Advancing Healthcare with Machine Learning

Research Talk, HOPPR.AI

14 Feb 2025

Asad Aali
Research Scientist
Stanford University
asadaali@stanford.edu

HOPPR

Stanford University

About Me

- Research Scientist at Stanford University
  - Lab: Machine Intelligence for Medical Imaging (MIMI)
  - Advisor: Akshay Chaudhari
- Passionate about developing machine learning algorithms for healthcare applications
- Research Interests:
  - Machine Learning
  - Foundation Models
  - Healthcare

Personal Journey

1. Bachelors in Accounting & Finance from LUMS, Pakistan
2. Three years work experience at:
    a. Solutions Consultant - EZOfficeInventory (Tech startup in Austin, TX)
    b. Data Analyst - Plutus21 Capital (Hedge fund startup in Dallas, TX)
3. Master's in Information Technology from UT Austin
    a. Focus: Applied machine learning
    b. Internship: Dell Technologies
4. Master's in Electrical and Computer Engineering from UT Austin
    a. Focus: Biomedical imaging, inverse problems, deep learning
    b. Internship: Amazon
5. Research Scientist at Stanford University
    a. Focus: Machine learning, foundation models, healthcare

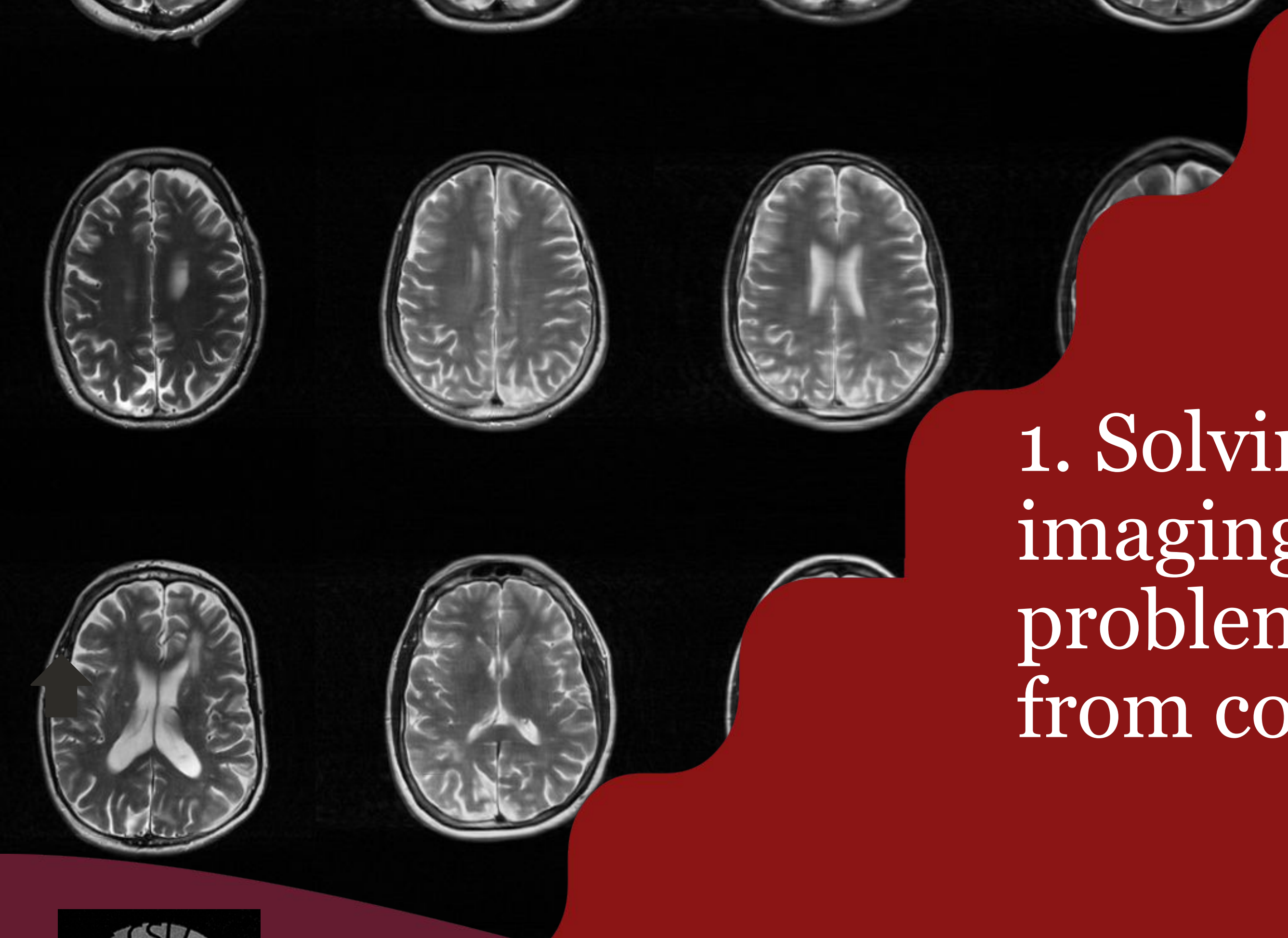
Plan for Today

1 Solving medical imaging inverse problems by learning from corrupted data

2 Optimizing LLM performance in clinical documentation tasks

3 Detecting underdiagnosed medical conditions via opportunistic imaging

Stanford University



1. Solving medical imaging inverse problems by learning from corrupted data

Relevant Publications:

1. Solving inverse problems with generative priors learned from noisy data
   a. Poster presentation, IEEE Asilomar 2023

2. GSURE Denoising enables training of higher quality generative priors for accelerated Multi-Coil MRI Reconstruction
   a. Oral presentation, ISMRM 2024

3. Ambient Diffusion Posterior Sampling: Solving Inverse Problems with Diffusion Models Trained on Corrupted Data
   a. Poster presentation, ICLR 2024

4. Enhancing Deep Learning-Driven Multi-Coil MRI Reconstruction via Self-Supervised Denoising
   a. Currently in review

Motivation

- Deep Diffusion Probabilistic (Generative) Models are powerful tools for accelerated MRI reconstruction
  - ✔ Exploit large training databases
  - ✔ Decouples from the forward model

Song *NeurIPS* (2019), Kingma *ICLR* (2014), Goodfellow *NeurIPS (2014)*

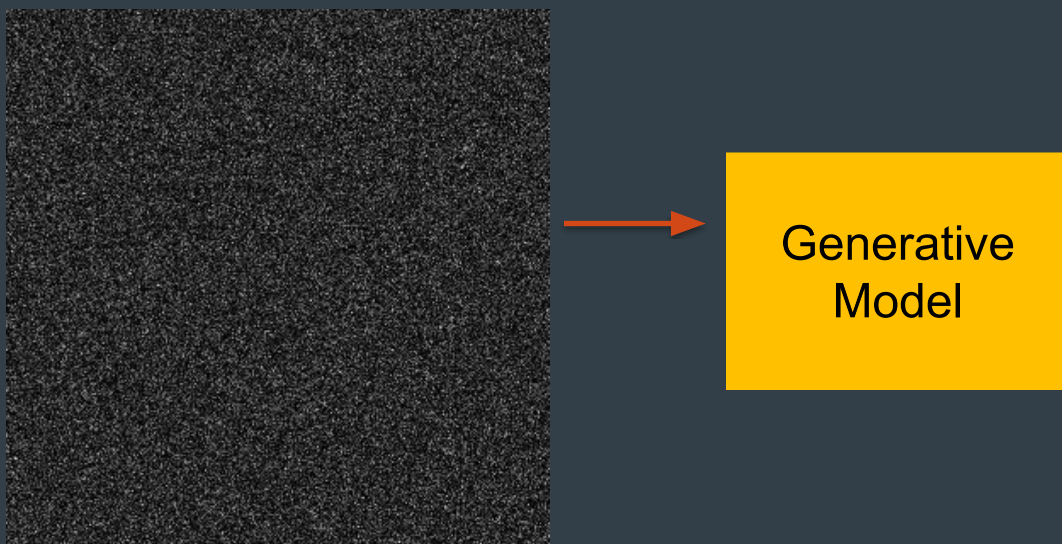
Motivation

- Generative models learn priors for MR images.

Motivation

- Generative models learn priors for MR images.

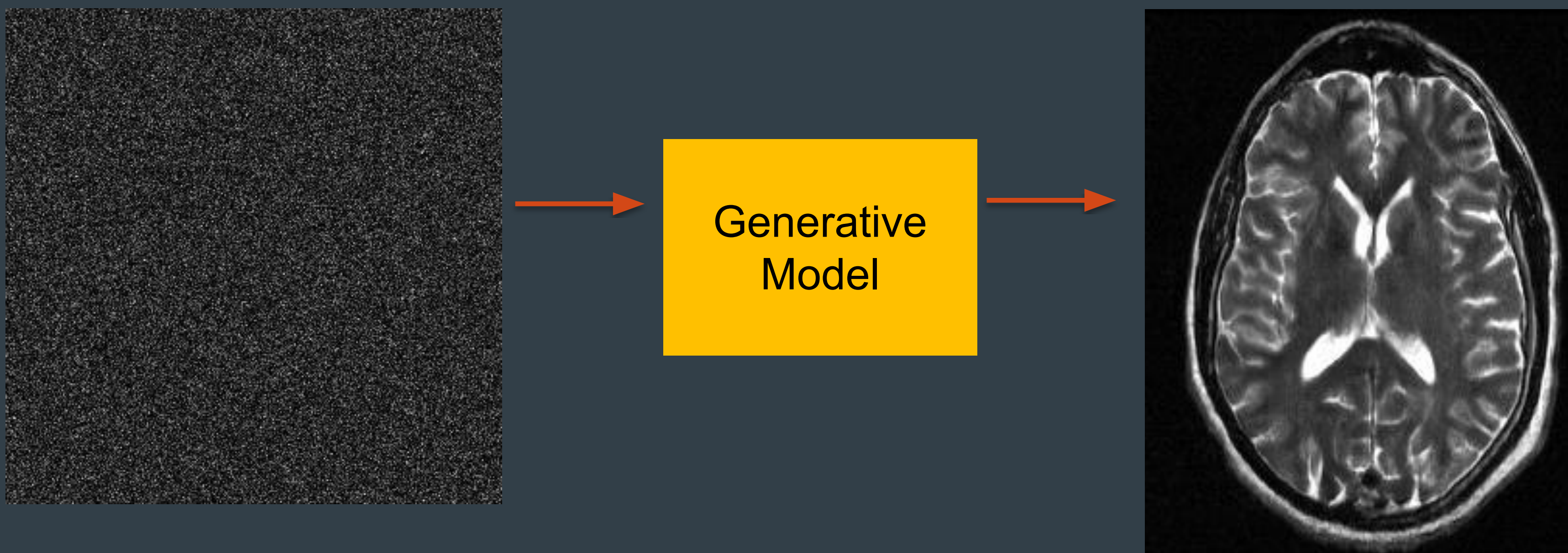
Sample from Gaussian Distribution



Motivation

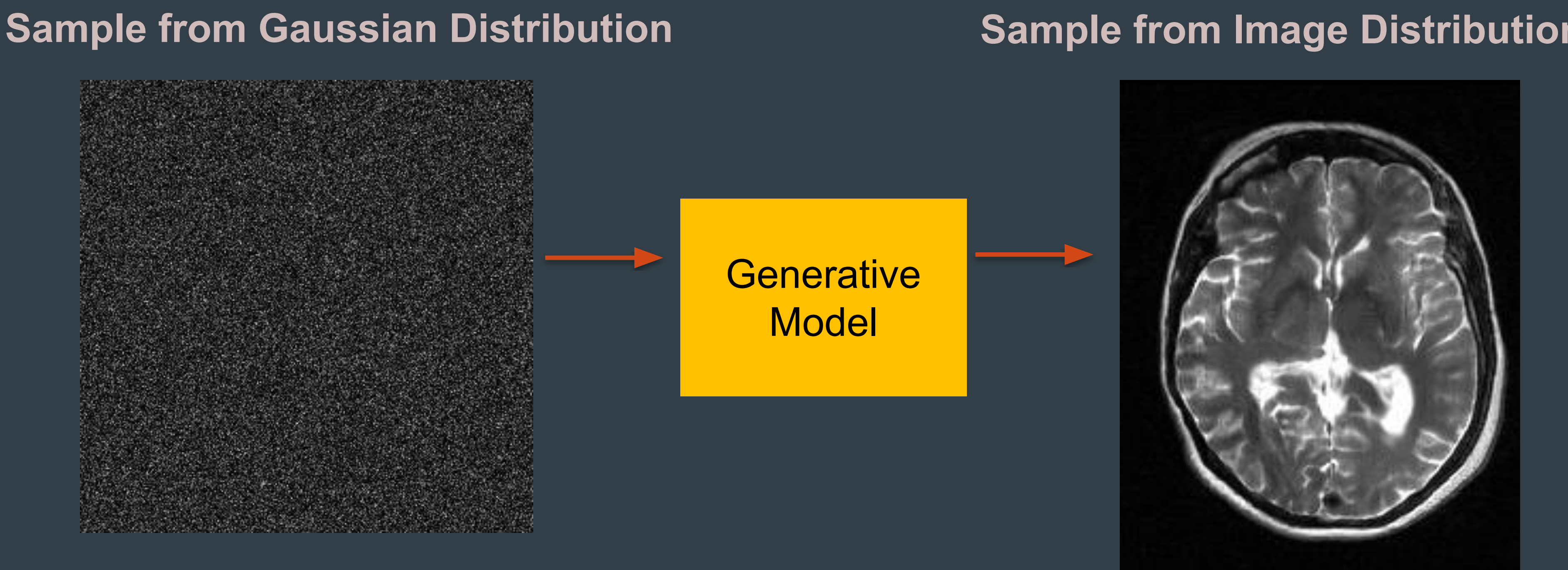
- Generative models learn priors for MR images.

Sample from Gaussian Distribution **Sample from Image Distribution**



Motivation

- Generative models learn priors for MR images.



Motivation

- Generative models learn priors for MR images.

Sample from Gaussian Distribution

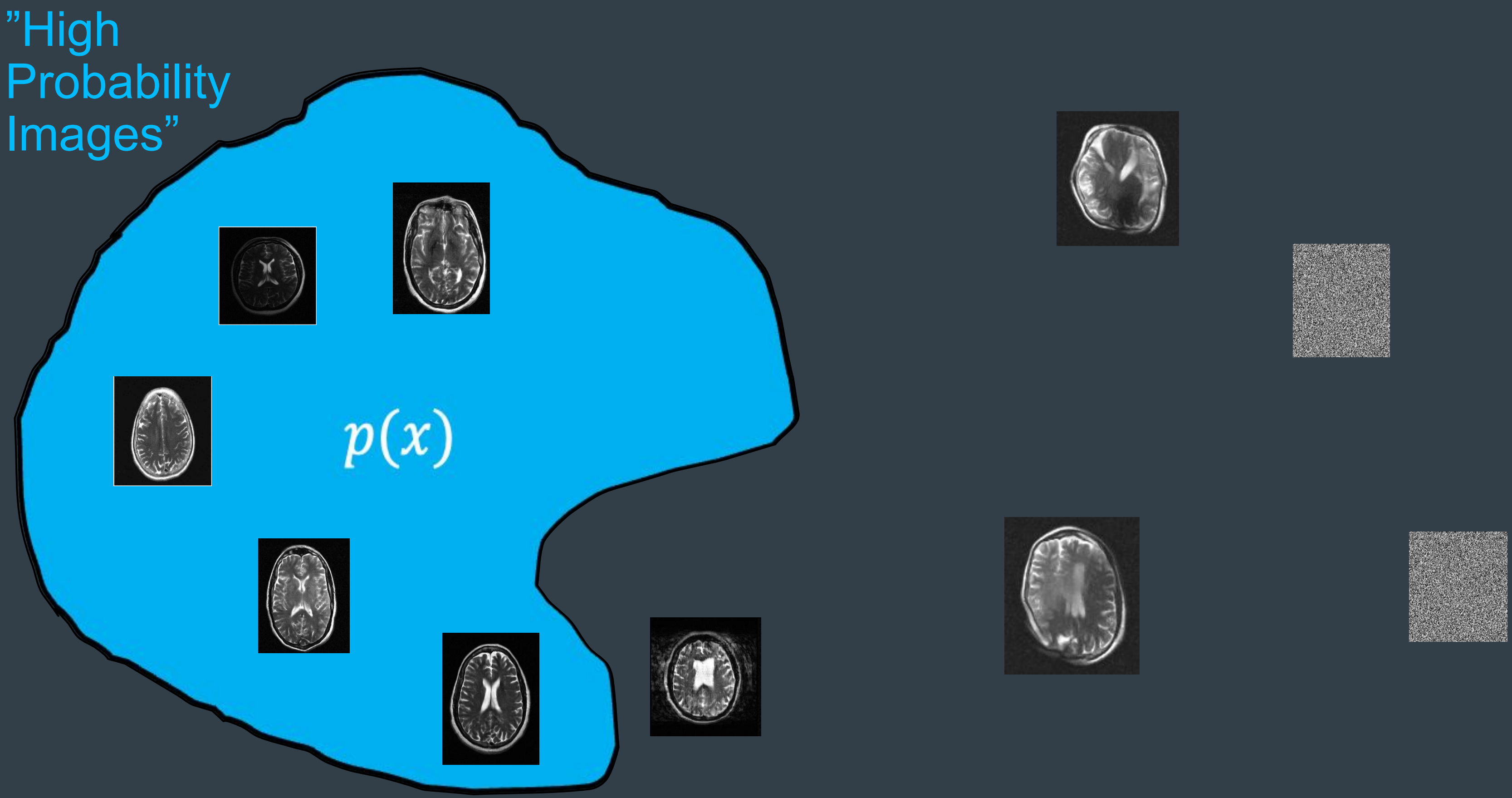
$$\sim \mathcal{N}(z) \rightarrow \boxed{\text{Generative Model}} \rightarrow \sim p(x)$$

Sample from Image Distribution

Image Prior

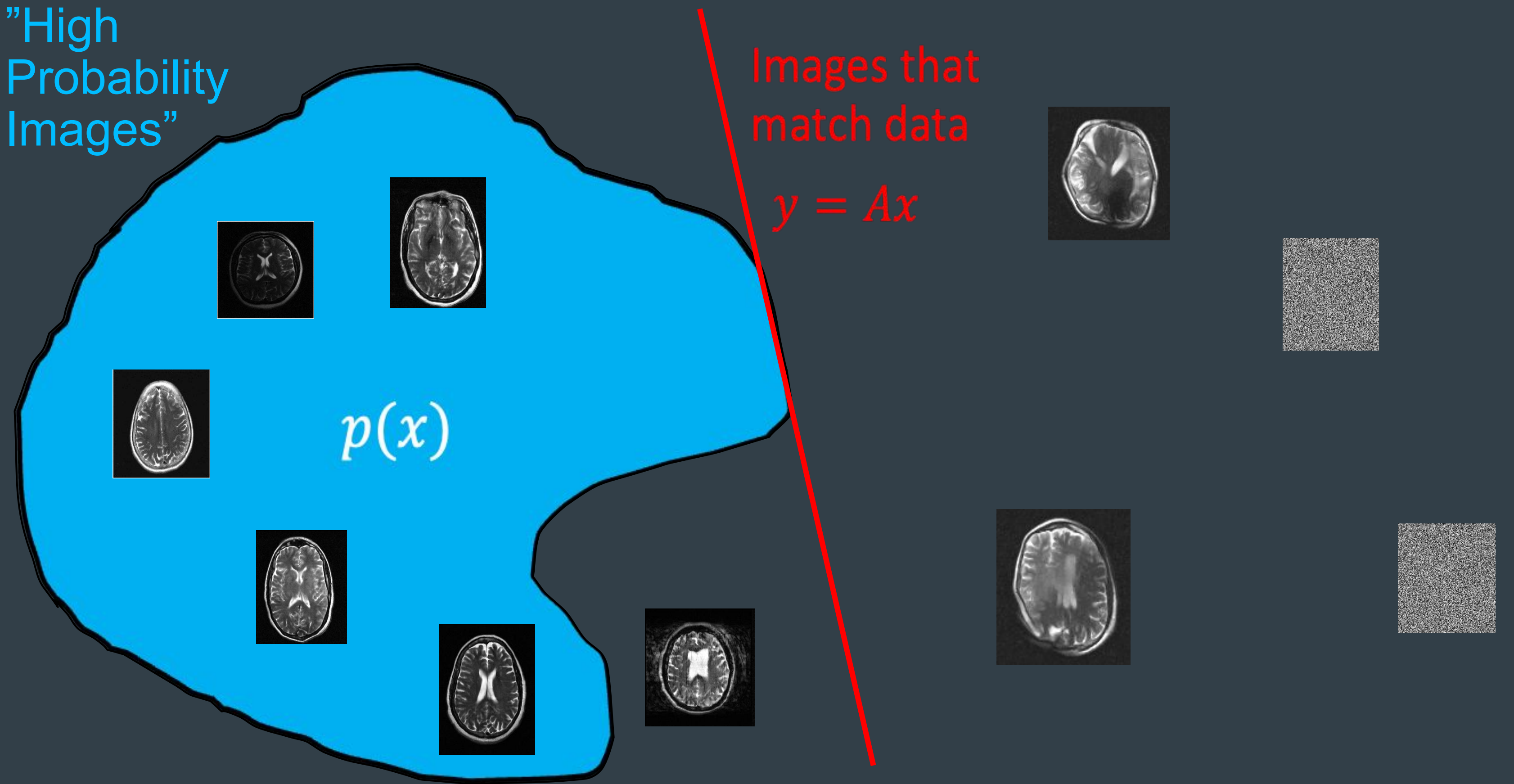
Motivation

- Generative Models to guide accelerated MRI reconstructions.



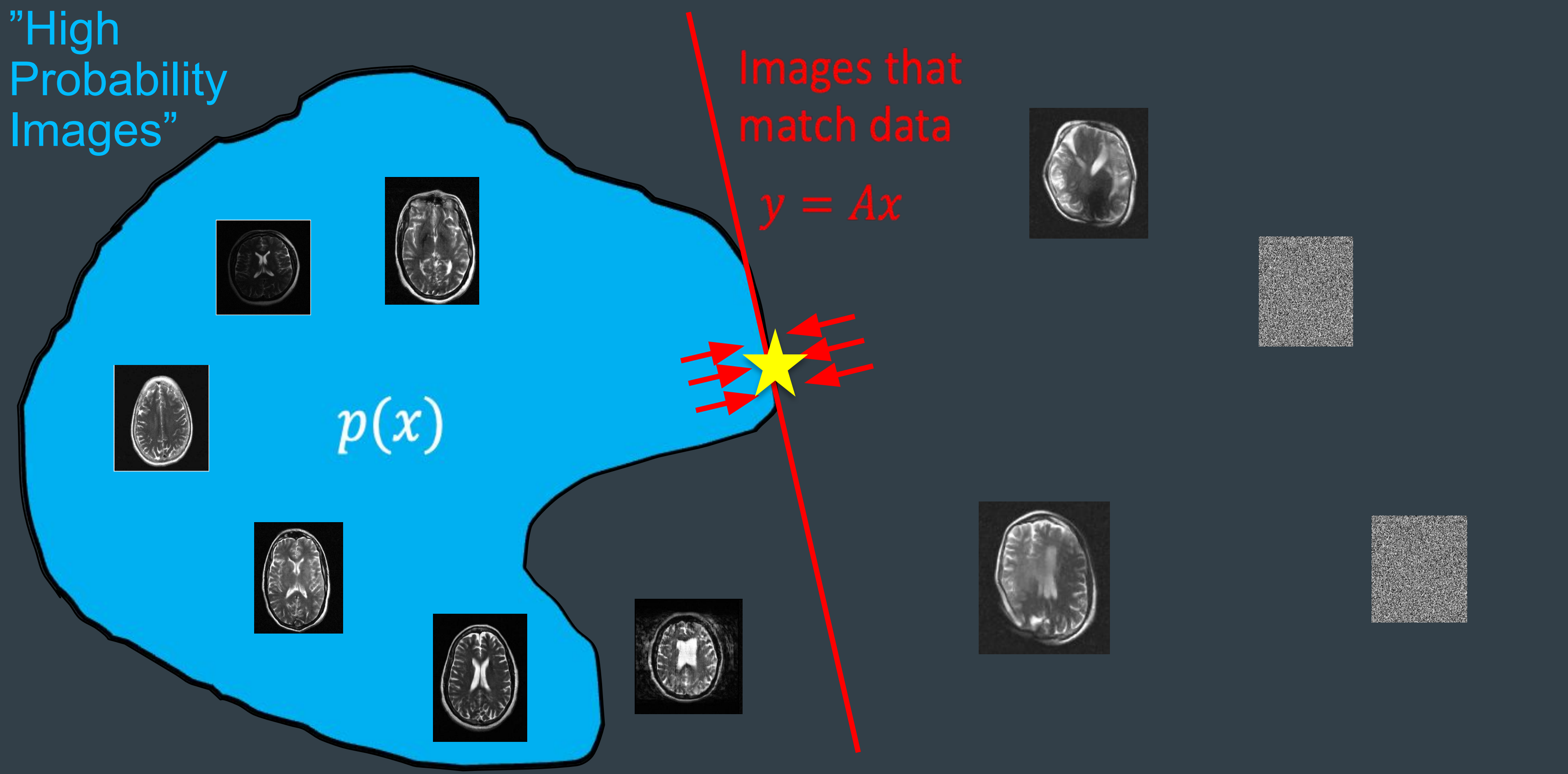
Motivation

- Generative Models to guide accelerated MRI reconstructions.



Motivation

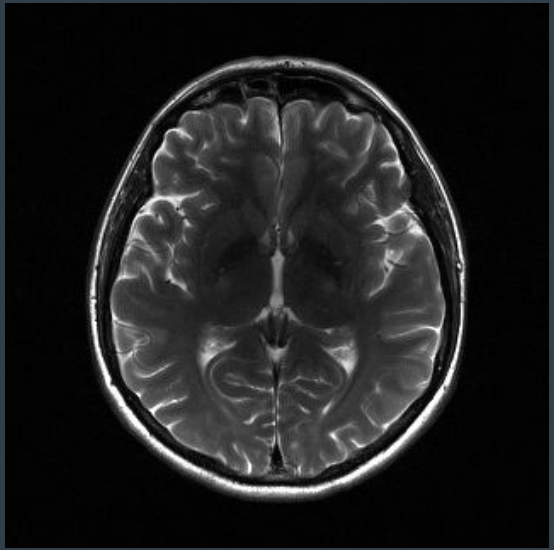
- Generative Models to guide accelerated MRI reconstructions.



Motivation

- Generative Models rely on large amounts of ***high-quality data***.
- MRI data are ***inherently noisy***[1,2], multi-coil k-space.

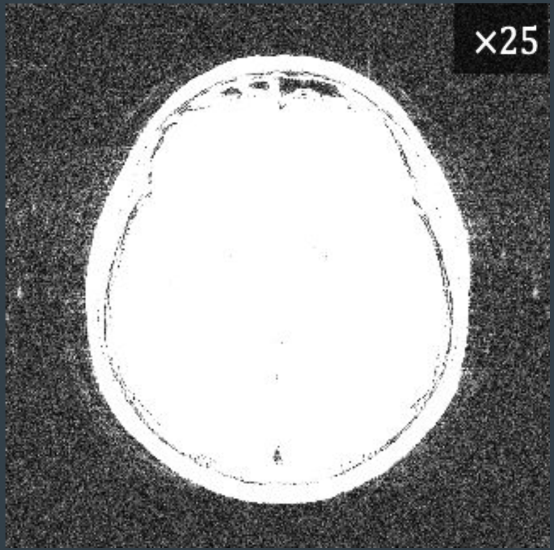
~~Training Dataset~~ Processed Dataset



Motivation

- Generative Models rely on large amounts of ***high-quality data***.
- MRI data are ***inherently noisy***[1,2], multi-coil k-space.

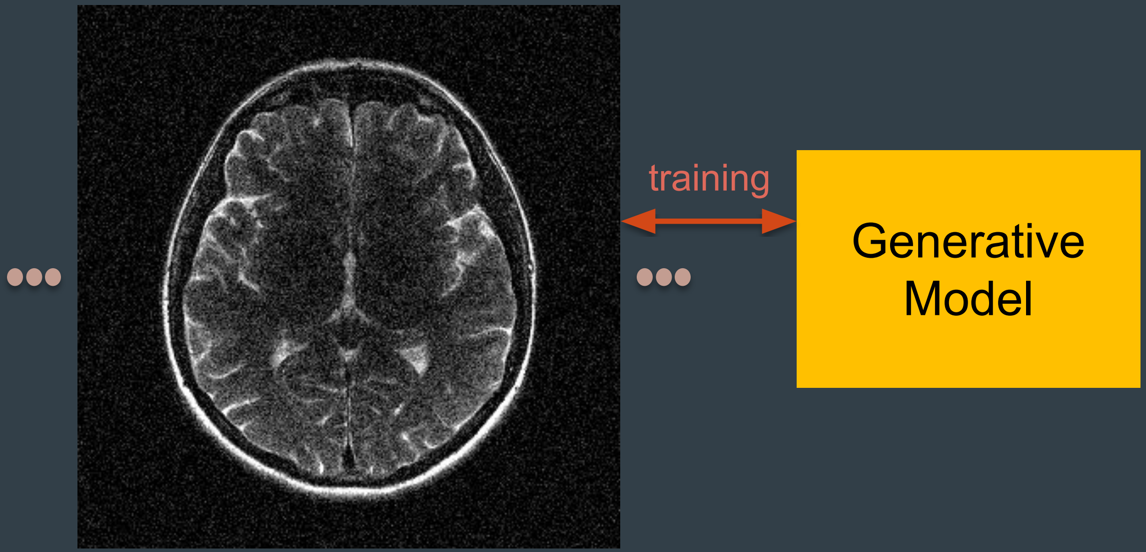
~~Training Dataset~~ Processed Dataset



Motivation

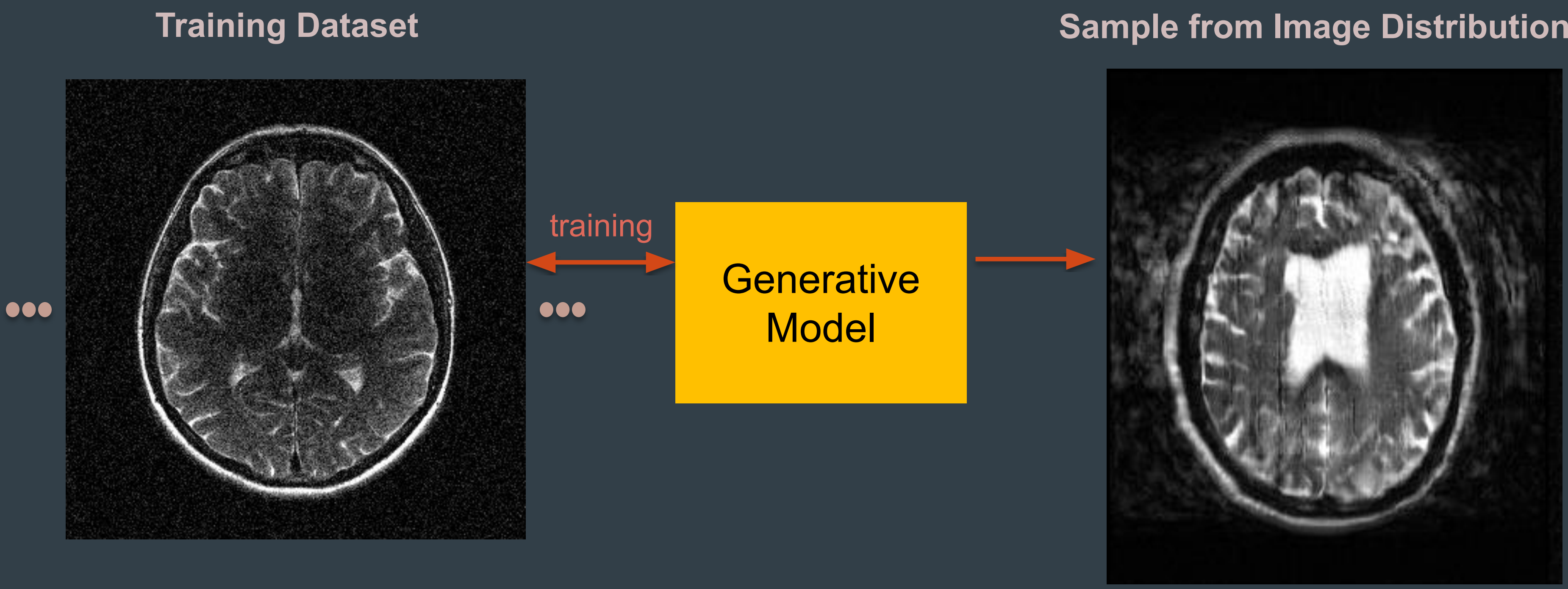
- Training generative models with noisy datasets leads to a poor prior.
- **Reconstruction performance depends on accuracy of priors**

Training Dataset



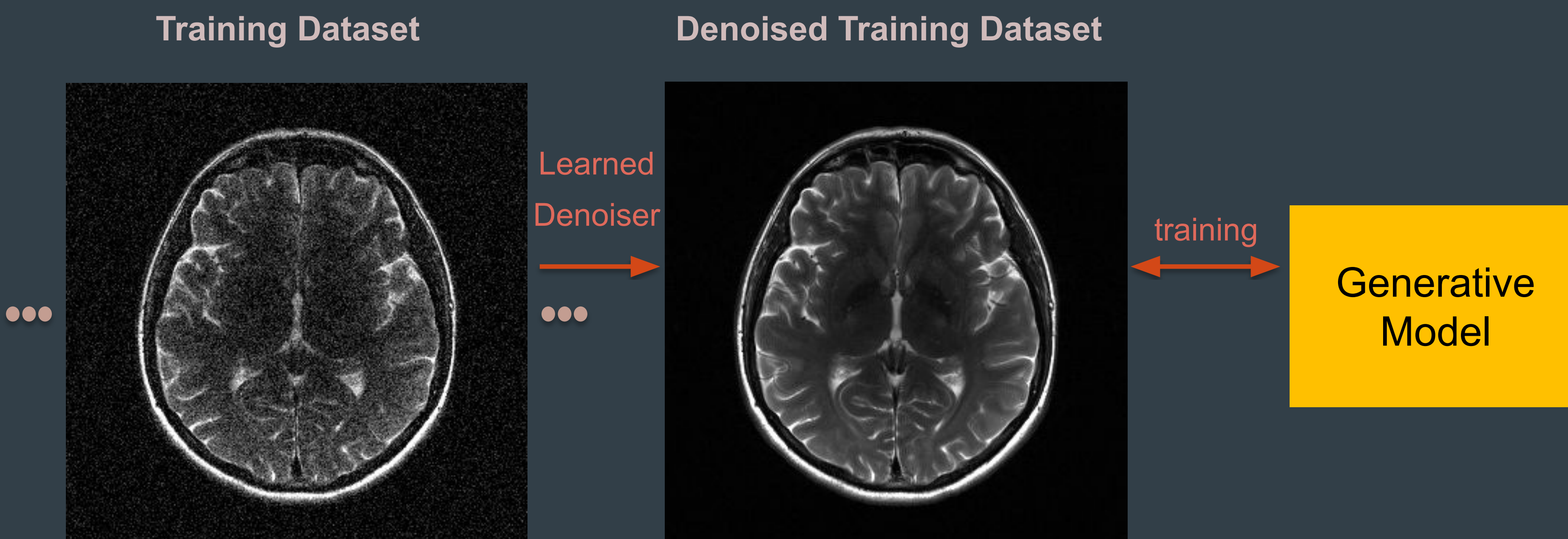
Motivation

- Training generative models with noisy datasets leads to a poor prior.
- **Reconstruction performance depends on accuracy of priors**



Purpose

- Learn model to denoise dataset before training generative models



Purpose

Training a denoiser without access clean training samples.

Investigate the **effectiveness** of self-supervised **denoising** as a pre-processing step to learning **generative priors** for accelerated MRI reconstruction

Problem Formulation

Problem Formulation

Goal: learn the clean distribution using fully sampled, noisy multi-coil MRI data

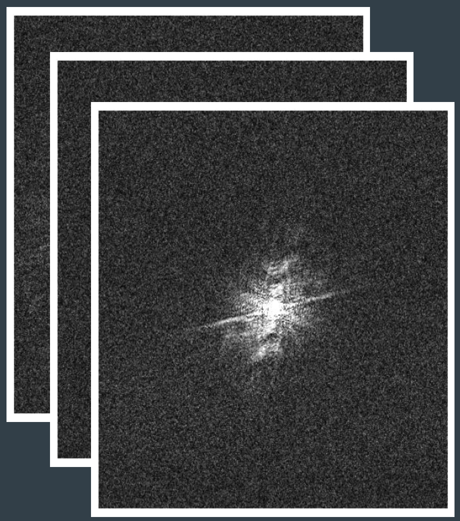
$$y = FSx + noise$$

Problem Formulation

Goal: learn the clean distribution using fully sampled, noisy multi-coil MRI data

$$y = FSx + noise$$

y



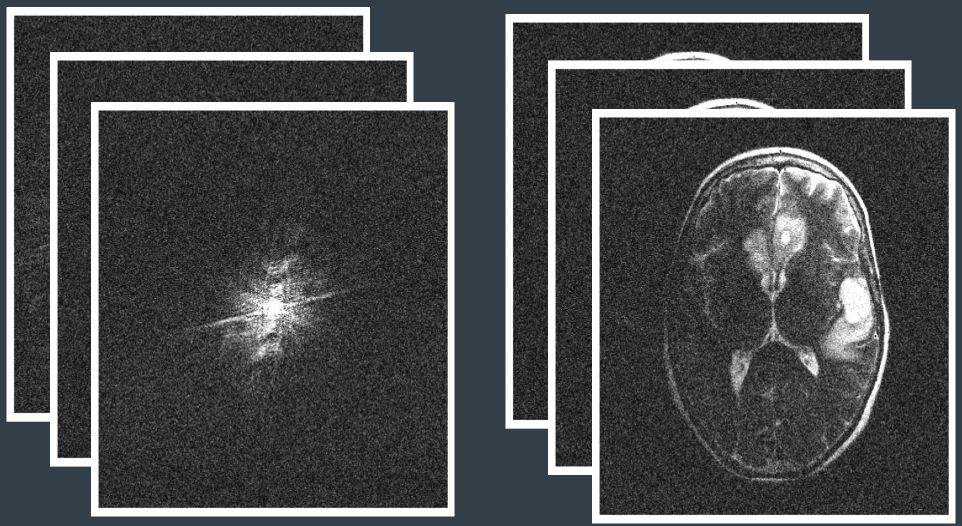
Original K-Space

Problem Formulation

Goal: learn the clean distribution using fully sampled, noisy multi-coil MRI data

$$y = FSx + noise$$

$y \longrightarrow F^{\mathrm{H}}y$



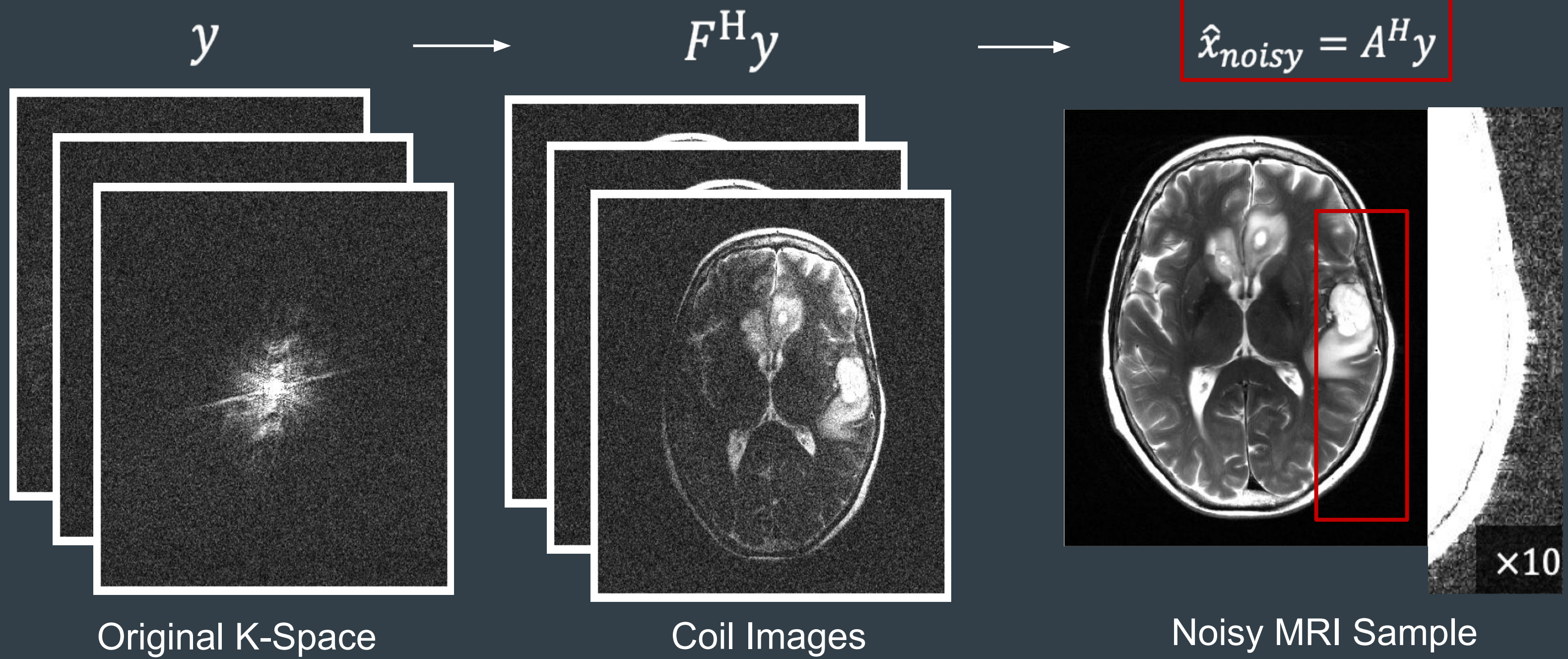
Original K-Space

Coil Images

Problem Formulation

Goal: learn the clean distribution using fully sampled, noisy multi-coil MRI data

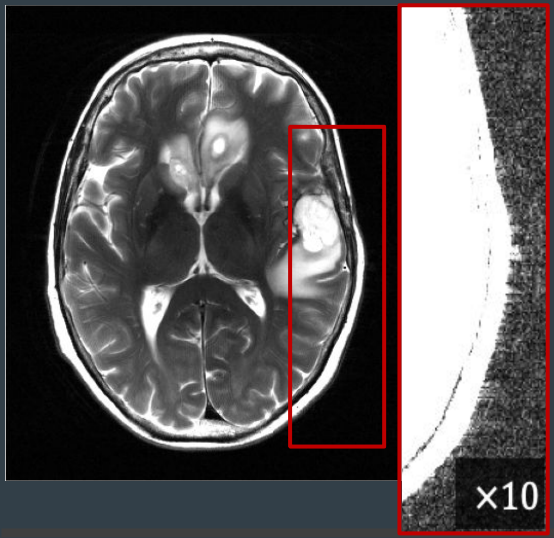
$$y = FSx + noise$$



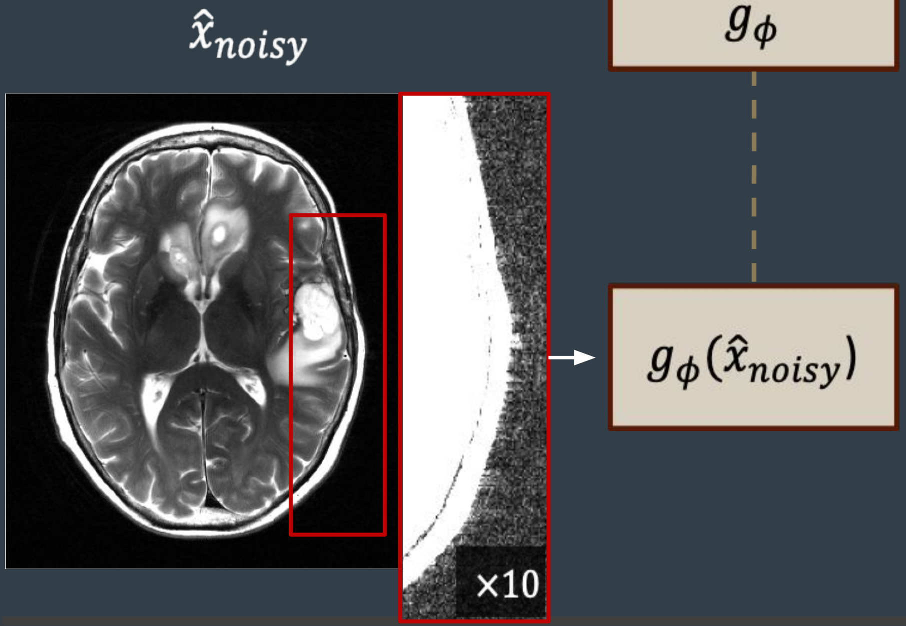
Proposed Methods

Proposed Methods

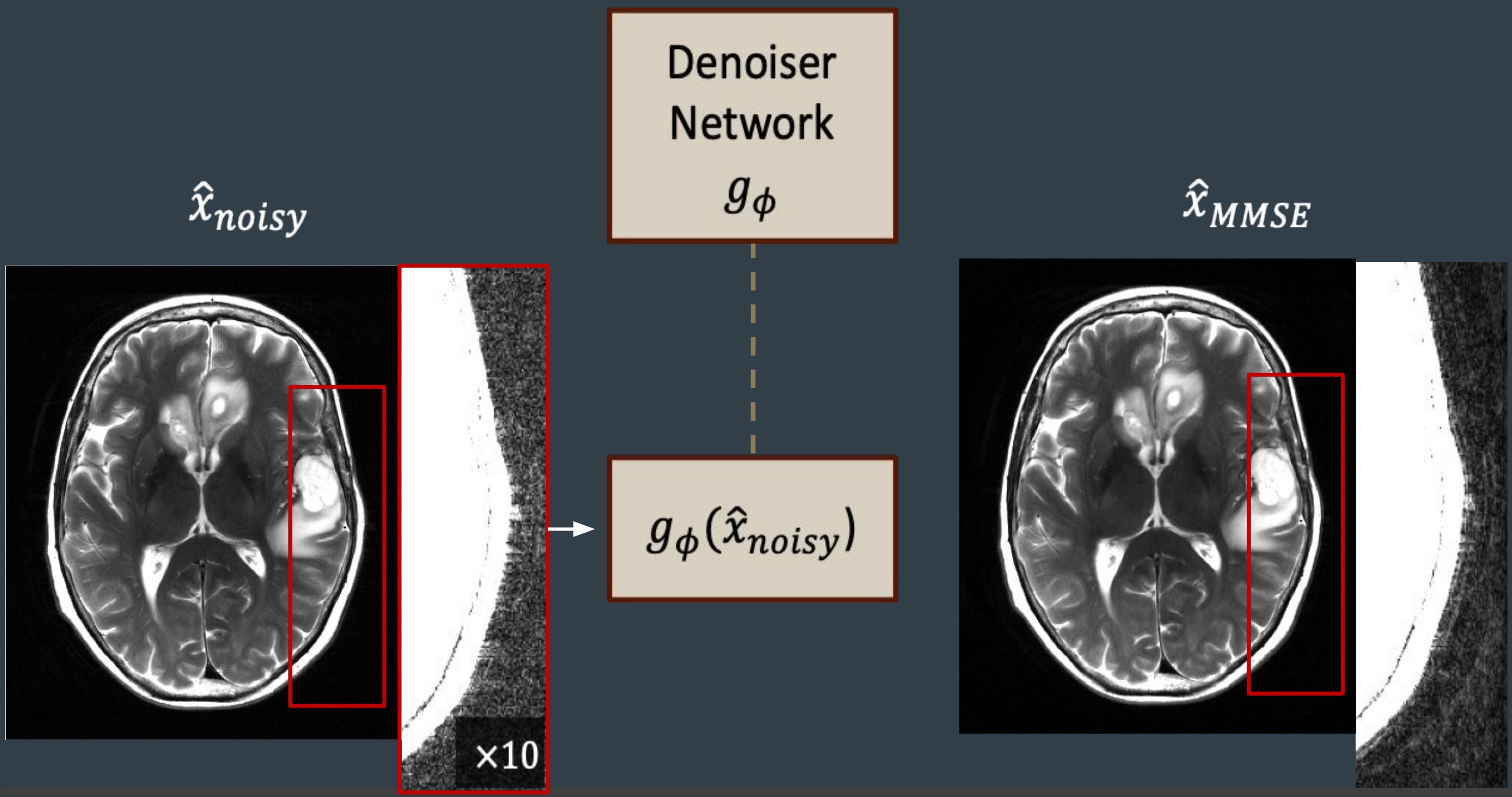
$\hat{x}_{noisy}$



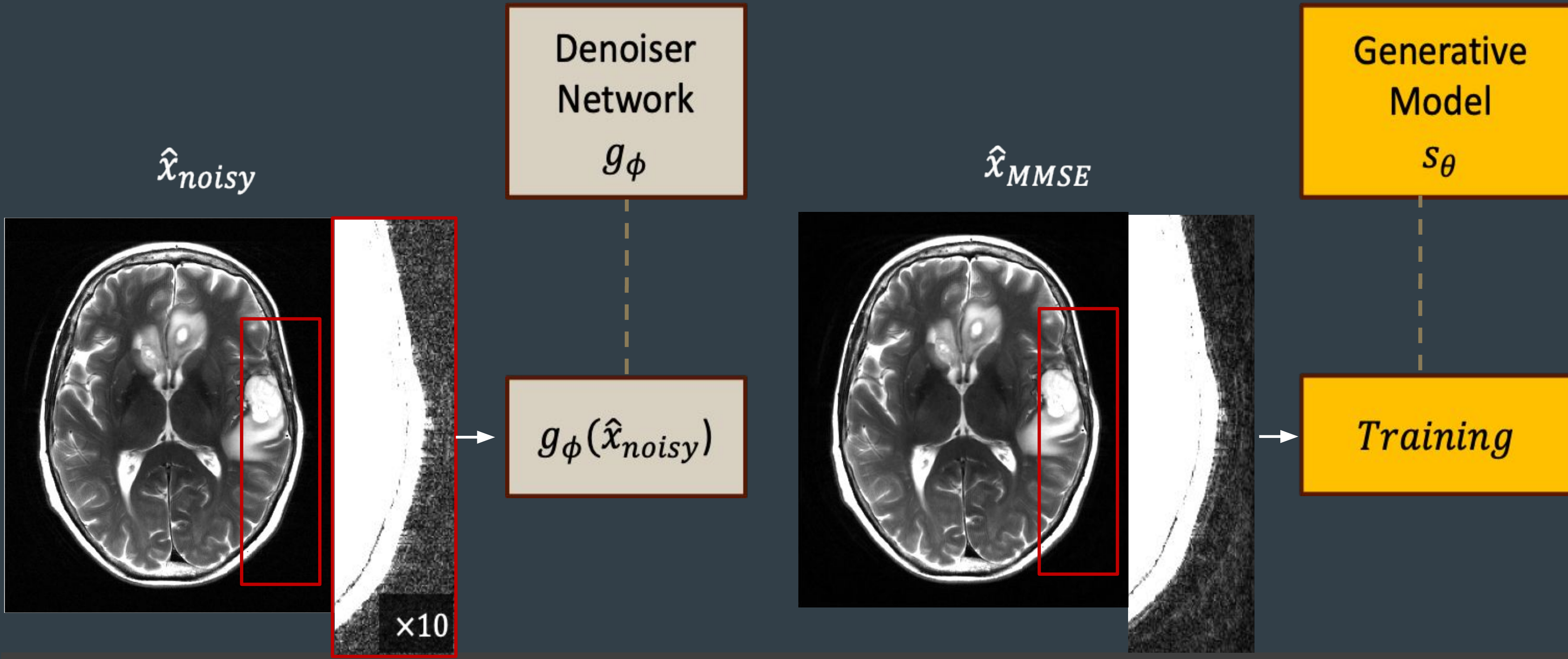
Proposed Methods



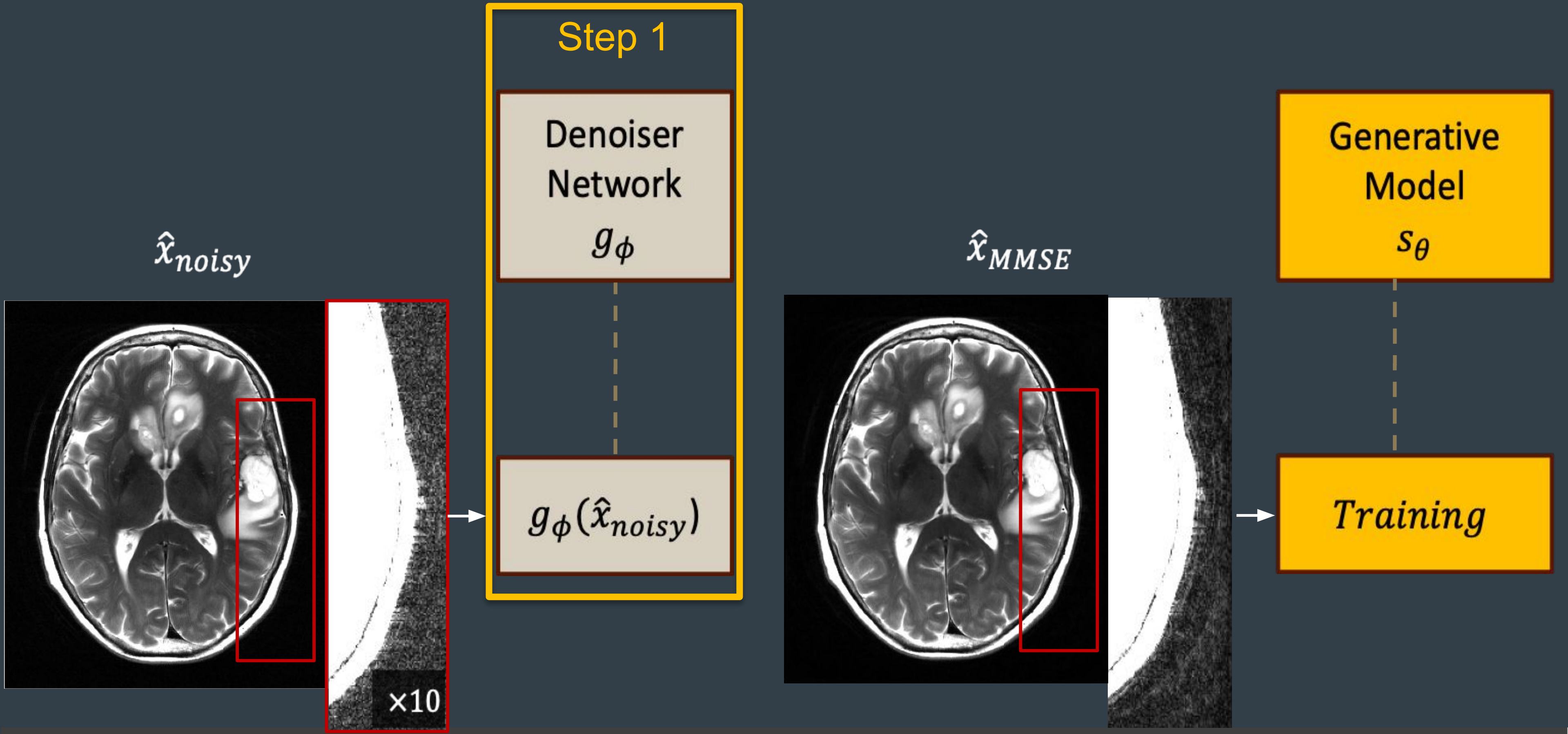
Proposed Methods



Proposed Methods



Proposed Methods



Self-Supervised Denoising

Training a denoiser with only access to noisy data

A is a Linear Forward Operator (Fully-Sampled) -> **GSURE**[1,2,3]

$$y = FSx + noise$$

[1]Soltanayev, *NeurIPS,* 2018, [2]Eldar, *IEEE Transactions on Signal Processing,* 2008, [3]Kawar, *TMLR,* 2023

Generalized SURE (GSURE) Basics

- GSURE[1]: Self-supervised denoising technique, only need access to:
  - $\hat{x}_{noisy}$ → **Noisy Samples**
  - Noise **Covariance Matrix**

- An unbiased estimate of the MSE

$$E\left[\mathrm{L}_{GSURE}\right] = E\left\|g_{\phi}\left(\hat{x}_{noisy}\right) - x\right\|$$

[1]Eldar, *IEEE Transactions on Signal Processing, 2008*

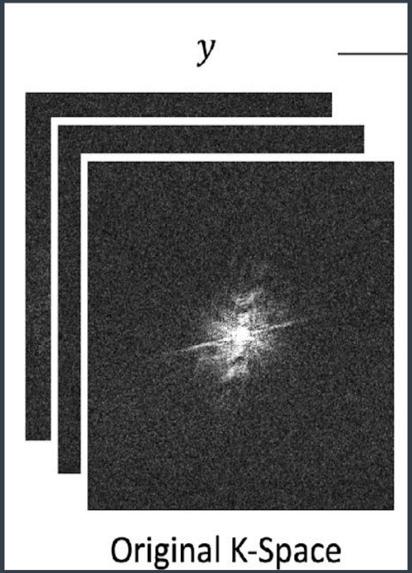
Pre-Processing Noisy Dataset

- GSURE loss requires the **noise covariance** matrix
- **Pre-whitening** (noise covariance $= I$) makes computation relatively straight-forward

Kellman *MRM* (2005)

Pre-Processing Noisy Dataset

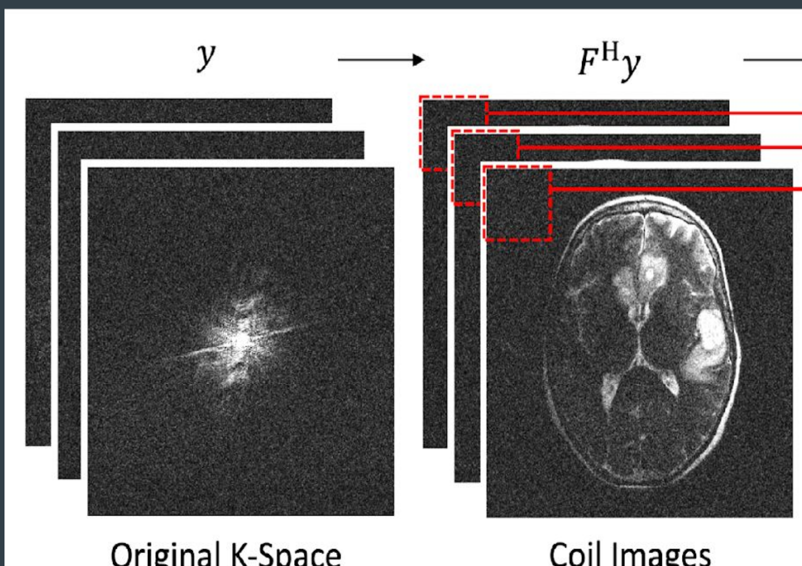
- GSURE loss requires the **noise covariance** matrix
- **Pre-whitening** (noise covariance = I) makes computation relatively straight-forward



Kellman *MRM* (2005)

Pre-Processing Noisy Dataset

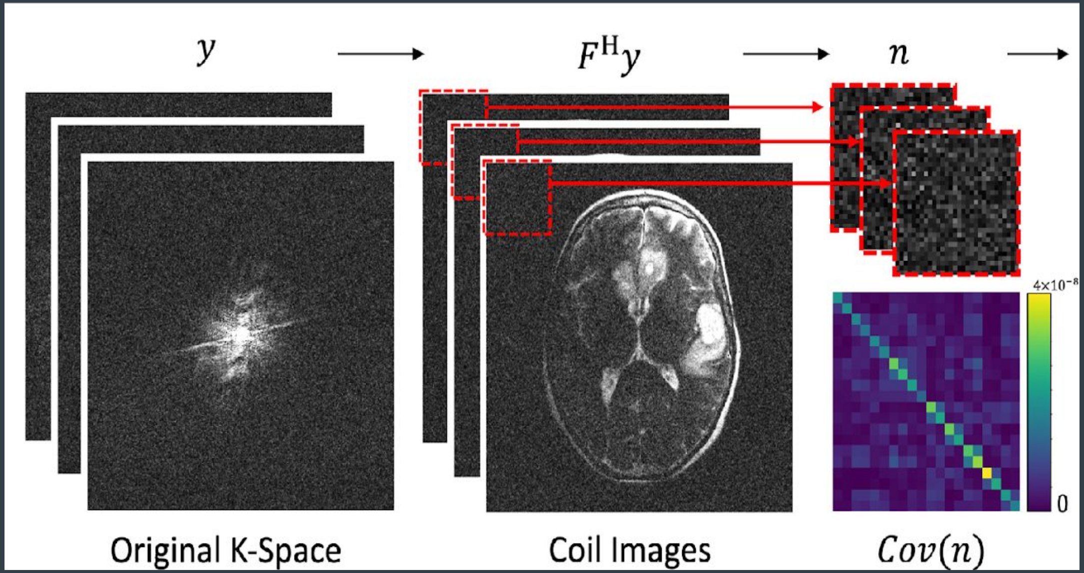
- GSURE loss requires the **noise covariance** matrix
- **Pre-whitening** (noise covariance = I) makes computation relatively straight-forward



Kellman *MRM* (2005)

Pre-Processing Noisy Dataset

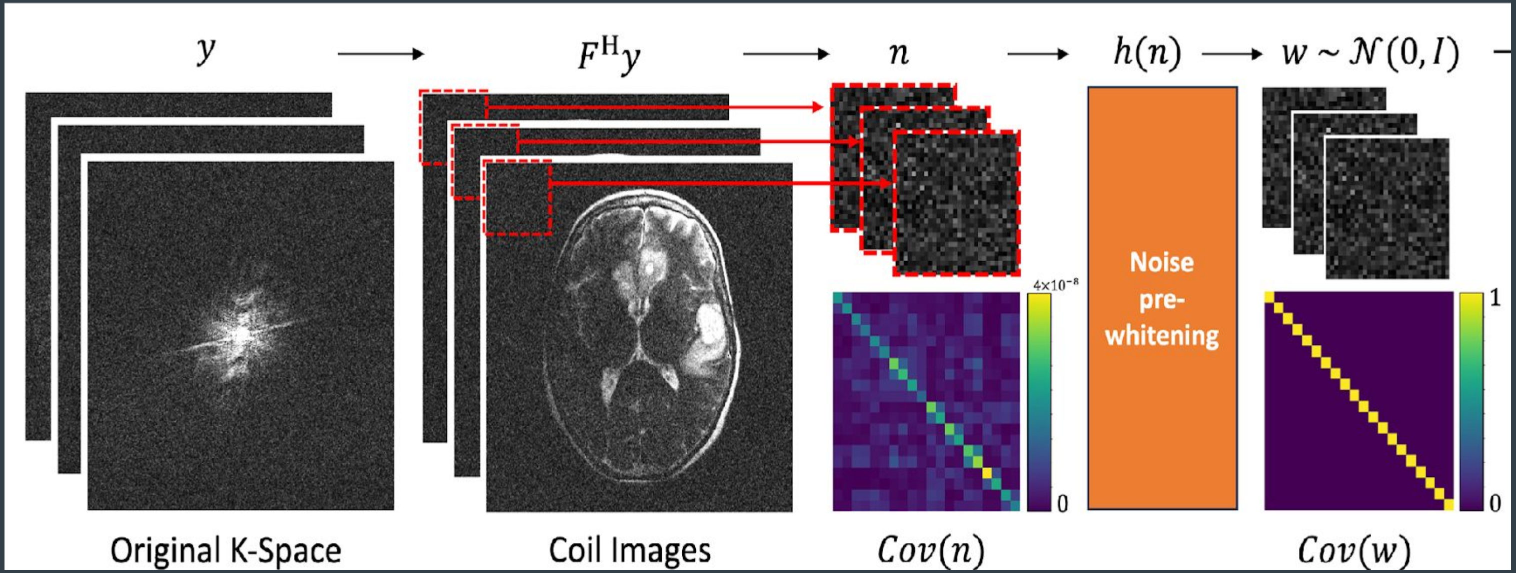
- GSURE loss requires the **noise covariance** matrix
- **Pre-whitening** (noise covariance = I) makes computation relatively straight-forward



Kellman *MRM* (2005)

Pre-Processing Noisy Dataset

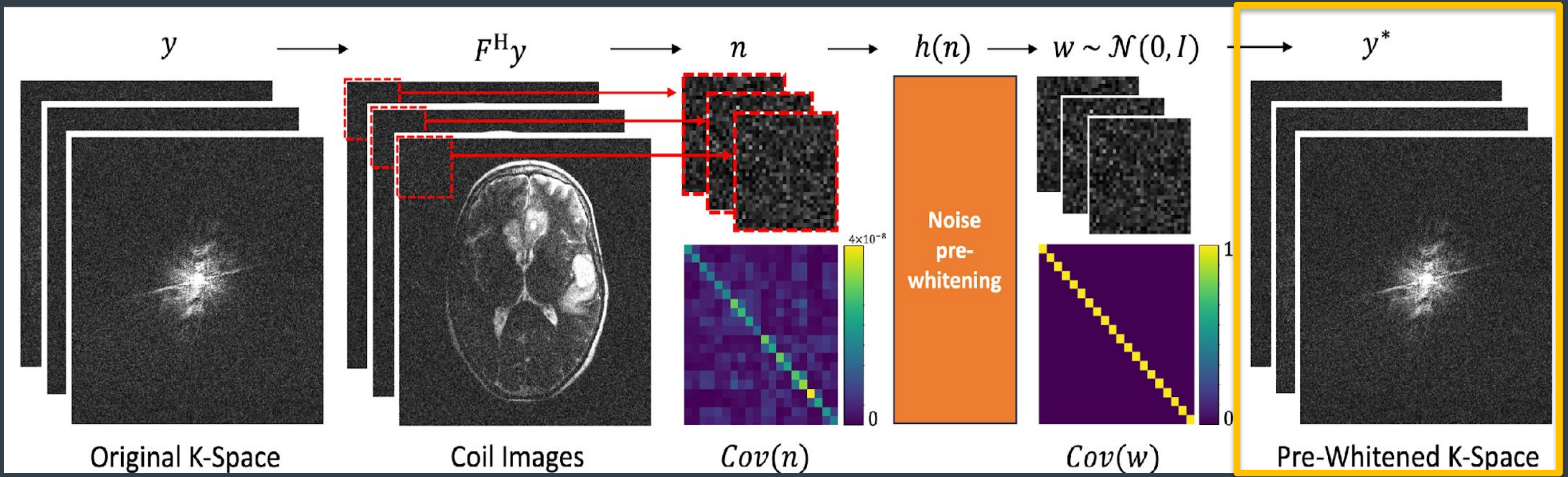
- GSURE loss requires the **noise covariance** matrix
- **Pre-whitening** (noise covariance = I) makes computation relatively straight-forward



Kellman *MRM* (2005)

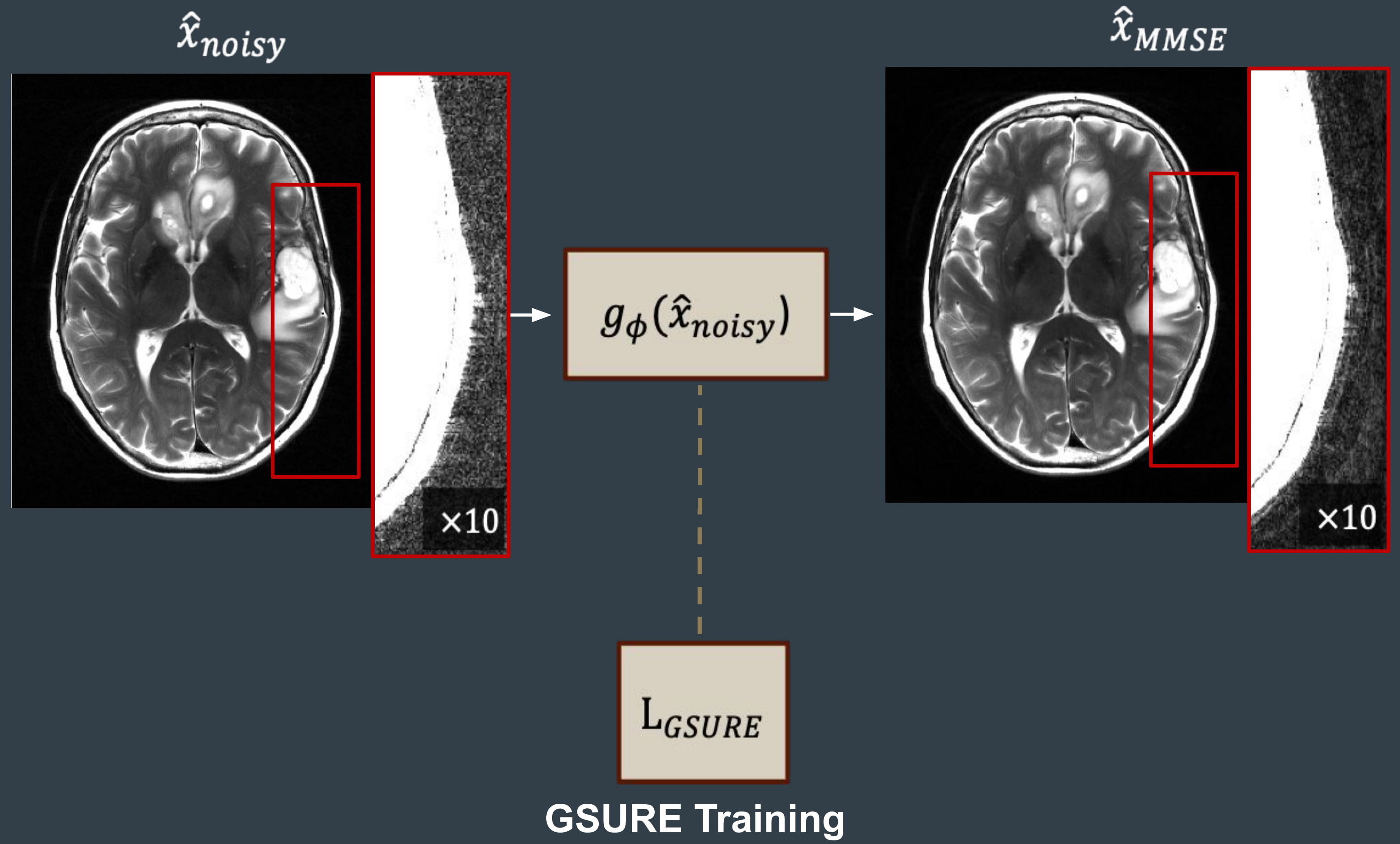
Pre-Processing Noisy Dataset

- GSURE loss requires the **noise covariance** matrix
- **Pre-whitening** (noise covariance = I) makes computation relatively straight-forward

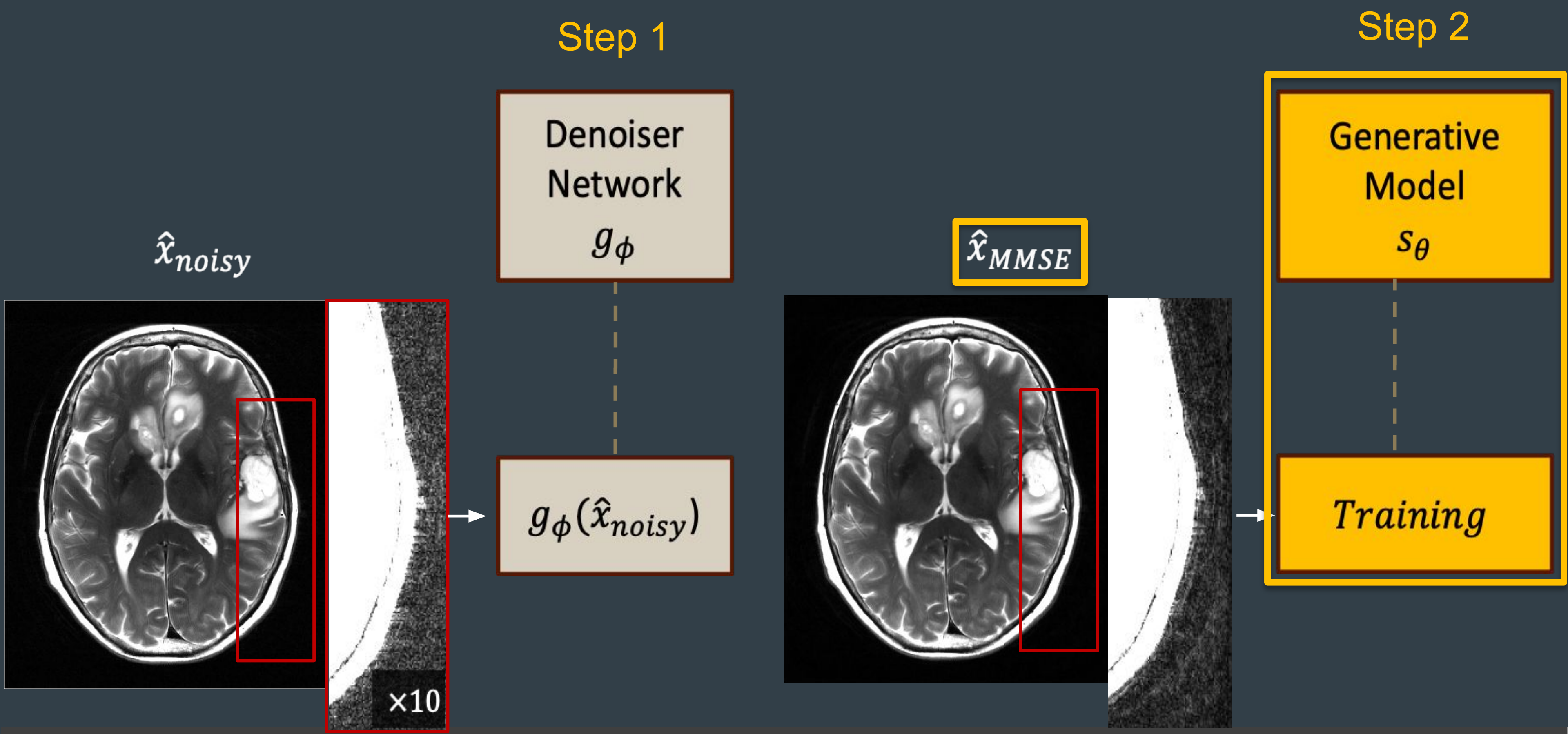


Kellman *MRM* (2005)

GSURE Denoising - Summary



Proposed Methods



Diffusion Probabilistic (Generative) Model Details

- Elucidating the Design Space of Diffusion-Based Generative Models (EDM)[1]

- Posterior sampling (MRI reconstruction) with Diffusion Posterior Sampling[6]

[1]Karras, *Neurips* 2022 [2]Song, *UAI,* 2018 [3]Vincent, *MIT Press* 2011 [4]Song, *NeurIPS* 2019 [5]Dhariwal, *NeurIPS* 2021 [6]Chung, *ICLR* 2023

Experiments

1. Evaluation of **Self-Supervised Denoising** (GSURE)
2. **Prior sampling** performance of score models trained on Noisy vs. GSURE Denoised data
3. **Accelerated MRI Reconstruction** performance of score models trained on Noisy vs. GSURE Denoised data

Experiments

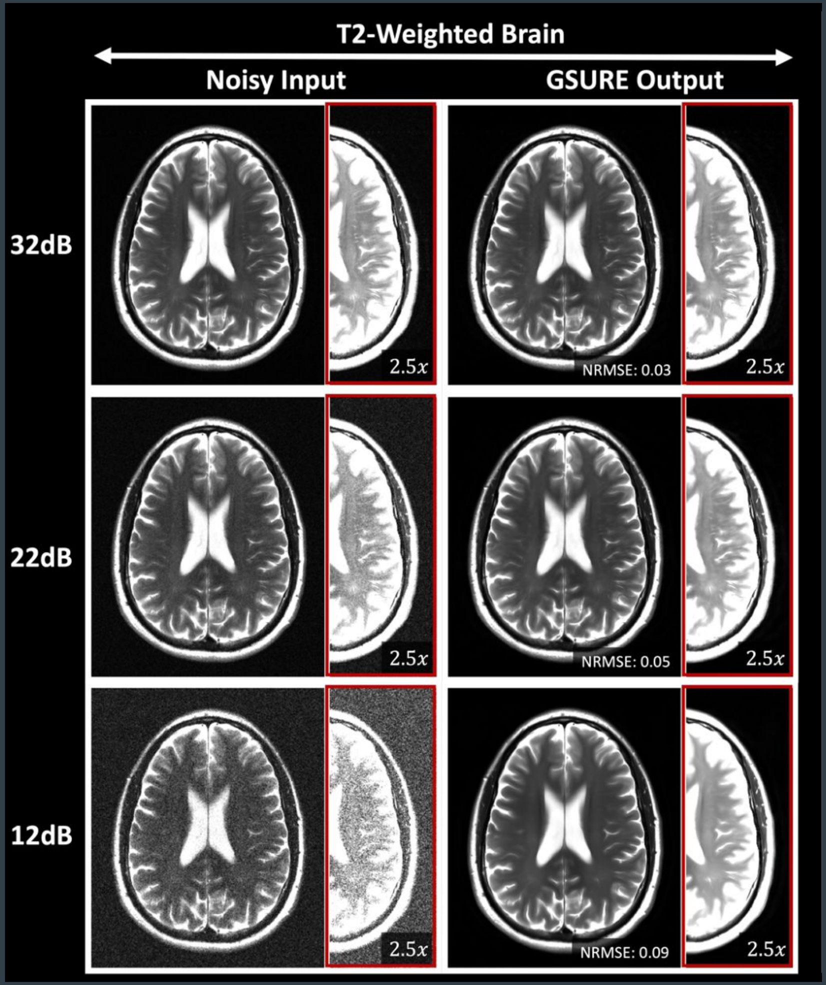
1. Evaluation of **Self-Supervised Denoising** (GSURE)

Experimental Details:

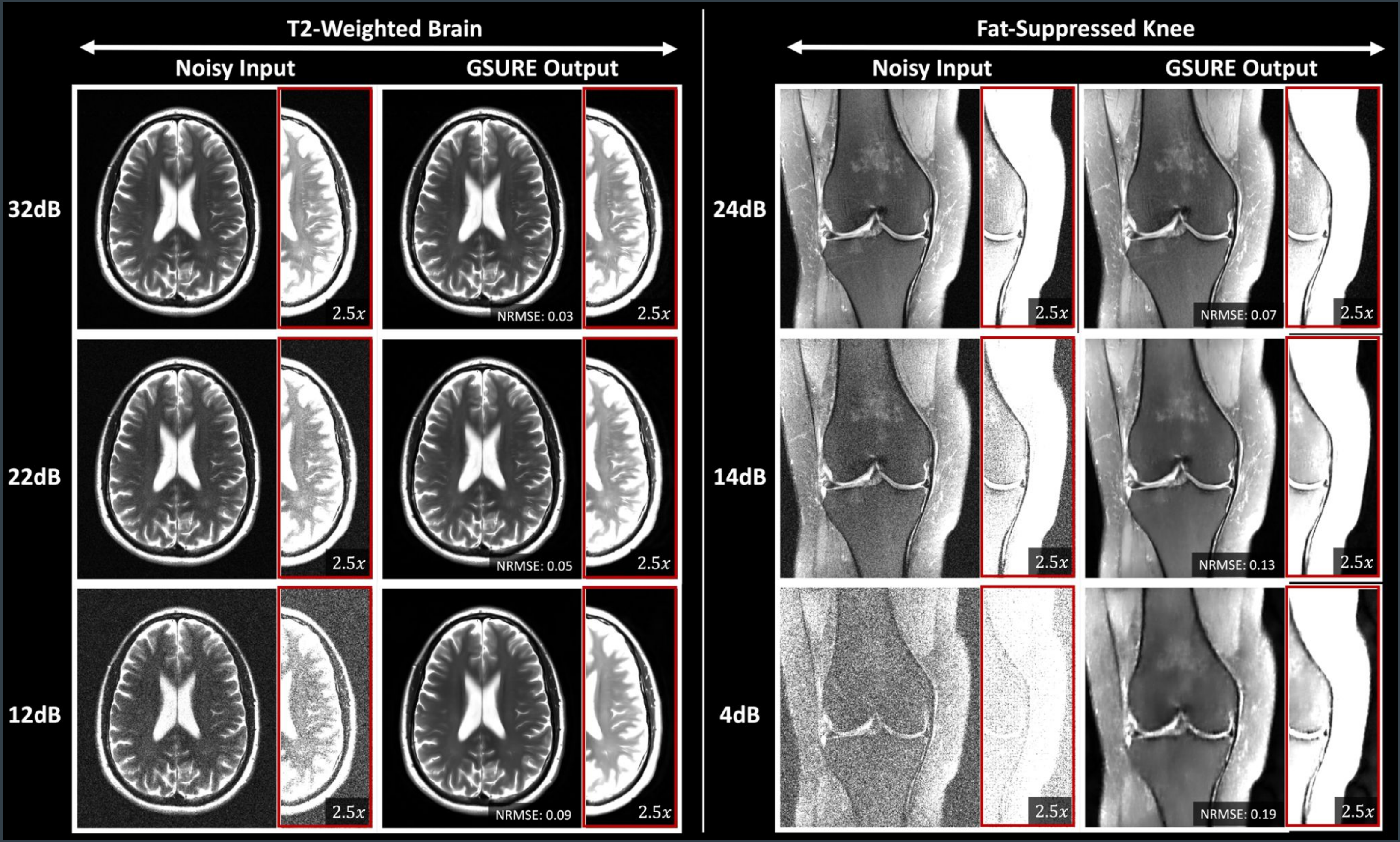
- Brain:
  - 10,000 2D T_2-weighted brain samples
- Knee:
  - 2,000 2D fat-suppressed knee

- Learned Denoiser Architecture: EDM (Karras *NeurIPS 2022*)

Denoising Performance



Denoising Performance



Results

1. Evaluation of **Self-Supervised Denoising** (GSURE)
2. **Prior sampling** performance of score models trained on Noisy vs. GSURE Denoised data
3. **Accelerated MRI Reconstruction** performance of score models trained on Noisy vs. GSURE Denoised data

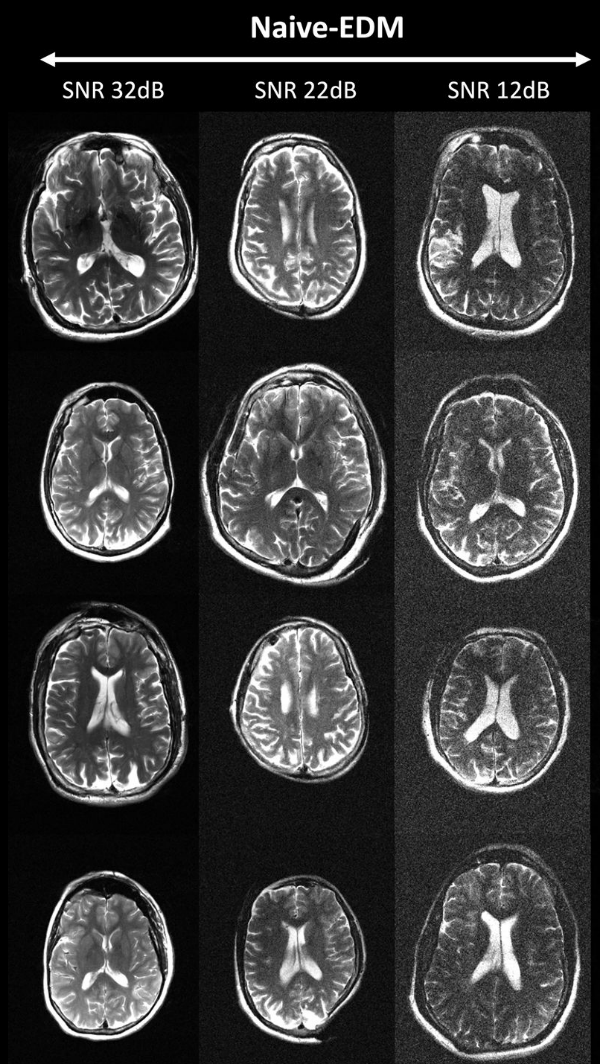
Results

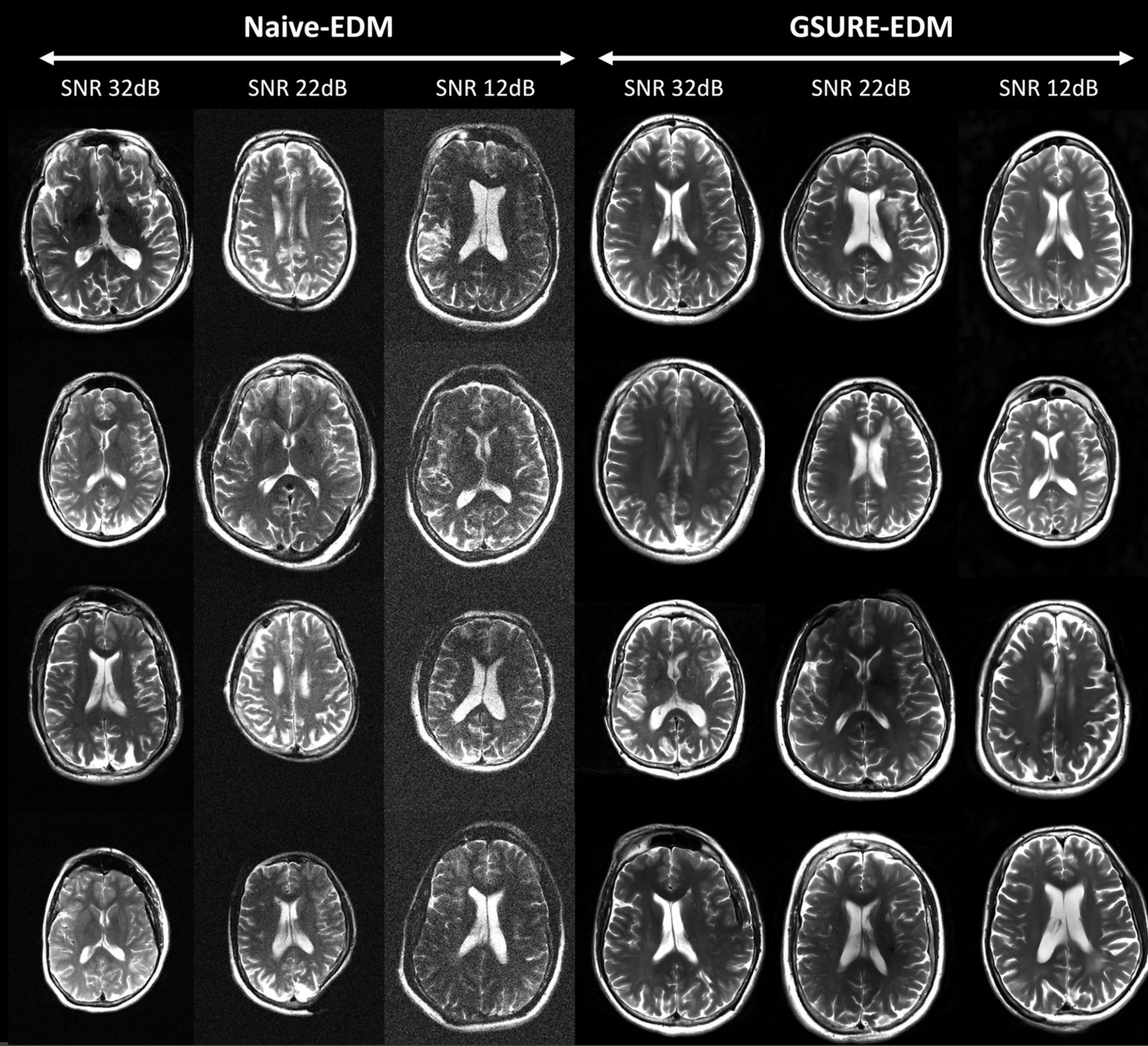
1. Evaluation of **Self-Supervised Denoising** (GSURE)
2. **Prior sampling** performance of score models trained on Noisy vs. GSURE Denoised data
3. **Posterior Reconstruction** performance of score models trained

Experimental Details:

- EDM Model Trained with on noisy and denoised versions of the 10,000 sample T_2 Brain dataset.

- EDM Architecture: EDM (Song *Karras 2022*)


Naive-EDM
SNR 32dB
SNR 22dB
SNR 12dB



Naive-EDM
GSURE-EDM
SNR 32dB
SNR 22dB
SNR 12dB
SNR 32dB
SNR 22dB
SNR 12dB

Results

1. Evaluation of **Self-Supervised Denoising** (GSURE)
2. **Prior sampling** performance of score models trained on Noisy vs. GSURE Denoised data
3. **Accelerated MRI Reconstruction** performance of score models trained on Noisy vs. GSURE Denoised data

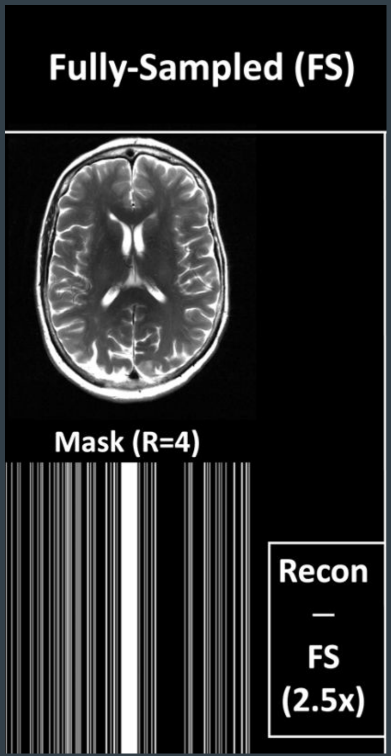
Results

1. Evaluation of **Self-Supervised Denoising** (GSURE)
2. **Prior sampling** performance of score models trained on Noisy vs. GSURE Denoised data
3. **Accelerated MRI Reconstruction** performance of score models trained on Noisy vs. GSURE Denoised data

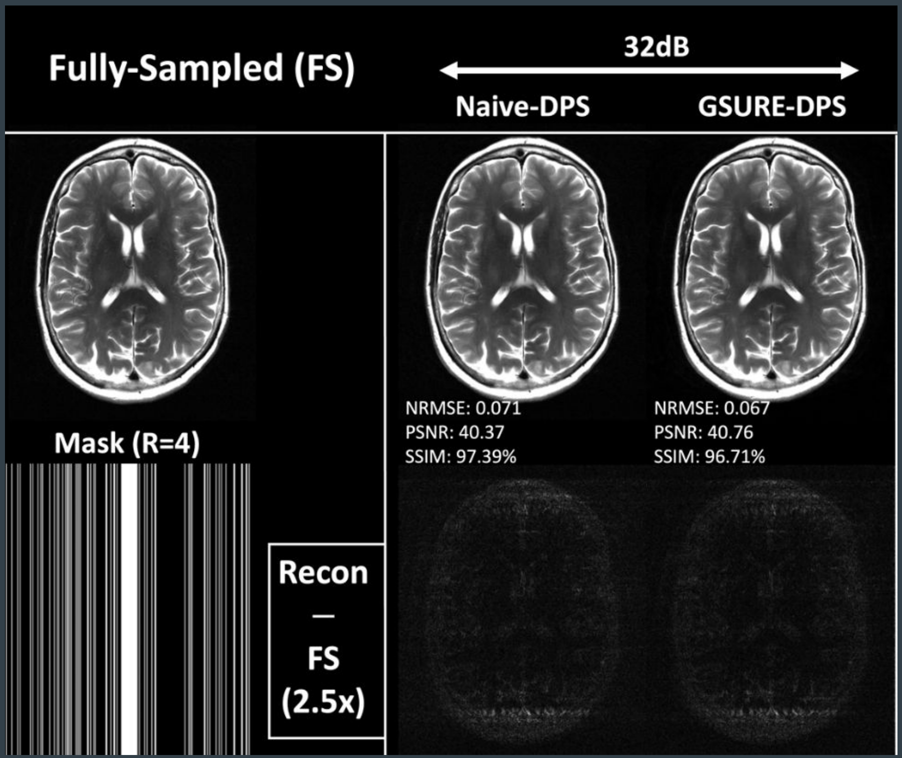
Experimental Details

- 100 retrospectively under-sampled 2D T_2 Brain validation samples

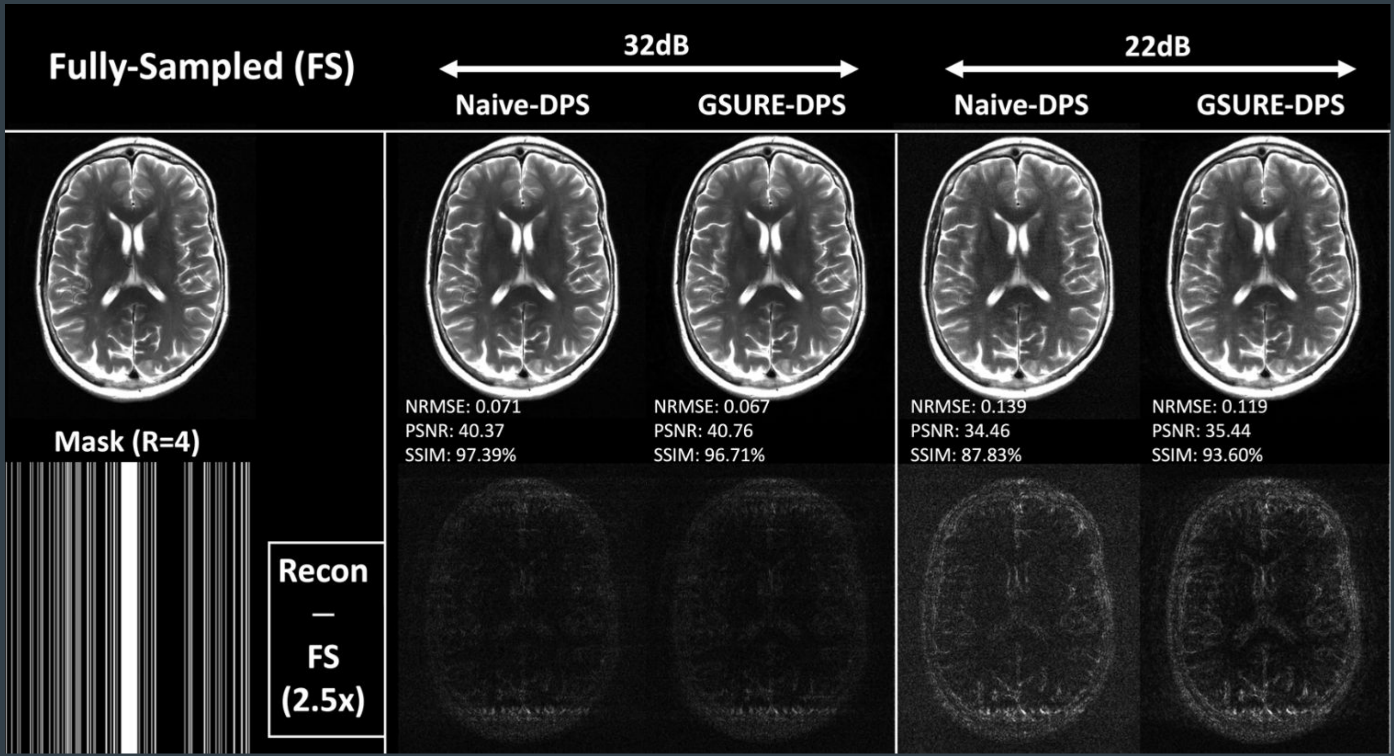
Posterior Sampling $x \sim p(x|y)$



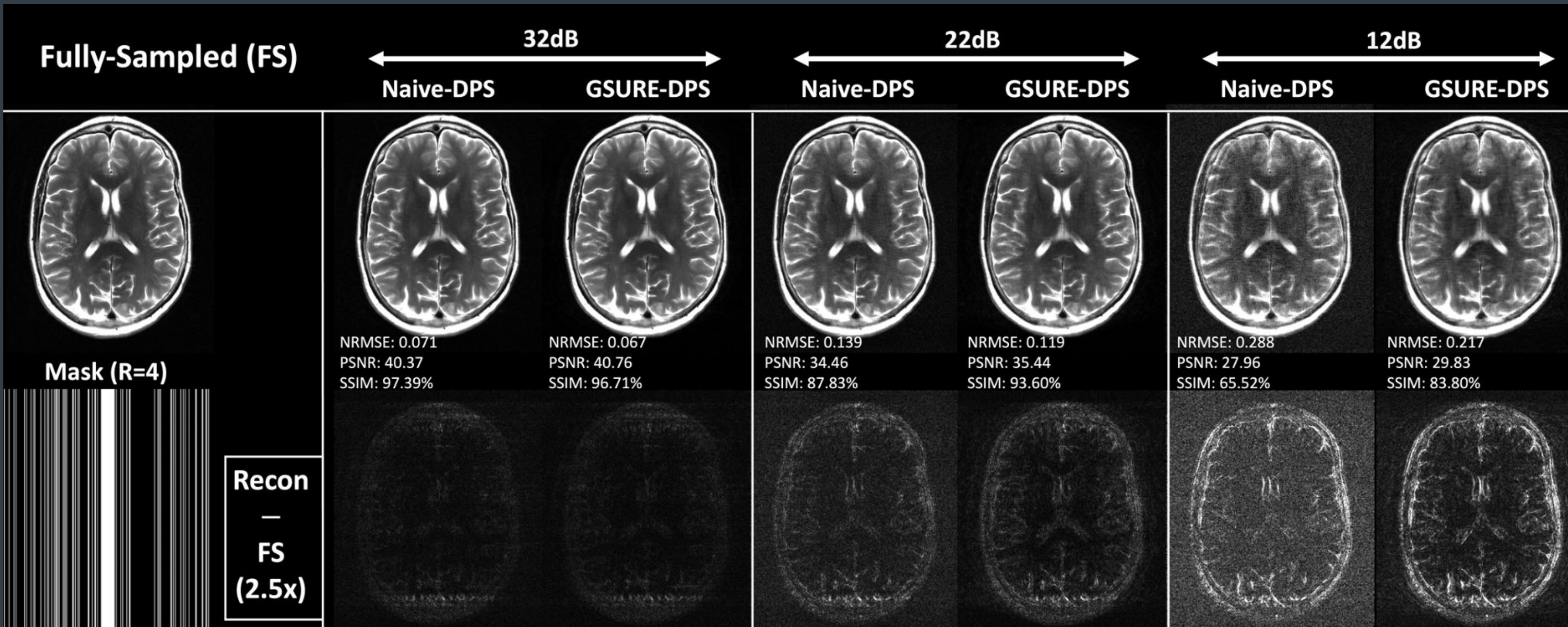
Posterior Sampling $x \sim p(x|y)$



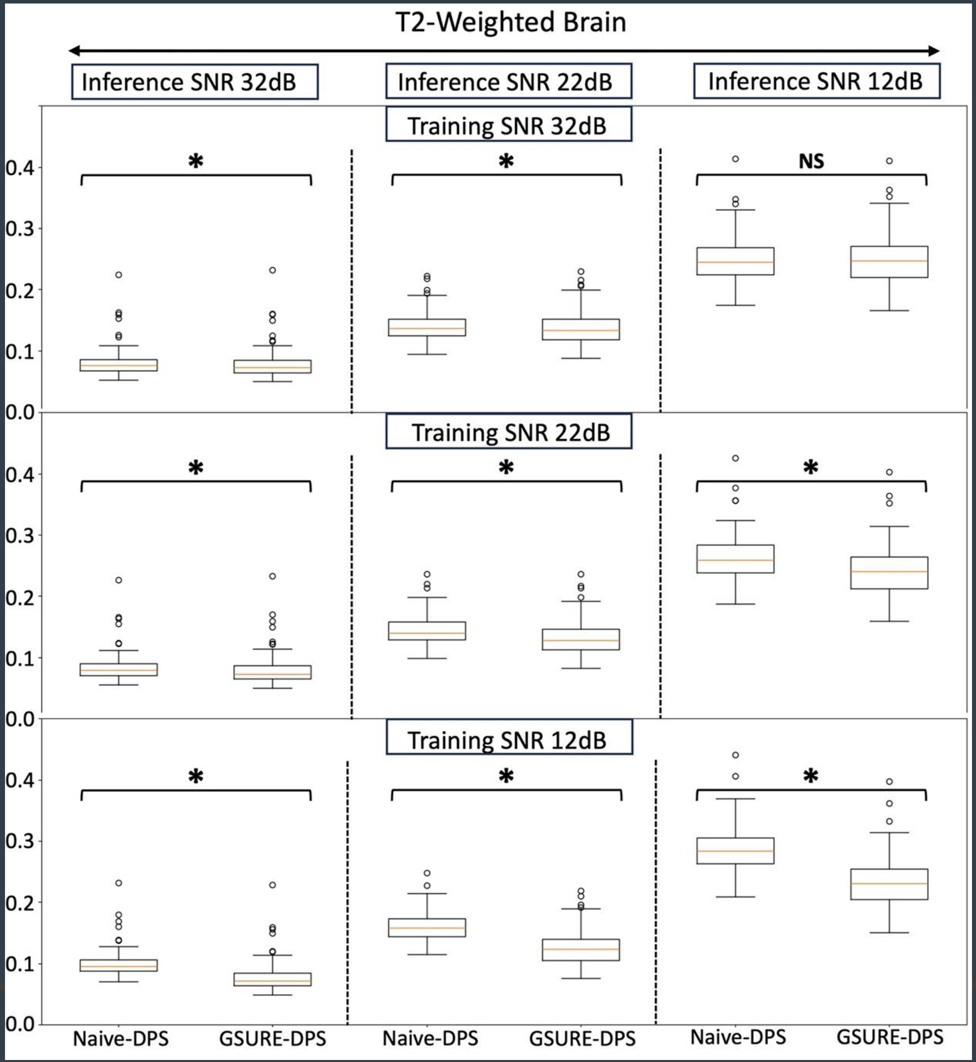
Posterior Sampling $x \sim p(x|y)$



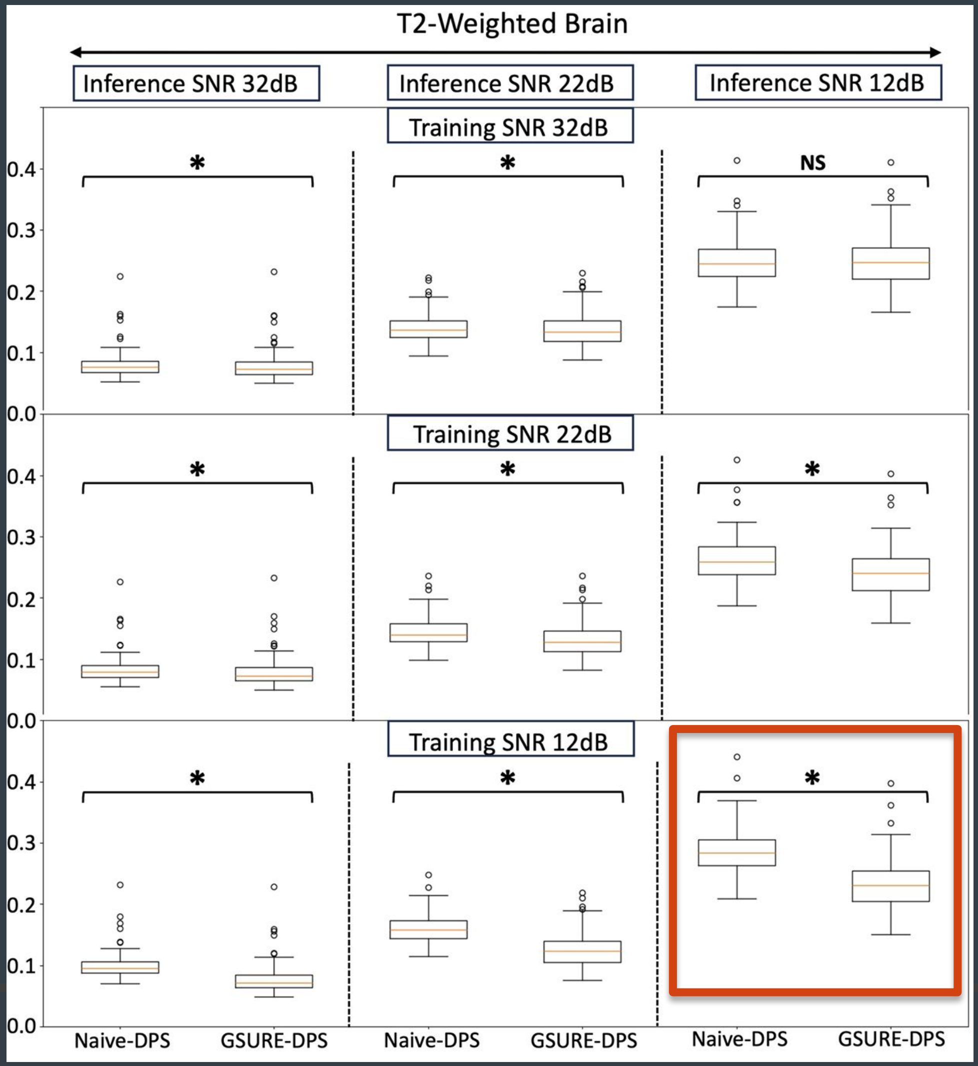
Posterior Sampling $x \sim p(x|y)$



Accelerated MRI Reconstruction with Posterior Sampling $x \sim p(x|y)$



Accelerated MRI Reconstruction with Posterior Sampling $x \sim p(x|y)$



Discussion and Conclusion

1. GSURE Denoising as a pre-processing step helps train more **accurate priors** which are better **inverse problem solvers** than naïve training.
2. The benefit of denoising is more visible in **lower SNR** settings
3. Important to investigate tradeoff between noise and distortion
4. Applicable to other learning settings (e.g. end-to-end methods)

Future Works

A is a Linear Forward Operator (**Fully-Sampled**) -> GSURE[1,2,3]

$$y = FSx + noise$$

[1]Soltanayev, *NeurIPS,* 2018, [2]Eldar, *IEEE Transactions on Signal Processing,* 2008, [3]Kawar, *TMLR,* 2023, [4]Aali, AmbientDPS, *Arxiv*, 2024

Future Works

A is a Linear Forward Operator (**Fully-Sampled**) -> GSURE[1,2,3]

$$y = FSx + noise$$

Assume A is a **Low-Rank** Forward Operator[4]

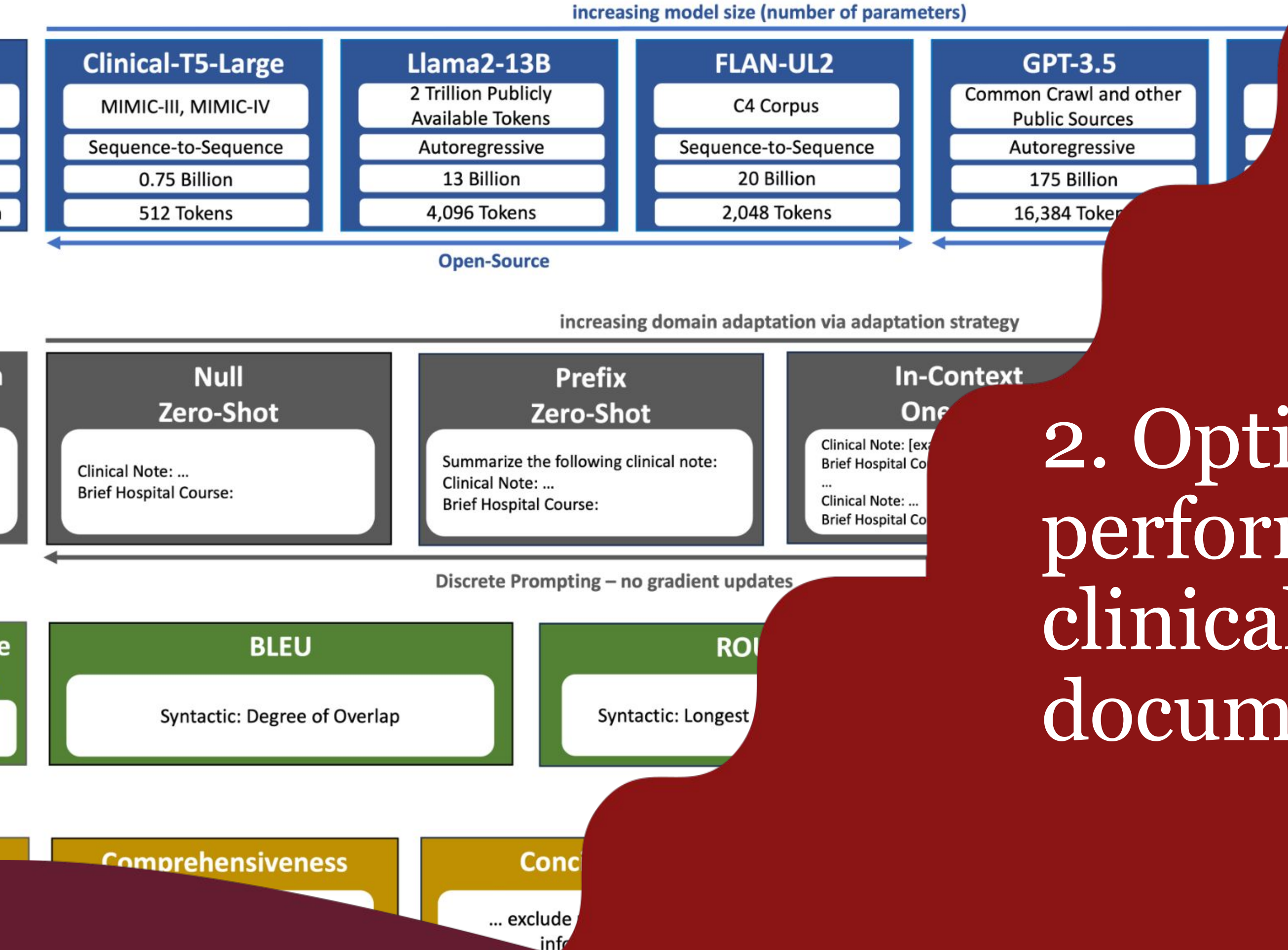
$$y = PFSx + noise$$

Ambient Diffusion Posterior Sampling: Solving Inverse Problems with Diffusion Models Trained on Corrupted Data

Asad Aali, Giannis Daras, Brett Levac, Sidharth Kumar, Alex Dimakis, Jon Tamir

Published: 22 Jan 2025, Last Modified: 11 Feb 2025 ICLR 2025 Poster Everyone Revisions BibTeX

Stanford University



2. Optimizing LLM performance in clinical documentation tasks

Motivation

1. Health Care providers at One Medical need to manually look through hundreds of clinical documents

2. Surfacing the most relevant clinical data can be accomplished with text summarization

3. This can allow for better **health outcomes** as it helps providers:
   a. Save valuable **time**
   b. Build a **deeper connection** with patients

Published in JAMIA

MIMIC-IV-BHC - Sample

Table 1. a) A sample of our novel pre-processed clinical notes dataset, extracted from raw MIMIC-IV notes.

Input	**Example**
SEX	F
SERVICE	SURGERY
ALLERGIES	No Known Allergies
CHIEF COMPLAINT	Splenic laceration
MAJOR PROCEDURE	NONE
HISTORY OF PRESENT ILLNESS	s/p routine colonoscopy this morning with polypectomy (report not available) ...
PAST MEDICAL HISTORY	Mild asthma, hypothyroid
FAMILY HISTORY	Non-contributory
PHYSICAL EXAM	Gen: Awake and alert CV: RRR Lungs: CTAB Abd: Soft, nontender, nondistended
PERTINENT RESULTS	03:45 PM BLOOD WBC-5.5 RBC-3.95 Hgb-14.1 ...
MEDICATIONS ON ADMISSION	1. Levothyroxine Sodium 100 mcg PO DAILY 2. Flovent HFA (fluticasone) ...
DISCHARGE DISPOSITION	Home
DISCHARGE DIAGNOSIS	Splenic laceration
DISCHARGE CONDITION	Mental Status: Clear and coherent. Level of Consciousness: Alert and interactive ...
DISCHARGE INSTRUCTIONS	You were admitted to ... in the intensive care unit for monitoring after a ...

MIMIC-IV-BHC - Sample

Table 1. a) A sample of our novel pre-processed clinical notes dataset, extracted from raw MIMIC-IV notes.

Input	Example
SEX	F
SERVICE	SURGERY
ALLERGIES	No Known Allergies
CHIEF COMPLAINT	Splenic laceration
MAJOR PROCEDURE	NONE
HISTORY OF PRESENT ILLNESS	s/p routine colonoscopy this morning with polypectomy (report not available) ...
PAST MEDICAL HISTORY	Mild asthma, hypothyroid
FAMILY HISTORY	Non-contributory
PHYSICAL EXAM	Gen: Awake and alert CV: RRR Lungs: CTAB Abd: Soft, nontender, nondistended
PERTINENT RESULTS	03:45 PM BLOOD WBC-5.5 RBC-3.95 Hgb-14.1 ...
MEDICATIONS ON ADMISSION	1. Levothyroxine Sodium 100 mcg PO DAILY 2. Flovent HFA (fluticasone) ...
DISCHARGE DISPOSITION	Home
DISCHARGE DIAGNOSIS	Splenic laceration
DISCHARGE CONDITION	Mental Status: Clear and coherent. Level of Consciousness: Alert and interactive ...
DISCHARGE INSTRUCTIONS	You were admitted to ... in the intensive care unit for monitoring after a ...

Output	Example
BRIEF HOSPITAL COURSE	Ms. ... was admitted to ... on .... After getting a colonoscopy and polypectomy, she ...

MIMIC-IV-Ext-BHC: Labeled Clinical Notes Dataset for Hospital Course Summarization

Asad Aali, Dave Van Veen, Yamin Arefeen, Jason Hom, Christian Bluethgen, Eduardo Pontes Reis, Sergios Gatidis, Namuun Clifford, Joseph Daws, Arash Tehrani, Jangwon Kim, Akshay Chaudhari

1. A curated collection of preprocessed and labeled clinical notes derived from the MIMIC-IV-Note database.

2. To facilitate development and training of machine learning models focused on summarizing brief hospital courses (BHC)

3. 270,033 meticulously cleaned and standardized clinical notes containing an average token length of 2,267

4. Preprocessing pipeline employed uses regular expressions to address common issues in the raw clinical text

[Published on PhysioNet]

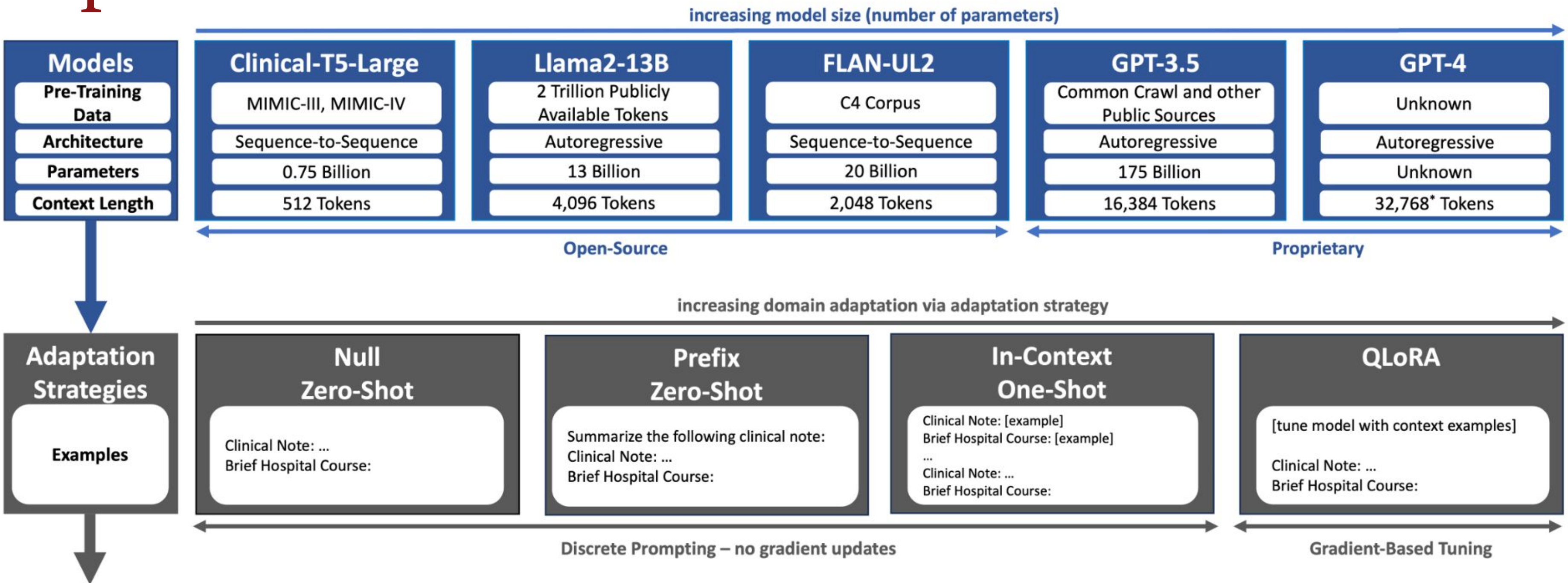
Pipeline

increasing model size (number of parameters)

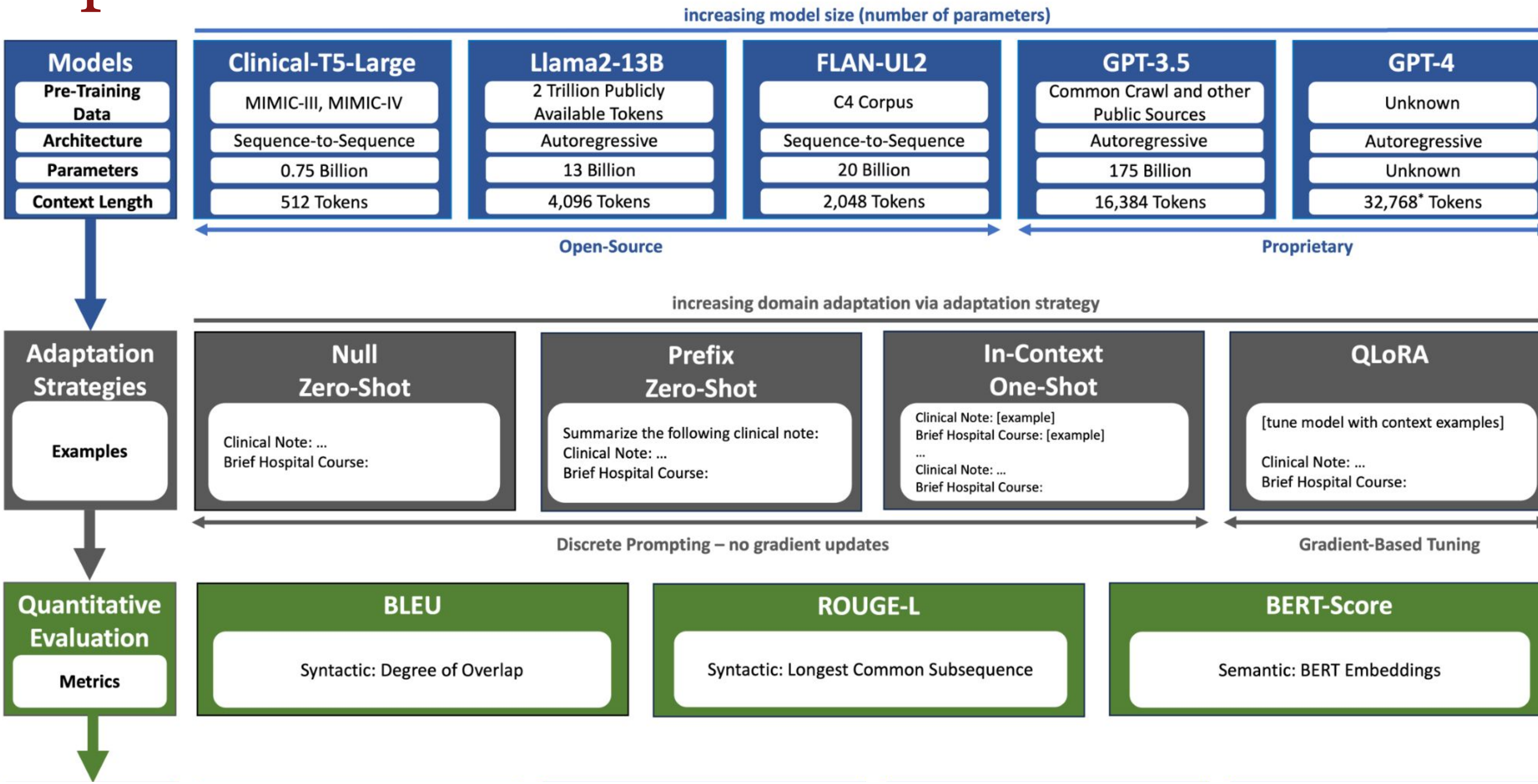
Models	Clinical-T5-Large	Llama2-13B	FLAN-UL2	GPT-3.5	GPT-4
Pre-Training Data	MIMIC-III, MIMIC-IV	2 Trillion Publicly Available Tokens	C4 Corpus	Common Crawl and other Public Sources	Unknown
Architecture	Sequence-to-Sequence	Autoregressive	Sequence-to-Sequence	Autoregressive	Autoregressive
Parameters	0.75 Billion	13 Billion	20 Billion	175 Billion	Unknown
Context Length	512 Tokens	4,096 Tokens	2,048 Tokens	16,384 Tokens	32,768* Tokens

Open-Source (Clinical-T5-Large, Llama2-13B, FLAN-UL2) | Proprietary (GPT-3.5, GPT-4)

Pipeline

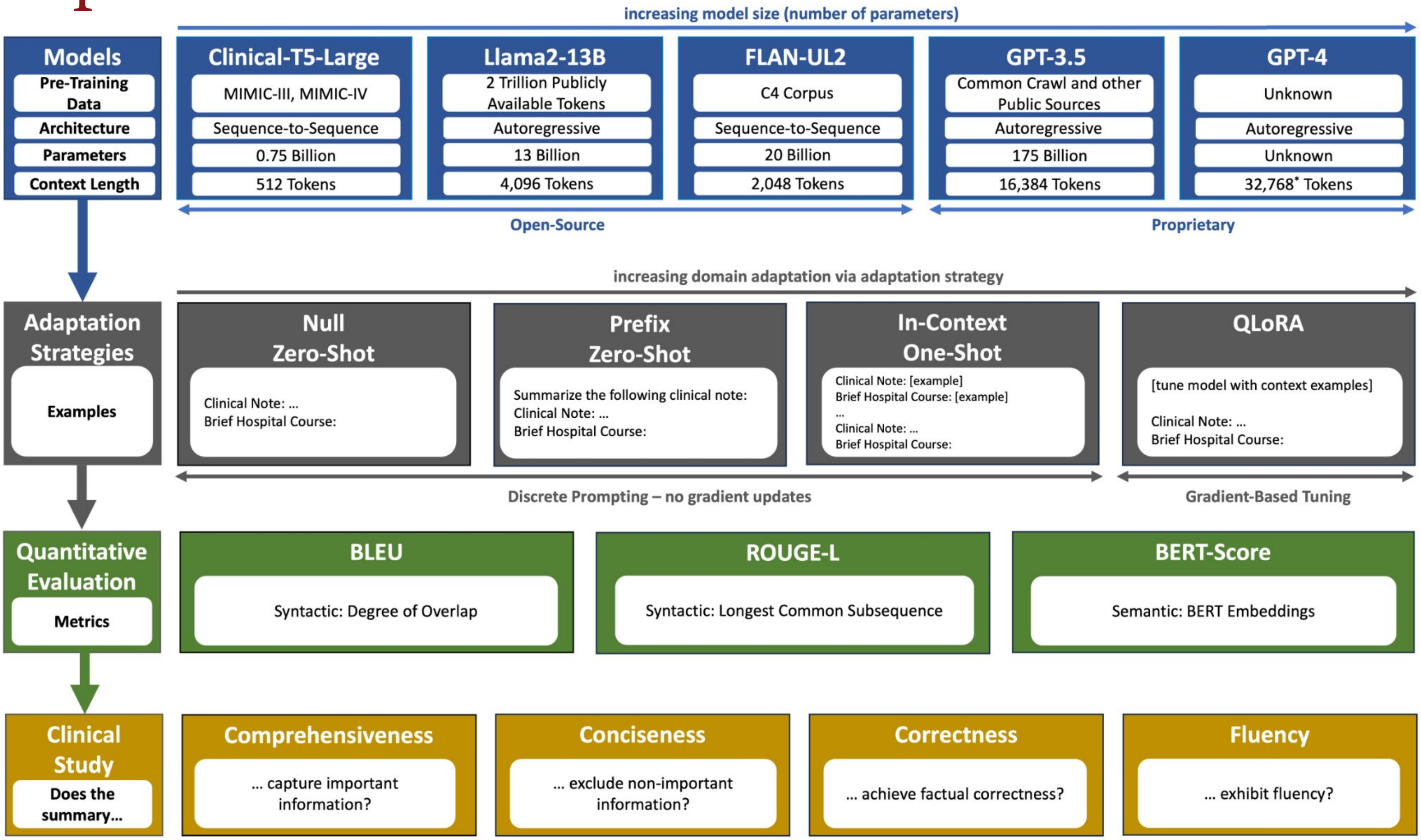

increasing model size (number of parameters)
Models
Pre-Training Data
Architecture
Parameters
Context Length
Clinical-T5-Large
MIMIC-III, MIMIC-IV
Sequence-to-Sequence
0.75 Billion
512 Tokens
Llama2-13B
2 Trillion Publicly Available Tokens
Autoregressive
13 Billion
4,096 Tokens
FLAN-UL2
C4 Corpus
Sequence-to-Sequence
20 Billion
2,048 Tokens
GPT-3.5
Common Crawl and other Public Sources
Autoregressive
175 Billion
16,384 Tokens
GPT-4
Unknown
Autoregressive
Unknown
32,768* Tokens
Open-Source
Proprietary
increasing domain adaptation via adaptation strategy
Adaptation Strategies
Examples
Null Zero-Shot
Clinical Note: ...
Brief Hospital Course:
Prefix Zero-Shot
Summarize the following clinical note:
Clinical Note: ...
Brief Hospital Course:
In-Context One-Shot
Clinical Note: [example]
Brief Hospital Course: [example]
...
Clinical Note: ...
Brief Hospital Course:
QLoRA
[tune model with context examples]
Clinical Note: ...
Brief Hospital Course:
Discrete Prompting – no gradient updates
Gradient-Based Tuning

Pipeline



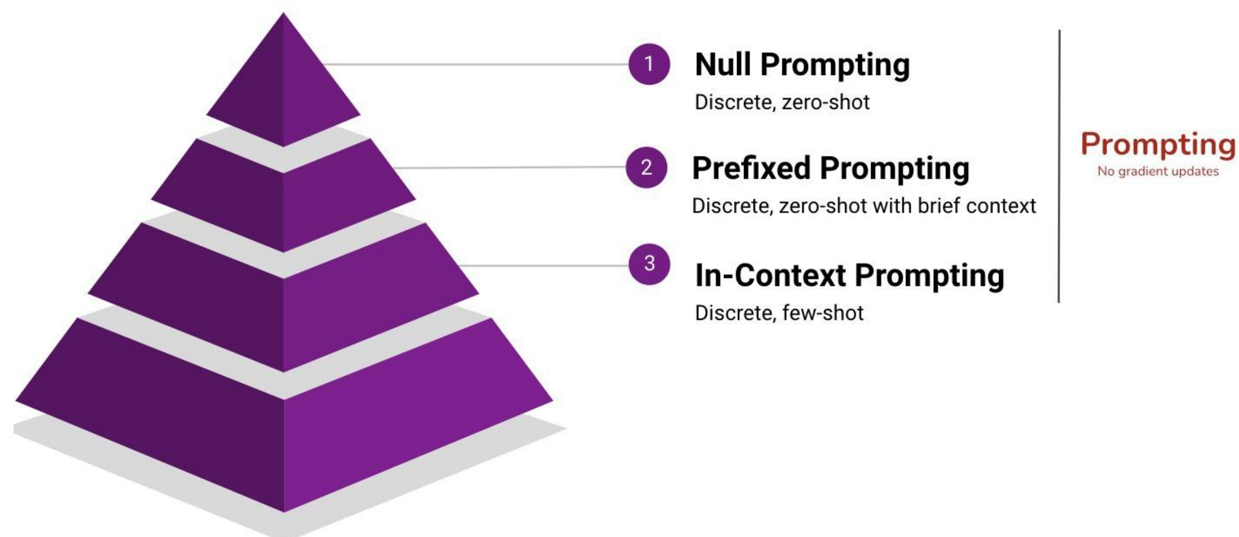
increasing model size (number of parameters)
Models
Pre-Training Data
Architecture
Parameters
Context Length
Clinical-T5-Large
MIMIC-III, MIMIC-IV
Sequence-to-Sequence
0.75 Billion
512 Tokens
Llama2-13B
2 Trillion Publicly Available Tokens
Autoregressive
13 Billion
4,096 Tokens
FLAN-UL2
C4 Corpus
Sequence-to-Sequence
20 Billion
2,048 Tokens
GPT-3.5
Common Crawl and other Public Sources
Autoregressive
175 Billion
16,384 Tokens
GPT-4
Unknown
Autoregressive
Unknown
32,768* Tokens
Open-Source
Proprietary
increasing domain adaptation via adaptation strategy
Adaptation Strategies
Examples
Null Zero-Shot
Clinical Note: ...
Brief Hospital Course:
Prefix Zero-Shot
Summarize the following clinical note:
Clinical Note: ...
Brief Hospital Course:
In-Context One-Shot
Clinical Note: [example]
Brief Hospital Course: [example]
...
Clinical Note: ...
Brief Hospital Course:
QLoRA
[tune model with context examples]
Clinical Note: ...
Brief Hospital Course:
Discrete Prompting – no gradient updates
Gradient-Based Tuning
Quantitative Evaluation
Metrics
BLEU
Syntactic: Degree of Overlap
ROUGE-L
Syntactic: Longest Common Subsequence
BERT-Score
Semantic: BERT Embeddings

Pipeline

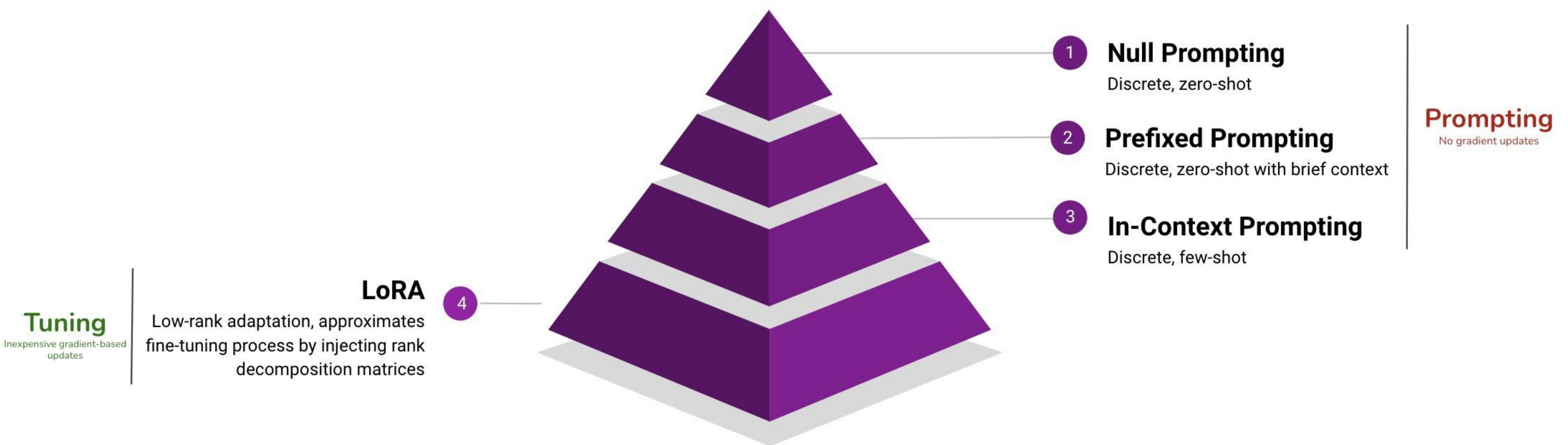


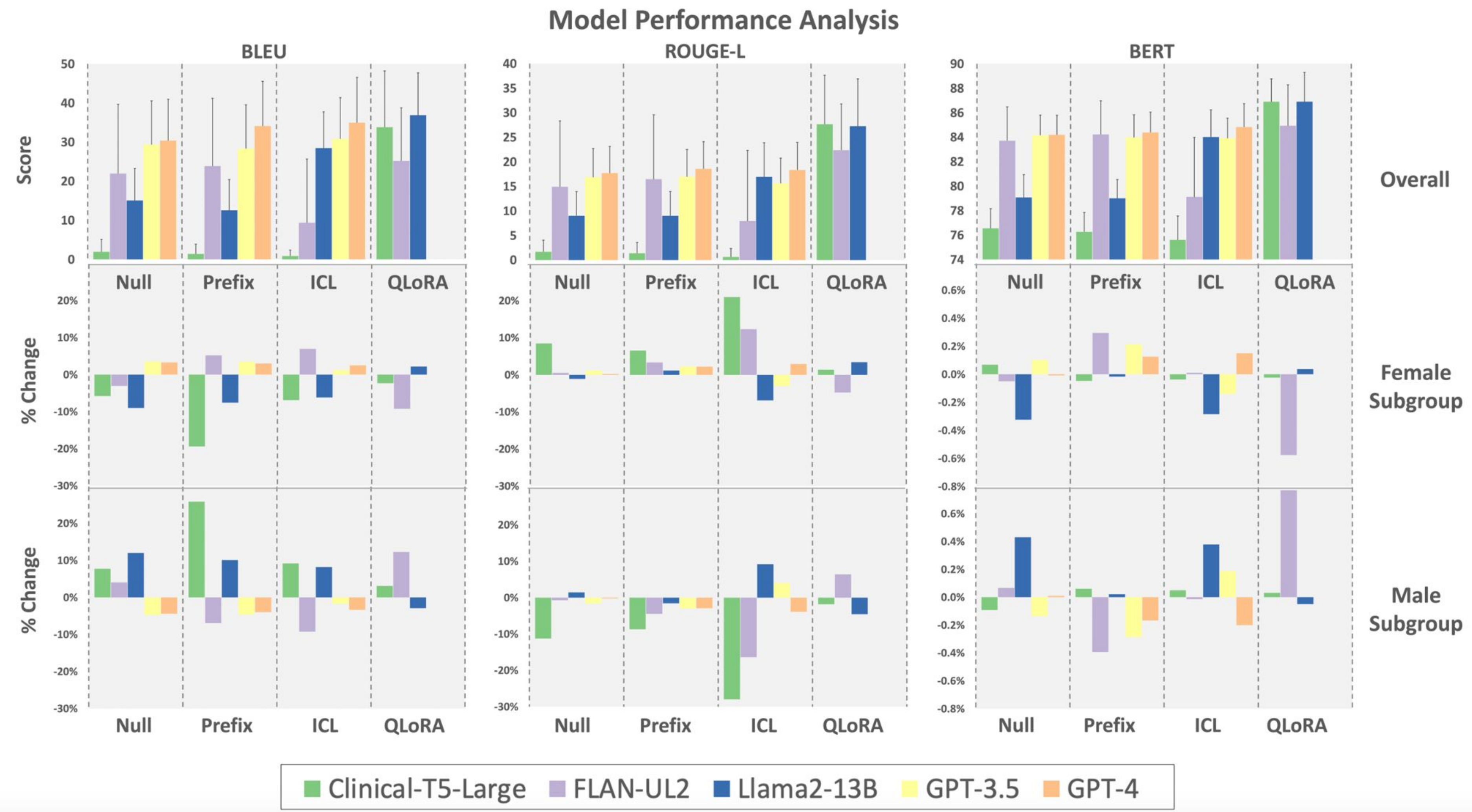
increasing model size (number of parameters)
Models
Pre-Training Data
Architecture
Parameters
Context Length
Clinical-T5-Large
MIMIC-III, MIMIC-IV
Sequence-to-Sequence
0.75 Billion
512 Tokens
Llama2-13B
2 Trillion Publicly Available Tokens
Autoregressive
13 Billion
4,096 Tokens
FLAN-UL2
C4 Corpus
Sequence-to-Sequence
20 Billion
2,048 Tokens
GPT-3.5
Common Crawl and other Public Sources
Autoregressive
175 Billion
16,384 Tokens
GPT-4
Unknown
Autoregressive
Unknown
32,768* Tokens
Open-Source
Proprietary
increasing domain adaptation via adaptation strategy
Adaptation Strategies
Examples
Null Zero-Shot
Clinical Note: ...
Brief Hospital Course:
Prefix Zero-Shot
Summarize the following clinical note:
Clinical Note: ...
Brief Hospital Course:
In-Context One-Shot
Clinical Note: [example]
Brief Hospital Course: [example]
...
Clinical Note: ...
Brief Hospital Course:
QLoRA
[tune model with context examples]
Clinical Note: ...
Brief Hospital Course:
Discrete Prompting – no gradient updates
Gradient-Based Tuning
Quantitative Evaluation
Metrics
BLEU
Syntactic: Degree of Overlap
ROUGE-L
Syntactic: Longest Common Subsequence
BERT-Score
Semantic: BERT Embeddings
Clinical Study
Does the summary...
Comprehensiveness
... capture important information?
Conciseness
... exclude non-important information?
Correctness
... achieve factual correctness?
Fluency
... exhibit fluency?

Overview of Adaptation Methods

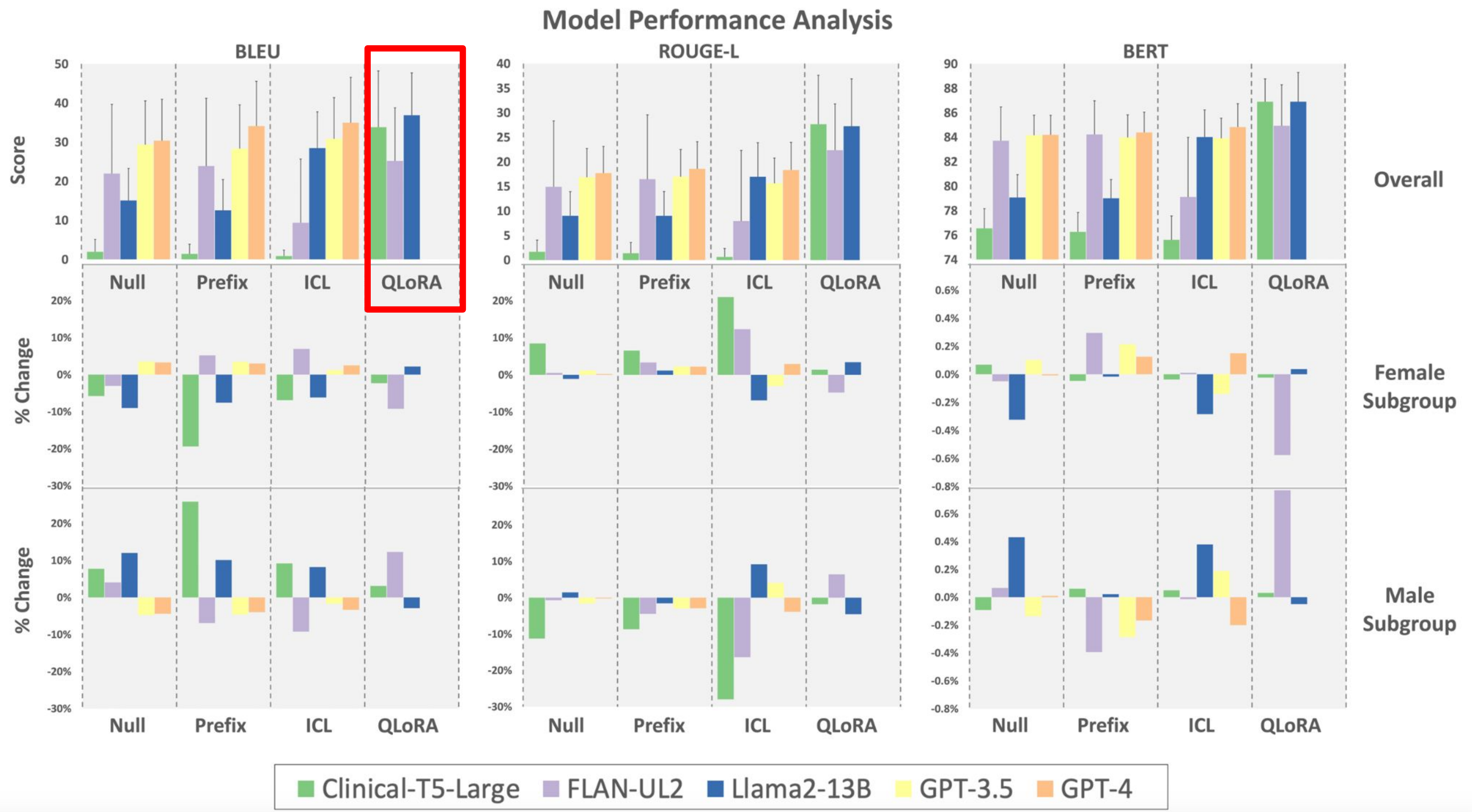


Overview of Adaptation Methods



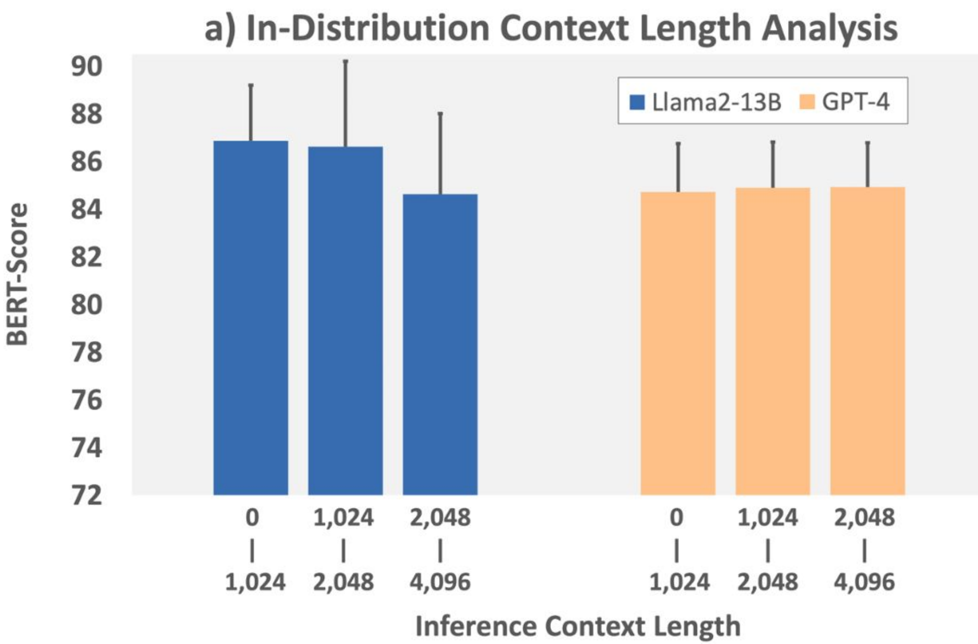


Model Performance Analysis
BLEU
ROUGE-L
BERT
Score
% Change
% Change
Overall
Female Subgroup
Male Subgroup
Null
Prefix
ICL
QLoRA
Clinical-T5-Large
FLAN-UL2
Llama2-13B
GPT-3.5
GPT-4

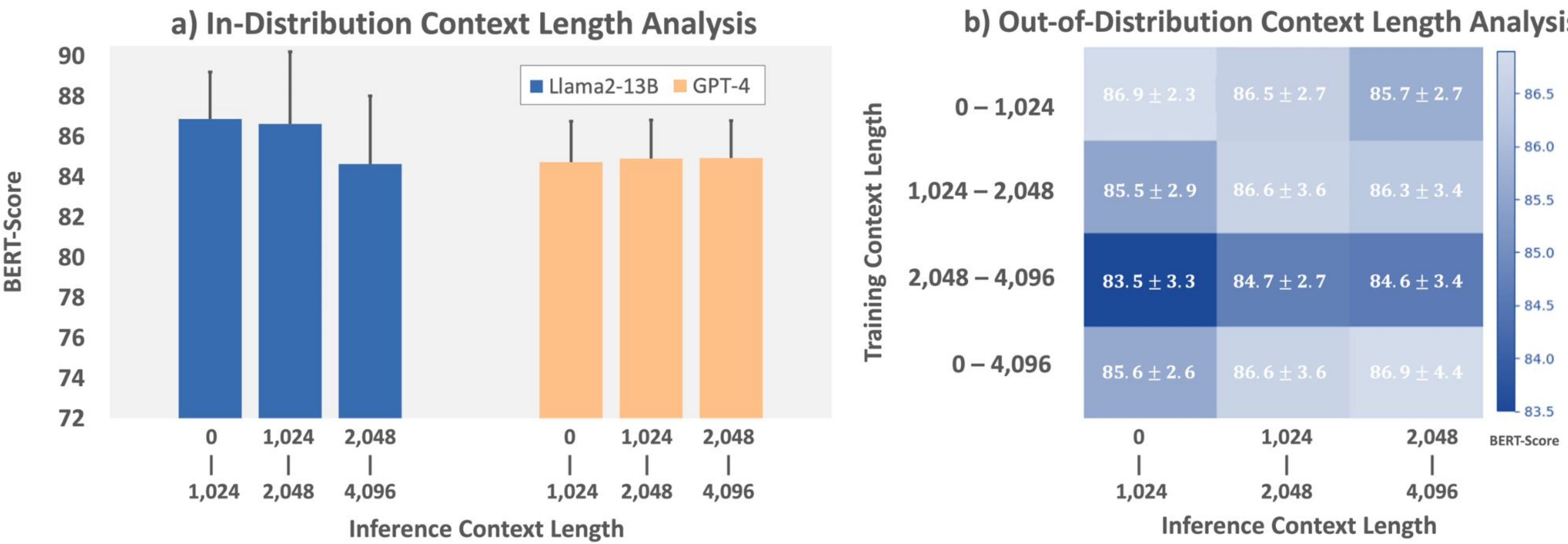


Model Performance Analysis
BLEU
ROUGE-L
BERT
Score
% Change
% Change
Overall
Female Subgroup
Male Subgroup
Null
Prefix
ICL
QLoRA
Clinical-T5-Large
FLAN-UL2
Llama2-13B
GPT-3.5
GPT-4

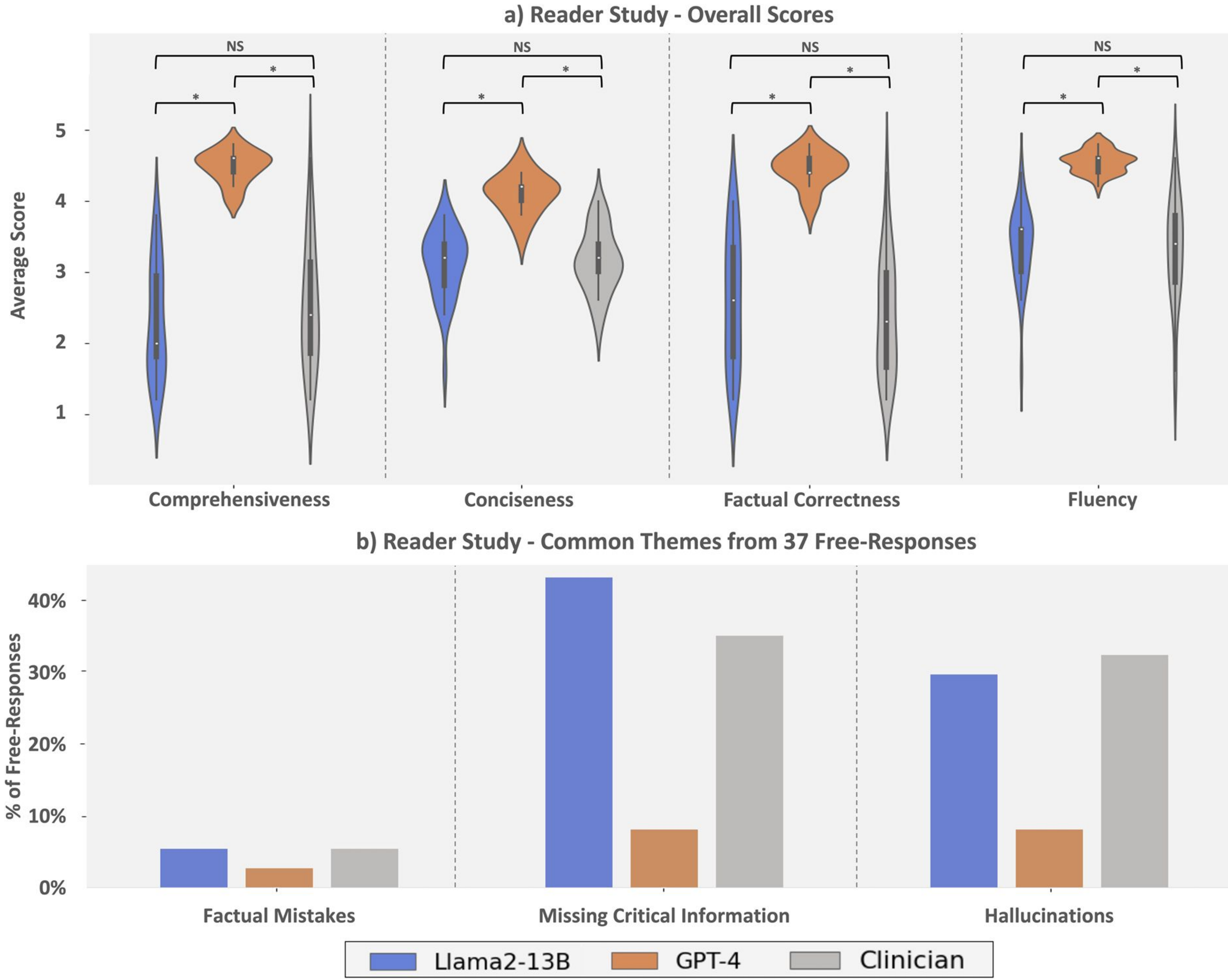
Context Length Analysis



Context Length Analysis



a) Reader Study - Overall Scores
NS
*
*
NS
*
*
NS
*
*
NS
*
*
Average Score
5
4
3
2
1
Comprehensiveness
Conciseness
Factual Correctness
Fluency



a) Reader Study - Overall Scores
NS
*
*
Average Score
5
4
3
2
1
Comprehensiveness
Conciseness
Factual Correctness
Fluency
b) Reader Study - Common Themes from 37 Free-Responses
% of Free-Responses
40%
30%
20%
10%
0%
Factual Mistakes
Missing Critical Information
Hallucinations
Llama2-13B
GPT-4
Clinician

Summarization Example

Expertise	*You are an expert medical professional*
Instruction	*Summarize the clinical note into a brief hospital course*

Summarization Example

Expertise	*You are an expert medical professional*
Instruction	*Summarize the clinical note into a brief hospital course*
In-Context Example	*Use the examples to guide word choice* *input: {example clinical note}* *summary: {example bhc}*

Summarization Example

Expertise

You are an expert medical professional

Instruction

Summarize the clinical note into a brief hospital course

In-Context Example

Use the examples to guide word choice
input: {example clinical note}
summary: {example bhc}

Clinical Note Input

SEX: F

SERVICE: OBSTETRICS/GYNECOLOGY

ALLERGIES: No Known Allergies / Adverse Drug Reactions

ATTENDING: ___.

CHIEF COMPLAINT: bleeding in pregnancy

MAJOR SURGICAL OR INVASIVE PROCEDURE: None

HISTORY OF PRESENT ILLNESS: ___ G4P0 (h/o) TAB x 3 @ ___ admitted with vaginal bleeding that started 4 days prior.

PAST MEDICAL HISTORY: abnormal pap smears anxiety depression warts colposcopy, LEEP ___ TAB x 3 marginal cord insert fibroadenoma of the breast

SOCIAL HISTORY: ___

FAMILY HISTORY: noncontributory

PHYSICAL EXAM: VS: 98.3, 109/69, 75, 20, O2 97% Gen: NAD Resp: No evidence of respiratory distress Abd: Soft, non-tender Ext: No lower extremity edema Date: ___ Time: 09: 00 FHT: 130s/mod variability/+accels/-decels (? quick deep variable x 1, assoc w/ loss of pickup); otherwise reactive Toco: rare ctx

PERTINENT RESULTS: ___ 05: 10PM WBC-9.3 RBC-4.24 HGB-13.7 HCT-39.3 MCV-93 MCH-32.3* MCHC-34.9 RDW-13.0 RDWSD-43.7 ___ 05: 10PM PLT COUNT-229 ___ 05: 10PM ___ PTT-28.6 ___ ___ 05: 10PM ___

MEDICATIONS ON ADMISSION: PNV

DISCHARGE MEDICATIONS: 1. Citalopram 20 mg PO QHS 2. Prenatal Vitamins 1 TAB PO DAILY

DISCHARGE DISPOSITION: Home

DISCHARGE DIAGNOSIS: Marginal cord insertion Vaginal bleeding in pregnancy

DISCHARGE CONDITION: Mental Status: Clear and coherent. Level of Consciousness: Alert and interactive. Activity Status: Ambulatory - Independent.

FOLLOWUP INSTRUCTIONS: ___

DISCHARGE INSTRUCTIONS: Please continue pelvic rest. Avoid heavy lifting or strenuous activity. Otherwise normal activity.

Summarization Example

Expertise

You are an expert medical professional

Instruction

Summarize the clinical note into a brief hospital course

In-Context Example

Use the examples to guide word choice
input: {example clinical note}
summary: {example bhc}

Clinical Note Input

SEX: *F*

SERVICE: *OBSTETRICS/GYNECOLOGY*

ALLERGIES: *No Known Allergies / Adverse Drug Reactions*

ATTENDING: *___.*

CHIEF COMPLAINT: *bleeding in pregnancy*

MAJOR SURGICAL OR INVASIVE PROCEDURE: *None*

HISTORY OF PRESENT ILLNESS: *___ G4P0 (h/o) TAB x 3 @ ___ admitted with vaginal bleeding that started 4 days prior.*

PAST MEDICAL HISTORY: *abnormal pap smears anxiety depression warts colposcopy, LEEP ___ TAB x 3 marginal cord insert fibroadenoma of the breast*

SOCIAL HISTORY: *___*

FAMILY HISTORY: *noncontributory*

PHYSICAL EXAM: *VS: 98.3, 109/69, 75, 20, O2 97% Gen: NAD Resp: No evidence of respiratory distress Abd: Soft, non-tender Ext: No lower extremity edema Date: ___ Time: 09: 00 FHT: 130s/mod variability/+accels/-decels (? quick deep variable x 1, assoc w/ loss of pickup); otherwise reactive Toco: rare ctx*

PERTINENT RESULTS: *___ 05: 10PM WBC-9.3 RBC-4.24 HGB-13.7 HCT-39.3 MCV-93 MCH-32.3* MCHC-34.9 RDW-13.0 RDWSD-43.7 ___ 05: 10PM PLT COUNT-229 ___ 05: 10PM ___ PTT-28.6 ___ ___ 05: 10PM ___*

MEDICATIONS ON ADMISSION: *PNV*

DISCHARGE MEDICATIONS: *1. Citalopram 20 mg PO QHS 2. Prenatal Vitamins 1 TAB PO DAILY*

DISCHARGE DISPOSITION: *Home*

DISCHARGE DIAGNOSIS: *Marginal cord insertion Vaginal bleeding in pregnancy*

DISCHARGE CONDITION: *Mental Status: Clear and coherent. Level of Consciousness: Alert and interactive. Activity Status: Ambulatory - Independent.*

FOLLOWUP INSTRUCTIONS: *___*

DISCHARGE INSTRUCTIONS: *Please continue pelvic rest. Avoid heavy lifting or strenuous activity. Otherwise normal activity.*

Summary 1 Clinician

Pt was Rh positive and admitted for vaginal bleeding. She had negative cultures (GBS and yeast, GC/CT/trich). She had spotting from admission until ___. She was monitored until ___. She was discharged home in a stable continue without any evidence of vaginal bleeding or spotting.

Summary 2 GPT-4

Ms. ___ was admitted under the care of Dr. ___ due to vaginal bleeding during her pregnancy. She has a history of abnormal pap smears, anxiety, depression, and fibroadenoma of the breast. No surgical procedures were performed during her stay. Her condition was monitored and managed with prenatal vitamins and Citalopram for her anxiety. She was discharged in a clear and coherent mental state, able to move independently. She was instructed to continue pelvic rest and avoid heavy lifting or strenuous activity. She will continue her care at home.

Summarization Example

Expertise

You are an expert medical professional

Instruction

Summarize the clinical note into a brief hospital course

In-Context Example

Use the examples to guide word choice
input: {example clinical note}
summary: {example bhc}

Clinical Note Input

SEX: *F*

SERVICE: *OBSTETRICS/GYNECOLOGY*

ALLERGIES: *No Known Allergies / Adverse Drug Reactions*

ATTENDING: *___.*

CHIEF COMPLAINT: *bleeding in pregnancy*

MAJOR SURGICAL OR INVASIVE PROCEDURE: *None*

HISTORY OF PRESENT ILLNESS: *___ G4P0 (h/o) TAB x 3 @ ___ admitted with vaginal bleeding that started 4 days prior.*

PAST MEDICAL HISTORY: *abnormal pap smears anxiety depression warts colposcopy, LEEP ___ TAB x 3 marginal cord insert fibroadenoma of the breast*

SOCIAL HISTORY: *___*

FAMILY HISTORY: *noncontributory*

PHYSICAL EXAM: *VS: 98.3, 109/69, 75, 20, O2 97% Gen: NAD Resp: No evidence of respiratory distress Abd: Soft, non-tender Ext: No lower extremity edema Date: ___ Time: 09: 00 FHT: 130s/mod variability/+accels/-decels (? quick deep variable x 1, assoc w/ loss of pickup); otherwise reactive Toco: rare ctx*

PERTINENT RESULTS: *___ 05: 10PM WBC-9.3 RBC-4.24 HGB-13.7 HCT-39.3 MCV-93 MCH-32.3* MCHC-34.9 RDW-13.0 RDWSD-43.7 ___ 05: 10PM PLT COUNT-229 ___ 05: 10PM ___ PTT-28.6 ___ ___ 05: 10PM ___*

MEDICATIONS ON ADMISSION: *PNV*

DISCHARGE MEDICATIONS: *1. Citalopram 20 mg PO QHS 2. Prenatal Vitamins 1 TAB PO DAILY*

DISCHARGE DISPOSITION: *Home*

DISCHARGE DIAGNOSIS: *Marginal cord insertion Vaginal bleeding in pregnancy*

DISCHARGE CONDITION: *Mental Status: Clear and coherent. Level of Consciousness: Alert and interactive. Activity Status: Ambulatory - Independent.*

FOLLOWUP INSTRUCTIONS: *___*

DISCHARGE INSTRUCTIONS: *Please continue pelvic rest. Avoid heavy lifting or strenuous activity. Otherwise normal activity.*

Summary 1 Clinician

Pt was Rh positive and admitted for vaginal bleeding. She had negative cultures (GBS and yeast, GC/CT/trich). She had spotting from admission until ___. She was monitored until ___. She was discharged home in a stable continue without any evidence of vaginal bleeding or spotting.

Summary 2 GPT-4

Ms. ___ was admitted under the care of Dr. ___ due to vaginal bleeding during her pregnancy. She has a history of abnormal pap smears, anxiety, depression, and fibroadenoma of the breast. No surgical procedures were performed during her stay. Her condition was monitored and managed with prenatal vitamins and Citalopram for her anxiety. She was discharged in a clear and coherent mental state, able to move independently. She was instructed to continue pelvic rest and avoid heavy lifting or strenuous activity. She will continue her care at home.

Reader Feedback

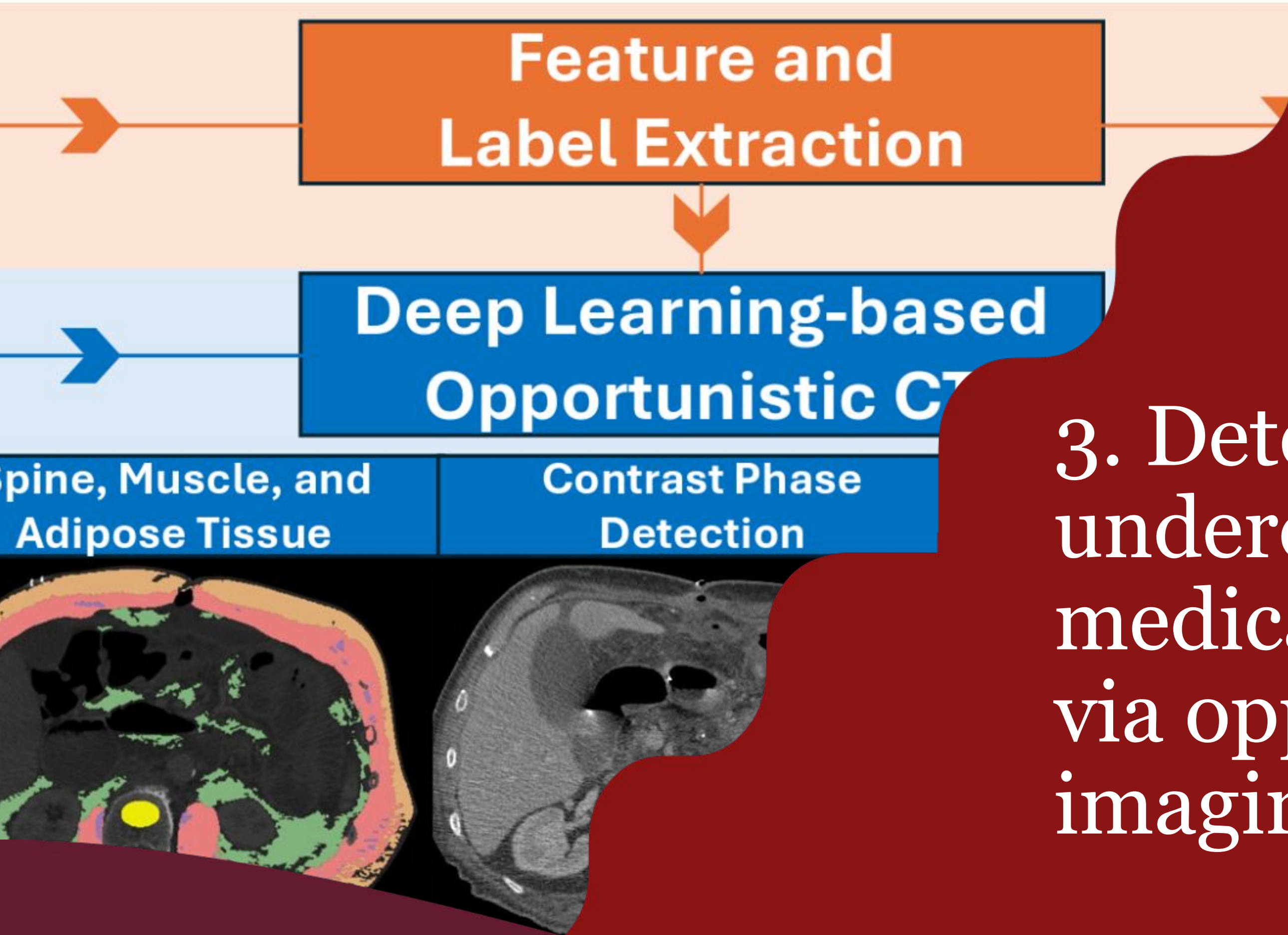
Summary A (Clinician) contains multiple factual mistakes (serial ultrasounds, no evidence of vaginal bleeding, closed cervix, negative culture). It seems to contain information not at all present in the actual clinical note. Summary B (GPT-4) failed to mention a summary of the patients labs or vital signs, but otherwise looks great.

Conclusions

1. Adapted **open-source models can match** the quality of clinician-written summaries

2. Adapted **proprietary models can outperform** the quality of clinician-written summaries

3. Adapted LLMs for summarization have the potential to:
   a. streamline documentation
   b. reduce errors
   c. enhance clinical workflows
   d. improve patient safety



3. Detecting underdiagnosed medical conditions via opportunistic imaging

Motivation

1. **Abdominal computed tomography** (CT) scans are frequently performed in clinical settings.

2. Opportunistic CT involves **repurposing routine CT** images to extract diagnostic information

3. This study utilizes deep learning methods to promote **accurate diagnosis and clinical documentation.**

4. We analyze **2,674 inpatient CT scans** to identify discrepancies between **imaging phenotypes** and corresponding documentation in **radiology reports** and **ICD coding**.

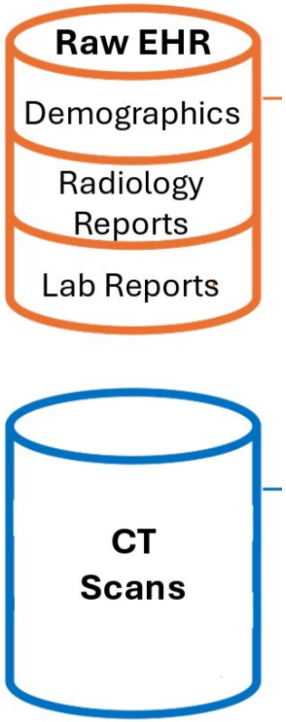
Detecting Underdiagnosed Medical Conditions with Deep Learning-Based Opportunistic CT Imaging

Asad Aali, MS[1], Andrew Johnston, MD, MBA[1], Louis Blankemeier, MS[1], Dave Van Veen, PhD[1], Laura T Derry, MD, MBA[1], David Svec, MD, MBA[1], Jason Hom, MD[1,*], Robert D. Boutin, MD[1,*], Akshay S. Chaudhari, PhD[1,*]

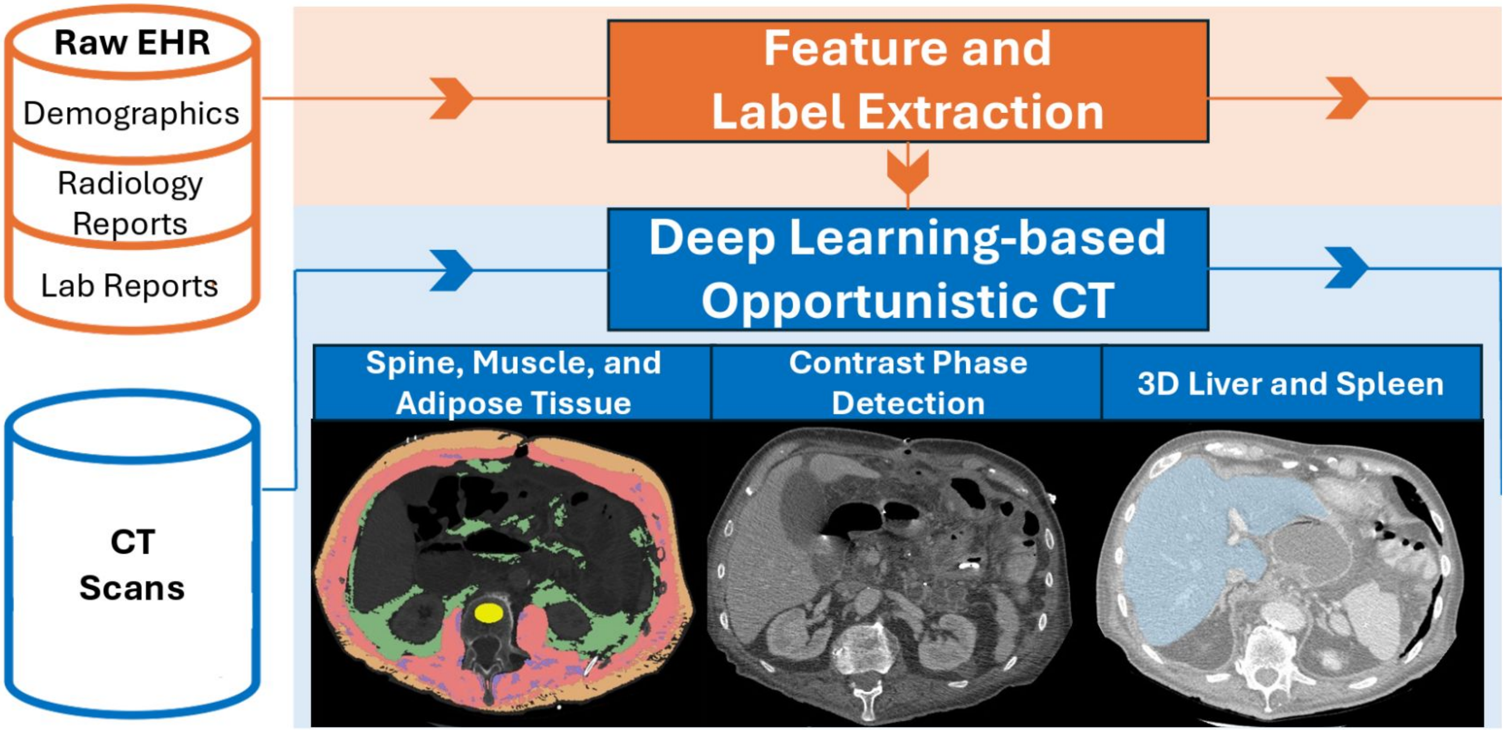
[1]Stanford University
Stanford, CA, USA

Available on ArXiv

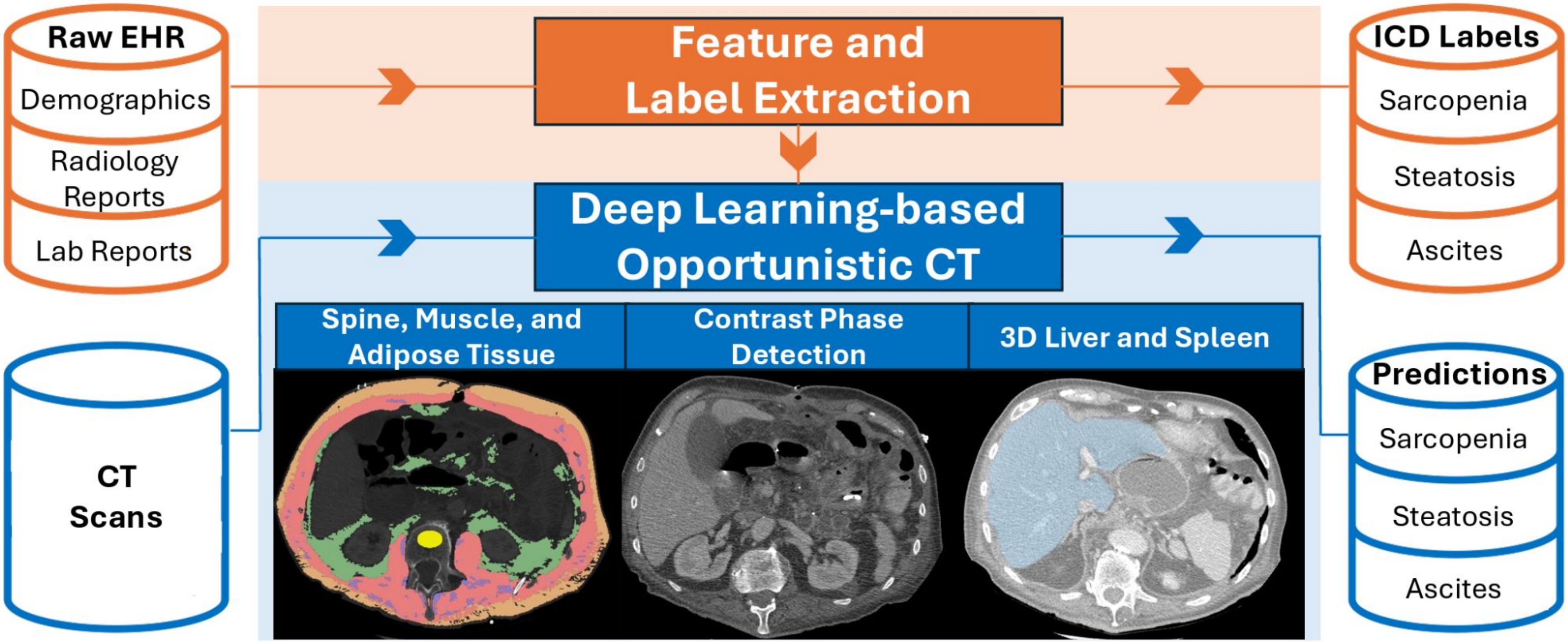
Pipeline


Raw EHR
Demographics
Radiology Reports
Lab Reports
CT Scans

Pipeline

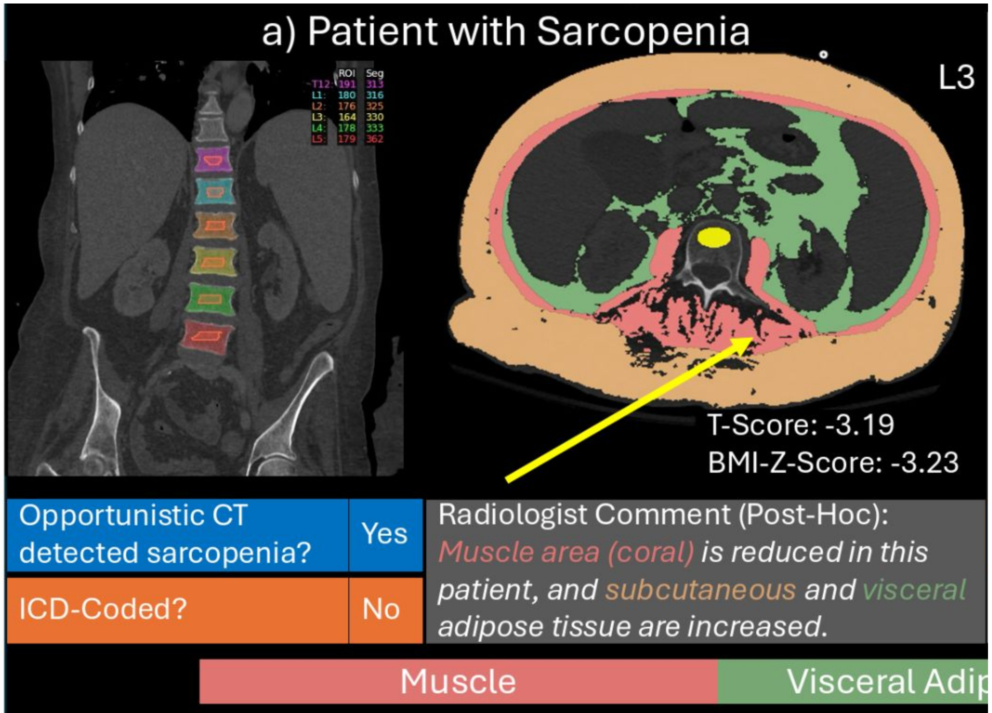

Raw EHR
Demographics
Radiology Reports
Lab Reports
CT Scans
Feature and Label Extraction
Deep Learning-based Opportunistic CT
Spine, Muscle, and Adipose Tissue
Contrast Phase Detection
3D Liver and Spleen

Pipeline



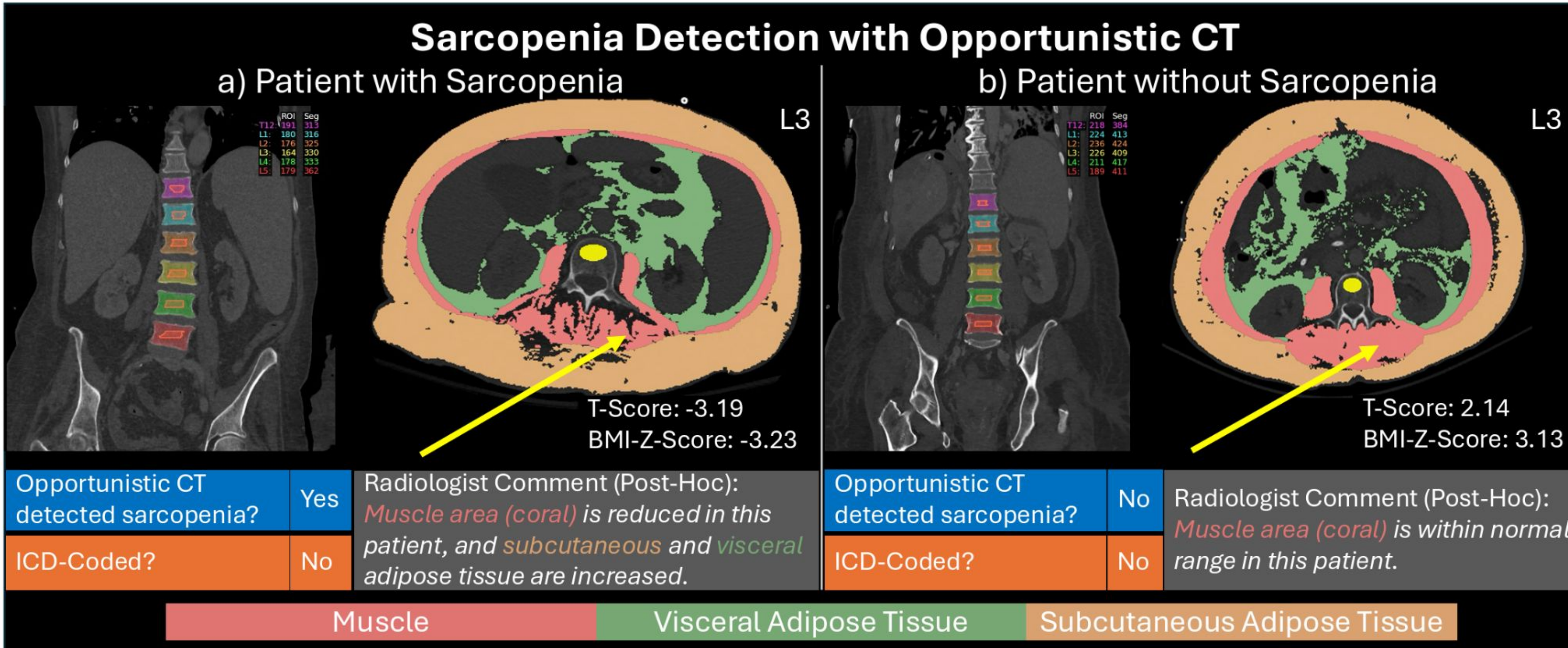
Raw EHR
Demographics
Radiology Reports
Lab Reports
Feature and Label Extraction
ICD Labels
Sarcopenia
Steatosis
Ascites
Deep Learning-based Opportunistic CT
Spine, Muscle, and Adipose Tissue
Contrast Phase Detection
3D Liver and Spleen
CT Scans
Predictions
Sarcopenia
Steatosis
Ascites

Sarcopenia



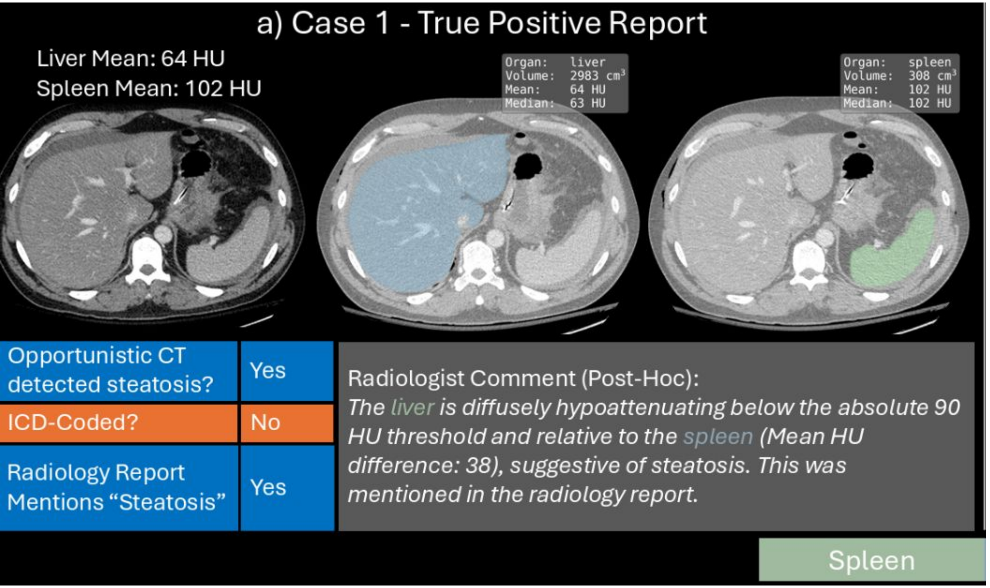
a) Patient with Sarcopenia
ROI Seg
T12: 191 313
L1: 180 316
L2: 176 325
L3: 164 330
L4: 178 333
L5: 179 362
L3
T-Score: -3.19
BMI-Z-Score: -3.23
Opportunistic CT detected sarcopenia?
Yes
ICD-Coded?
No
Radiologist Comment (Post-Hoc):
Muscle area (coral) is reduced in this patient, and subcutaneous and visceral adipose tissue are increased.
Muscle
Visceral Adip

Sarcopenia

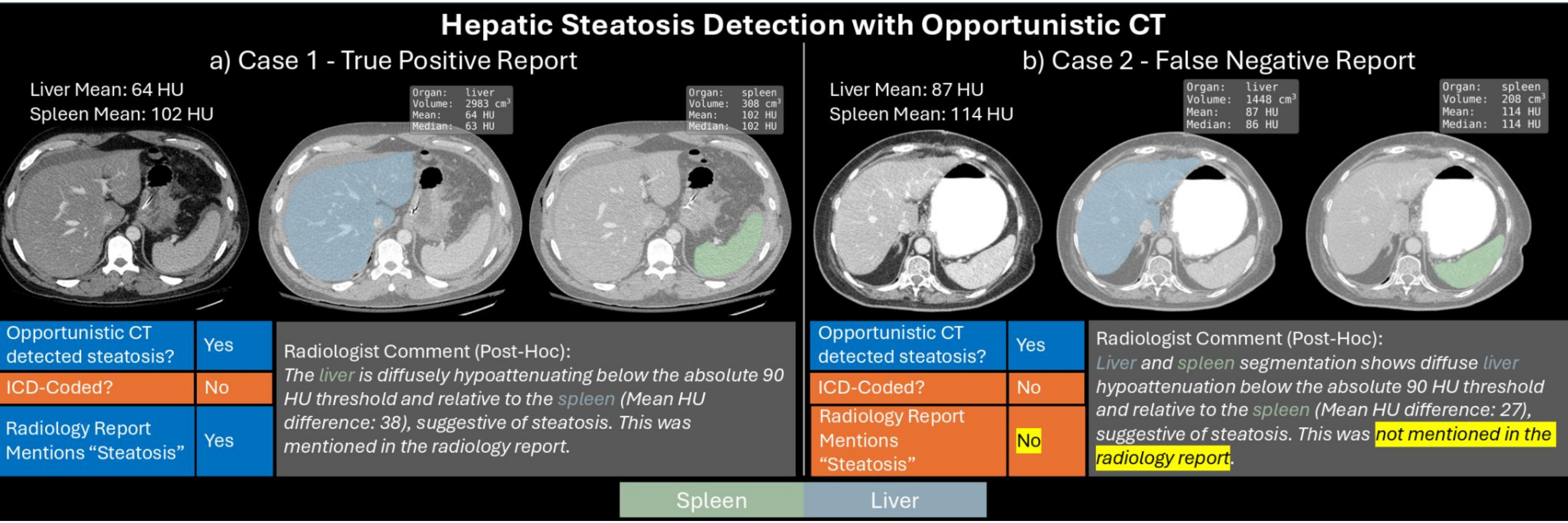


Sarcopenia Detection with Opportunistic CT
a) Patient with Sarcopenia
b) Patient without Sarcopenia
L3
T-Score: -3.19
BMI-Z-Score: -3.23
L3
T-Score: 2.14
BMI-Z-Score: 3.13
Opportunistic CT detected sarcopenia?
Yes
ICD-Coded?
No
Radiologist Comment (Post-Hoc): Muscle area (coral) is reduced in this patient, and subcutaneous and visceral adipose tissue are increased.
Opportunistic CT detected sarcopenia?
No
ICD-Coded?
No
Radiologist Comment (Post-Hoc): Muscle area (coral) is within normal range in this patient.
Muscle
Visceral Adipose Tissue
Subcutaneous Adipose Tissue

Hepatic Steatosis



Hepatic Steatosis



Overlap in Steatosis Detection

Table 2: Overlap in steatosis detection using: a) Liver HU, b) Liver-Spleen HU, c) Radiology Reports, d) ICD coding.

Liver ≤ 90 HU	Liver-Spleen ≤ -19 HU	Radiology Report	ICD-Coding	Count	%
Yes	Yes	Yes	Yes	5	0.2%
			No	58	2.5%
		No	Yes	1	0.0%
			No	68	3.0%
Yes	No	Yes	Yes	1	0.0%
			No	16	0.7%
		No	Yes	0	0.0%
			No	85	3.7%
No	Yes	Yes	Yes	2	0.1%
			No	33	1.5%
		No	Yes	6	0.3%
			No	211	9.3%
No	No	Yes	Yes	2	0.1%
			No	52	2.3%
		No	Yes	11	0.5%
			No	1,724	75.8%
Total				**2,275**	**100.0%**

Overlap in Steatosis Detection

Table 2: Overlap in steatosis detection using: a) Liver HU, b) Liver-Spleen HU, c) Radiology Reports, d) ICD coding.

Liver ≤ 90 HU	Liver-Spleen ≤ -19 HU	Radiology Report	ICD-Coding	Count	%
Yes	Yes	Yes	Yes	5	0.2%
			No	58	2.5%
		No	Yes	1	0.0%
			No	68	3.0%
Yes	No	Yes	Yes	1	0.0%
			No	16	0.7%
		No	Yes	0	0.0%
			No	85	3.7%
No	Yes	Yes	Yes	2	0.1%
			No	33	1.5%
		No	Yes	6	0.3%
			No	211	9.3%
No	No	Yes	Yes	2	0.1%
			No	52	2.3%
		No	Yes	11	0.5%
			No	1,724	75.8%
Total				**2,275**	**100.0%**

Conclusions

1. We demonstrate the potential of deep learning-based **opportunistic CT** in **improving the detection and coding** of medical conditions.

2. Found substantial discrepancies b/w condition prevalence and coding:
   a. Sarcopenia: Out of scans diagnosed through opportunistic imaging, only **0.5% scans were ICD-coded**
   b. Hepatic Steatosis: Out of scans diagnosed through opportunistic imaging or radiology reports, only **3.2% scans were ICD-coded**
   c. Ascites: Out of scans diagnosed with ascites through opportunistic imaging or radiology reports, only **30.7% scans were ICD-coded**

Thank You