Deep Learning for solving ill-posed Problems in Medical Image Analysis

The goal is to learn f s.t. $f(data) =$ *useful* ***output***

SIAM Conference on Imaging Science (IS22)

March 22 - Jin Keun Seo (Yonsei University, Korea)

Inverse Problems in Medical Imaging

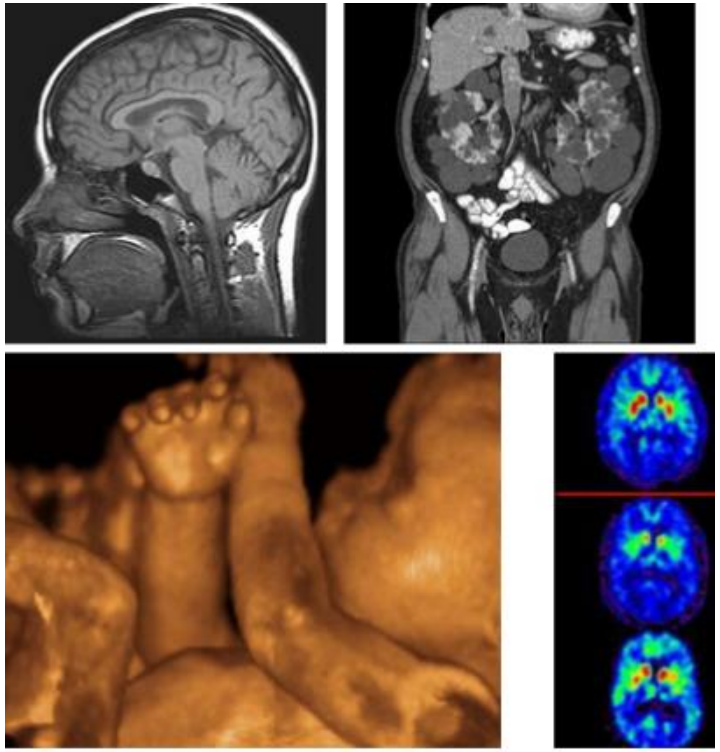
Medical imaging techniques have advanced to improve our ability to visualize internal information of the human body .

In CT& MRI, tomographic images are obtained by solving the inverse problem

$$g\,(x) = b$$

Forward Model   Image   data

Image : x



Data Acquisition : b

$$f = g^{-1}$$

ill-posed inverse problems arising in medicine

Solve $$g(x) = b$$

Forward model **image** **data**

Increase the resolution as high as possible

Reduce data as little as possible

The goal is to provide high resolution images (e.g., CT & MRI), while optimizing data collection in terms of **minimal time, cost-effectiveness, and low invasiveness.**

This willingness leads to **ill-posed inverse problem** in the classical sense.

Linearization

Actual models of CT & MRI are somewhat non-linear ($\frac{\partial g}{\partial x}$ depends on x) due to the interaction between the applied energy and tissue. Linearization is used for **robust reconstruction** and to guarantee **repeatability**.

Solve $g(\boldsymbol{x}) = \boldsymbol{b}$

Linearization

(# rows << # columns)

$$A \boldsymbol{x} \approx \boldsymbol{b}$$

Forward Matrix

Medical image

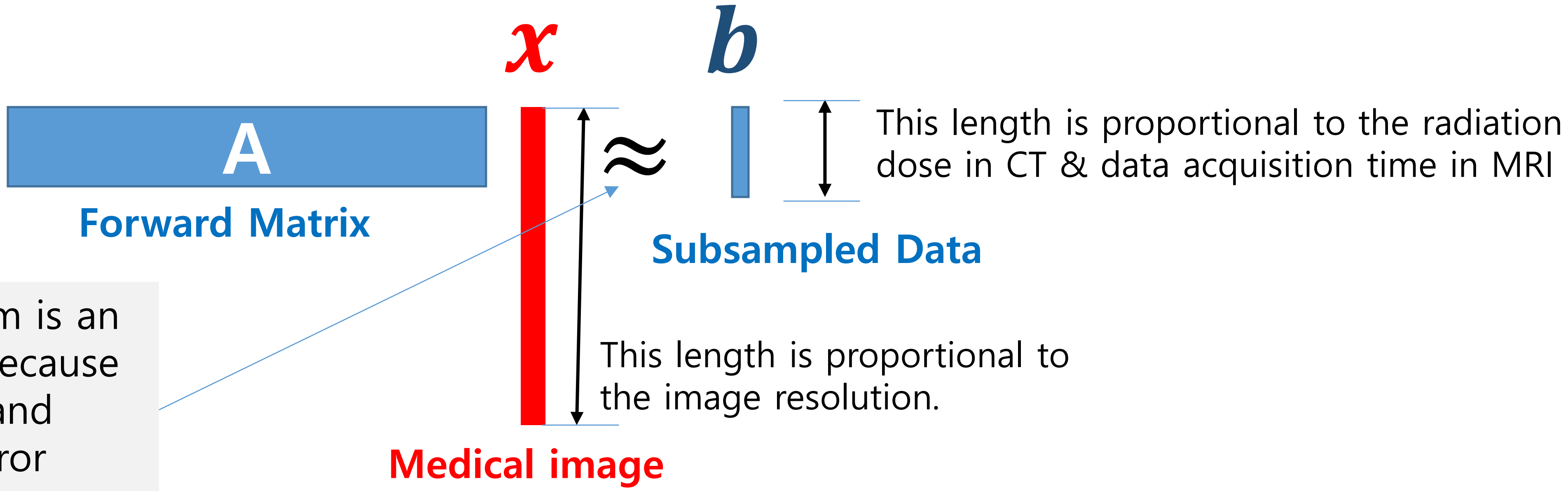
Measured Data

- Matrix **A** is determined by the data sampling strategy $\boldsymbol{b}$.
- The $j-th$ column of A represents the change of $\boldsymbol{b}$ due to the perturbation of $j-th$ element of $\boldsymbol{x}$.

The goal is to reduce measurement data as little as possible while maintaining high resolution images.

✓ **The goal is to make $\dfrac{\text{\# of equations}}{\text{\# of unknowns}}$ as small as possible.**

Example: Reduce radiation dose in CT & data acquisition time in MRI.



This linear system is an approximation because modeling error and measurement error always exist.

Example 1: **Low-dose Dental Cone-Beam CT** (1/3)

The goal is to provide high-resolution imaging with the lowest possible radiation dose & minimal equipment and maintenance costs.

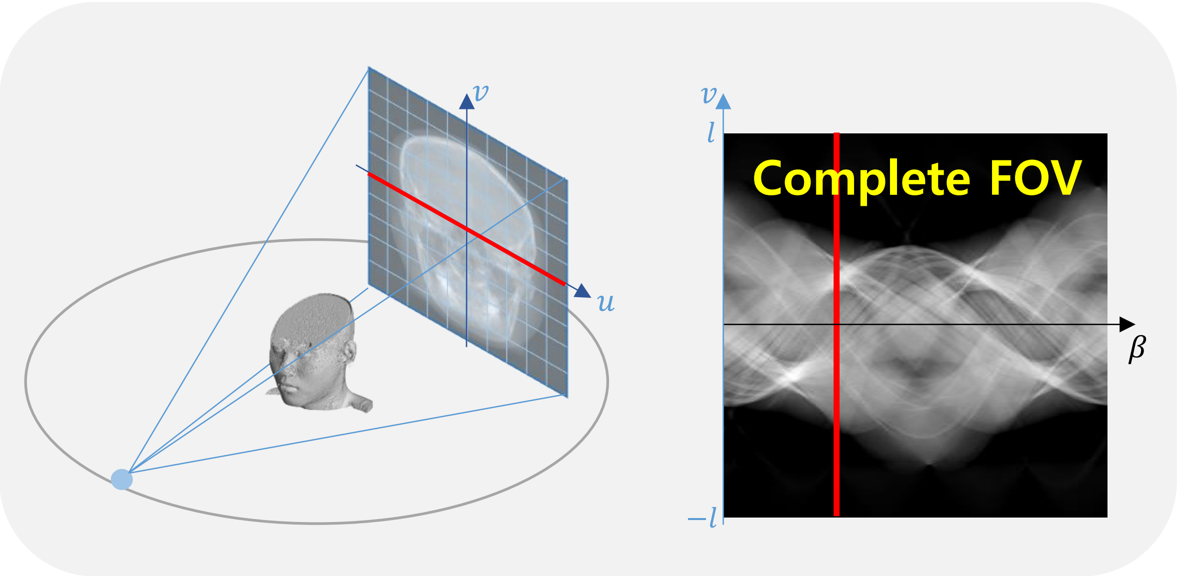
A $\boldsymbol{x}$ ≈ $\boldsymbol{b}$

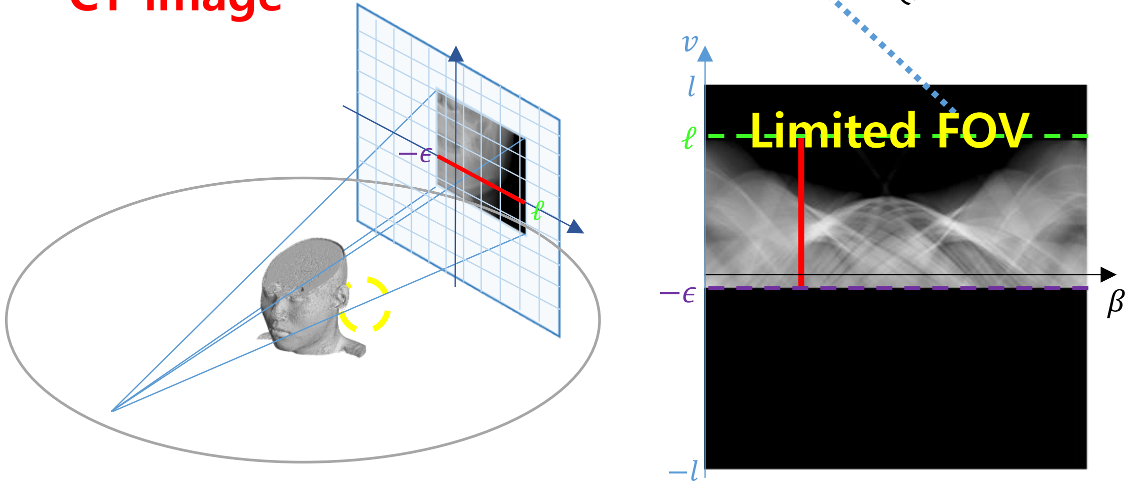
Radon Transform

CT image

Sinogram Data

Reduce data as little as possible



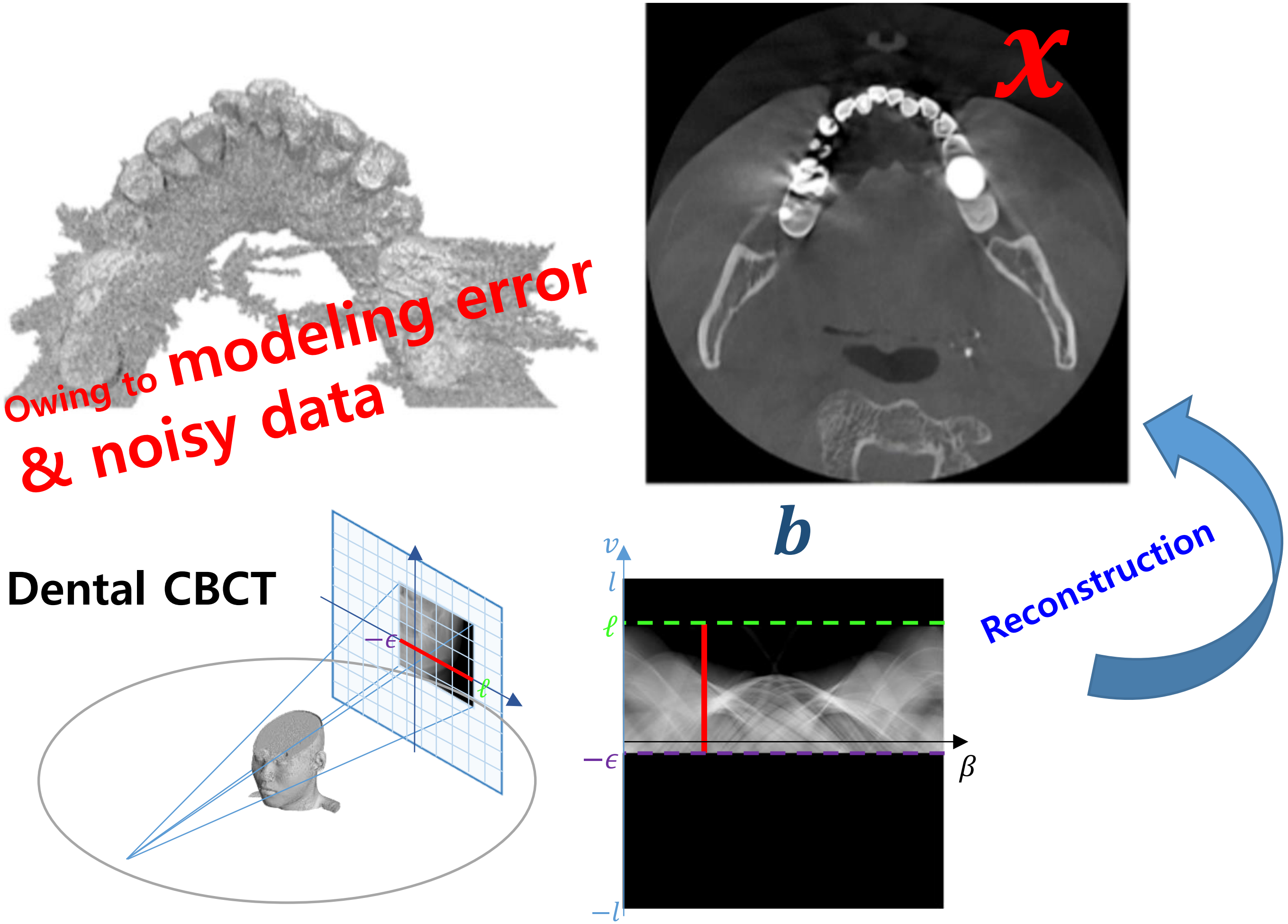


Most Dental CBCTs use flat-panel offset detectors.

Example 1: Low-dose Dental Cone-Beam CT (2/3)

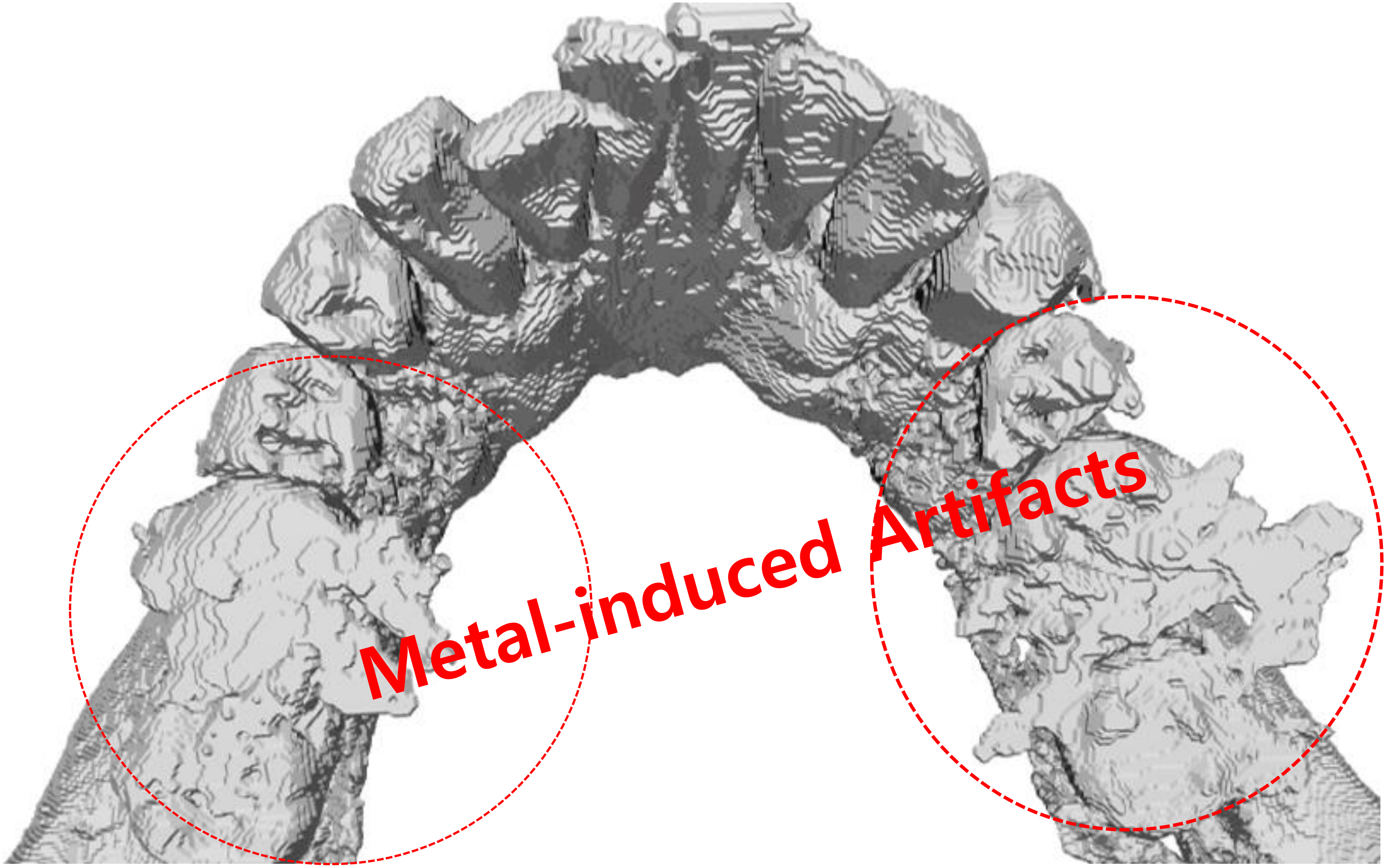
In dental CBCT, **metal-induced artifacts** are common. Image reconstruction using low-dose truncated data tends to be severely degraded by various artifacts and noise.

Dental CBCT	MDCT
· Circular cone-beam	· Helical cone-beam
· Scan time: 8-24sec	· Scan time< 1sec
· FOV truncation	· No FOV truncation
· Offset detector	· No Offset detector
· Price > \$ 0.1 million	· Price > \$ 1 million
· Low X-ray dose	· High X-ray dose



Example 1: **Low-dose Dental Cone-Beam CT** (3/3)

The presence of metal implants in CT scans causes severe discrepancy in the model assumption of the X-ray data, resulting in metal artifacts. The challenging problem is getting rid of these artifacts.

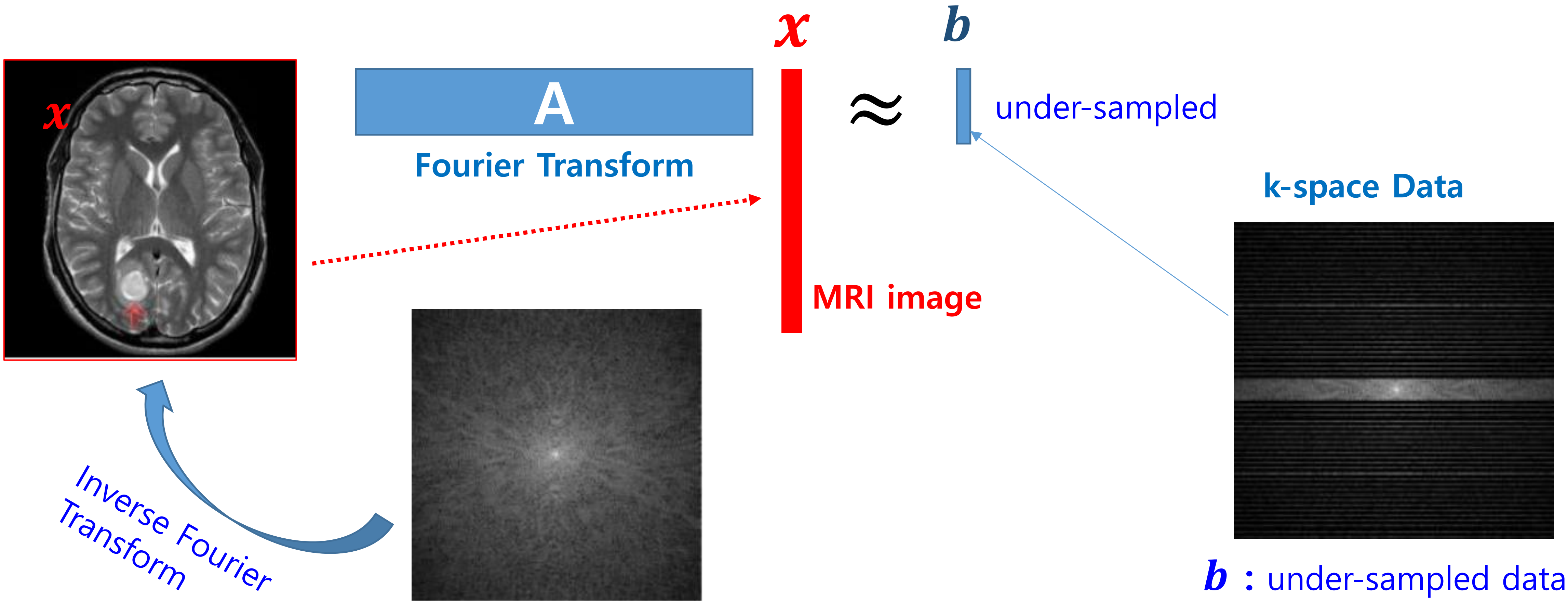


[My Experience]

- ✓ Attempting to incorporate nonlinearity into the model actually tends to do more harm than good.
- ✓ Repeatability is important in medicine.

Example 2: **Compressed Sensing MRI** (1/3)

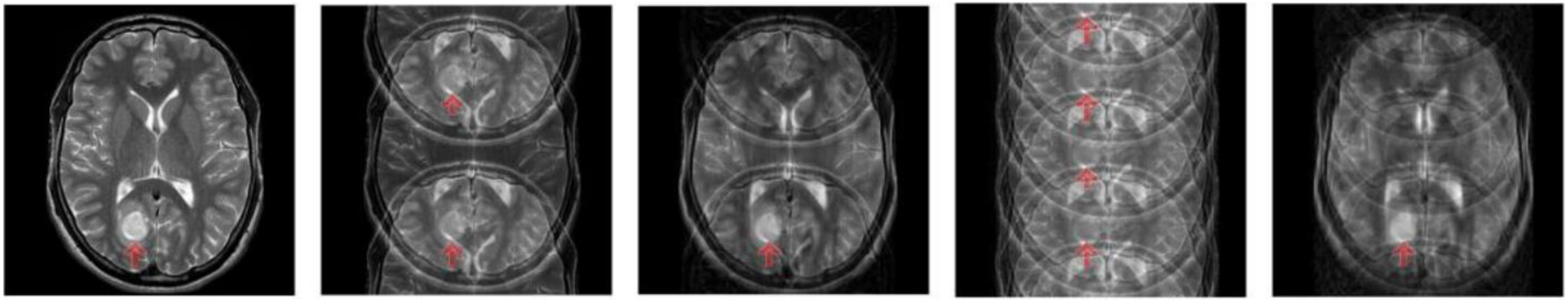
MRI has limited clinical application due to the **long data collection time**.
The goal is to reconstruct x from **minimally sampled data** b that violate the Nyquist criterion .



b_{full} **:** fully sampled data satisfying Nyquist criterion

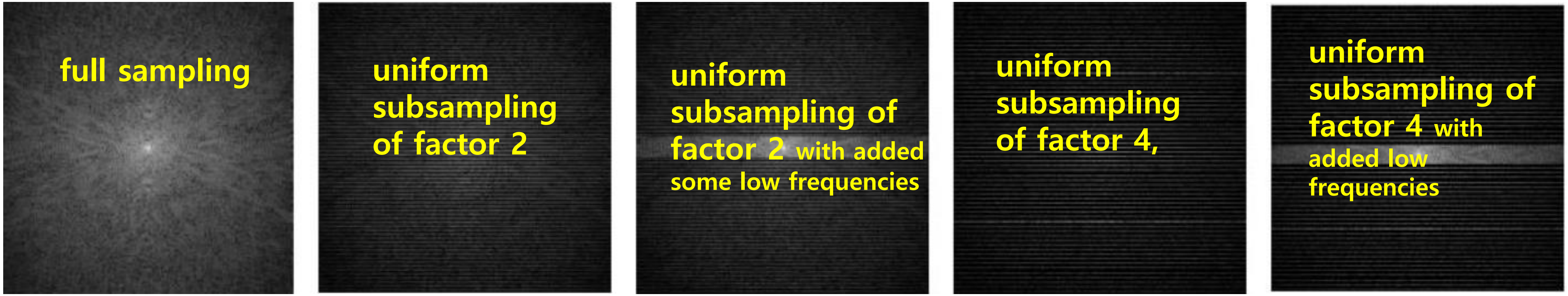
b **:** under-sampled data violating Nyquist criterion

Example 2: **Undersampled MRI** (2/3)



Inverse Fourier Transform

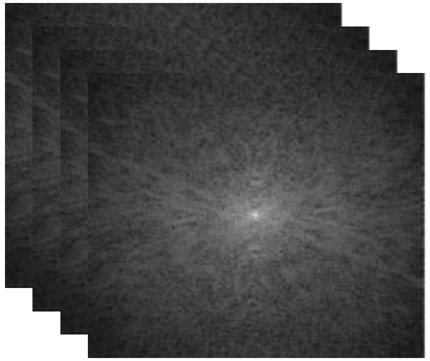
According to the Poisson summation formula, the inverse Fourier transform of the uniformly subsampled data with factor 4 provides an four-folded image. Unfolding is a highly non-linear problem that is difficult to deal with. This is why random sampling patterns were used for accelerated MRI strategies (Donoho, Candes, Tao, Rustig).



Example 2: **Undersampled MRI** (3/3)

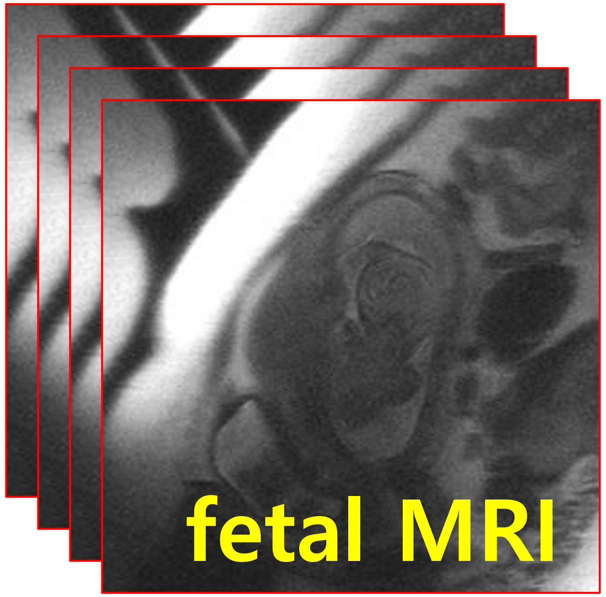
For **fetal MRI**, a fast MRI is needed to deal with **motion artifacts.**

A full sampling that meets the Nyquist criteria can take approximately 10-30 minutes.



Inverse Fourier Transform

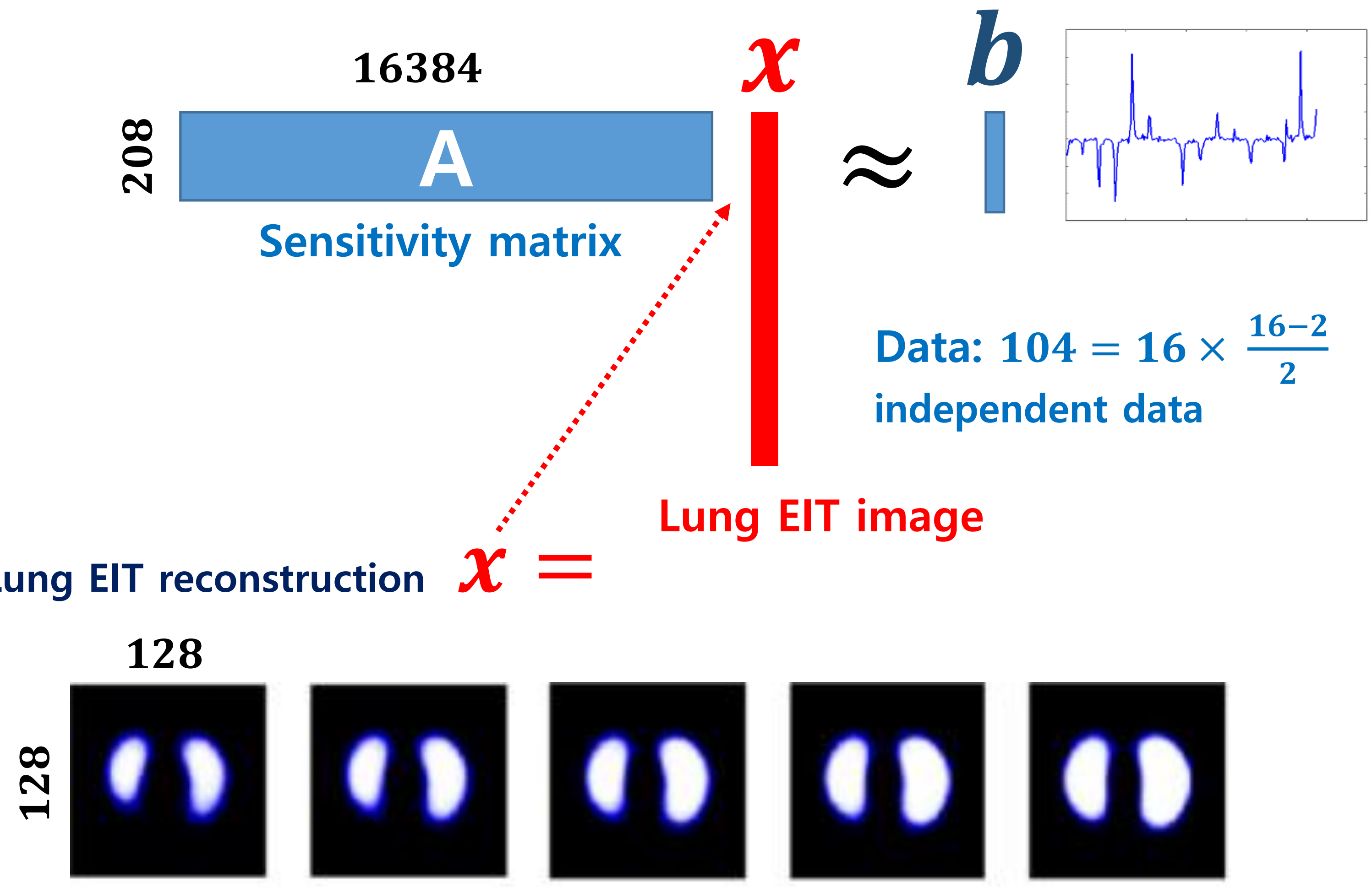
Due to the long scan time, the reconstructed image is severely degraded by motion artifacts.



Normally, the fetus is always moving. Hence, full sampling has a fundamental limitation of being contaminated with motion artifacts.

Example 3: **Electrical Impedance Tomography** (1/3)

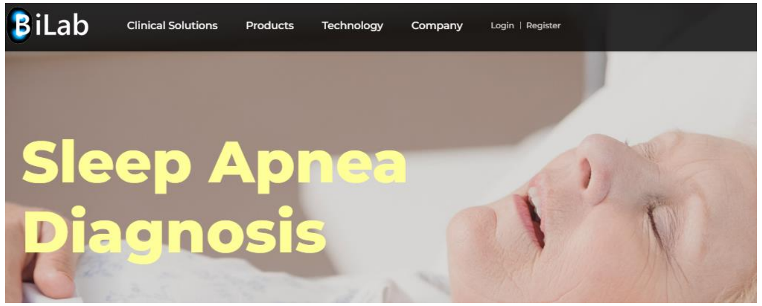
This is a very nonlinear ill-posed problem because we have no control over the path of the injection current, unlike CT & MRI. It is difficult to provide high-resolution images due to severe modeling errors and boundary uncertainties.



16 Channel EIT system

Example 3: **EIT** (2/3)

The goal is to provide useful images,
not high resolution images.

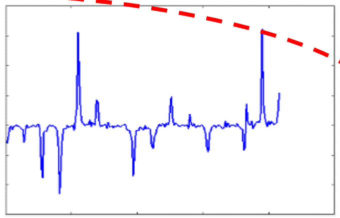


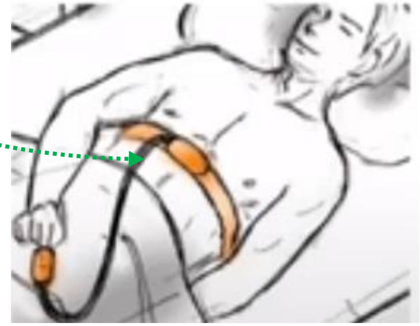
16384

208

$$\mathbf{A} = \frac{\partial b}{\partial x} \quad \text{(sensitivity matrix)}$$

x $\approx$ **b**



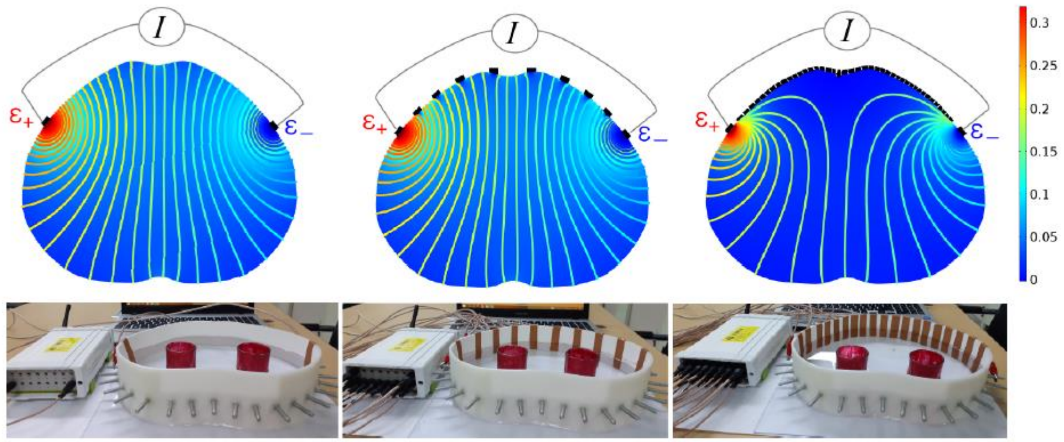


- ✓ BiLab provided a lung ventilation EIT system to diagnose sleep apnea.
- ✓ The most important issue for sale is the development of a **simple electrode belt.** Doctors are reluctant to use it because the electrode attachment method is inconvenient.

Example 3: **EIT** (3/3)

In my opinion, more data from many electrodes does not help with better image reconstruction due to various uncertainties and modeling errors.

- ✓ When many electrodes are attached to obtain a lot of data (Neuman-to-Dirichlet map) and the distance between the electrodes is narrow, the current flows only near the boundary, so the NtD data cannot reflect the internal conductivity distribution.



So far, **four-pair** electrode **bioimpedance** has been commercially successful.

How to solve
ill-posed inverse problems
in medicine

# of equations

$$A \quad \approx \quad b$$

Reduced # of measurement data

subsampled data

# of unknowns (# of pixels)

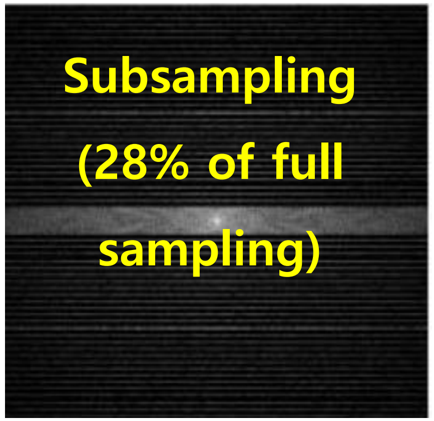
Due to the under-sampled data relative to the resolution of the solution, it is necessary to **impose specific prior knowledge** of the expected solutions.

How to solve the ill-posed problem $Ax = b$?

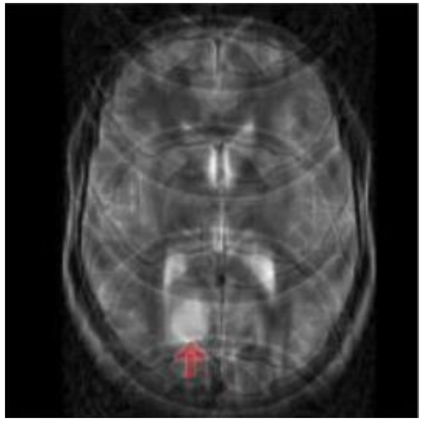
Example 2: Under-sampled MRI model

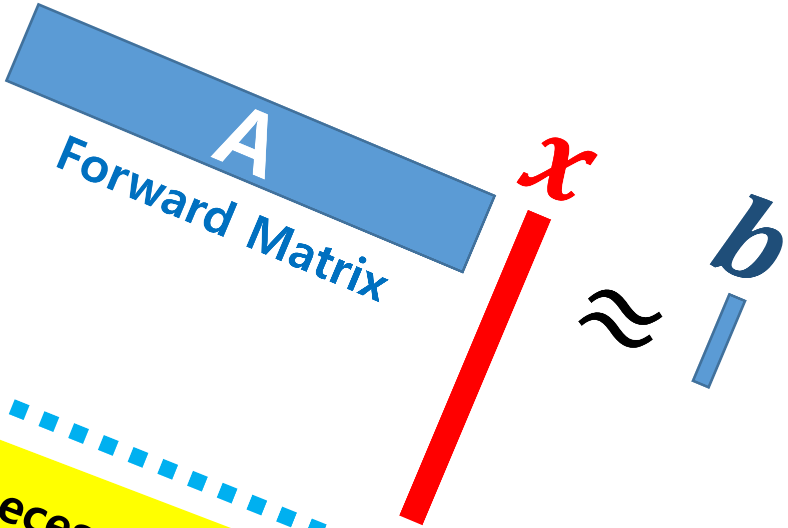
Using a minimally sampled data b, we want to reconstruct an image x similar to that obtained from the fully sampled k-space data b_{full}.

b

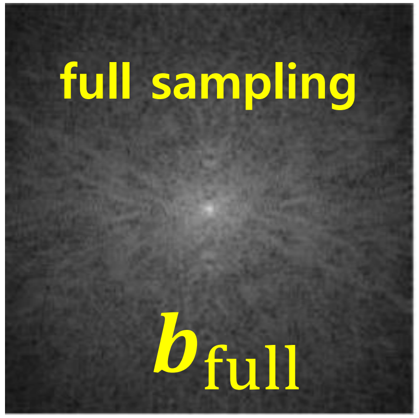


Inverse Fourier Transform

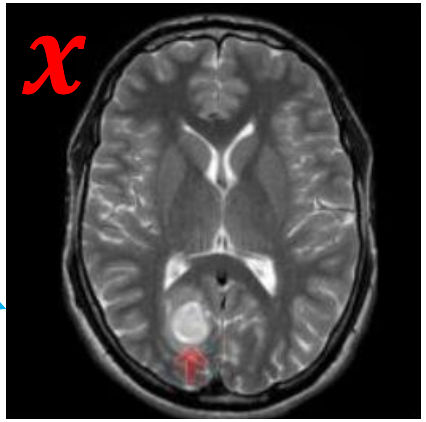




It is necessary to utilize prior information about the solution.



Inverse Fourier Transform



A necessary condition for solving the ill-posed problem $Ax = b$

To convert an ill-posed problem into a well-posed one,

1) need a suitable **data sampling strategy** involving A

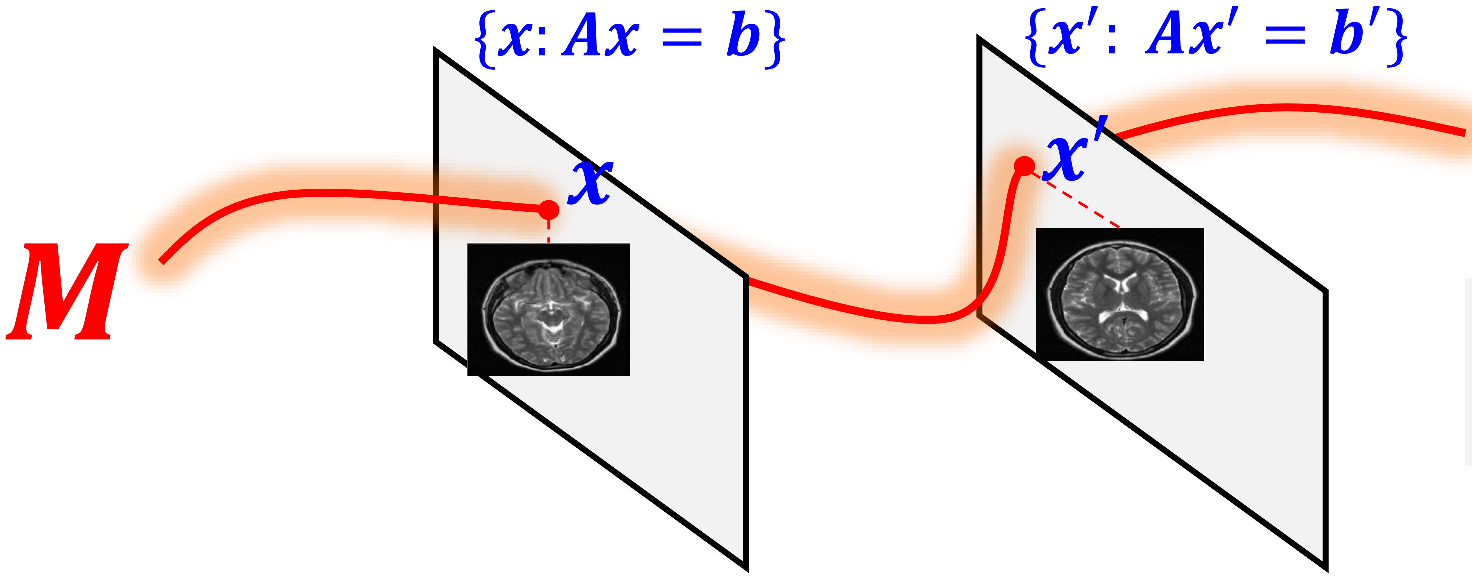
2) choose a **highly reduced solution space (or manifold),** denoted by M,

so that these choices allow to satisfy the **M-Restricted Isometry Property (RIP)** condition:

(Hyun etal 2021, Candes & Tao 2005)

$$\frac{1}{c}||x - x'|| \leq ||Ax - Ax'|| \leq c\,||x - x'||, \quad \forall\, x, x' \in M$$

Image Prior



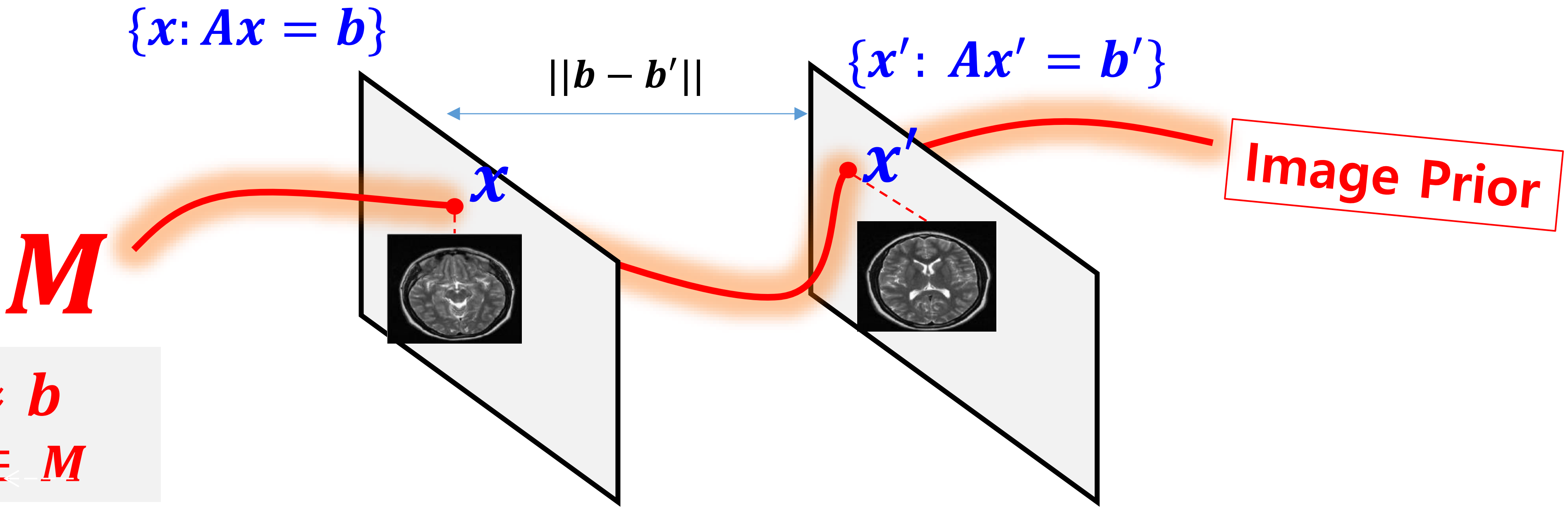
Solve $Ax \approx b$

subject to $x \in M$

What does M-RIP condition mean?

$$\frac{1}{c}||x - x'|| \le ||Ax - Ax'|| \le c\,||x - x'||, \quad \forall\, x, x' \in M$$

The Euclidean distance between data ($||b - b'||$) is comparable to the distance between images ($||x - x'||$) within the solution manifold M .



Solve $Ax \approx b$

subject to $x \in M$

If the **sampling strategy** (involving A) does not satisfy **M-RIP**, there is no way to solve $Ax = b$.

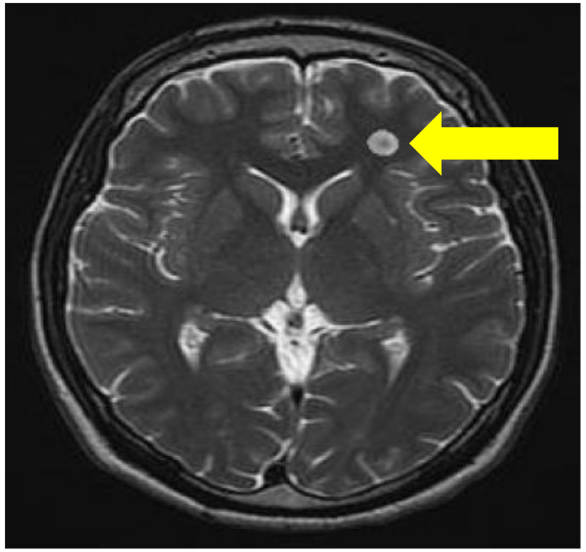
Example 2 (Undersampled MRI with **uniform subsampling** of factor 4);

We cannot solve $Ax = b$

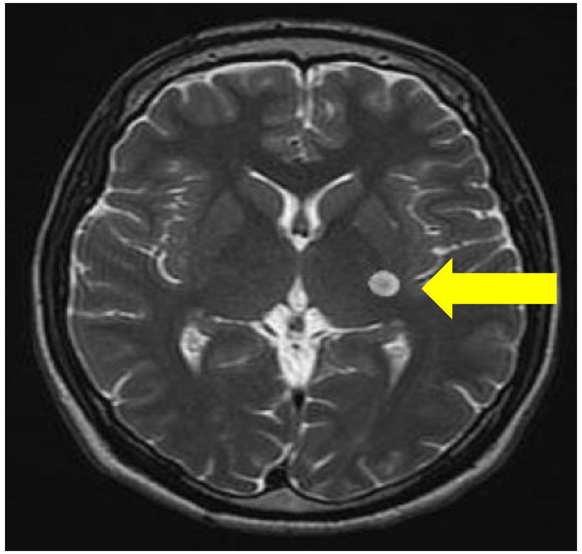
because $\exists\, \boldsymbol{x}, \boldsymbol{x}' \in \boldsymbol{M}$ **s.t.** $||\boldsymbol{x} - \mathbf{x}'|| > \mathbf{0} = ||\boldsymbol{Ax} - \boldsymbol{Ax}'||.$

This uniform subsampling does **NOT satisfy** $\boldsymbol{M}$**-RIP:** $\frac{1}{c}||\boldsymbol{x} - \boldsymbol{x}'|| \leq ||\boldsymbol{Ax} - \boldsymbol{Ax}'|| \leq \boldsymbol{c}\, ||\boldsymbol{x} - \boldsymbol{x}'||, \quad \forall\, \boldsymbol{x}, \boldsymbol{x}' \in \boldsymbol{M}$

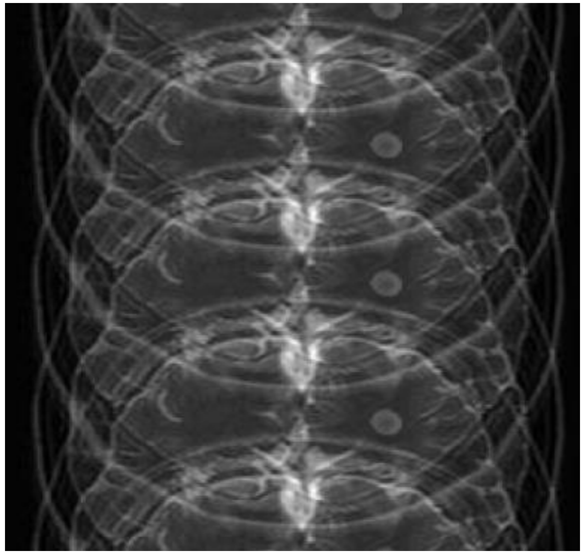
$\boldsymbol{x} \in \boldsymbol{M}$



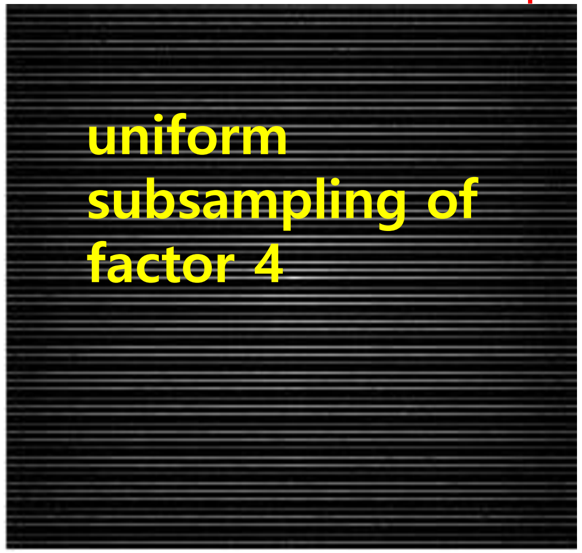
$\boldsymbol{x}' \in \boldsymbol{M}$



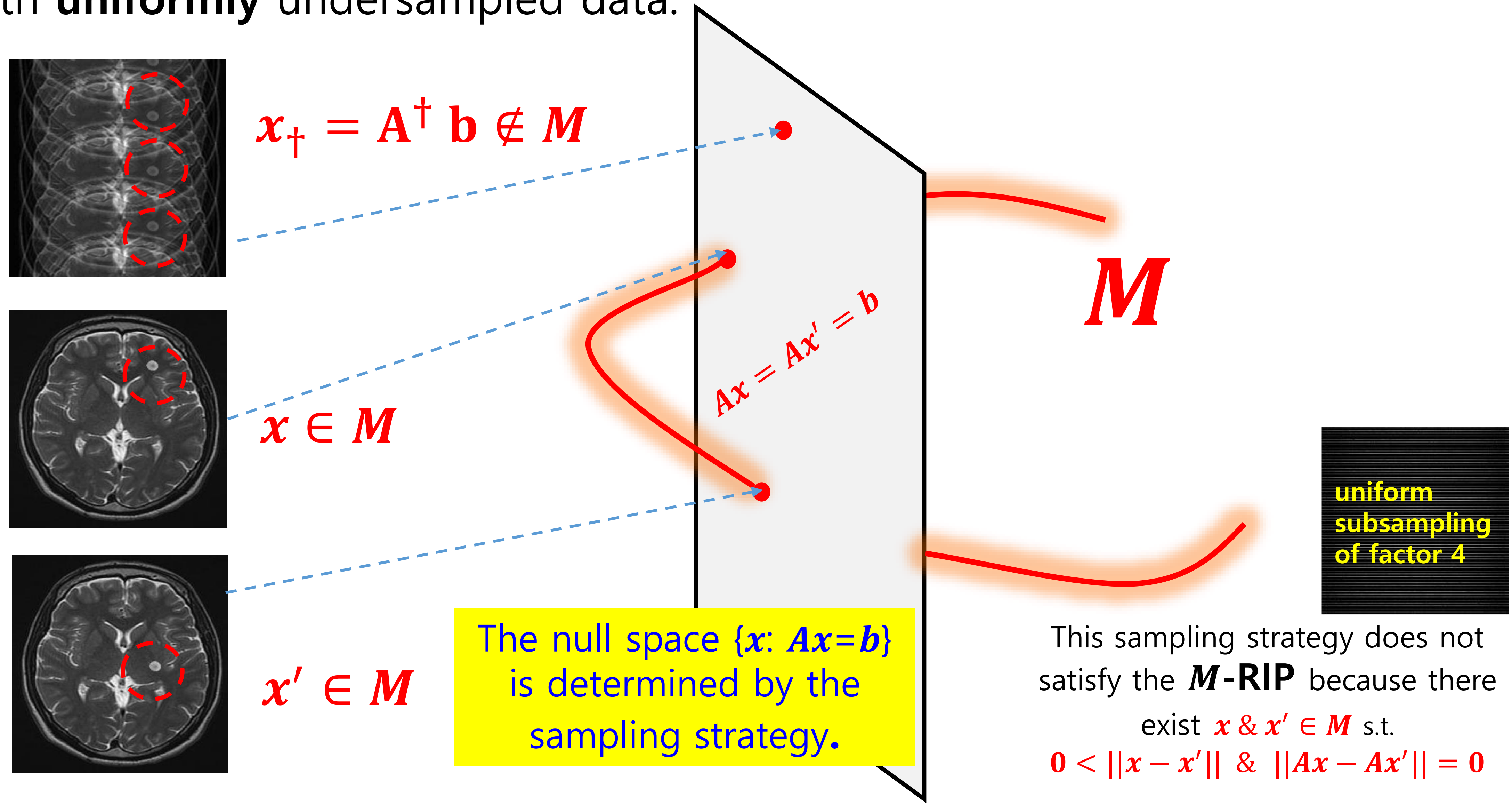
$\boldsymbol{x}_{\dagger} = \mathbf{A}^{\dagger}\,\mathbf{b} \notin \boldsymbol{M}$



$\boldsymbol{Ax} = \boldsymbol{Ax}' = \boldsymbol{Ax}_{\dagger}$



As shown in the figure below, it is not possible to know where the small tumor is with **uniformly** undersampled data.

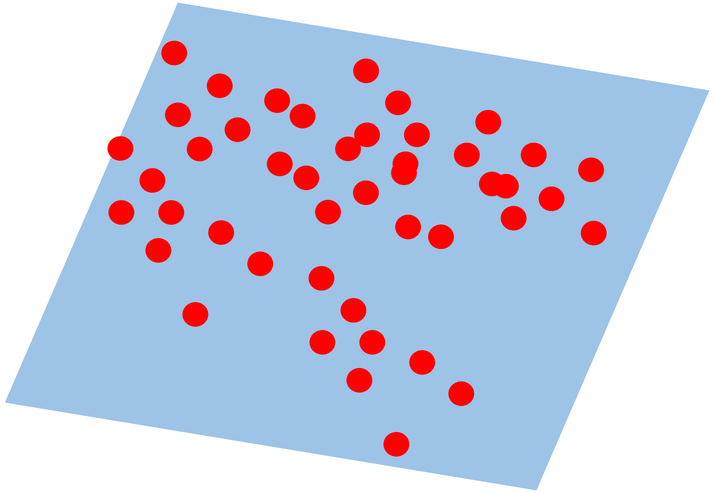


This sampling strategy does not satisfy the M-RIP because there exist x & $x' \in M$ s.t. $0 < ||x - x'||$ & $||Ax - Ax'|| = 0$

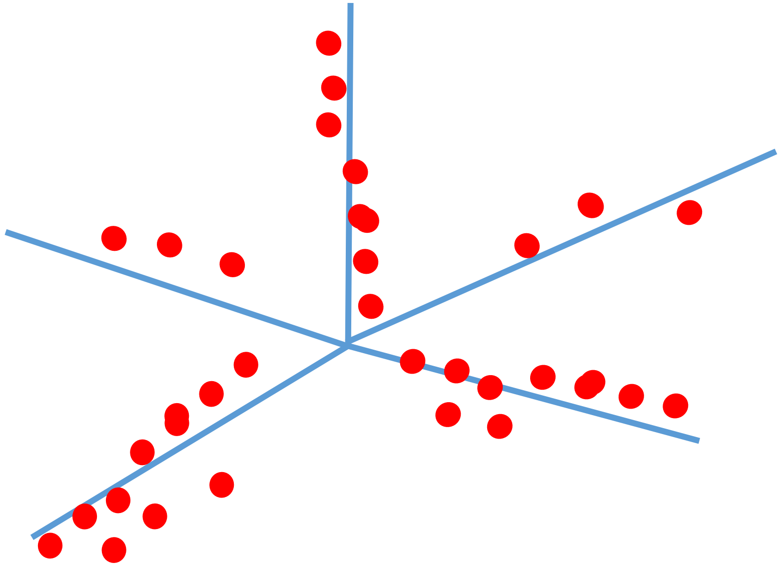
What is the solution space M?

What kind of prior information about the solution x constitutes M?

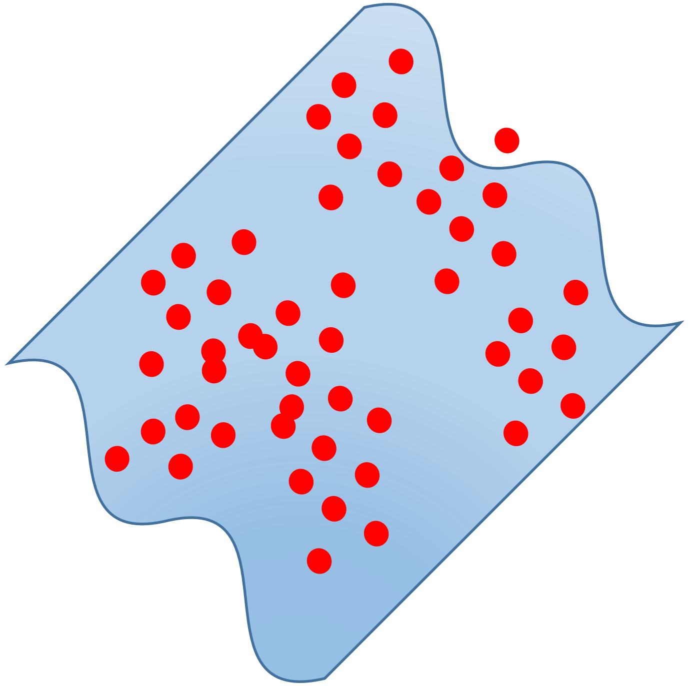
Linear regression
(PCA, truncated Fourier, Wavelet, Framelet, etc)



Piece-wise linear regression
Sparse Sensing (Total Variation, dictionary learning, sparse representation using wavelet, Framelet, etc)

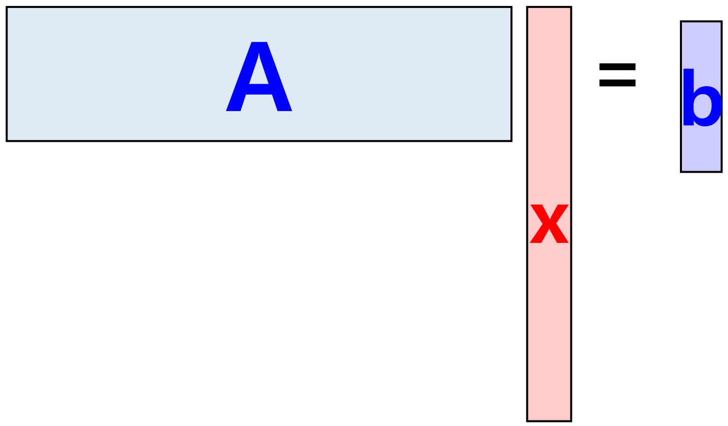


Non-Linear regression
(deep learning)



How can we impose prior information of target images?

Regularized Data Fitting
vs
Deep Learning



Two different approaches to impose prior information of target images

Solve $Ax \approx b$ subject to $x \in M$

Find $f: b \rightarrow x \in M$ s.t $Af(b) = b$

Regularized data fitting (Single Fidelity)

$$\mathrm{x} = \underset{x}{argmin} \sum_{\mathrm{n}} ||Ax - b||^2 + \lambda\, R(x)$$

$R(x)$ is regularization term such as

- $R(x) = ||\nabla x||_{\ell_1}$
- $R(x) = ||h||_{\ell_1}, \quad x = Dh$

Deep learning (Group Fidelity)

$$f = \underset{f \in \text{Neural Nets}}{argmin} \sum_{\mathrm{n}} ||x^{(n)} - f(b^{(n)})||^2$$

$\{(x^{(n)}, \mathrm{b}^{(n)}) : n = 1, \cdots, N\}$; **training data**

When f is learned, we get the solution directly from the input data.

$$f(b) = x$$

Regularized Data Fitting

$$A\,x = b$$

Compressive sensing (CS) methods have shown remarkable performance in image denoising.
But CS methods have limitations in medical imaging where small anomalous details are more important than overall features.

Image Prior

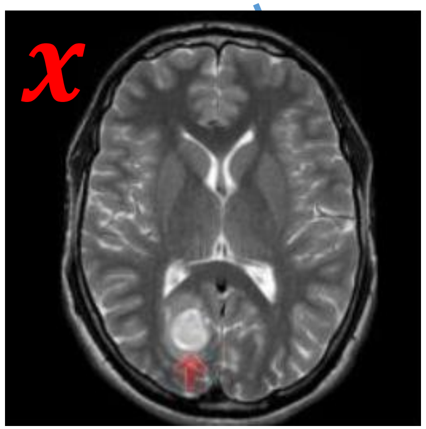
Regularization

$$f(\mathbf{b}) = \mathbf{Wh},\ \mathbf{h} = \underset{\mathbf{h}}{\operatorname{argmin}} \|A W \mathbf{h} - \mathbf{b}\|_{\ell^2}^2 + \lambda \|\mathbf{h}\|_{\ell^1}$$

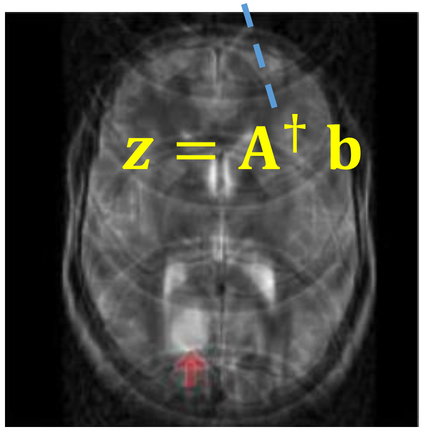
hope

data fidelity

Basis/axis



$$= \begin{bmatrix} w^{1,1} & \cdots & w^{1,k} \\ \vdots & \ddots & \vdots \\ w^{256^2,1} & \cdots & w^{256^2,k} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_k \end{bmatrix}$$

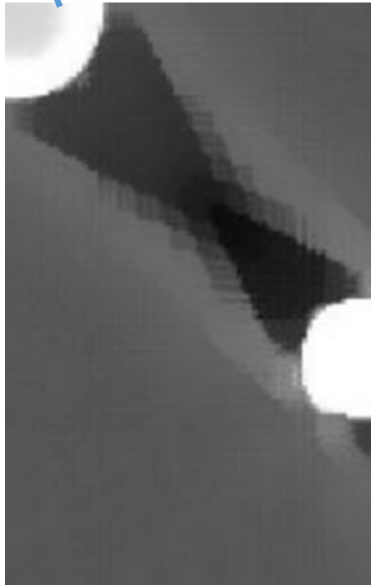


λ is a regularization parameter that controls the trade-off between data fidelity and the regularity enforcing the sparsity of **h**.

The term $\|\mathrm{h}\|_{\ell^1}$ is used to promote the sparse representation..

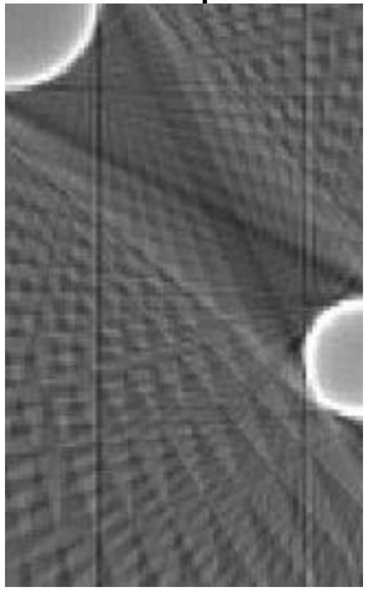
For example, Total Variation regularization method may **not selectively preserve small features** because it penalizes uniformly based on image gradient, regardless of image structure.

$$x = \underset{x}{\operatorname{argmin}} \|x - x_{\dagger}\|_{\ell^2}^2 + \lambda \|\nabla x\|_{\ell^1}$$

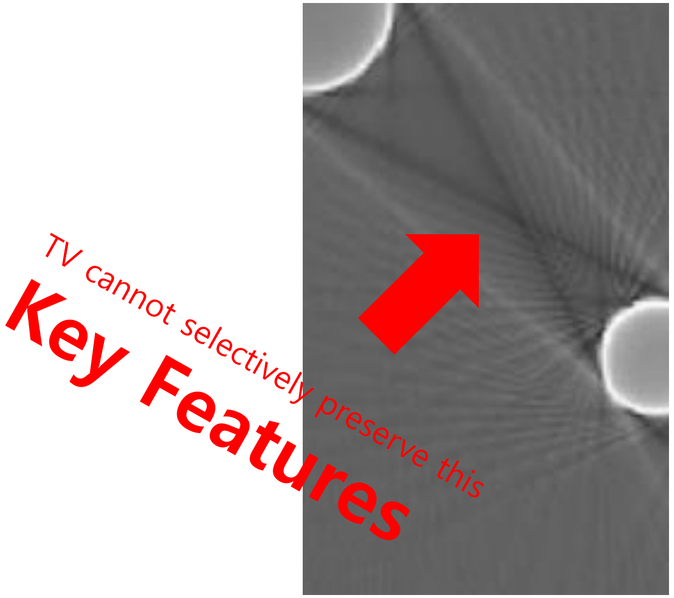


TV regularization

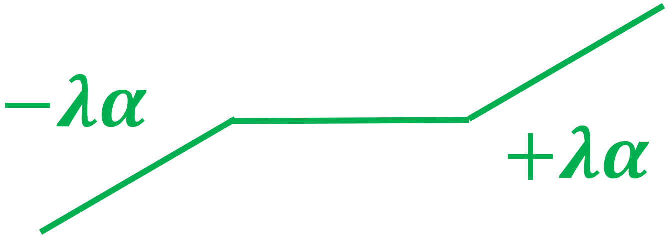
$$x_{\dagger} = \mathrm{A}^{\dagger}\,\mathbf{b}$$



True solution



TV removes everything within the shrinkage interval without exception.



- DL approach can **selectively** preserve fine features.

Deep learning (Group Fidelity)

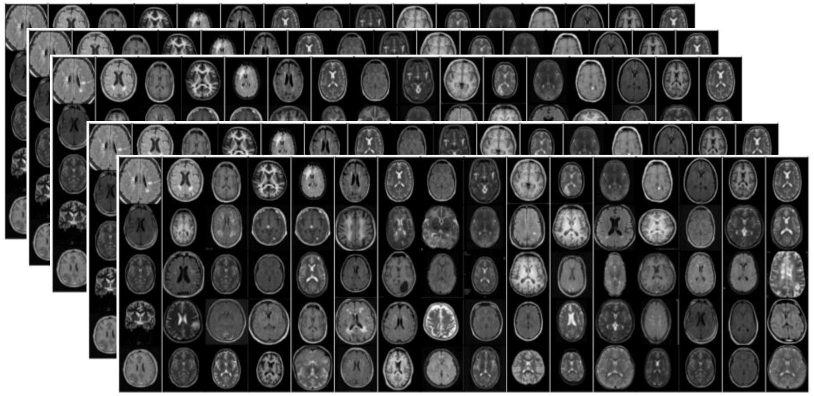
$$f = \underset{f \in \text{Neural Nets}}{argmin} \sum_{n} ||x^{(n)} - f(b^{(n)})||^2$$

$\{(x^{(n)}, \mathbf{b}^{(n)}) : n = 1, \cdots, N\}$ **; training data**

$$\underbrace{\{x^{(1)}, \cdots, x^{(N)}\}}_{\text{Training data}} \subset M \subset$$

Pixel-dimensional Euclidean space

Learning



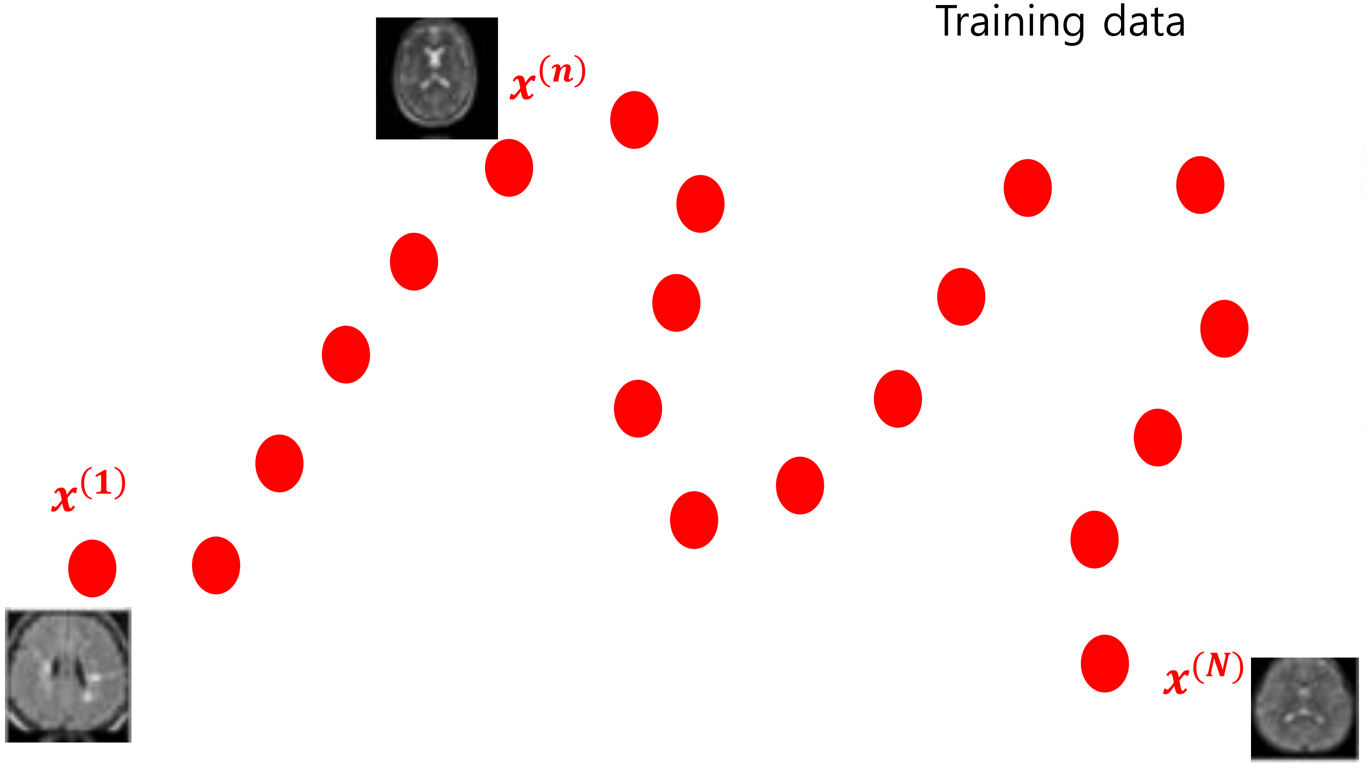
What is M ?

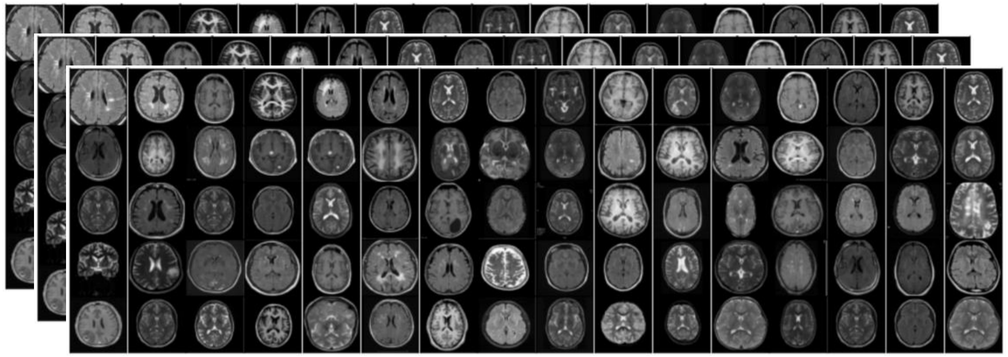
Medical image (e.g., 256 grayscale level, 300×300 size) can be regarded as **a point** $\mathbf{x} = (x_1, \cdots, x_{300^2})$ **in pixel-dimensional Euclidean space**, where x_j (j-th axis coordinate) corresponds to the grayscale intensity at the j-th pixel.

Training data are grayscale level images

$$\underbrace{\{\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(N)}\}}_{\text{Training data}} \subset \boldsymbol{M} \subset \underbrace{\{\mathbf{0}, \cdots, \mathbf{256}\}}_{\text{256 grayscale level}}{}^{\mathbf{300^2}}$$

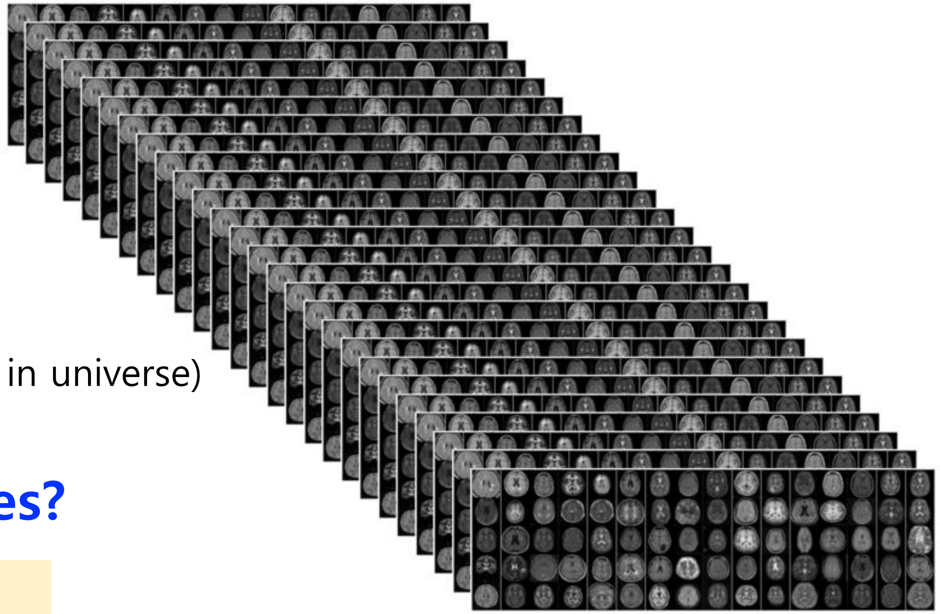
Number of pixels





What is M ? Even if tens of thousands of tomographic images are collected every day for 1000 years, it occupies a very small fraction of the pixel-dimensional Euclidean space.

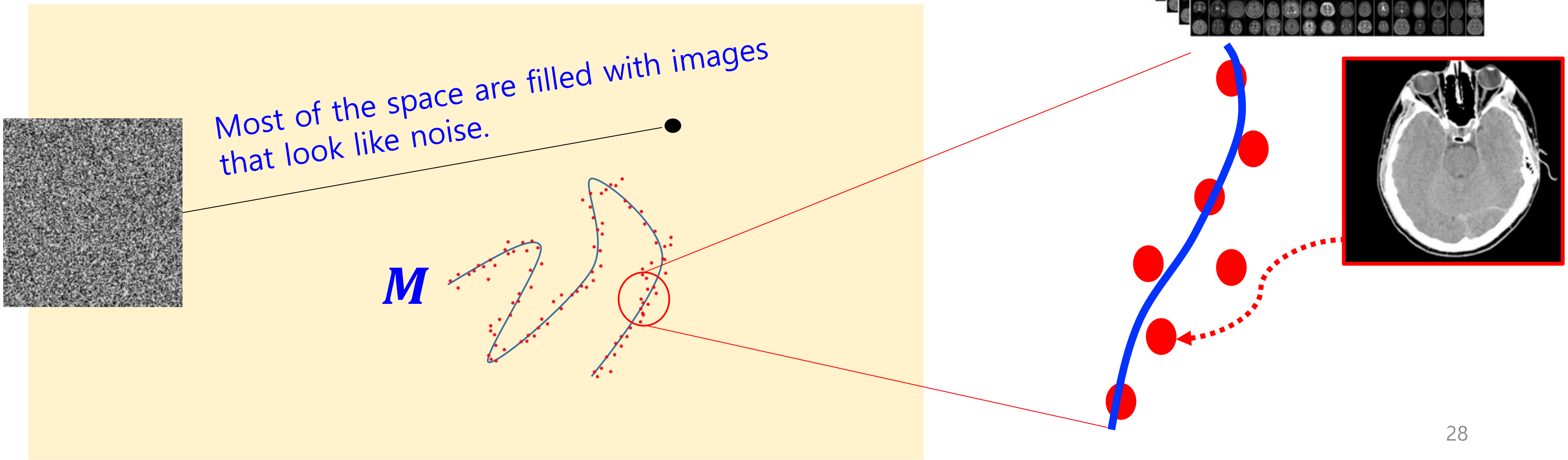
(much less than 0.0000000000000000000001%)



# of all possible images = $256^{300\times300} \approx \infty$

(256 grayscale level, 300×300 pixel size) (much greater than # of atoms in universe)

Challenging question: Can we find M from data samples?



Let us say that we have the following disentagled representation:

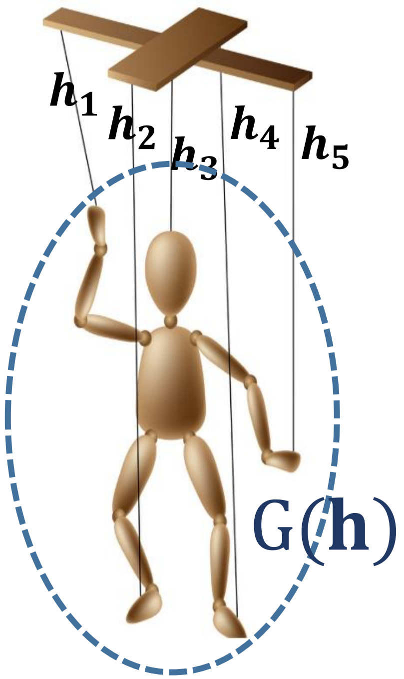
$\boldsymbol{M} = \{\, \boldsymbol{G(h)}: \boldsymbol{h} \in \boldsymbol{K} \}$ **with** $\boldsymbol{K} \subset \boldsymbol{R^k}$.

If $\boldsymbol{k} \leq$ **# rows** , the problem becomes well-posed.

$$\mathrm{G}(\,\mathbf{h}\,) = \boldsymbol{x}$$

Generator /decoder

Latent variable



Forward matrix $\boldsymbol{x}$ $\approx$ $\boldsymbol{b}$

Input Data

The latent variables can be regarded as strings connected to the marionette. The generator can be seen as realizing the movement of the marionette by pulling the strings.

- Many problems are ill-posed because the solution space is too large .

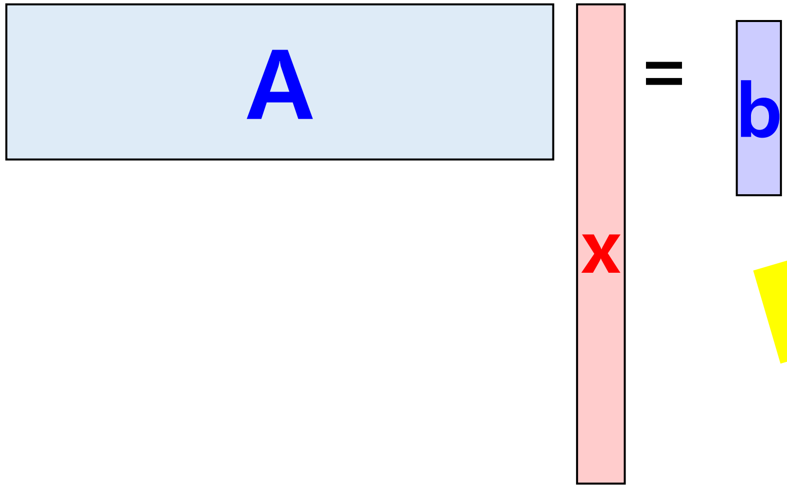
$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is nonlinear problem

if **dim** $(\mathrm{span}\{\partial_j G(h) : h \in K\}) > \# \, rows$.

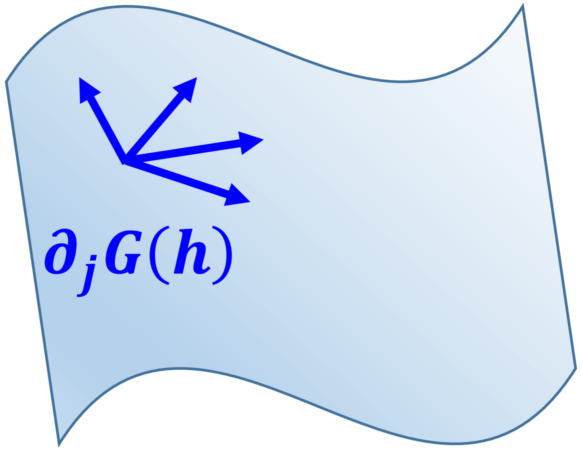
Tangent vectors on M

Solve $\boldsymbol{A}\boldsymbol{x} \approx \boldsymbol{b}$ subject to $\boldsymbol{x} \in \boldsymbol{M}$

$\boldsymbol{M} = \{\, G(h) : h \in K \}$



$\# \, \mathbf{b} \downarrow \;\rightarrow\;$ ***nonlinearity*** $\uparrow$



Chang Min Hyun, Seong Hyeon Baek, Mingyu Lee, Sung Min Lee, Jin Keun Seo, Deep Learning-Based Solvability of Underdetermined Inverse Problems in Medical Imaging, Medical Image Analysis

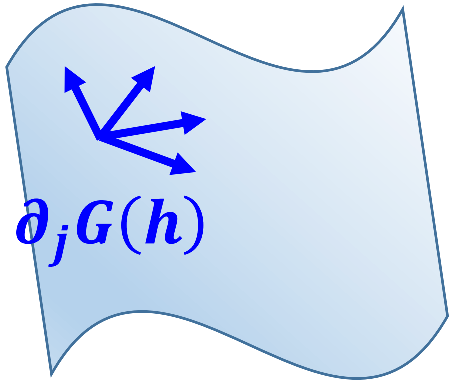
$f: b \to x$ is **nonlinear** if $\dim(\text{span}\{\partial_j G(h): h \in K\}) > \#\, rows$.

$f(b)$ is the solution of $Ax = b$

Proof:

- $b = Ax$ & $\mathbf{x} = \mathbf{G(h)}$ $\to$ $\forall h \in K$, $f(\mathbf{AG(h)}) = \mathbf{G(h)}$
- Assume $f: b \to x$ is linear to induce a contradiction. Then $\nabla f(\cdot)$ **is a constant matrix** &

$$\nabla f \mathbf{A}\, \nabla\mathbf{G(h)} = \nabla\mathbf{G(h)} \quad \text{for all } h \in K.$$



Hence, $\partial_j G(h) \in \mathbf{Eigen_1}(\nabla f \mathbf{A})$ **, the eigenspace of $\nabla f \mathbf{A}$ corresponding to the eigenvalue 1.**

Since $\mathbf{dim}\, [\mathbf{Eigen_1}(\nabla f \mathbf{A})] \le \mathbf{dim\ Range}\ A \le \#\, rows$,

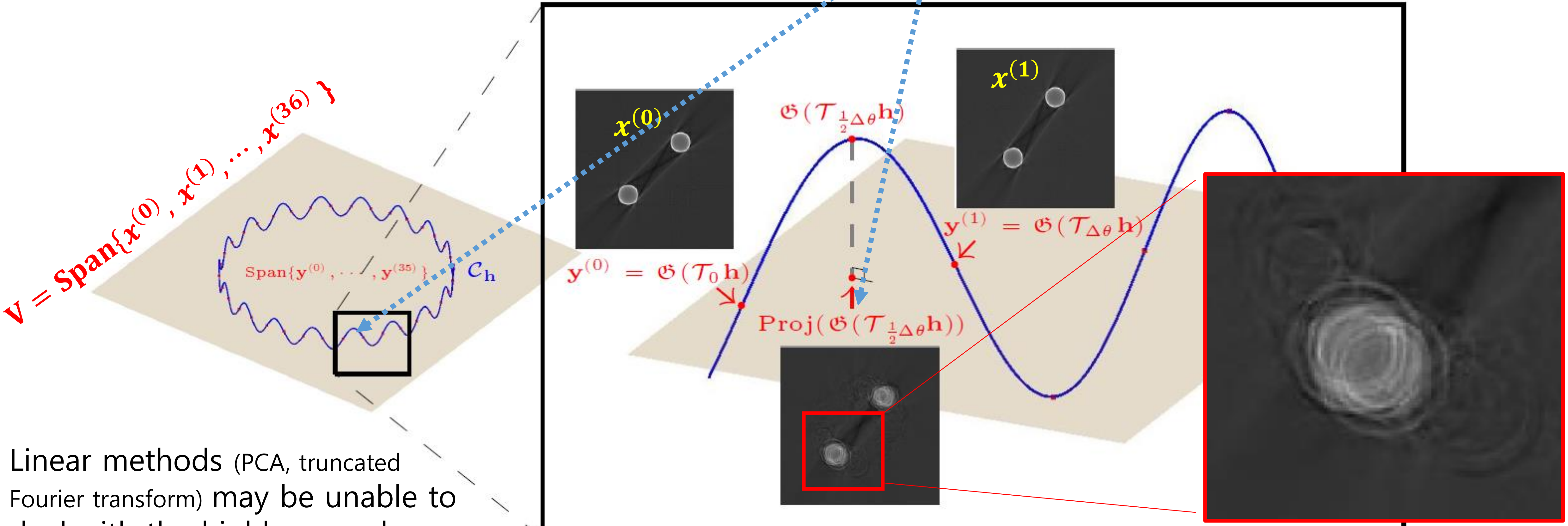
this **contradicts to the assumption** ($\dim(span\{\partial_j G(h): h \in K\}) > \#rows$).

Message: The degree of nonlinearity depends on # sampling of b & the degree of bending of the solution manifold M_{image} .

M is highly curved. Why?

Assume $M = \{ x^{(\theta)}: \theta \in [0,2\pi] \}$ where $x^{(\theta)}$ denotes the θ degree rotated image of $x^{(0)}$.

Let $\mathbf{V} = \mathbf{Span}\{x^{(0)}, x^{(\frac{\pi}{18})}, \cdots, x^{(2\pi)}\}$. Consider **the projection of $x^{(\frac{\pi}{36})}$ onto $\mathbf{V}$.**

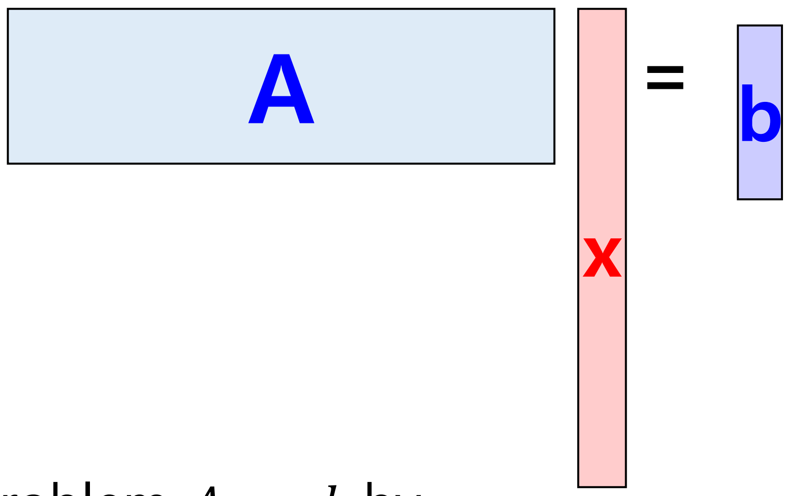


Linear methods (PCA, truncated Fourier transform) may be unable to deal with the highly curved solution manifold.

$\mathbf{P_V}\, x^{(0.5)} \approx \frac{1}{2}(x^{(0)} + x^{(1)})$ destroys the main characteristics of $x^{(0.5)}$.

The highly underdermined problems are **highly nonlinear!**

As the missing data increases, the nonlinearity of inpainting increases.



The degree of nonlinearity depends on # sampling of data b
& the degree of bending of the solution manifold.

- ✓ This is why it is difficult to solve highly underdetermined problem $Ax = b$ by conventional linear or piecewise-linear approaches.
- ✓ Deep learning techniques appear to handle nonlinear problems.

Deep Learning Approach

DL performance depends not only on the neural network architecture, but also on **the sampling strategy & the quality and quantity of training datasets**.

Learn f from **training data** $\{(x^{(n)}, b^{(n)}) : n = 1, \cdots, N\}$ by:

$$f = \underset{f \in \text{Neural Nets}}{\mathbf{argmin}} \sum_{\mathbf{n}} ||x^{(n)} - f(b^{(n)})||^2$$

DL performance depends on the **sampling strategy.**

The necessary condition for learning f is that

$$f(Ax) = x \quad \forall\, x \in \textbf{Image Manifold.}$$

M-RIP condition

Use training data $\{x^{(n)}: n = 1, \cdots, N\}$ to get prior knowledge.

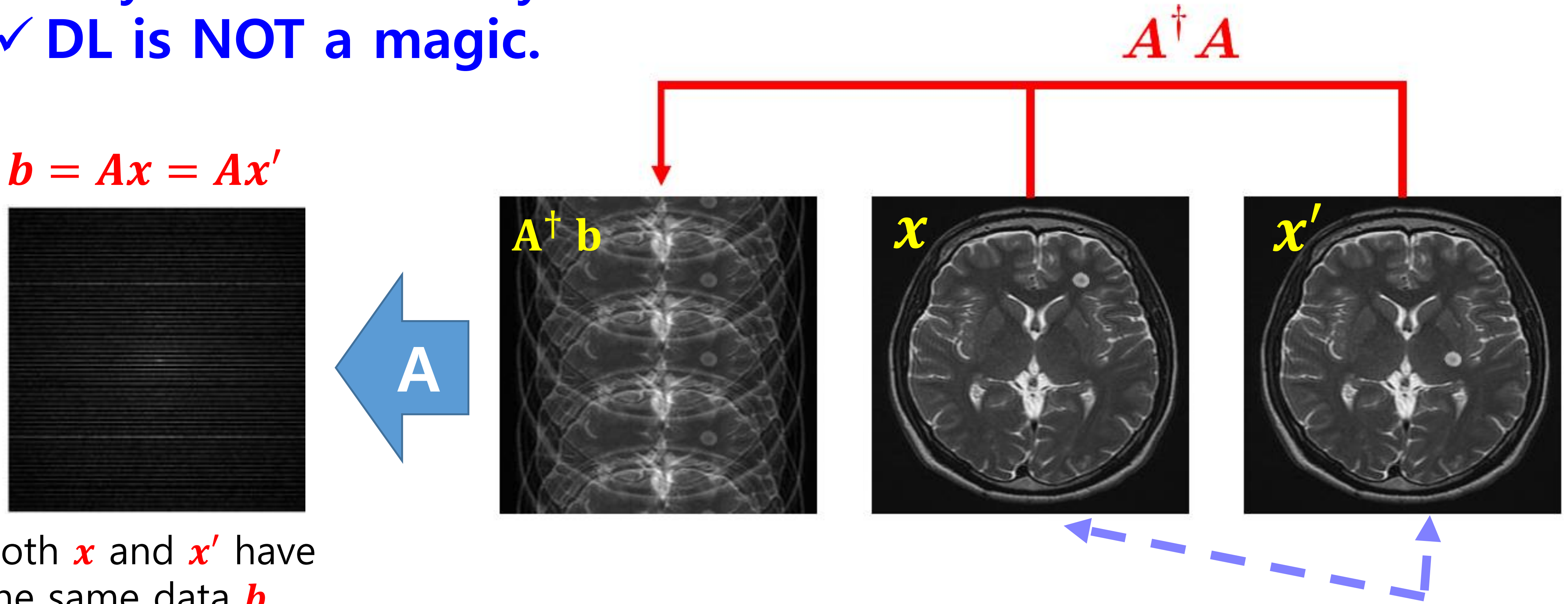
Undersmapled MRI

If we use **uniform subsampling** with factor 4,

it is difficult to learn f $s.t.$ $f(\mathbf{A}x) = \mathbf{x}$ $\forall\, \mathbf{x} \in$ **Image Manifold**

- ✓ **Why? It fails to satisfy M-RIP condition.**
- ✓ **DL is NOT a magic.**

$b = Ax = Ax'$

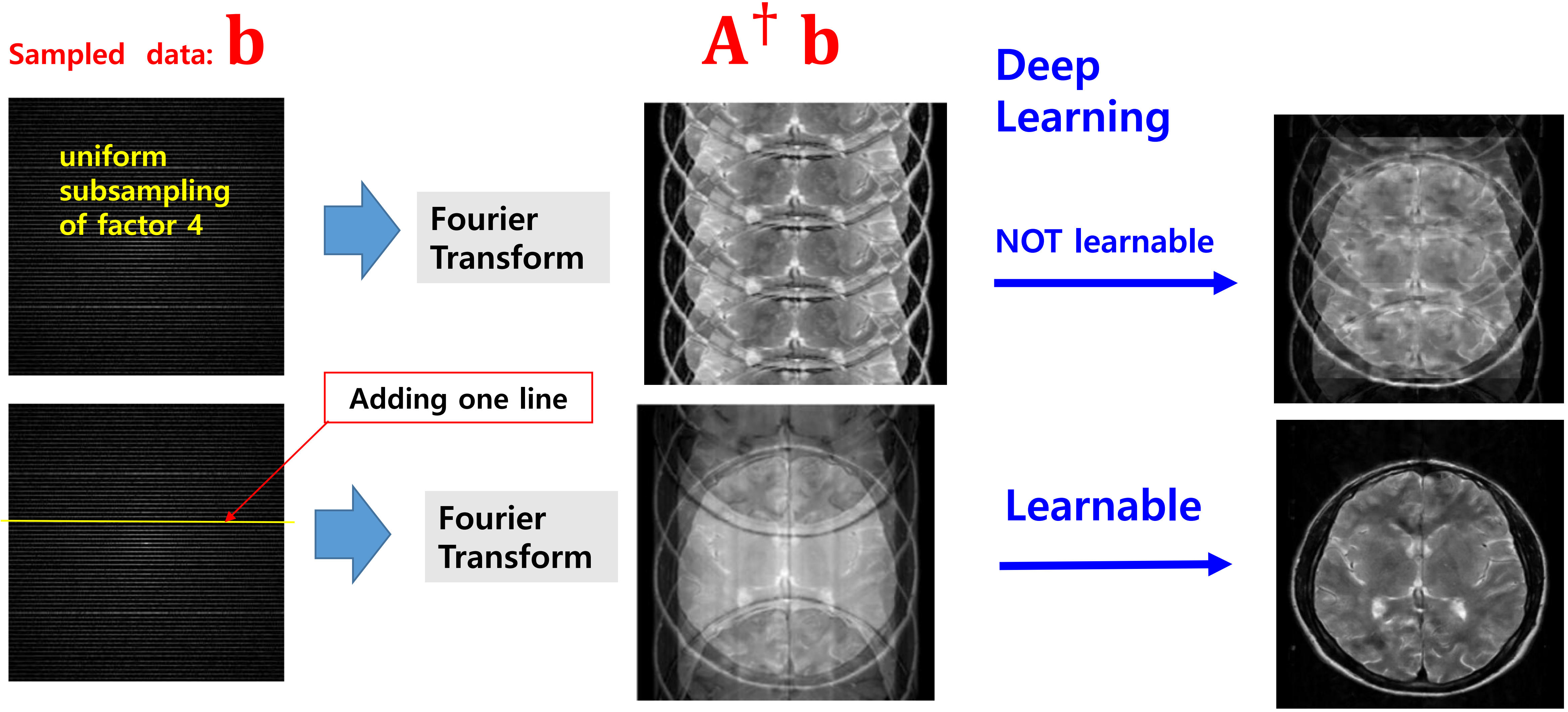


Both x and x' have the same data b.

Even deep learning is confused about which of the two (x or x') to restore.

A small change to the sampling strategy can lead to a dramatic improvement in learning.

✓ Adding a single phase encoding line can deal with position uncertainty to unfold the folded image.



DL performance depends on the **quality of training datasets.**

Deep learning-based Denoising

$$f = \underset{f \in \text{Neural Nets}}{\arg\min} \sum_{n} ||x^{(n)} - f(z^{(n)})||^2$$

z: Noisy input

x: Denoised output

Park, H. S., Baek, J., You, S. K., Choi, J. K., & Seo, J. K. (2019). Unpaired image denoising using a generative adversarial network in X-ray CT. *IEEE Access*, *7*, 110414-110425.

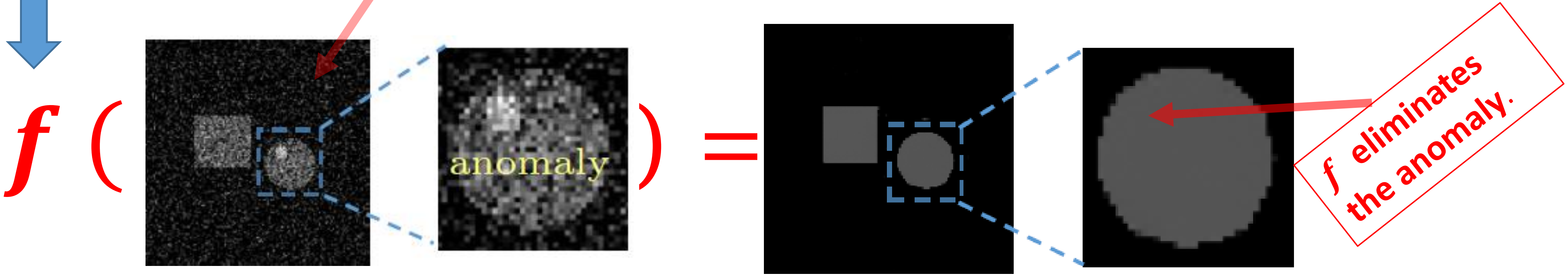
Training Sample 1

This training samples $\{\mathbf{x}^{(n)}: n = 1, \cdots, N\}$ consist of one disk and one rectangle with different sizes and positions

Consider image denoising problem:

$$f = \underset{f \in \text{Neural Nets}}{argmin} \sum_{n} ||x^{(n)} - f(z^{(n)})||^2$$

The test image contains a anomaly inside the disk, but **training data do not contain any anomaly inside disk and rectangle.**



This f can successfully remove noises on the test image but also remove small anomaly.

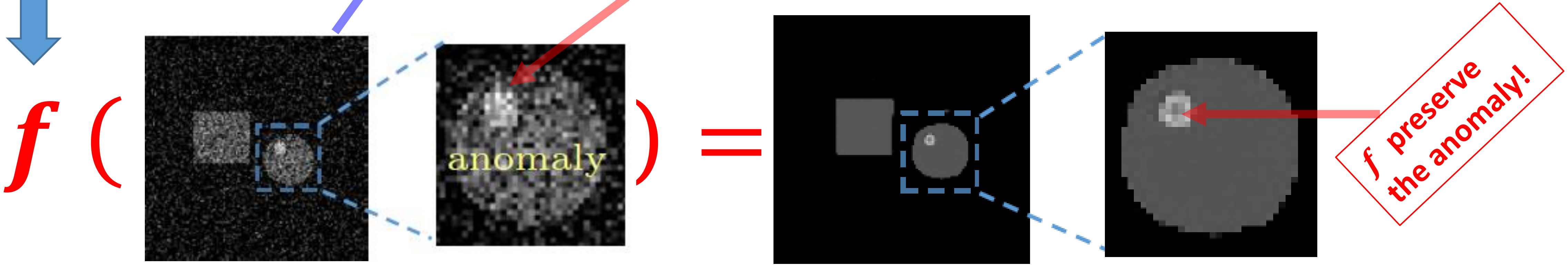
Training Sample 2:

This training samples $\{x^{(n)}: n = 1, \cdots, N\}$ contains small **anomaly inside rectangle.**

Learn Image denoising

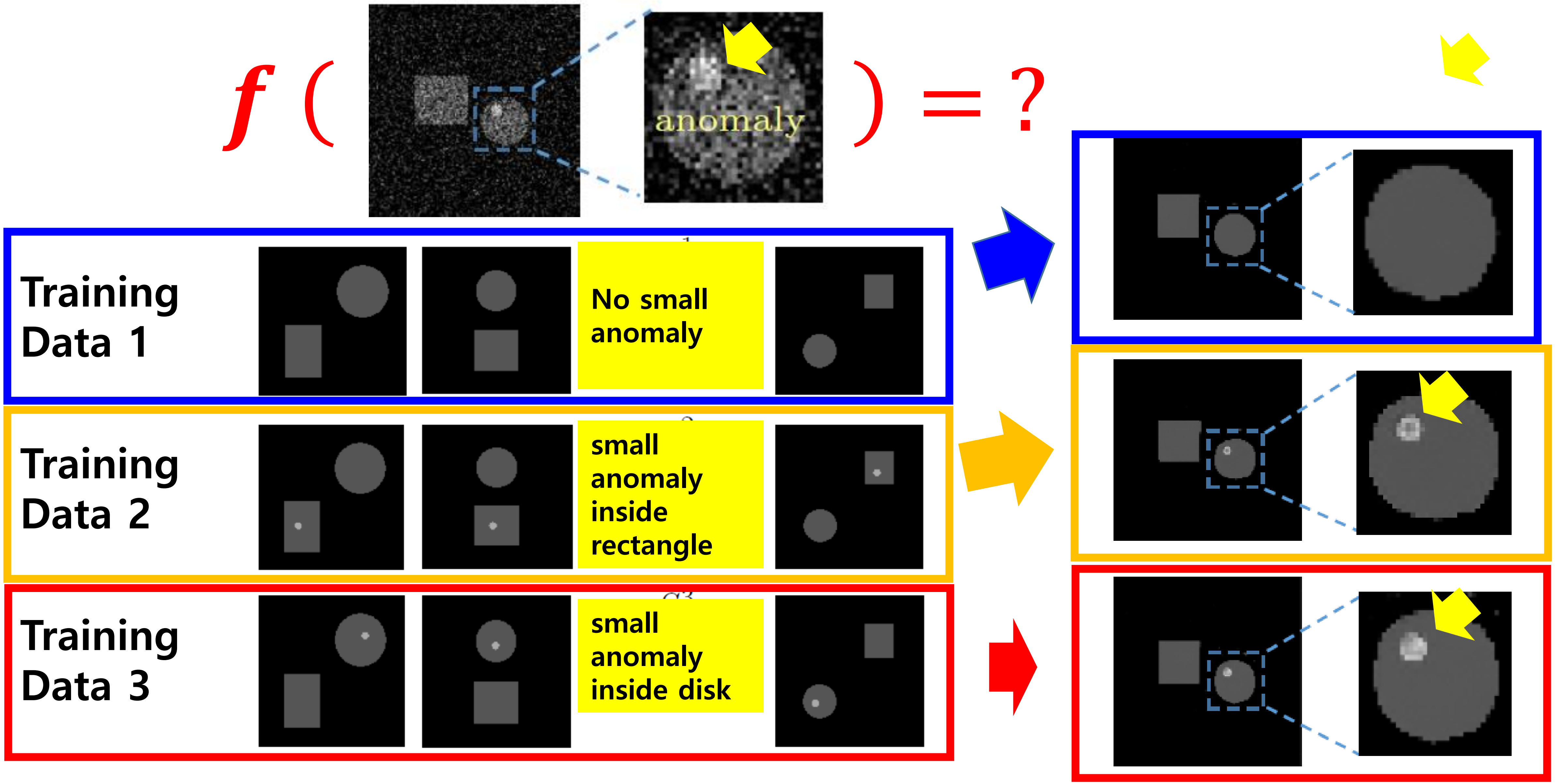
$$f = \underset{f \in \text{Neural Nets}}{\operatorname{argmin}} \sum_{n} ||x^{(n)} - f(z^{(n)})||^2$$

The test image contains anomaly inside the disk, but **training data do not contain the anomaly inside disk.**

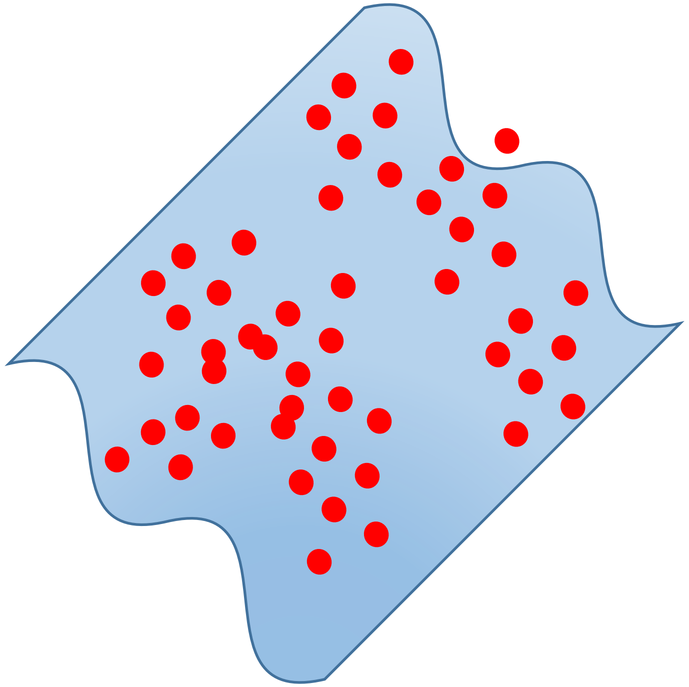


This f can successfully remove noises on the test image while preserving the small feature.

Impact of Training Data $\{(\mathrm{z}^{(n)}, \mathrm{x}^{(n)}): n = 1, \cdots, N\}$



An important open question is how **nonlinear dimensionality reduction** can be done .



A challenging problem is how to find a
low-dimensional representation from the training data.

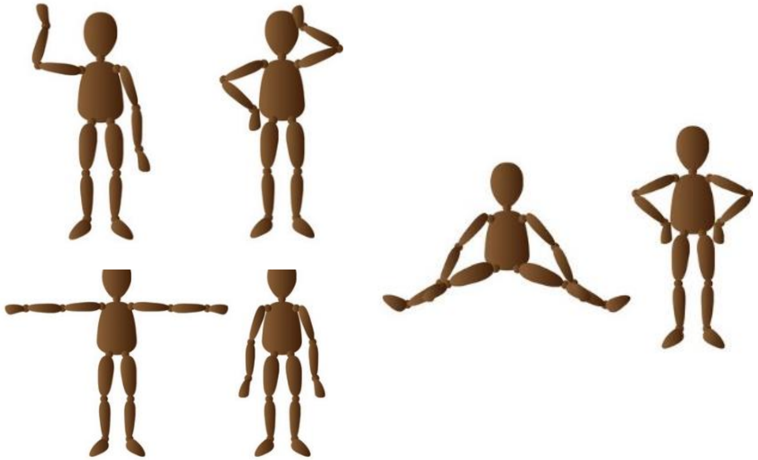
5 Latent variables

$$\mathbf{h} = (h_1, \cdots, h_5)$$

$$\Psi(\ \mathbf{h}\ ) =$$

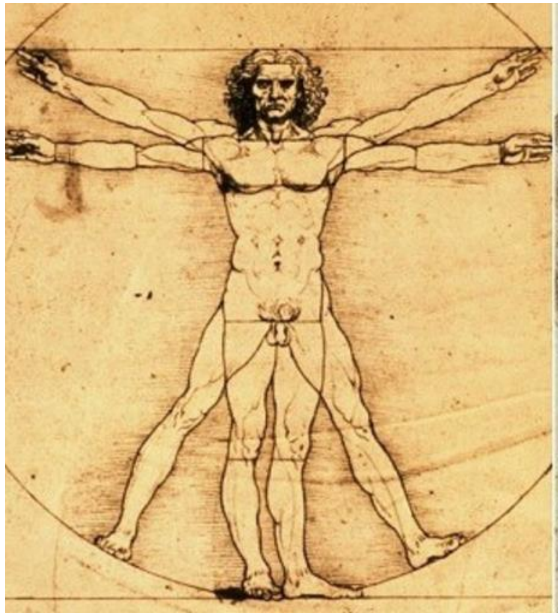
Generator /decoder

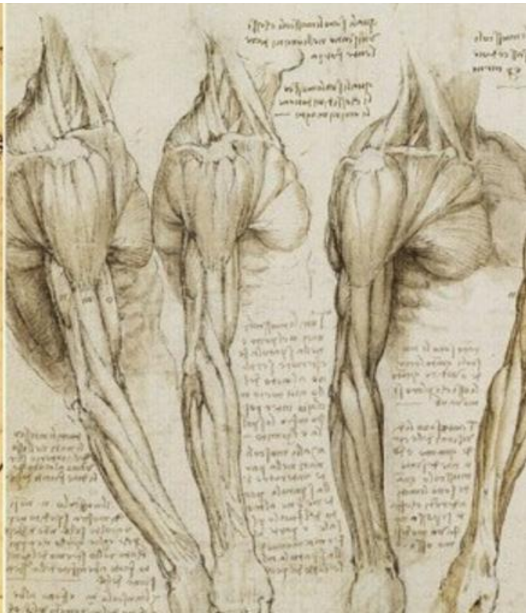
Latent variable



??

Can we find a **Disentangled expression** by extracting the underlying explanatory axis?

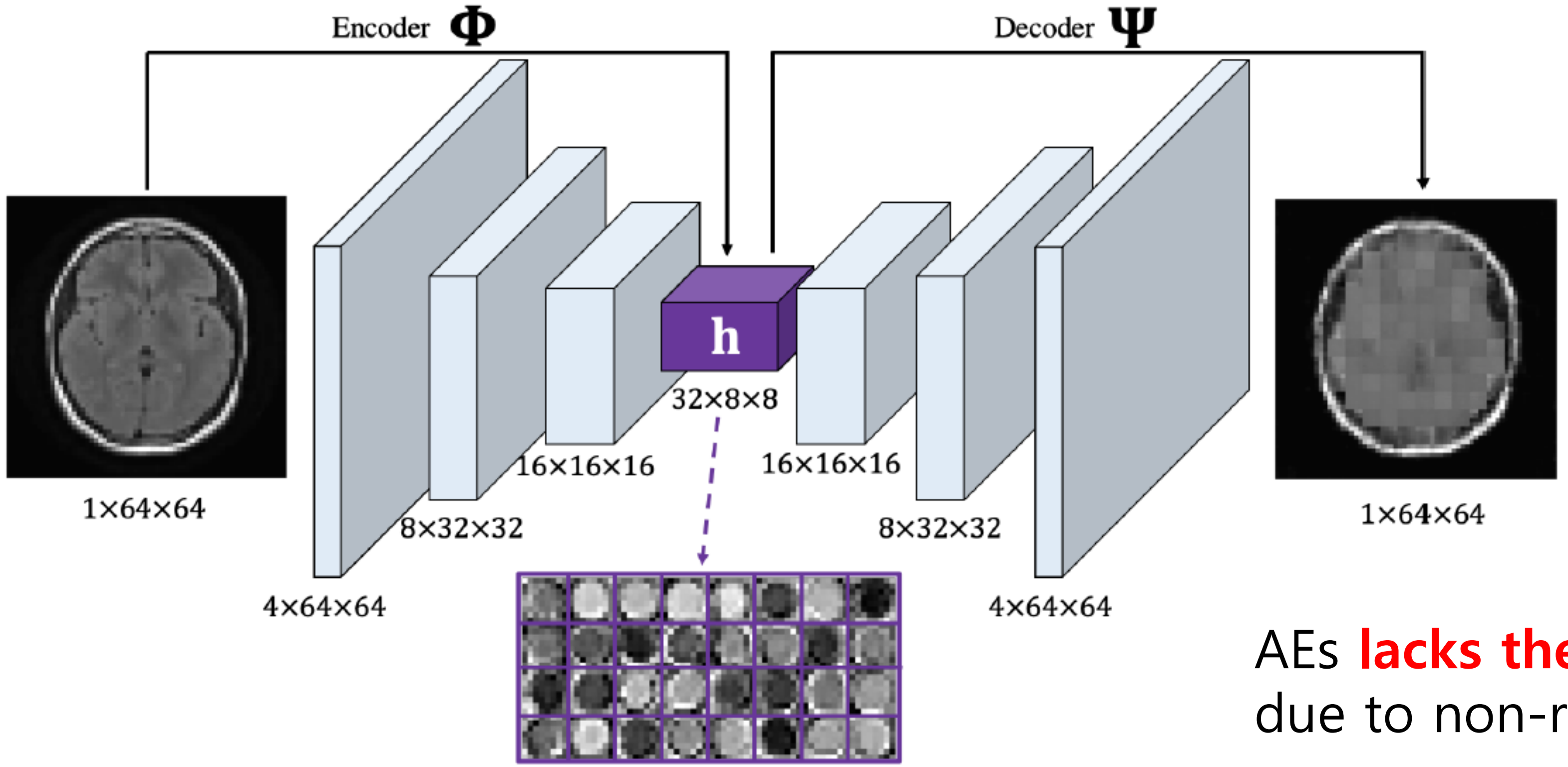




AutoEncoder (AE)

- ✓ It aims to generate the data manifold $M = \{\, \Psi(\mathrm{h}) : \mathrm{h} \in K \}$ by learning `hierarchical disentangled representation'.
- ✓ AE can be viewed as a non-linear extension of PCA.

$$\boldsymbol{Loss_{AE}\,(\Psi, \Phi) = E_{x \sim p_{data}(x)} \|\Psi \circ \Phi(x) - x\|^2}$$

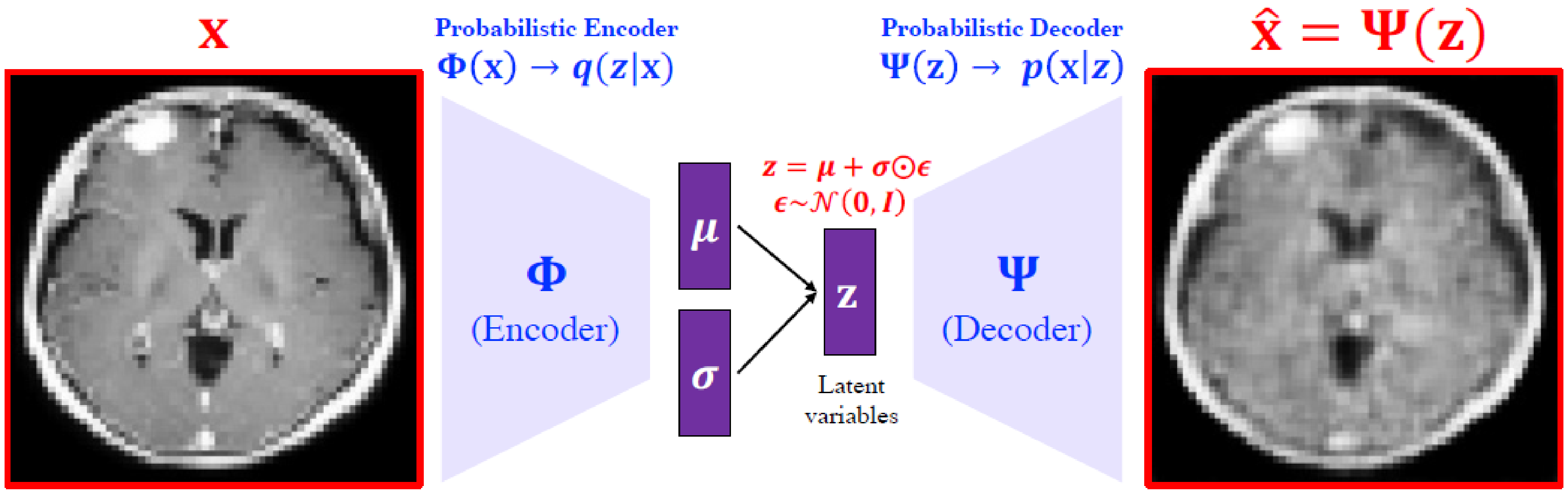


AEs do their best to learn (Φ, Ψ) with as little reconstruction loss as possible, rather than trying to organize the latent space well for generative purposes.

AEs **lacks the generalized capability** due to non-regularized latent space.

Variational AE (VAE)

- ✓ VAEs serve somewhat as a generative model.
- ✓ In VAE, the encoded latent variables are compressed and normalized to a normal distribution, to enable the generative process.



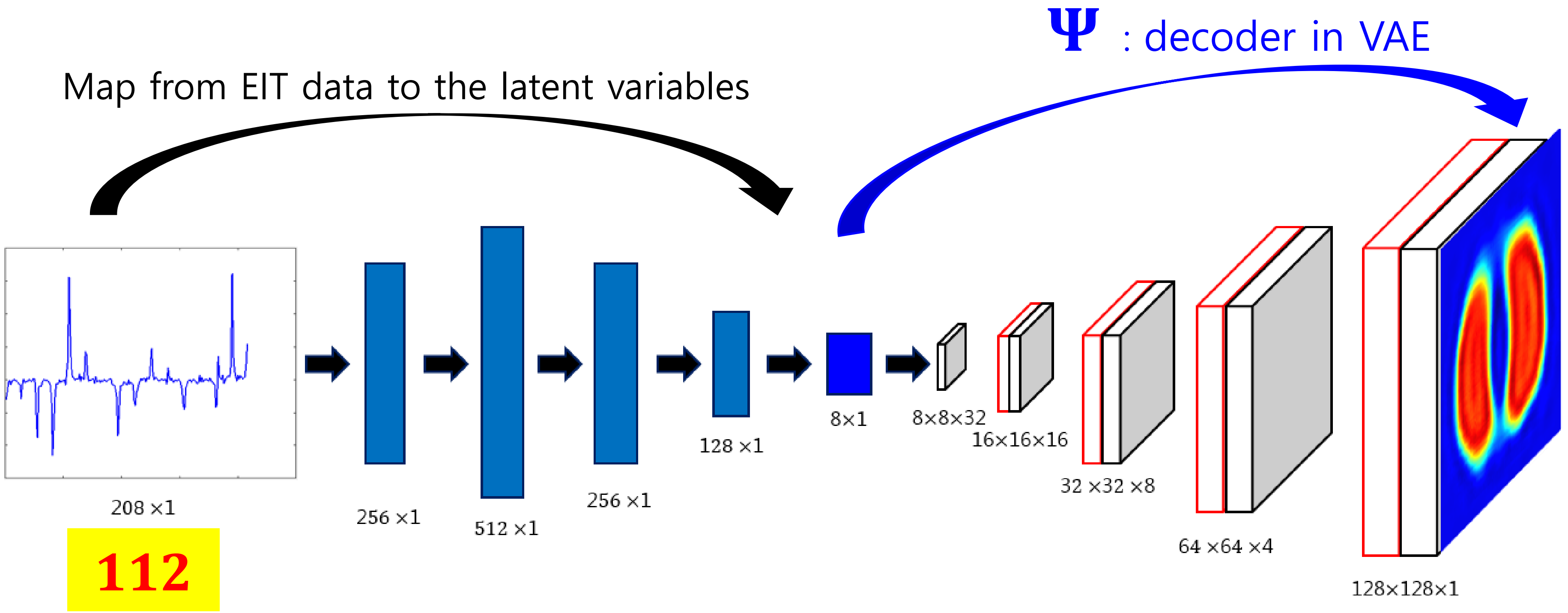
AE encodes the input as a single point, whereas VAE encodes the input as a distribution over the latent space.

$$Loss(\Phi, \Psi) = \sum_{n=1}^{N_{\text{data}}} \Big\{ \underbrace{D_{KL}\big[q(\mathbf{z}|\mathbf{x}^{(n)}) \parallel p(\mathbf{z})\big]}_{\text{Regularization}} - \underbrace{E_{q_\phi(\mathbf{z}|\mathbf{x}^{(n)})} \log p\big(\mathbf{x}^{(n)}|\mathbf{z}\big)}_{\text{Reconstruction loss}} \Big\}$$

Application of VAE

Electrical Impedance Tomography

- ✓ We use **VAE** to **disentangle** lung EIT images, so that lung EIT images are generated by $\boldsymbol{x} = \boldsymbol{\Psi}(\boldsymbol{h}),\ \boldsymbol{h} \in \boldsymbol{R}^8$.
- ✓ To solve $\mathrm{A}\boldsymbol{x} = \boldsymbol{b}$, instead of looking for images with **16384** pixels, we only need to find **8 latent variables**.



16384 pixels

The **ill-posed problem** *(16384 unknowns with 112 equations)* is turned into a **well-posed problem** *(8 unknowns with 112 equations).*

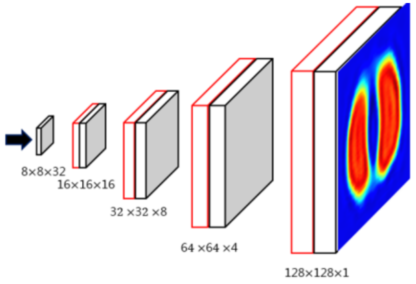
J.K. Seo, K.C. Kim, A. Jargal, K. Lee and B. Harrach, A learning-based method for solving ill-posed nonlinear inverse problems: a simulation : a simulation study of Lung EIT," SIAM Journal on Imaging Sciences (2019)

Using VAE, we represent lung impedance images (16384 pixels) by 8 latent variables.

$\Psi(\mathbf{h})$

$\mathbf{h} = (h_1, h_2, \cdots, h_8)$

Latent variables



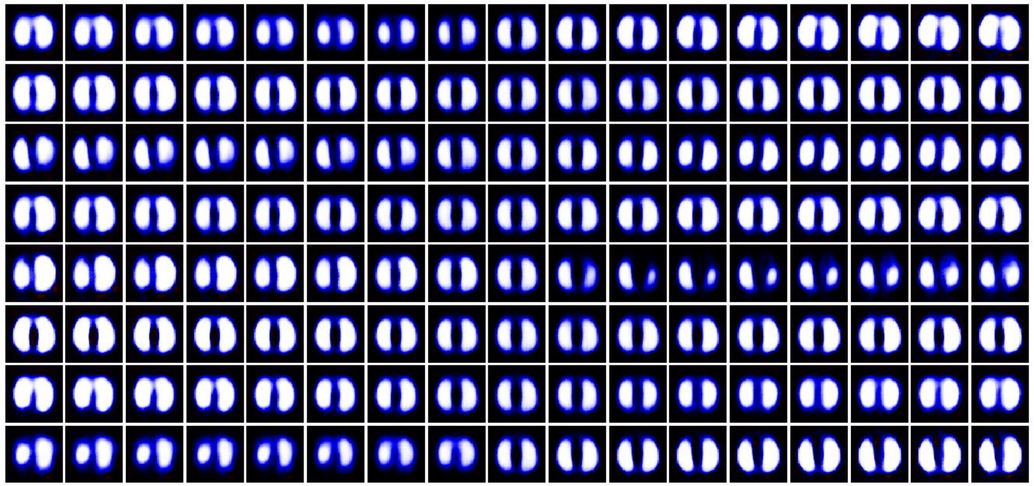
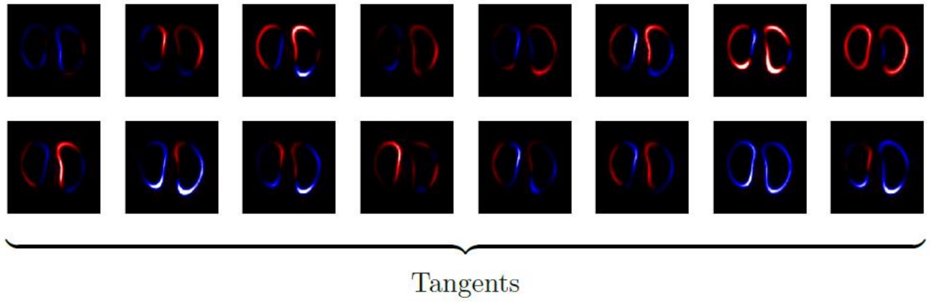
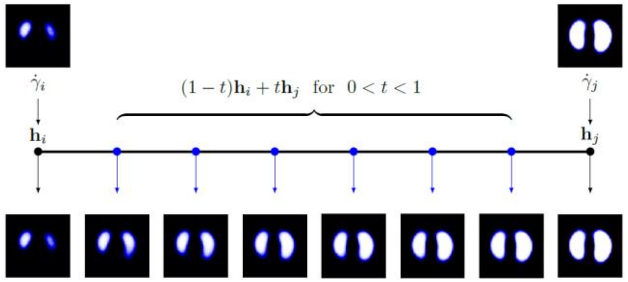


Image $\dot{\gamma} = \Psi(\mathbf{h})$

EIT lung ventiltional shape deformations

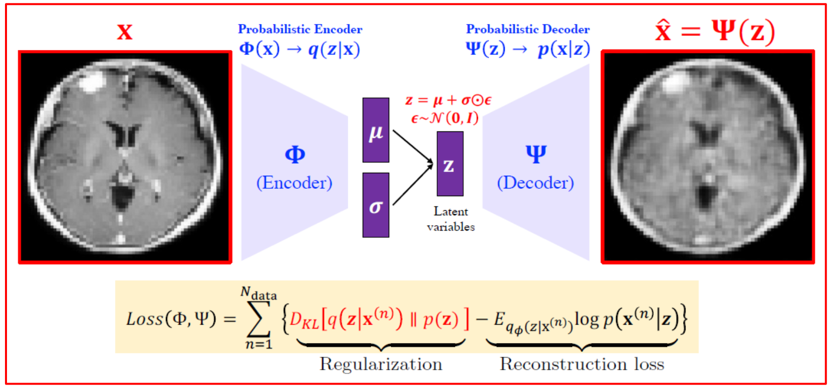


$\Psi(\mathbf{h_i})$ $\Psi(\mathbf{h}_j)$

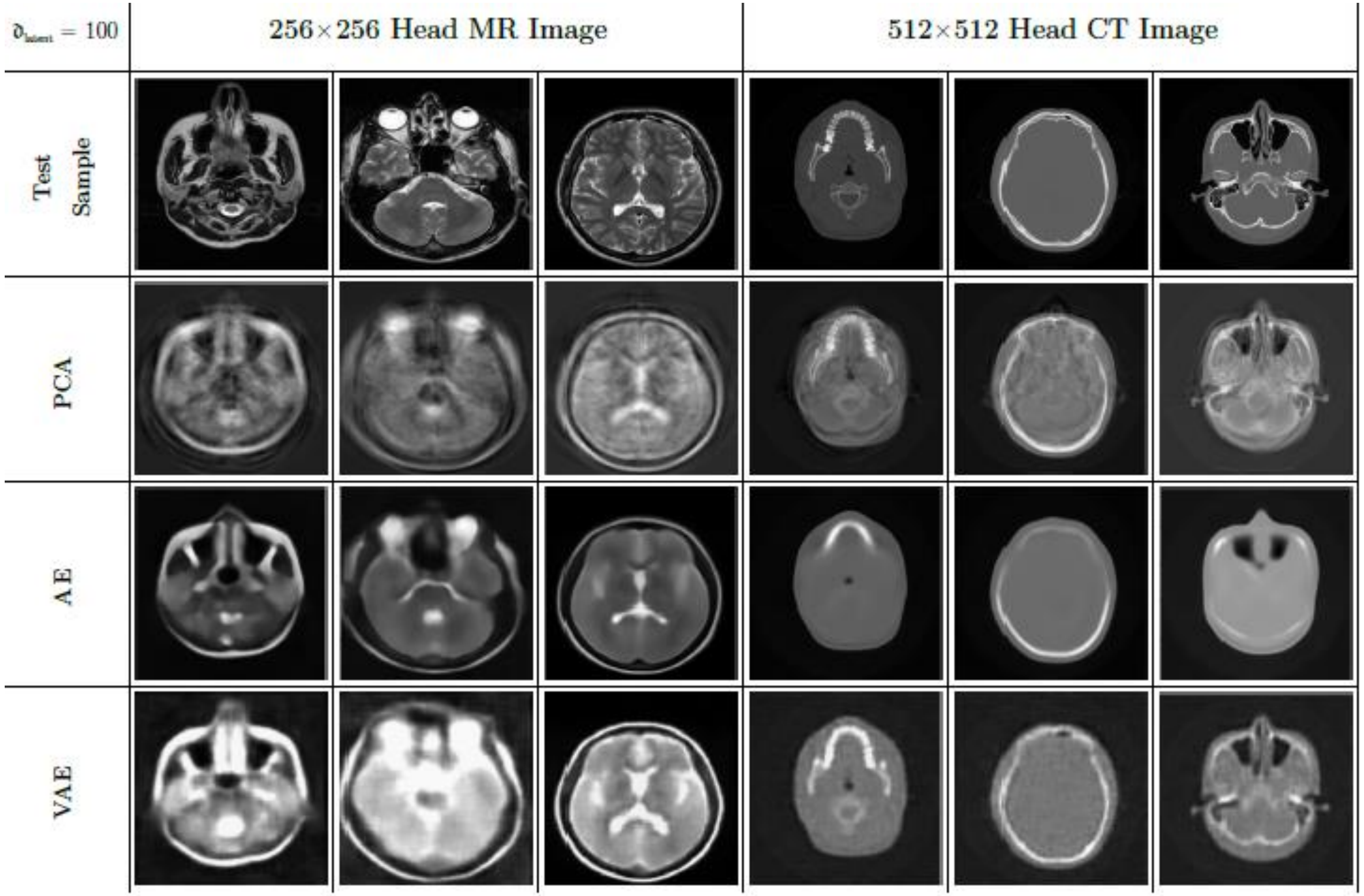


Unfortunately, for high-dimensional image,

VAE has limitations in that the image is blurred and small details are lost.

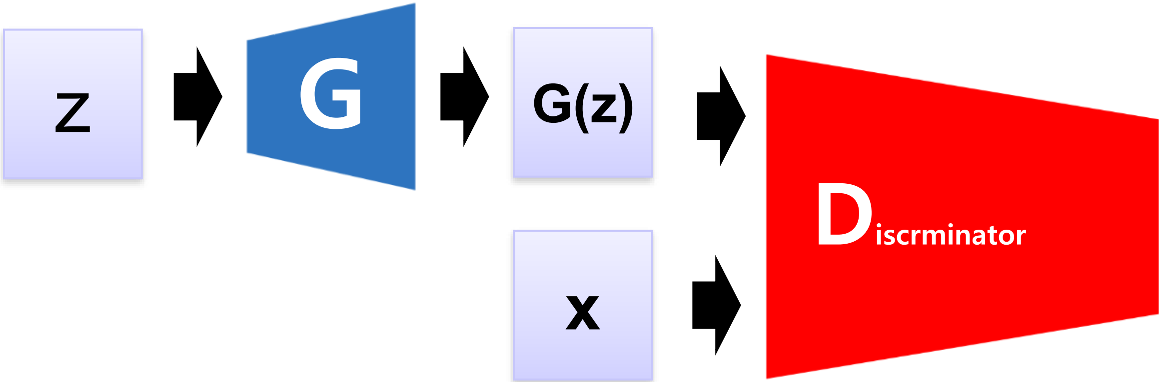


MRI and CT images
(high dimensional data: $512 \times 512 \times 400$ voxels)



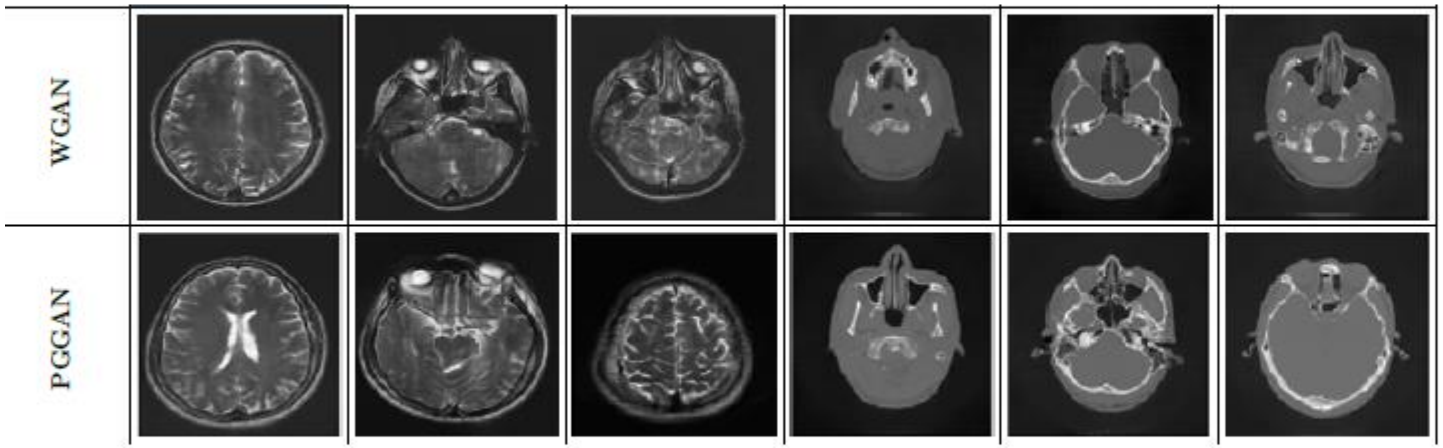
Challenging Issue: Low-dimensional representation of MRI and CT images

- ✓ **Generative Adversarial Networks(GANs)** have shown remarkable performance in generation of various realistic images.
- ✓ However, GANs have difficulties to learn **disentangled representation.**
- ✓ VAEs learns a bidirectional mapping(encoder and decoder), while GANs learn somewhat the unidirectional mapping (decoding) in high dimensional medical images



$$L_{GAN}(G,D) = E_{x\sim p_{data}(x)}[\log D(x)] + E_{z\sim p_z(z)}[\log(1 - D\circ G(z)]$$

The superiority of GANs in entertainment-related fields can be a disadvantage in the medical field, as it tends to reject the presence of small anomalies that are rarely seen, due to the strong punishment of the discriminator.



The Last Example: low-dose Cone Beam CT

Deep Learning-based 3D segmentation.

Input : z

Output : $\mathbf{x}$

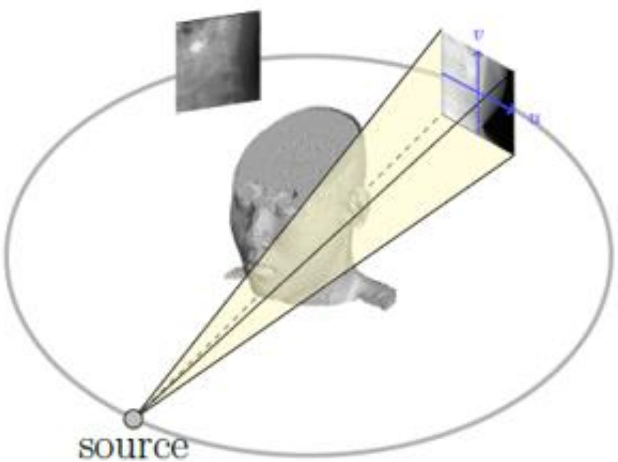
$f($  $) =$ 

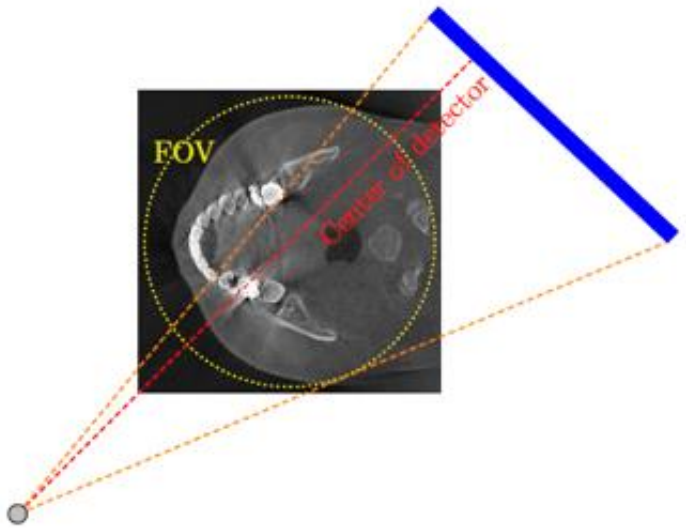
3D CBCT image

Target error : 0.2 mm

Low Dose Dental CBCT

CBCT images are affected by "offset detector, FOV truncation, low X-ray dose" and metal induced beam hardening, resulting in significant image noise and artifacts.



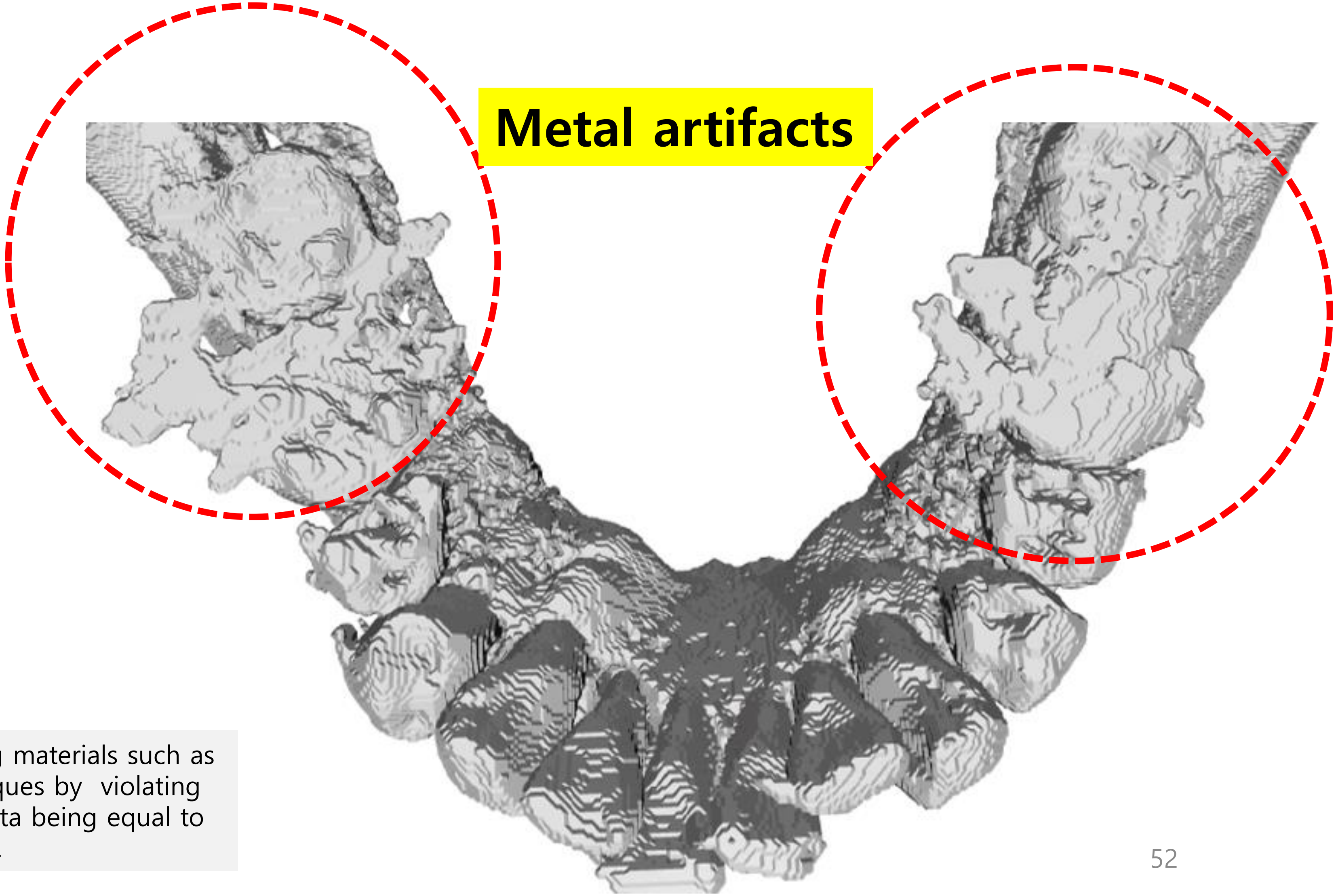


- **Dental CBCT**
  - Circular conebeam
  - Scan time: 8-24sec
  - Resolution < 0.2mm
  - FOV truncation
  - Offset detector
  - Price < $ 0.1 billion
  - Low X-ray dose

- **MDCT**
  - Helical conebeam
  - Scan time < 1sec
  - Resolution < 0.3mm
  - No FOV truncation
  - No Offset detector
  - Price > $ 1 billion
  - High X-ray dose

Even with modern deep learning techniques, it is difficult to perform accurate tooth segmentation on the metal-artifacts contaminated image.

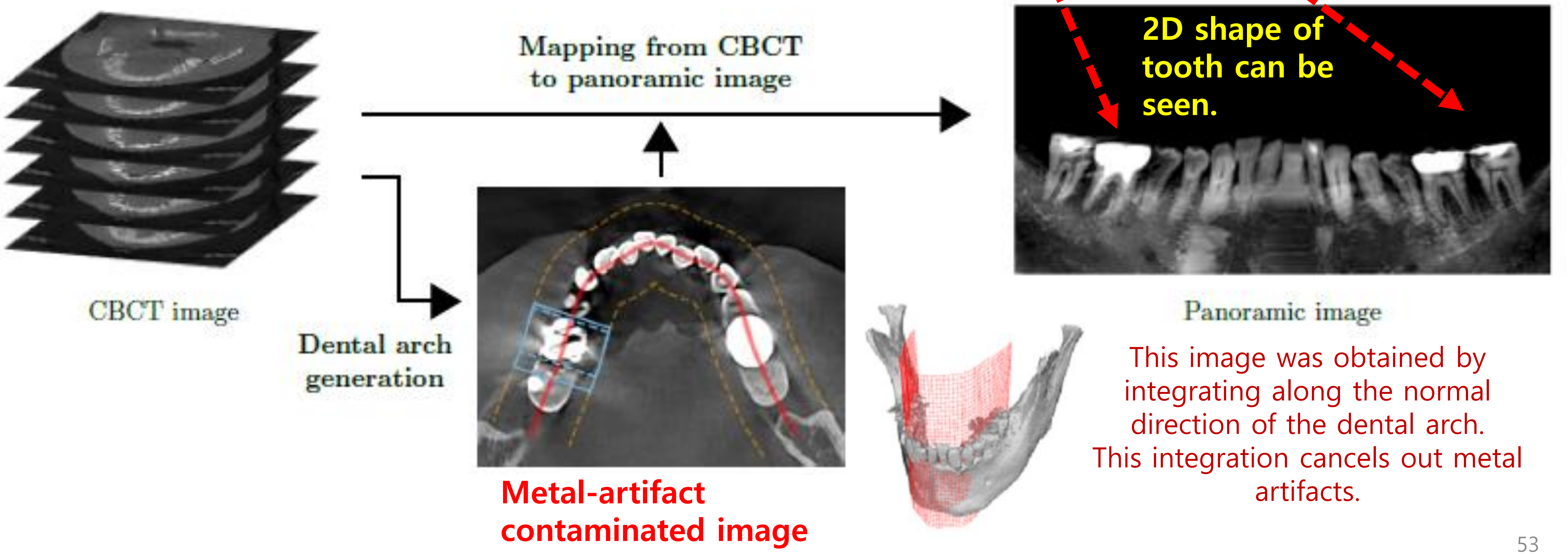
In dental CBCT, metal artifacts are common.



- **Modeling error:** The presence of highly attenuating materials such as metallic objects complicates reconstruction techniques by violating the forward model assumption of the sinogram data being equal to the Radon transform of an image.

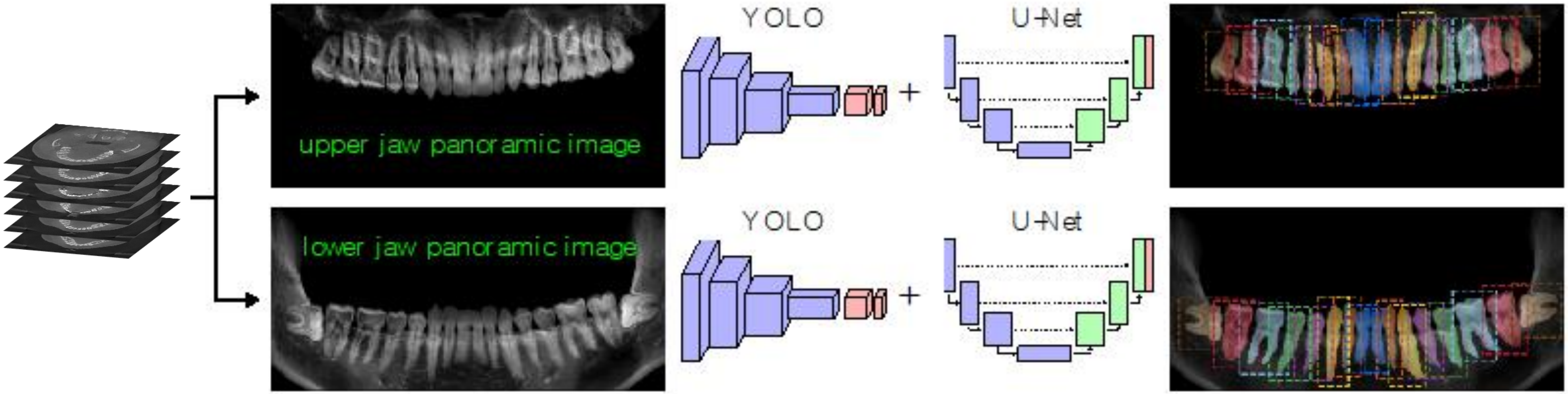
The key idea to overcome the difficulty is to get a **good prior knowledge** that is obtained by generating a clean **panoramic image** from the noisy CBCT image.

The panoramic image is not much affected by metal-related artifacts.



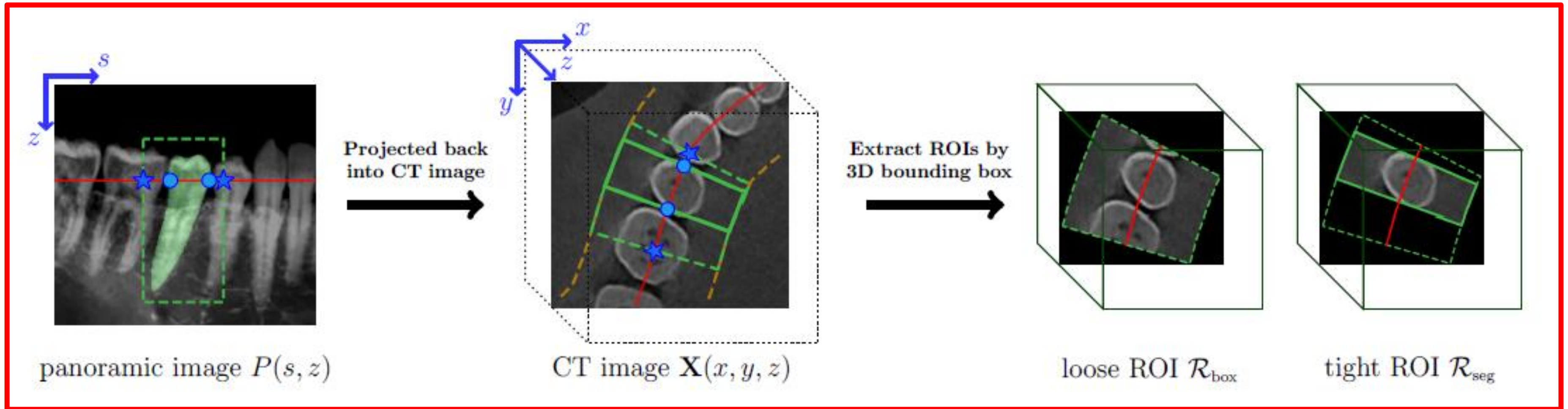
Get **prior knowledge of 3D teeth** from 2D tooth segmentation obtained from panoramic images.

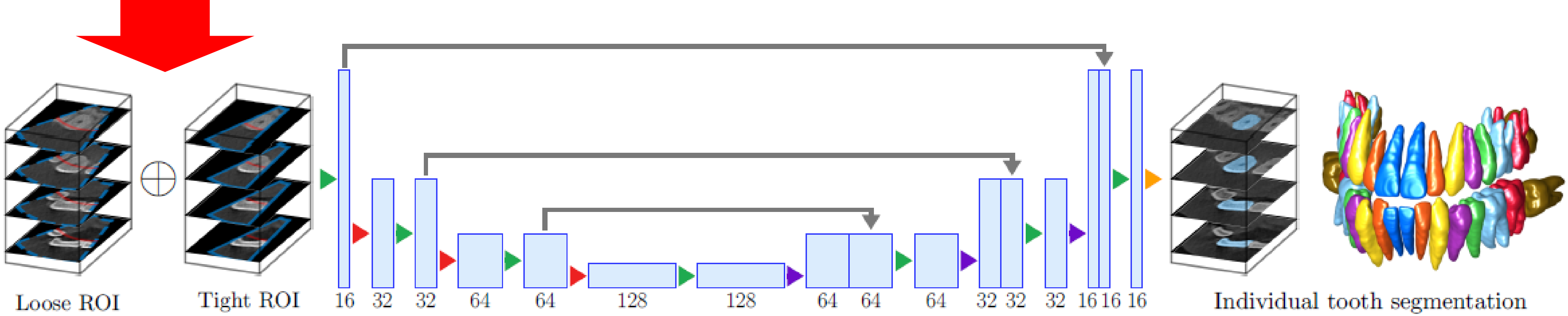
This 2D segmentation is used to find accurate 3D tooth ROIs and identify individual teeth.



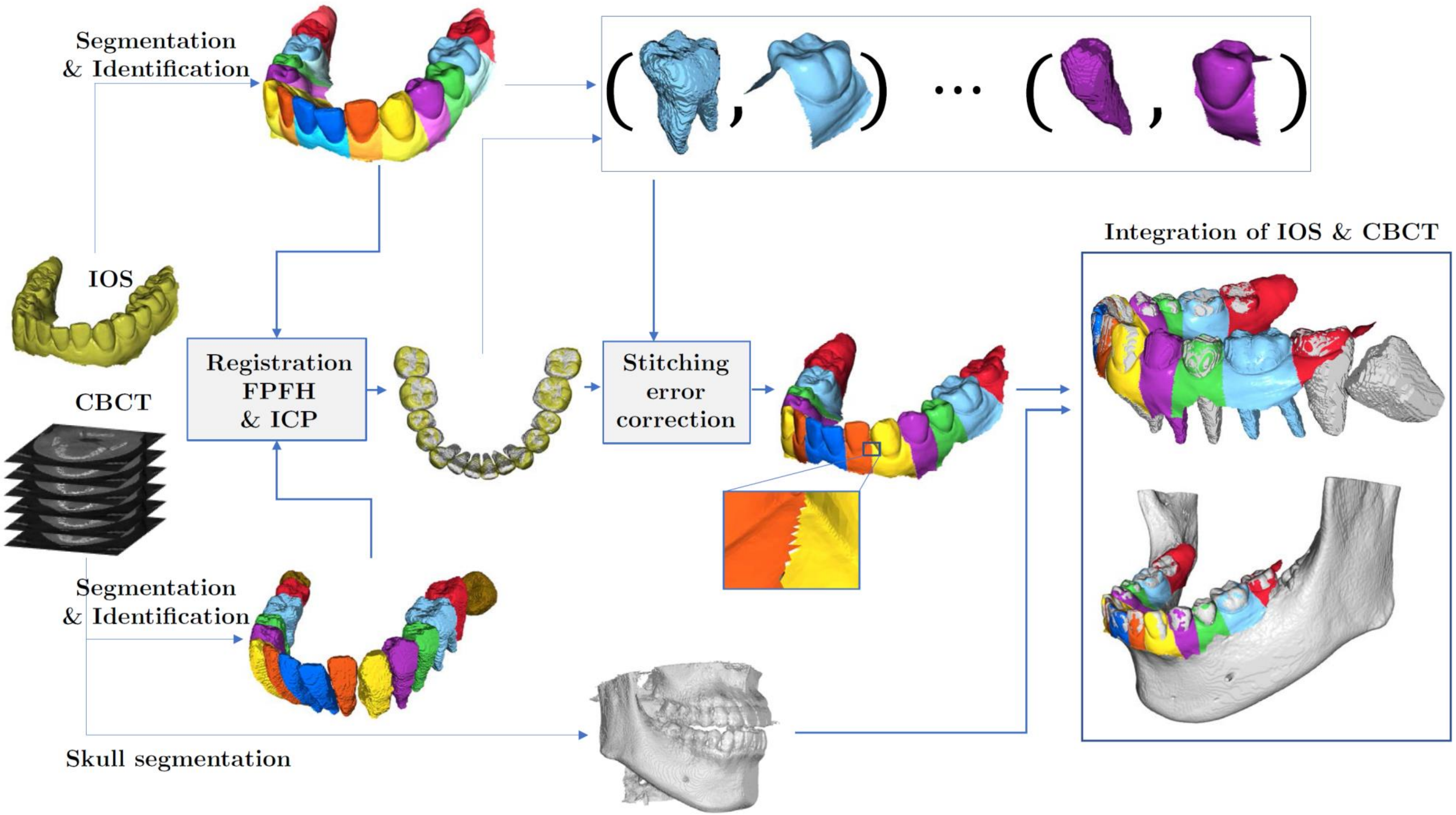
2D tooth segmentation provides a **deep learning friendly environment** for 3D tooth segmentation.

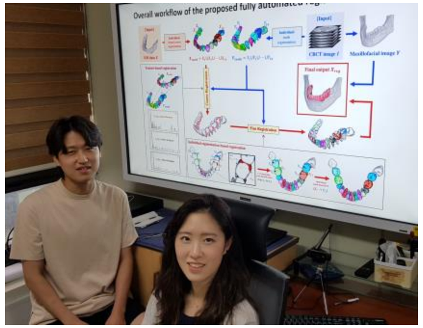
Tae Jun Jang, Kang Cheol Kim, Hyun Cheol Cho, and Jin Keun Seo, A fully automated method for 3D individual tooth identification and segmentation in dental CBCT, IEEE Transactions on Pattern Analysis and Machine Intelligence (2021)





Comment 1: The aforementioned tooth segmentation enables the fusion of CBCT and intraoral scans, eliminating the cumbersome procedure of conventional impressions.





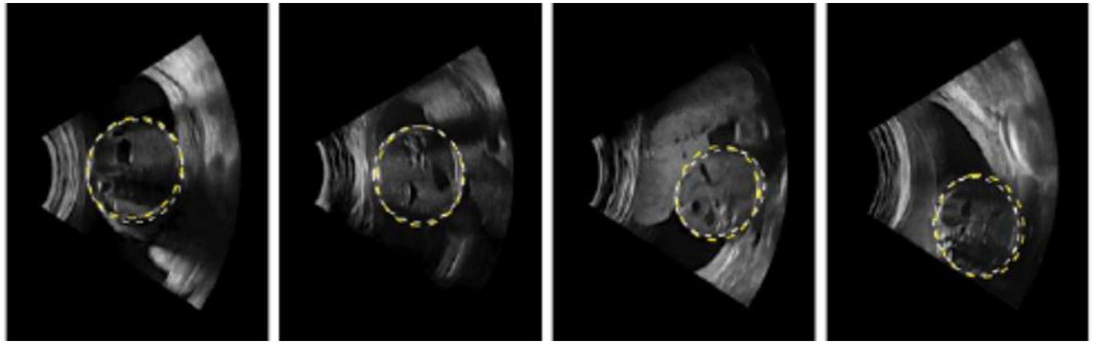
Error< 0.2mm

AI-based Digital Dentistry to improve workflow

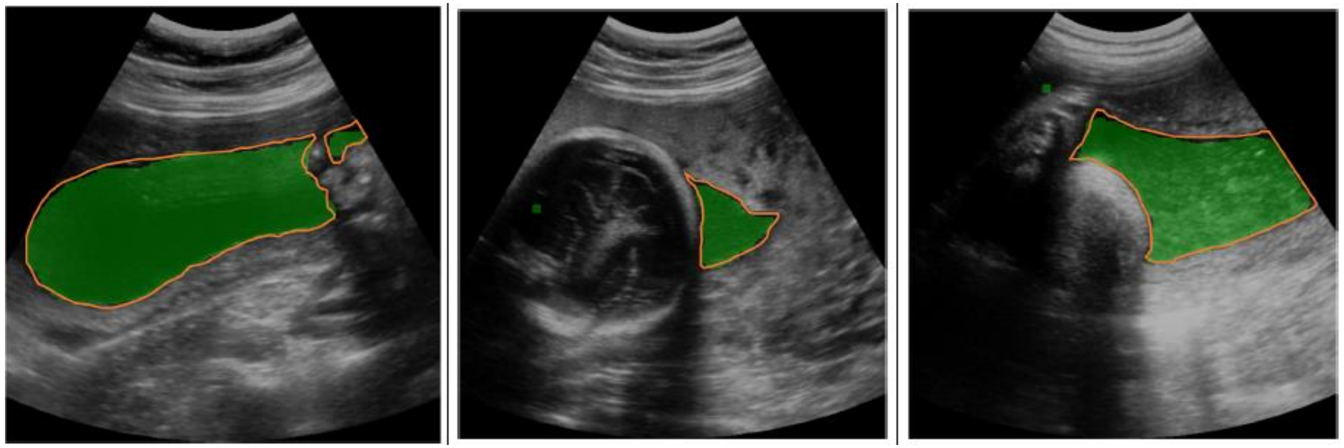
DLs appear to overcome limitations of existing mathematical methods in handling various complex problems in segmentation.

Comment 2: Paradigm shift in Segmentation

- ✓ Automated fetal biometric measurements for fetal ultrasound have been very difficult tasks for over 30 years, but recently some of these problems have been solved with DLs!



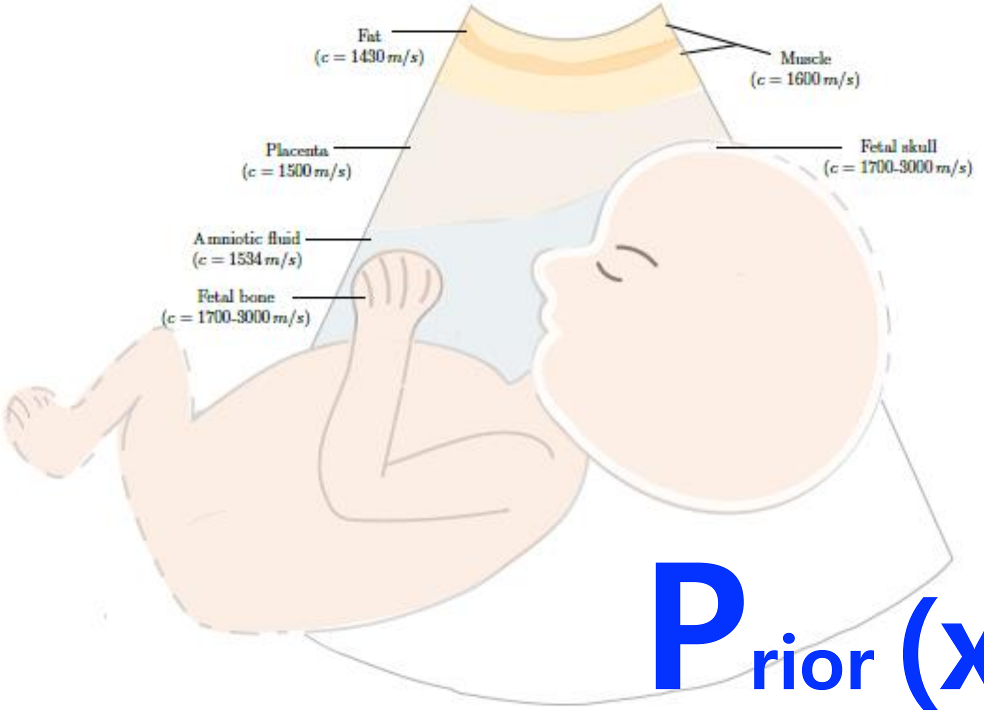
Kim, Bukweon; Kim, Kang Cheol; Park, Yejin; Kwon, Ja-Young; Jang, Jaeseong; Seo, Jin Keun*, Machine-learning-based Automatic Identification of Fetal Abdominal Circumference from Ultrasound Images, Physiological Measurement (2018)



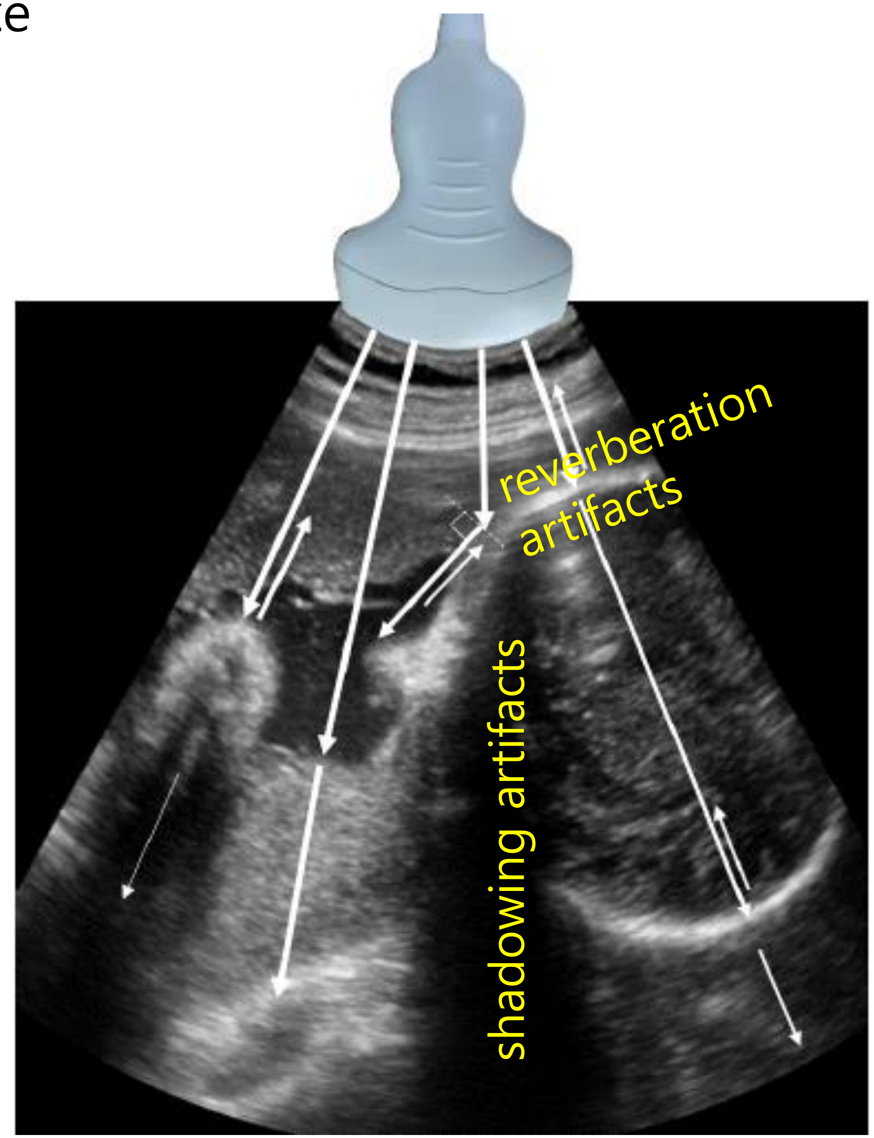
Hyun Cheol Cho, Siyu Sun, Chang Min Hyun, Ja-Young Kwon, Bukweon Kim, Yejin Park, Jin Keun Seo, Automated ultrasound assessment of amniotic fluid index using deep learning, Medical Image Analysis (2021)

Why did the DL methods achieve remarkable performance in image segmentation tasks?

- ✓ DLs can learn prior anatomical knowledge to analyze even heavily distorted images.
- ✓ DLs capture the spatial relationships between pixels to figure out local and global interconnections.



Prior (x)



Thank you!

- ✓ Medical imaging is in fact experiencing a paradigm shift due to a marked and rapid advance in deep learning techniques.
- ✓ However, there is a tremendous lack of a rigorous mathematical foundation which would allow us to understand the reasons why deep learning methods perform that well.
- ✓ Despite the lack of rigorous analysis in deep learning, recent rapid advances indicate that DL methodologies will see continued improvements in performance as training data and experience accumulate over time.

SAMSUNG
SCIENCE & TECHNOLOGY FOUNDATION