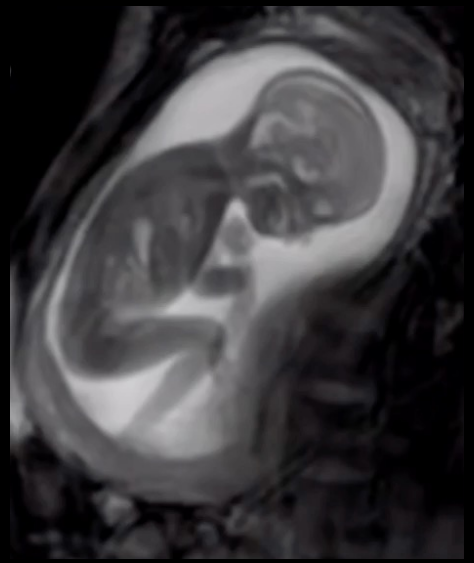
Courtesy of Lustig Lab



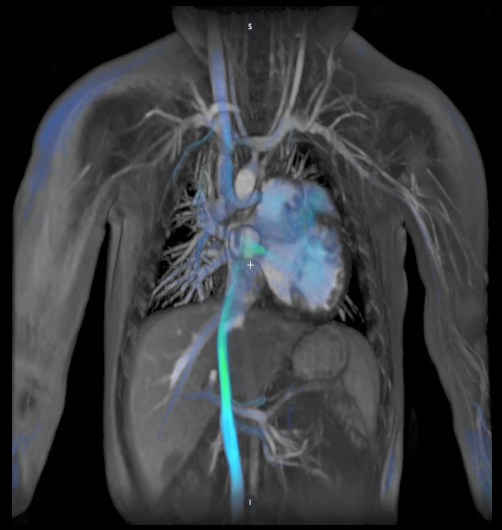
Courtesy of Lustig lab



Courtesy of Lustig lab



www.youtube.com/user/channelmum



Courtesy of Joseph Cheng

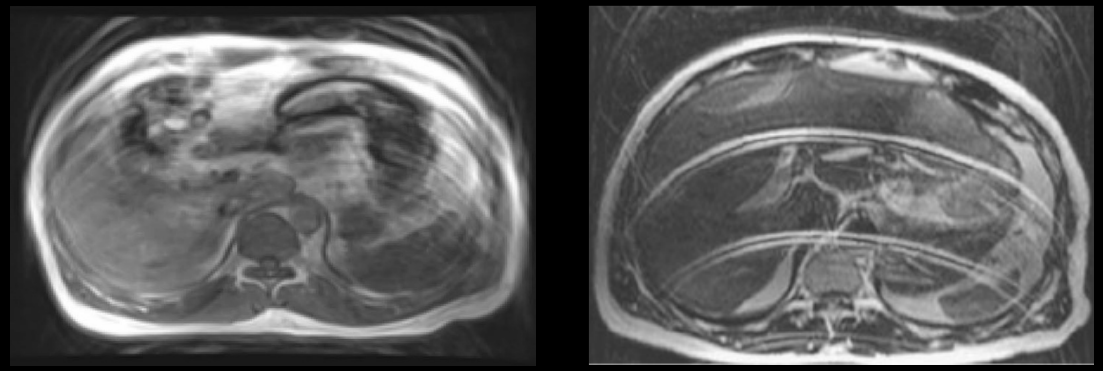
Introduction to MRI

MRI

Limitations:

- Long scan duration (30-60min)
- Sensitivity to motion artifacts
- Expensive

Motion Artifacts



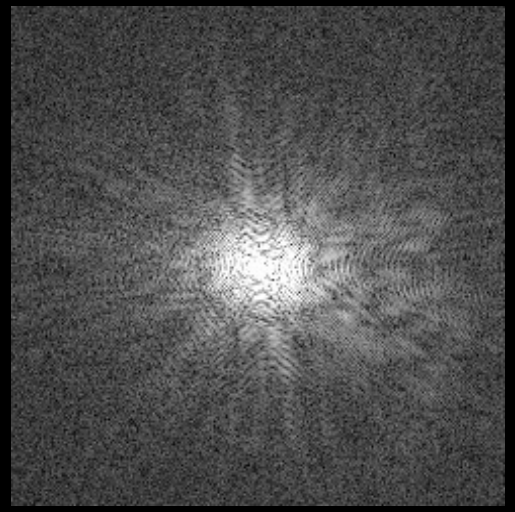
Zaitsev et al. MRM 2015

How can we make scans faster?

MRI: from sampling to images

MRI

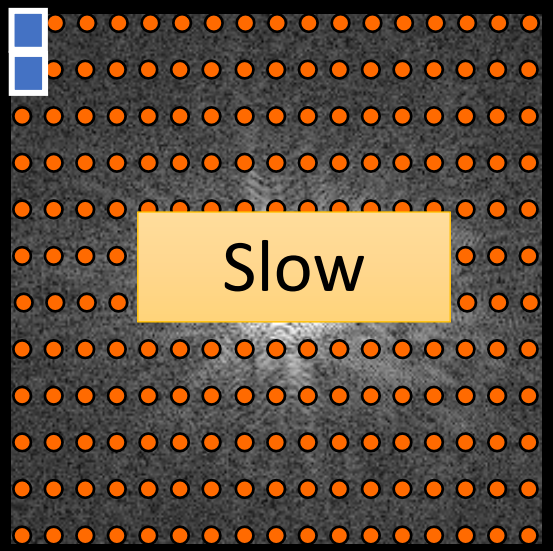
Fourier domain
k-space



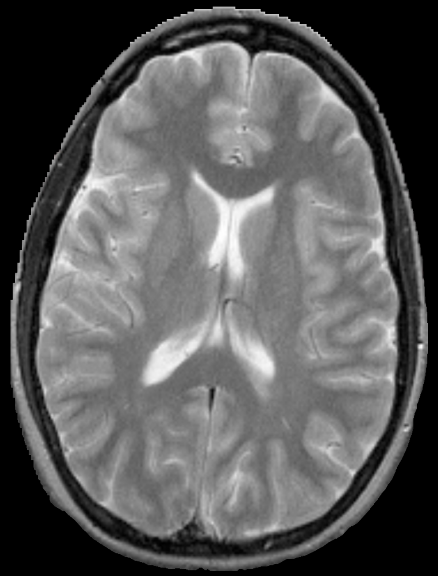
MRI: from sampling to images

MRI

Fourier domain
k-space



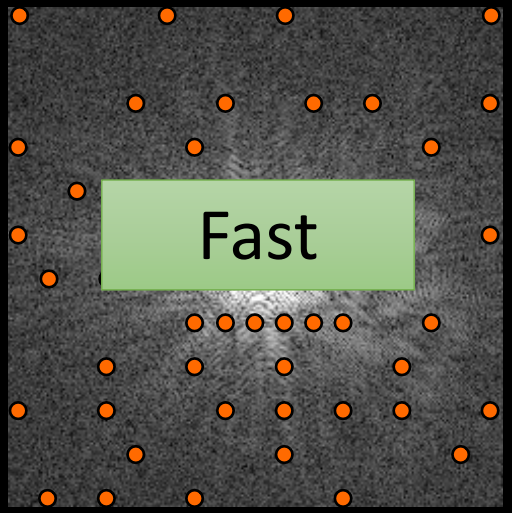
IFFT



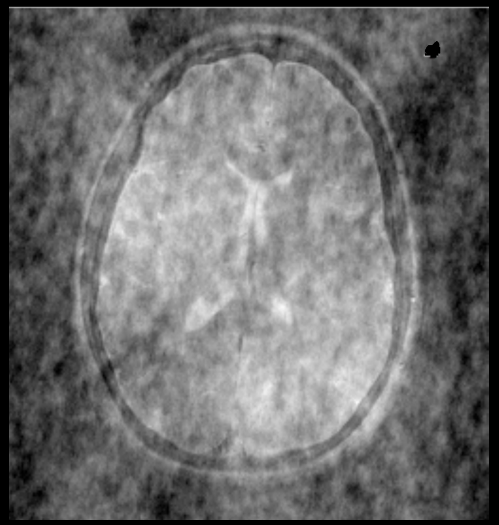
MRI: from sampling to images

MRI

Fourier domain
k-space



IFFT



MRI: from sampling to images

MRI

Fourier domain
k-space

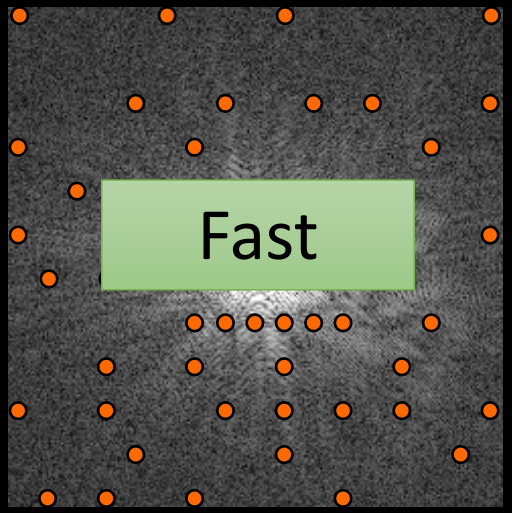
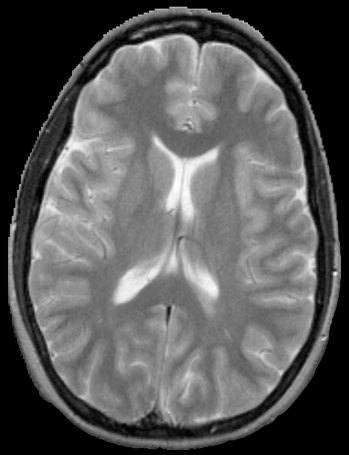


Image reconstruction algorithm



MRI: from sampling to images

MRI

Fourier domain
k-space

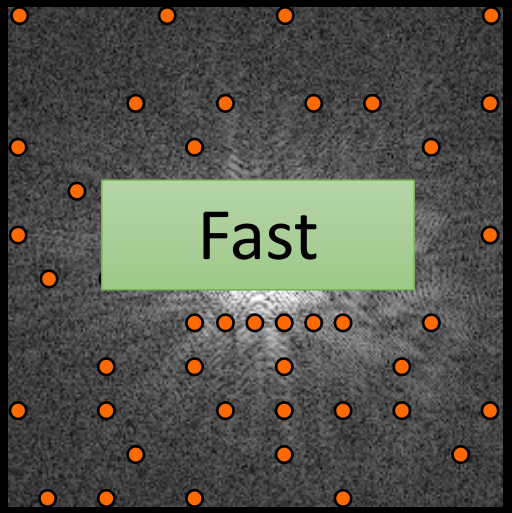
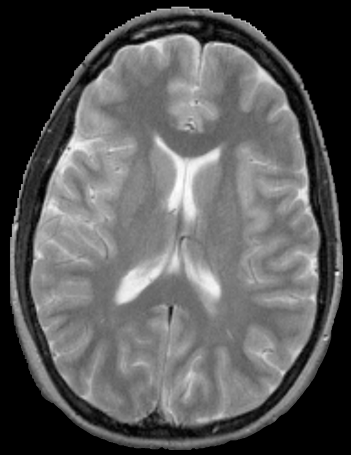


Image reconstruction algorithm

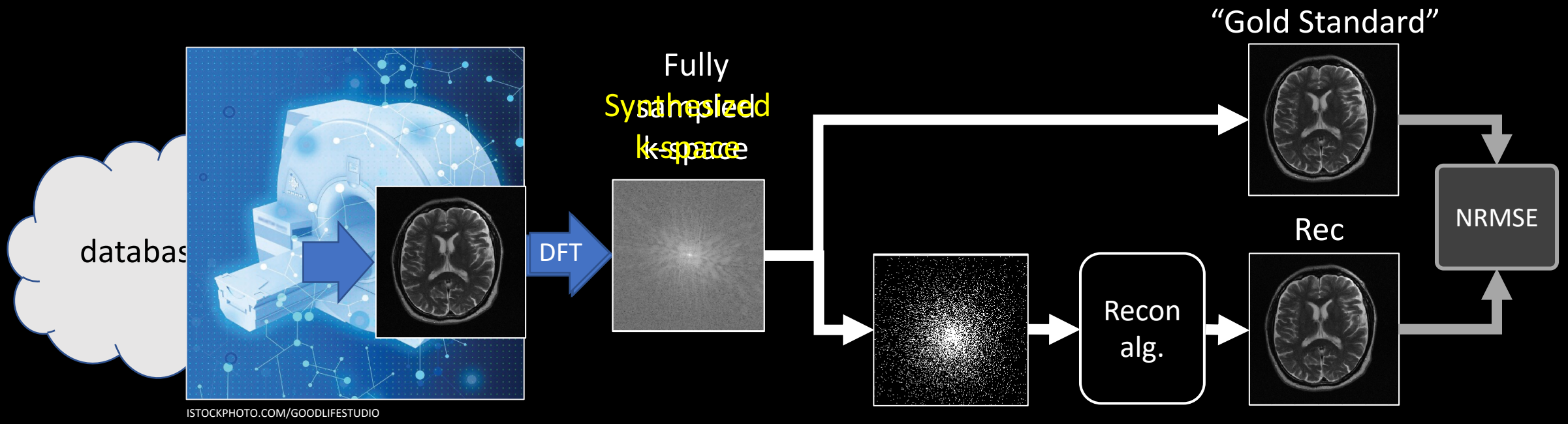


Parallel Imaging (1990s-2000s)
Compressed Sensing (2006-)
Machine Learning (2016-)

Training data?

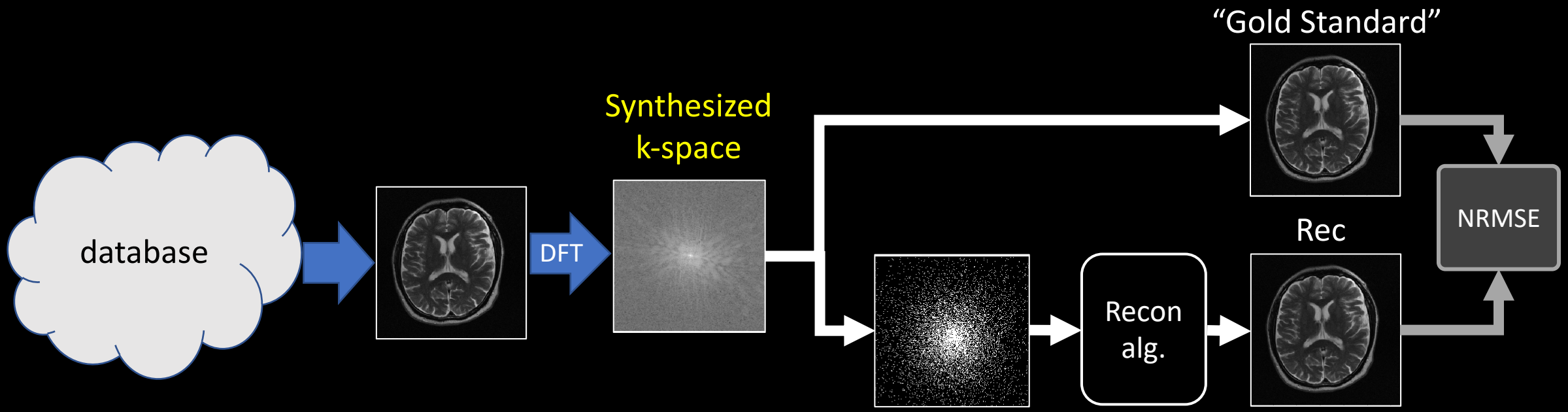
Data Crimes

How image reconstruction algorithms are developed



Shimron et al., *Implicit data crimes,* PNAS (2022)

How image reconstruction algorithms are developed



Shimron et al., *Implicit data crimes,* PNAS (2022)

How image reconstruction algorithms are developed

fastMRI

SKM-TEA

mridata.org

Calgary-Campinas Public Dataset

Hidden data processing

Raw

Blackbox

database

Multimodal Brain Tumor Segmentation Challenge 2018 BraTS

OASIS OPEN ACCESS SERIES OF IMAGING STUDIES

kaggle

Alzheimer's Dataset

IXI Dataset

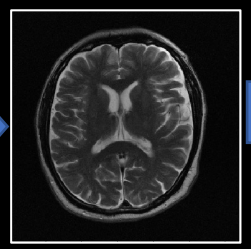
AccelMR a faster MRI challe

Human Connectom

Biased!

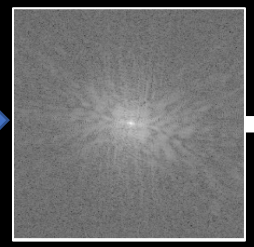
download

database

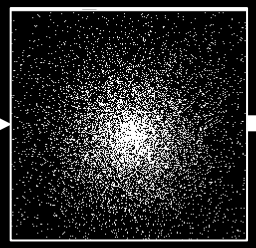


DFT

Synthesized k-space



"Gold Standard"



Recon alg.

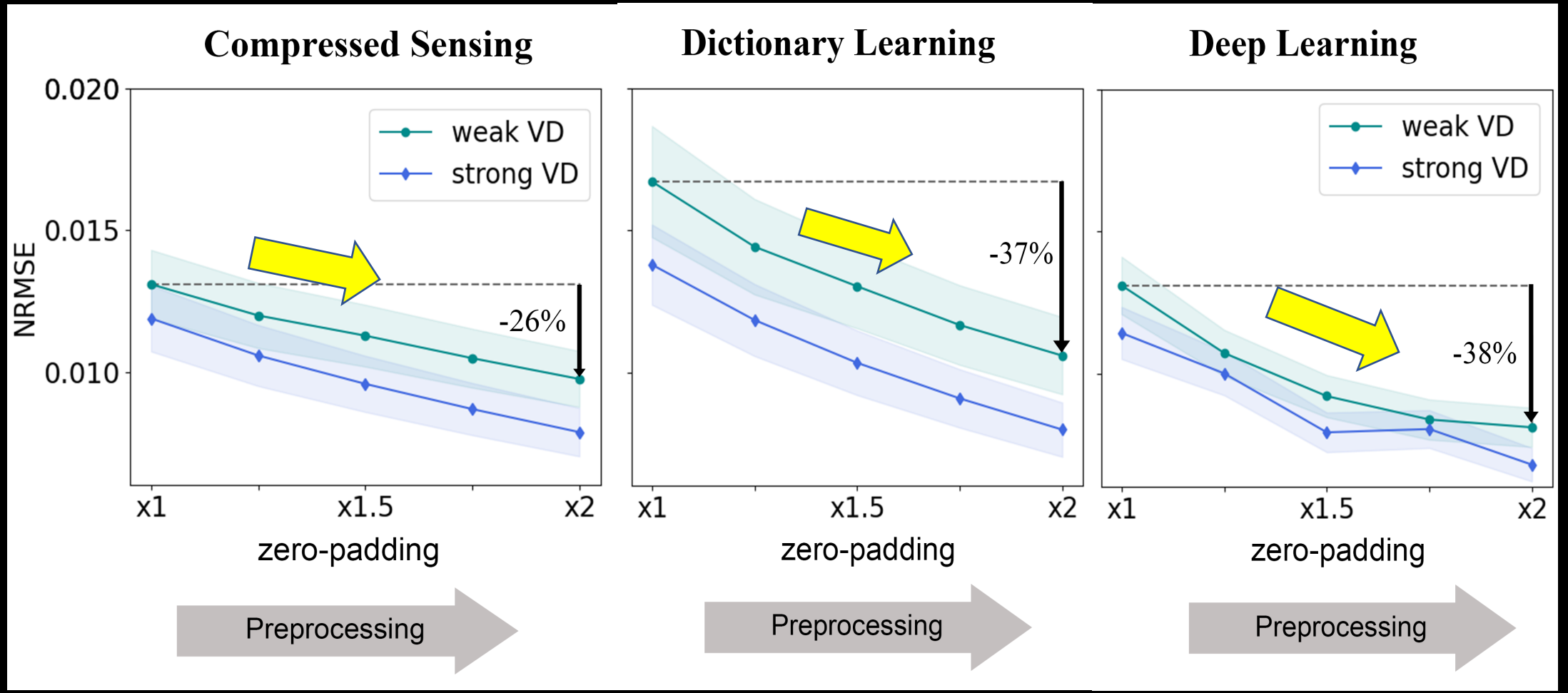
Rec

NRMSE

Shimron et al., *Implicit data crimes,* PNAS (2022)

Data Crimes: Implicit Bias of Reconstruction Algorithms

FastMRI Data

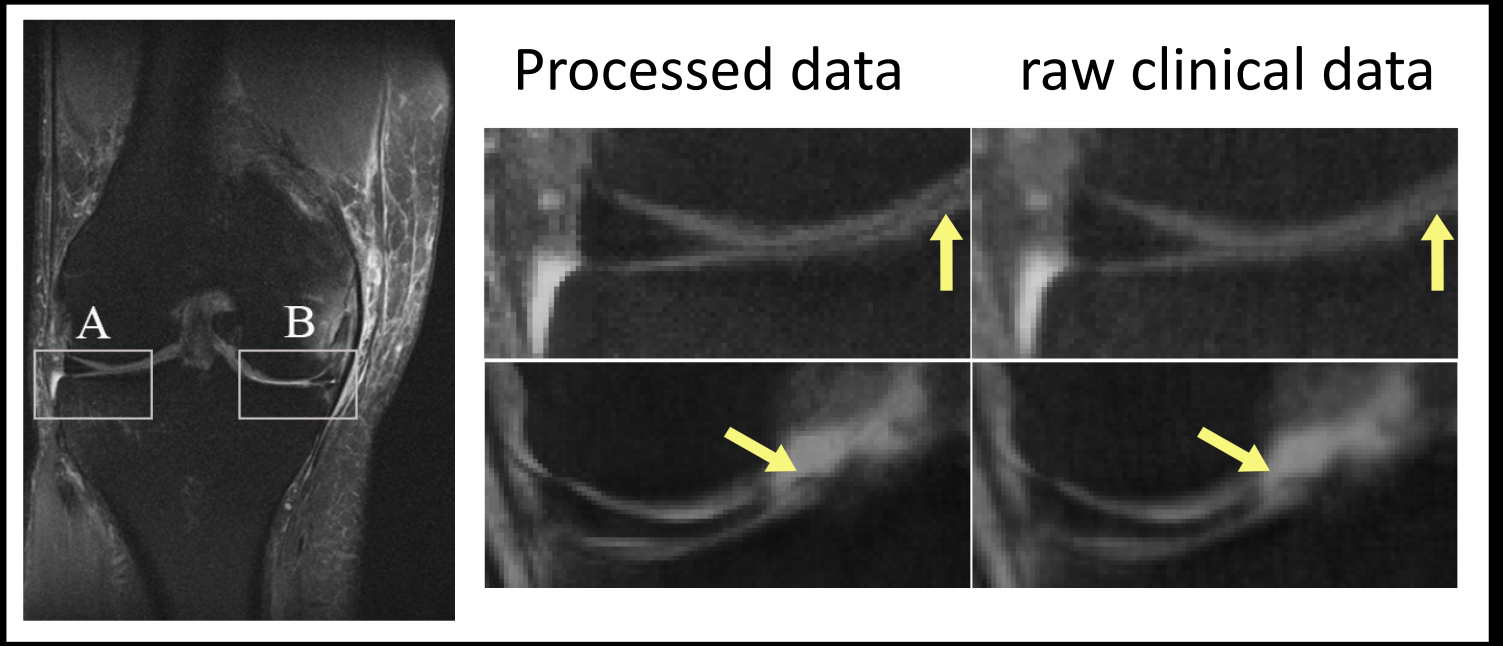


Shimron et al., *Implicit data crimes,* PNAS (2022)

Lustig et al., MRM (2007); Ravishankar et al, IEEE TMI (2010); Aggarwal et al. IEEE TMI (2019)

Can we train on processed data &
use the algorithms for clinical data?

Data Crimes: Implicit Bias of Reconstruction Algorithms

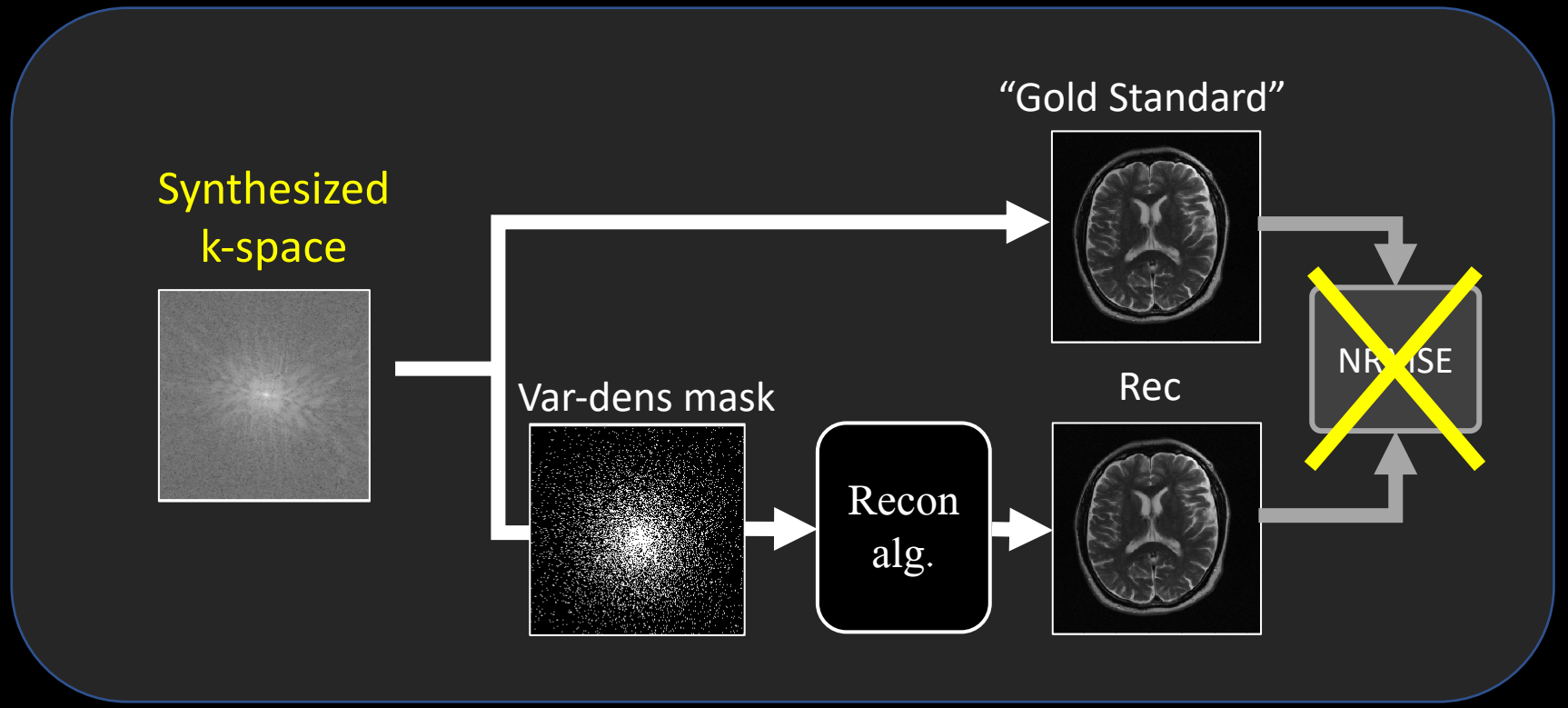


Shimron et al., *Implicit data crimes,* PNAS (2022)

Data Crimes: Implicit Bias of Reconstruction Algorithms

✗ Naïve use of Big Data can lead to biased results

✗ Error metrics - *blind* to the preprocessing



Shimron et al., *Implicit data crimes,* PNAS (2022)

Data Crimes: Implicit Bias of Reconstruction Algorithms

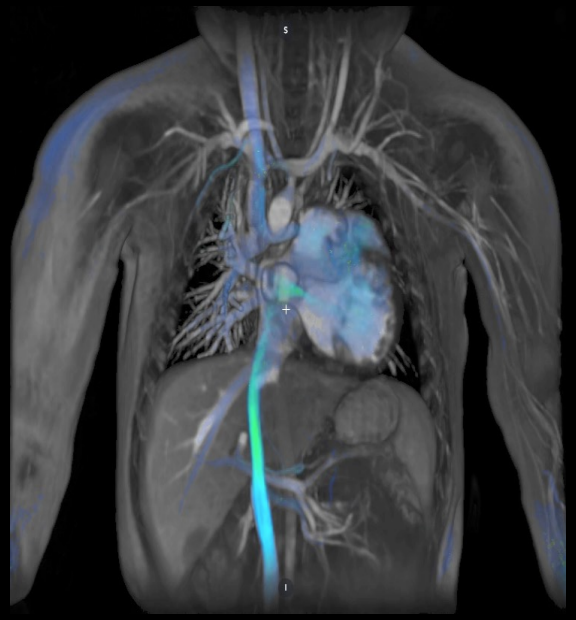
✗ Naïve use of Big Data can lead to biased results

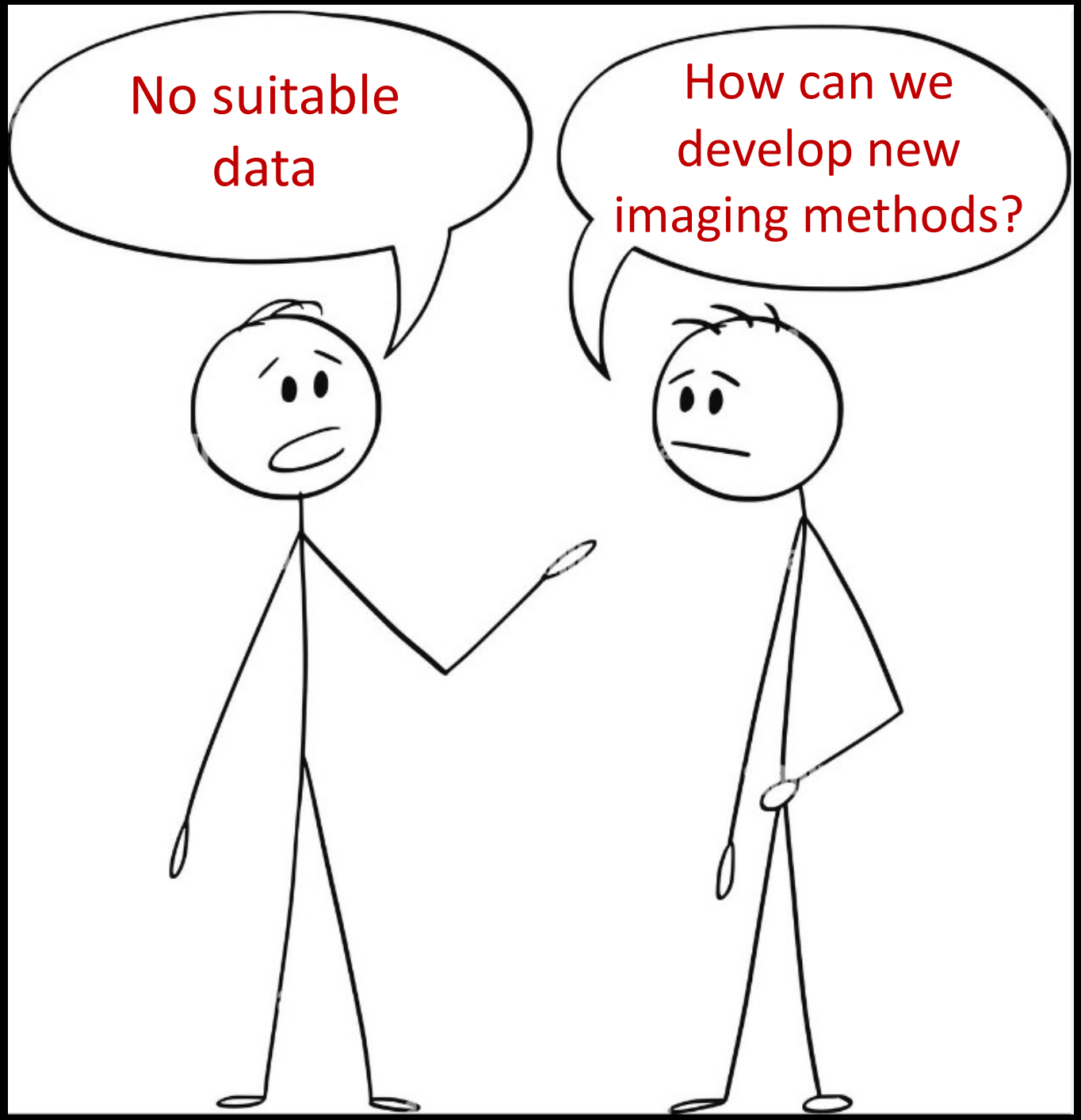
✗ Error metrics - *blind* to the preprocessing

✗ Algorithms trained on processed data could fail for clinical data - they miss important details

Shimron et al., *Implicit data crimes,* PNAS (2022)

Acquisition is too slow – full data is unavailable

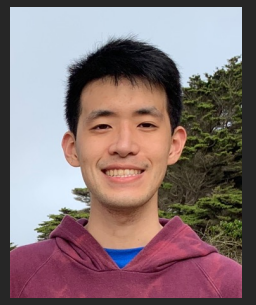




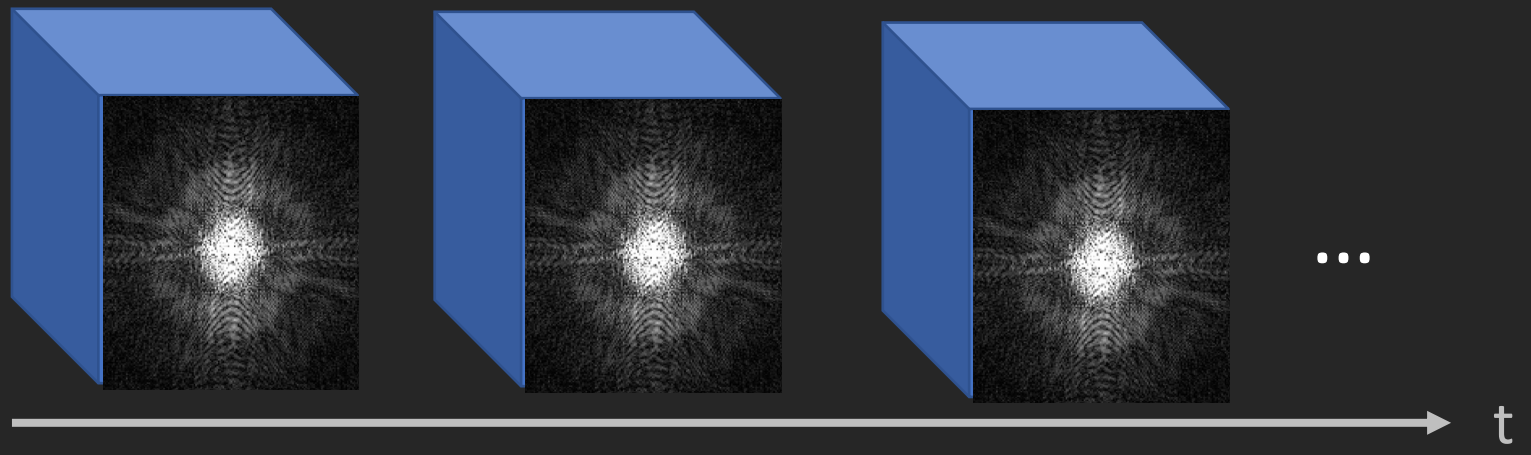
K-Band

K-band: Fast Acquisition & Self-supervised Reconstruction

Han Qi



Fred Wang

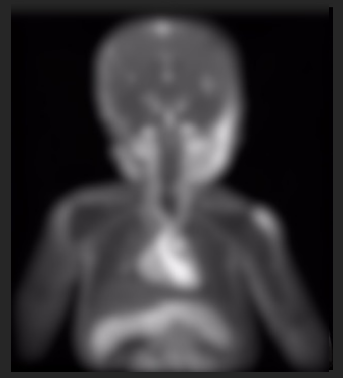


Video is courtesy of Lustig lab

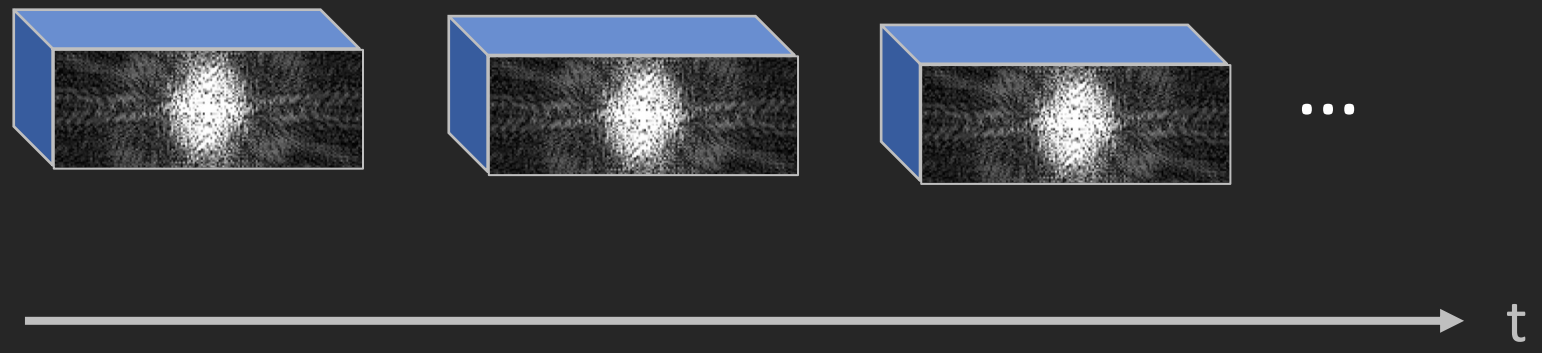
Wang et al., arXiv 2023, “k-band: self-supervised MRI reconstruction”

K-band: Fast Acquisition & Self-supervised Reconstruction

But gives data with
a limited resolution

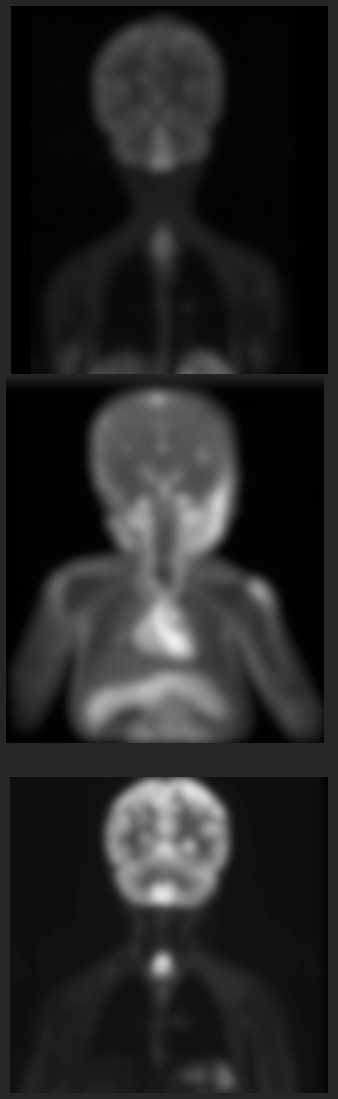


fast!



Wang et al., arXiv 2023, “k-band: self-supervised MRI reconstruction”

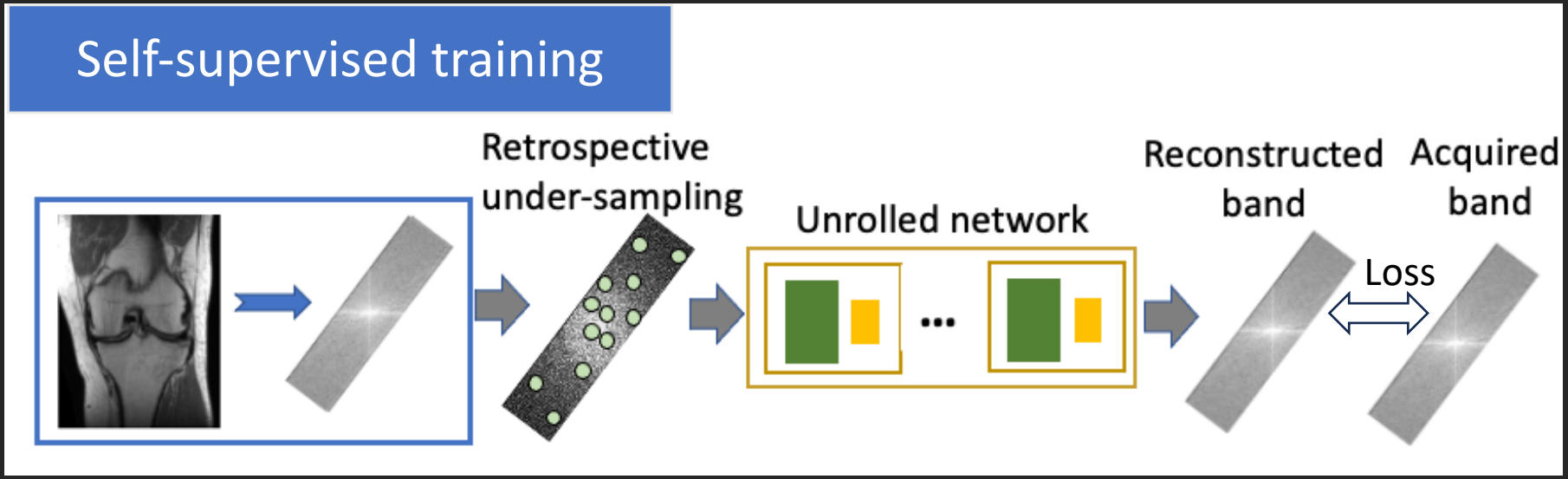
K-band: Fast Acquisition & Self-supervised Reconstruction



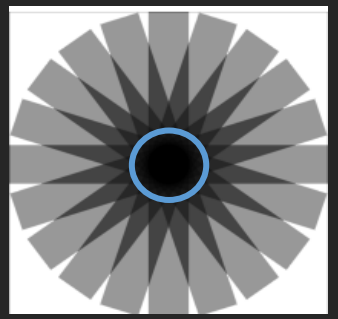
Can we use this limited-res data?

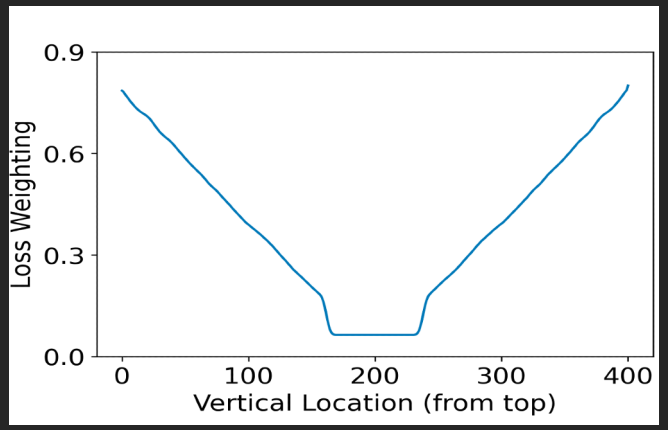
Wang et al., arXiv 2023, “k-band: self-supervised MRI reconstruction”

K-band: Fast Acquisition & Self-supervised Reconstruction



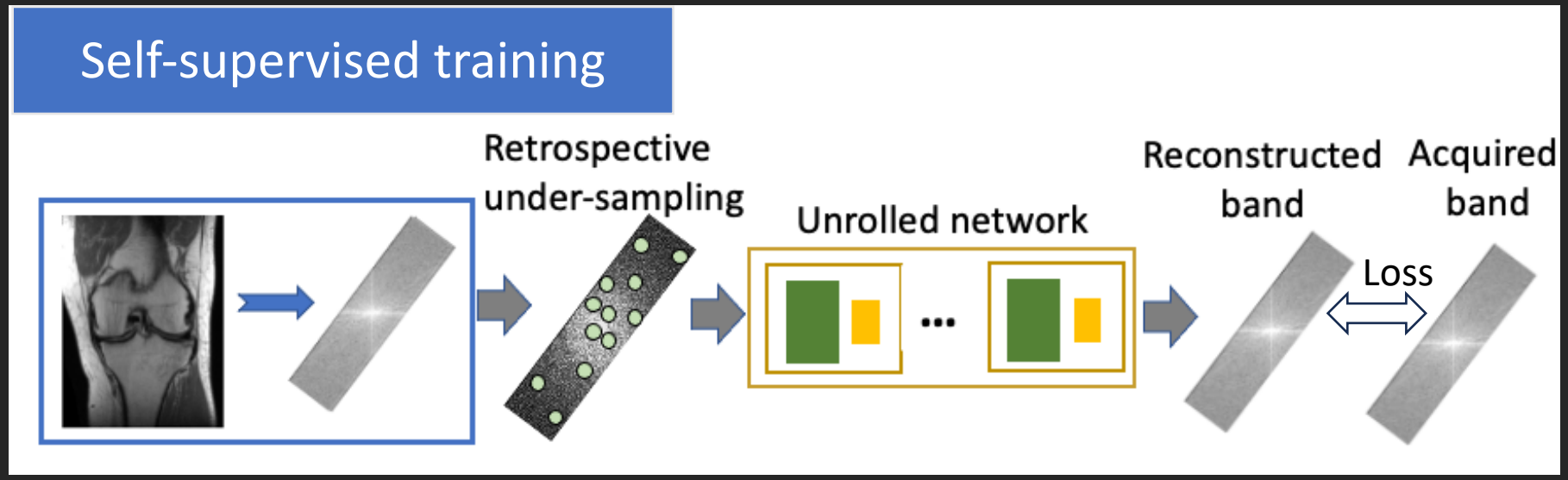
- S.G.D. over k-space subsets
- ✓ Convergence proof
- Loss-weighting mask

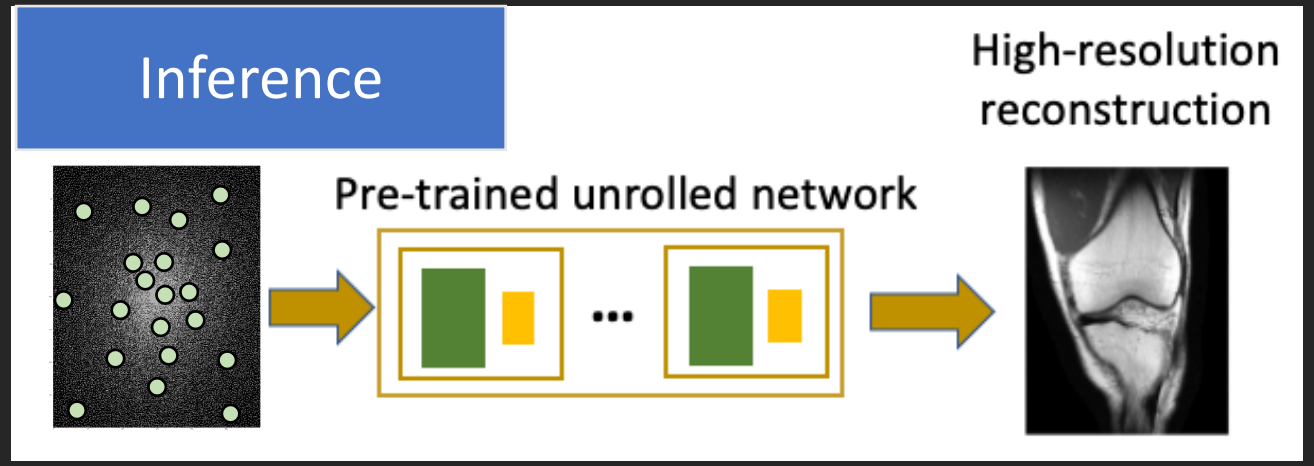




Wang et al., arXiv 2023, “k-band: self-supervised MRI reconstruction”

K-band: Fast Acquisition & Self-supervised Reconstruction





Training on partial, limited-res data

Test-time generalization to high-res data

Wang et al., arXiv 2023, “k-band: self-supervised MRI reconstruction”

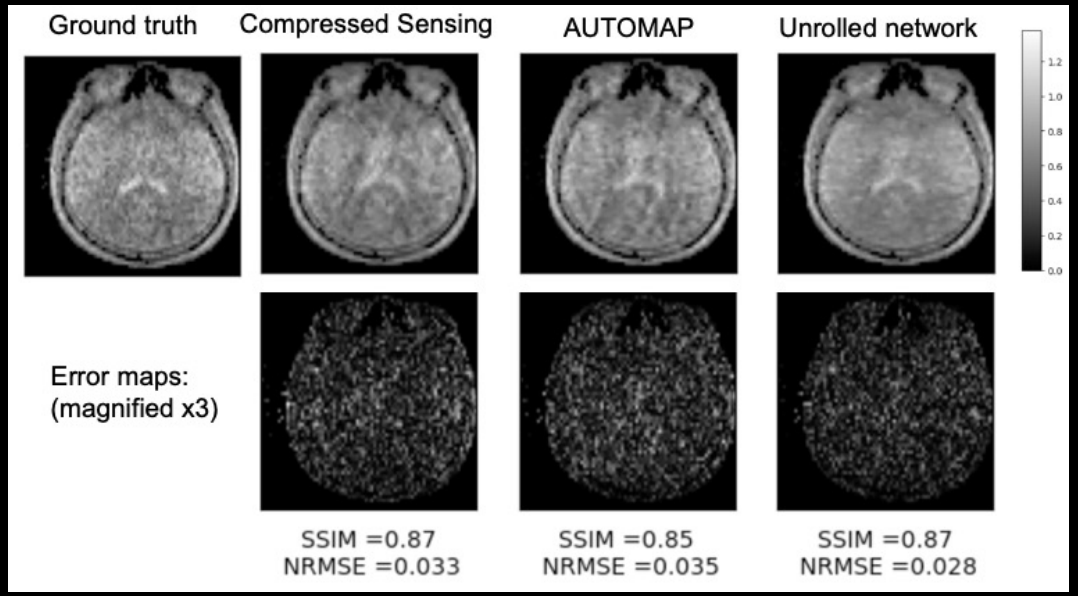
Low-field MRI: Accessible & affordable

HARVARD MEDICAL SCHOOL

Hyperfine.io

- ✓ Low-cost
- ✓ Portable
- - Low SNR
- - Long scans

We improve speed & SNR with AI



Waddington, Shimron, et al., ISMRM 2023

Berkeley | EECS
Electrical Engineering and Computer Sciences

BAIR
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

Stanford

HARVARD UNIVERSITY

LUMC

BME

Miki Lustig

Ke Wang

Alfredo De. Goyeneche

Shreyas Vasanawala

Ali B. Syed

Matthew Rosen

Andrew Webb

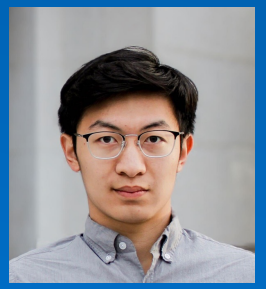
Jonathan Tamir

Haim Azhari

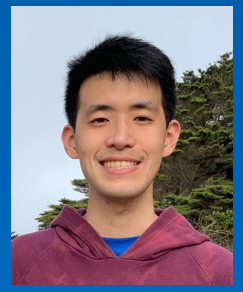
My team

@Berkeley

Han Cui


Boyuan Ma

Han Qi


Fred Wang

N. Devashar

@Technion

Tal Oved

Orel Tsioni

Our lab

Thank You