Deep Learning-based Solvability of Underdetermined Inverse problems in medicines

Lunar New Year 2020

Learn $f(data) =$ useful output

Jan 2020@ IPAM

Jin Keun Seo
with Chang Min Hyun & Seong Hyeon Baek
Yonsei Univ., Korea

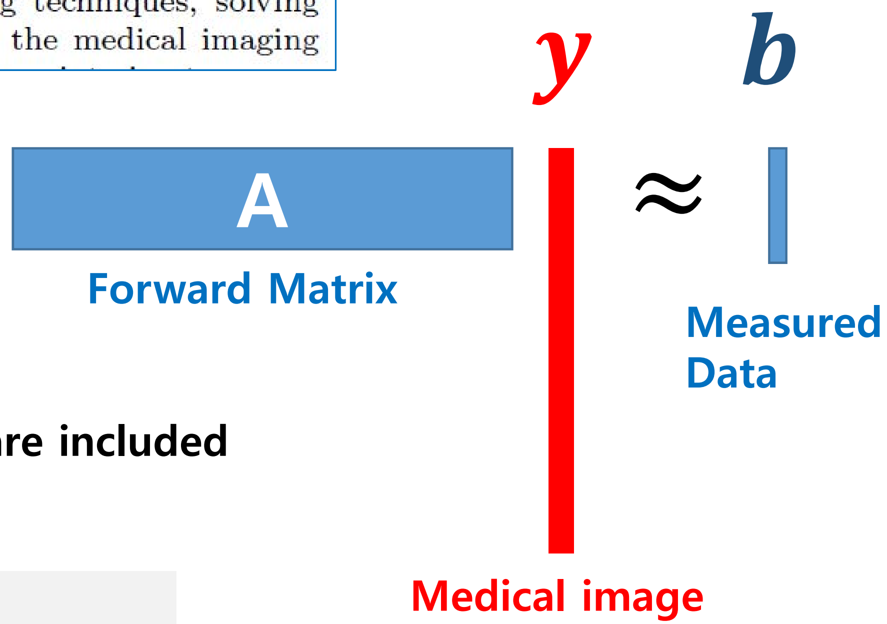
This talk is based on joint work with my PhD students.

DEEP LEARNING-BASED SOLVABILITY OF UNDERDETERMINED INVERSE PROBLEMS IN MEDICAL IMAGING*

CHANG MIN HYUN[†], SEONG HYEON BAEK[†], MINGYU LEE[†], SUNG MIN LEE[†], AND JIN KEUN SEO[†‡]

Abstract. Recently, with the significant developments in deep learning techniques, solving underdetermined inverse problems has become one of the major concerns in the medical imaging

Solve



In this talk, many of my personal opinions (not rigorous) are included to give an exaggerated emphasis on deep learning.

Learn $f(b) =$ *useful* ***output***

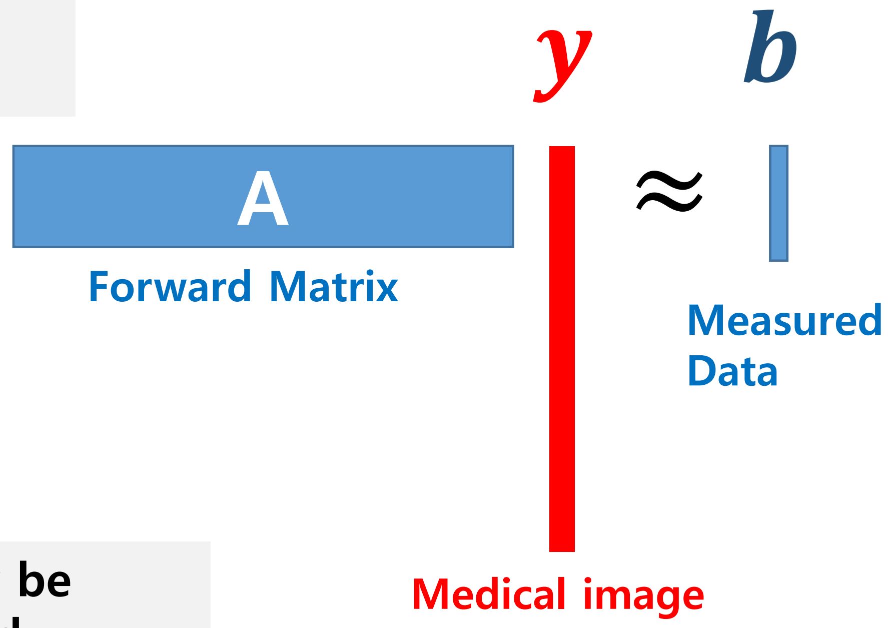
ill-posed inverse problems

Hadamard's well-posedness (excluding existence)

$Ay = b$ is well-posed if the following two conditions hold:

1) for each b, $Ay = b$ has a unique solution;
2) the solution is stable under perturbation of b.

Solve



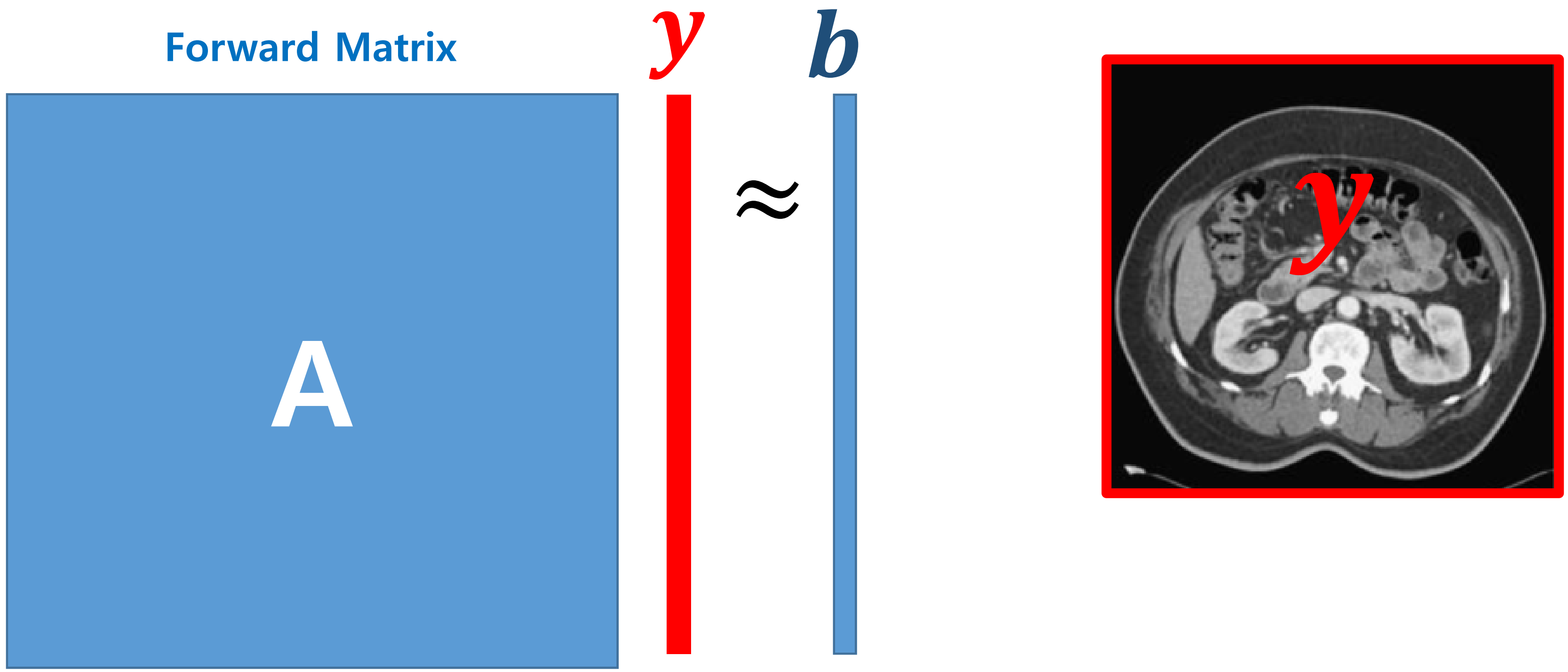
Learn $f(b) = y$

My personal opinion

- Whether or not a problem is well-posed may be dependent on how the solution is expressed.
- Many problems are ill-posed because we are overly ambitious or lacking in expressiveness.

Conventional CT and MRI data collections are designed
for the corresponding forward matrix A
to be well-expressed & to be reasonably complete.

# of equations (data) ≈ # of unknowns (pixels of image)



The classical principle that make problems well-posed is:
# of equations (number of samples) $\approx$ **# of unknowns** (number of pixels of image).

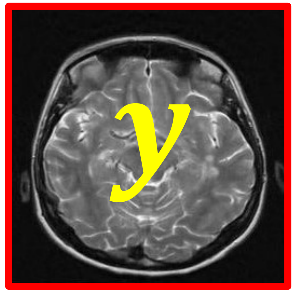
Tomography with Nyquist Sampling

- **MRI** measures approximately an image's Fourier transform. Nyquist sampling is required for the analytic reconstruction.

$$\# \textit{ pixels in image} \approx \# \text{ samplings in } \textit{k-space}$$

$A^{-1}($

$) =$

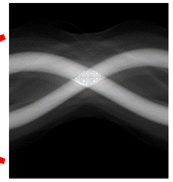


Inverse Fourier Transform

- **CT** measures approximately an image's Radon transform. According to Nyquist sampling & Fourier slice theorem,

$$\sqrt{\# \text{ pixels in image}} \approx \# \text{ projection angles}$$

$A^{-1}($



$) =$

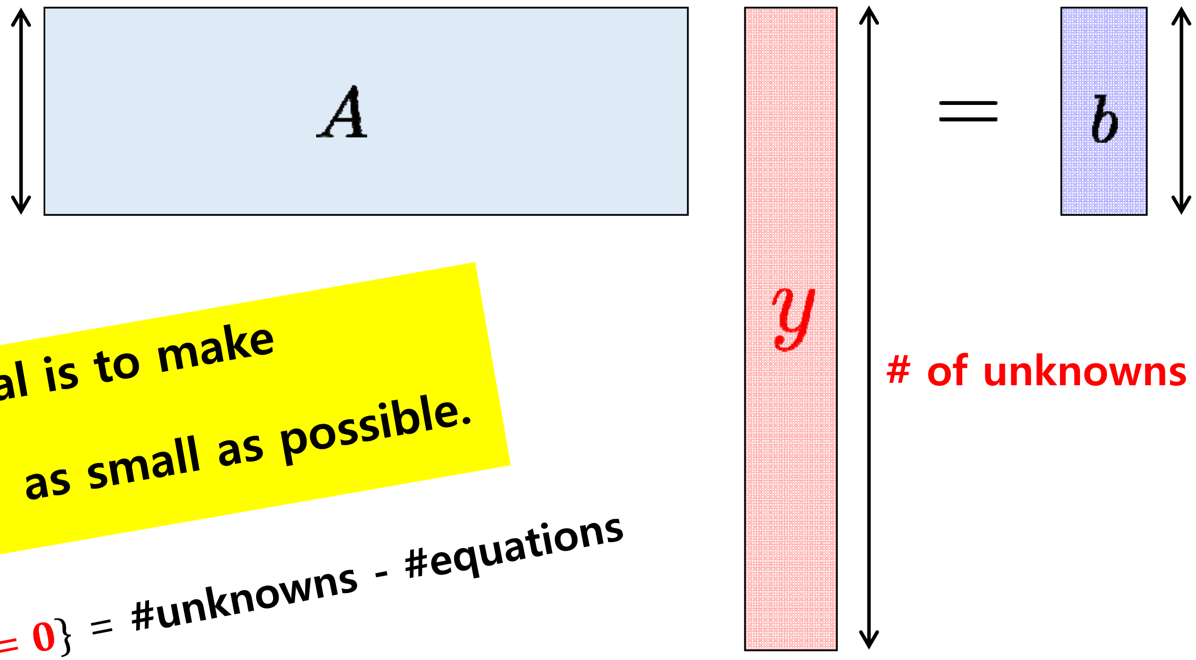
Data

Filtered Backprojection

Why do we pay attention to underdetermined problems in CT & MRI ?

(fewer equations than unknowns)

It is because of the great needs to reduce **radiation dose in CT** & **data acquisition time in MRI**.



# of equations

# of unknowns (# of pixels)

Reduced # of measurement data

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

$dim\{z\colon Az = 0\}$ = #unknowns - #equations

Solving $Ay = b$ is to find

A y = b

the reconstruction map $f: b \rightarrow y = A_{full}^{-1} b_{full}$.

Is it possible to solve it ?

- b_{full} denotes the "fully sampled" data (e.g, sinogram in CT and k-space data in MRI).
- $b = S_{ub} b_{full}$ where S_{ub} denotes a subsampling operator.

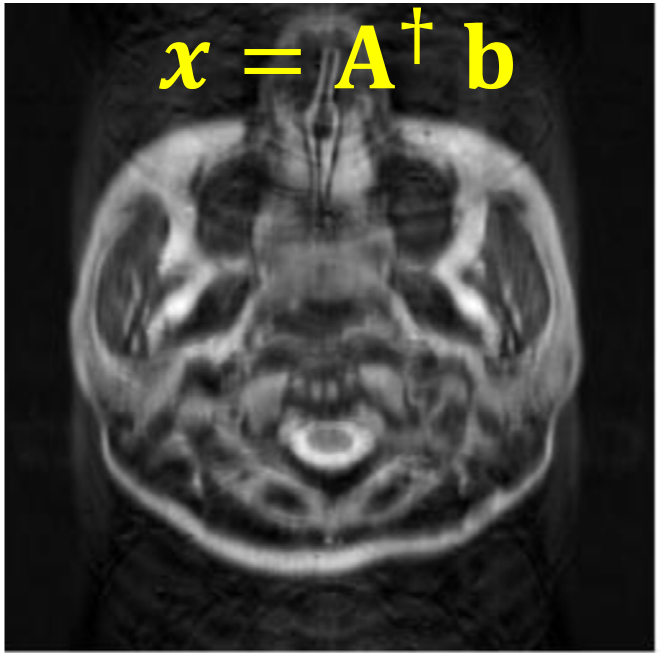
$$A = S_{ub} \; A_{full}$$

Subsampling operator

- A_{full} is discrete
  Fourier transform in MRI &
  Radon transform in CT.

Undersampled MRI problem

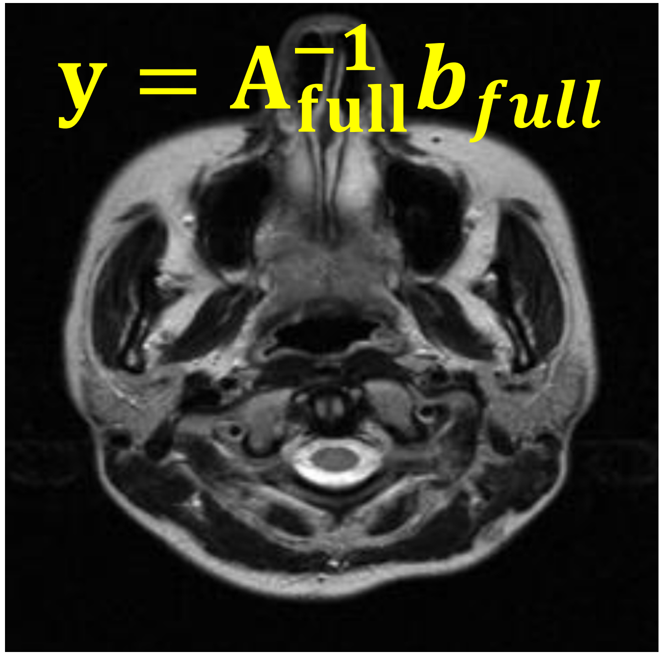
Is it possible?



Subsampling (30%)

$\mathbf{A}^{\dagger}$: Pseudo-Inverse of A.

$f: x \rightarrow y$



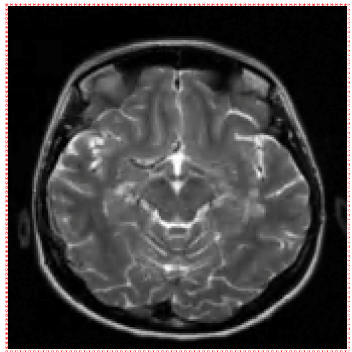
Full sampling

How to solve

Forward operator

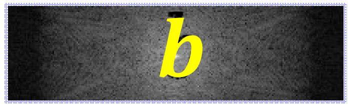
A

MR image



$=$

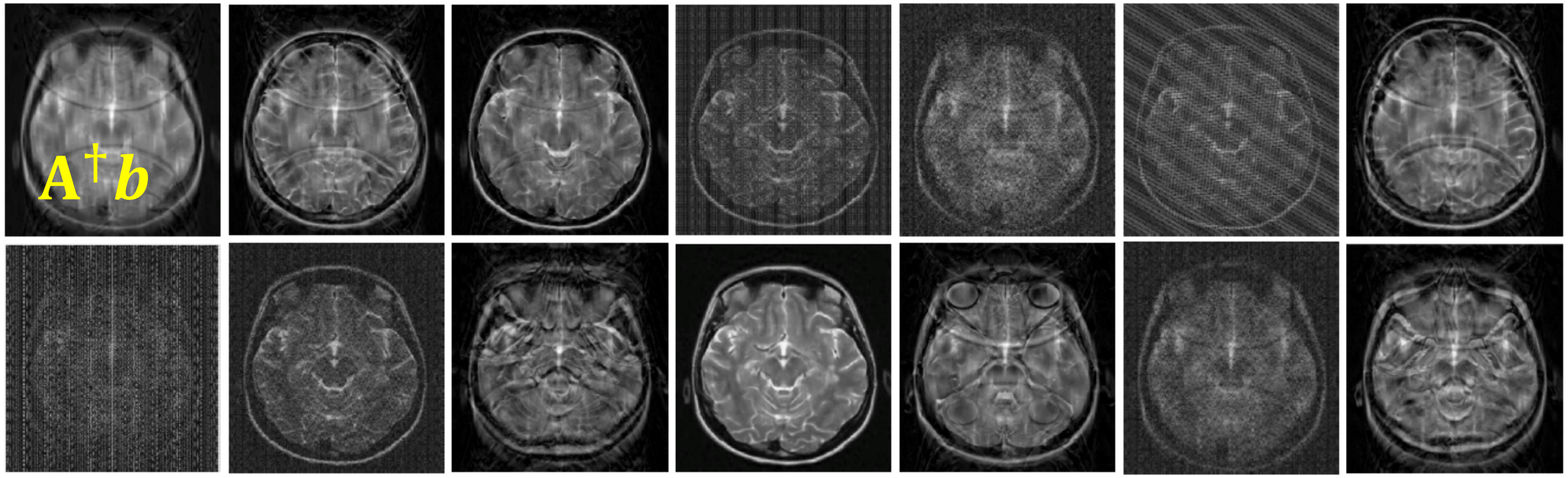
k-space data



Without imposing prior knowledge on the solution, this problem has infinitely many solutions.

✓ Need to choose one out of infinitely many images in $N_b(A) := \{ z : Az = b \}$.

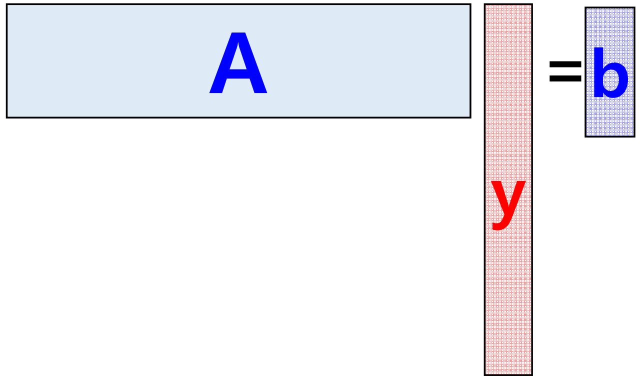
. .



. .

$dim\ N_b(A) = \#\ columns - \#\ rows$

Is it possible to find $f: A^{\dagger}b \rightarrow y = A_{full}^{-1} b_{full}$?



- ✓ Solving $Ay = b$ depends on an appropriate use of a priori information about medical CT or MRI images as solutions.
- ✓ We need to consider a constraint problem:

$$Ay = b \text{ subject to } y \in M \text{ (Solution Manifold)}$$

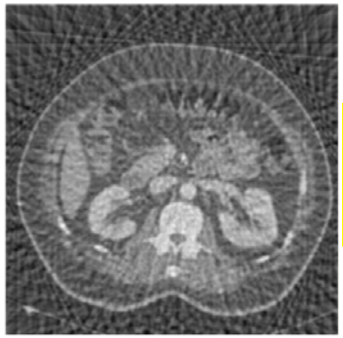
unknown

Example: sparse view CT

Pseudo-Inverse

$\in M$

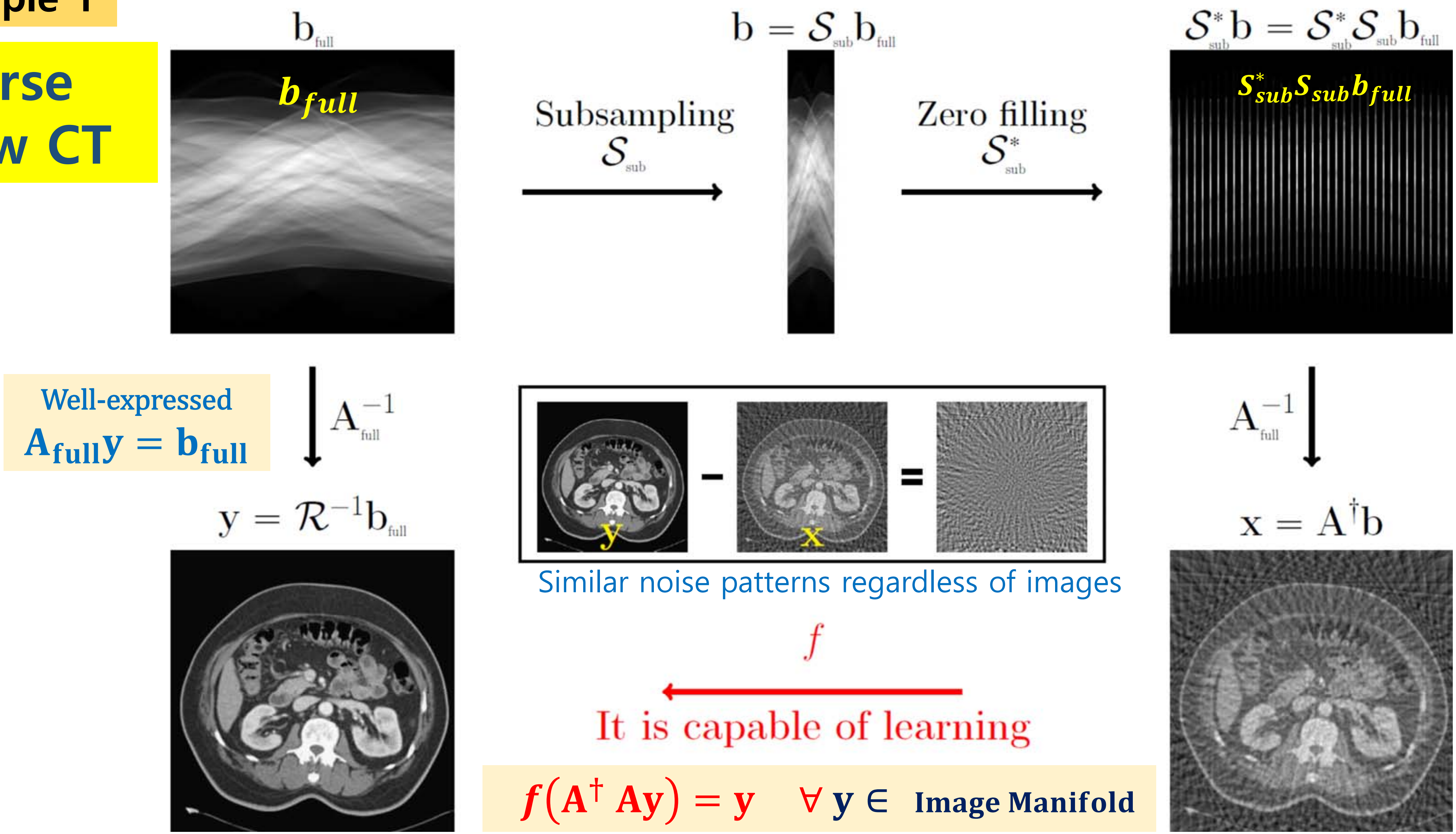
$$= A_{full}^{-1}\, b_{full} \neq A^{\dagger}b =$$



$\notin M$

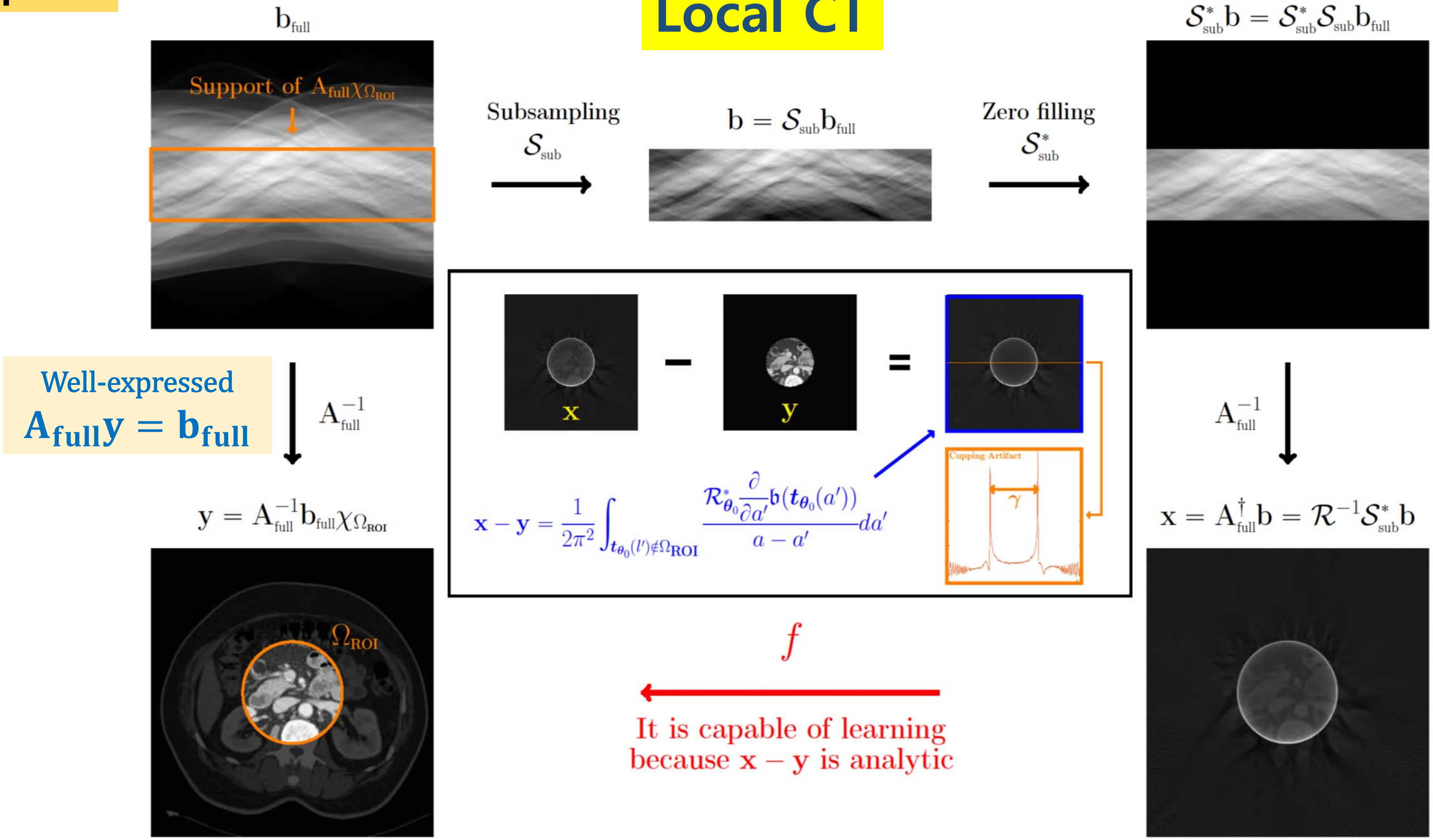
Example 1

Sparse View CT



Example 2

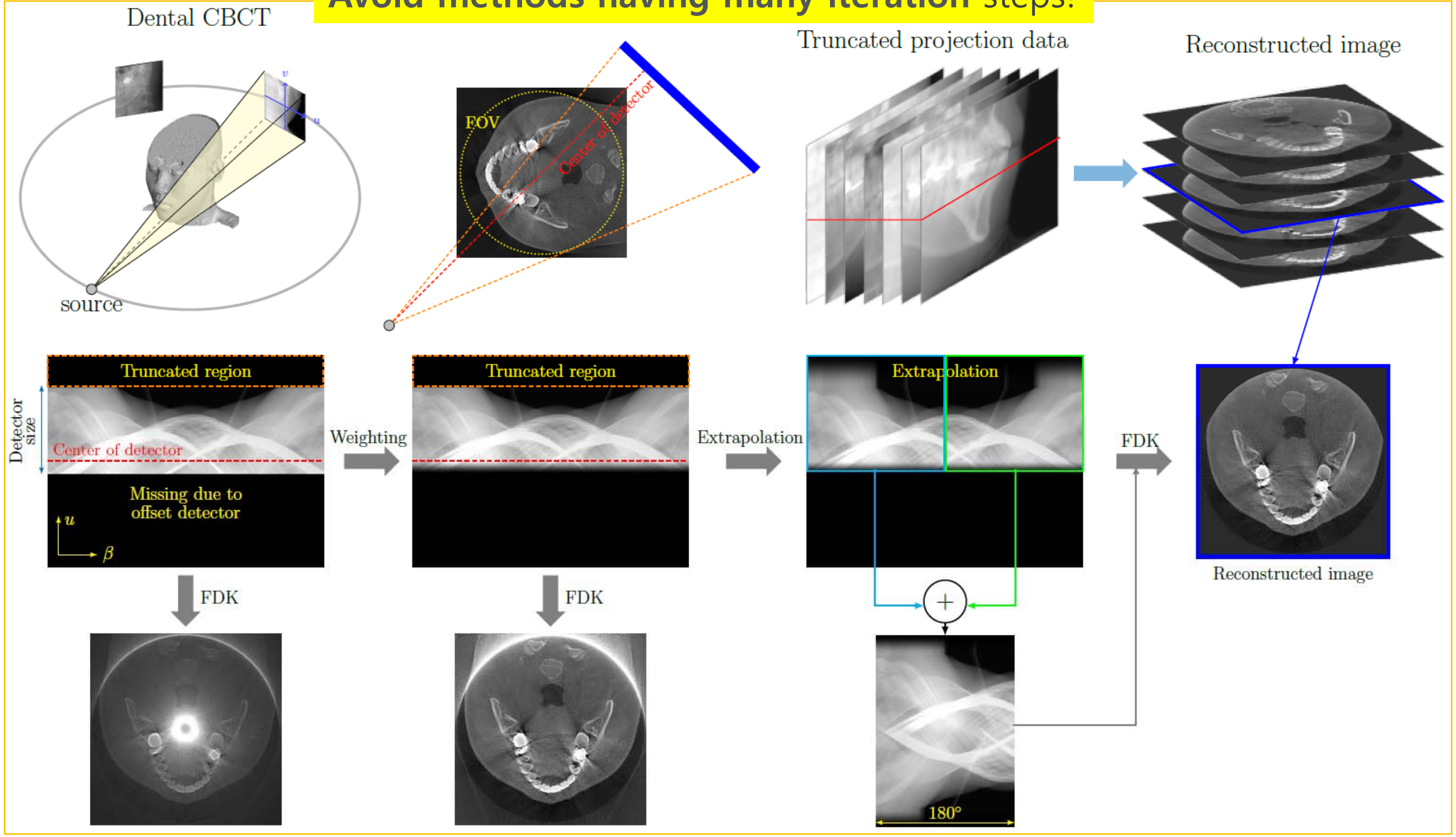
Local CT



Local CT

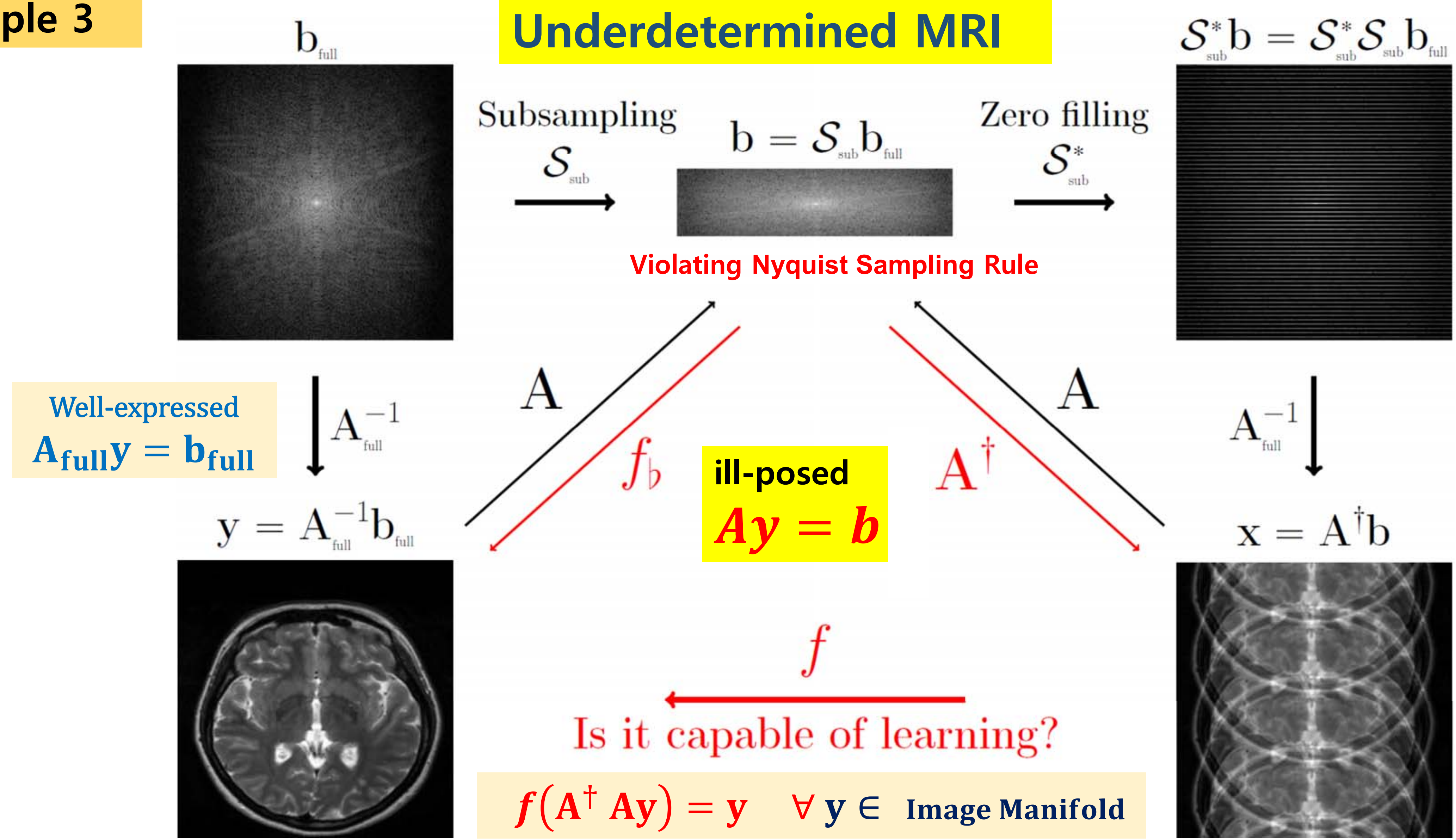
Dental CBCT: Need to develop a reconstruction method that addresses the problems caused by "Offset detector, FOV truncation, Low X-ray dose".

Avoid methods having many iteration steps!

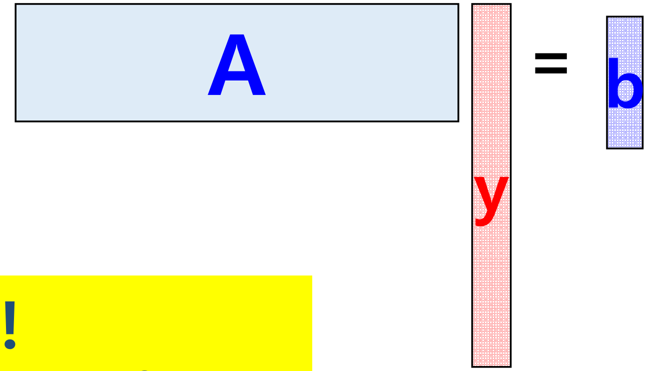


Example 3

Underdetermined MRI



Methods to solve the ill-posed problem



✓ **This is a highly nonlinear problem!**

The degree of nonlinearity depends on the sampling of data b and solution manifold.

✓ **Methods to impose Prior Knowledge on the solution**

Hand-made Sparse Sensing

- Use sparse representation of y
- Regularized data fitting method :

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

- Single data fidelity

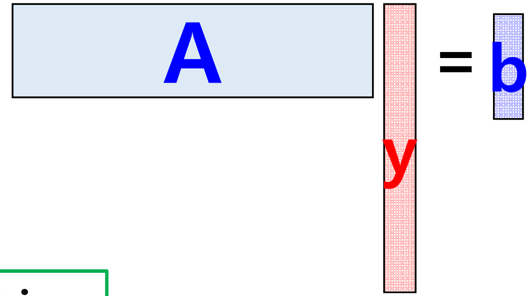
Machine-made Deep Regression

- Use training data $\{y^{(n)} : n = 1, \cdots, N\}$ to get the prior knowledge.
- Deep Learning :

$$f = \underset{f \in \text{Neural Nets}}{argmin} \sum_k ||y_k - f(x_k)||^2$$

- Group data fidelity

Comparison



- **Hand-made Sparse Sensing**

Image Prior

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

Single data fidelity

versus

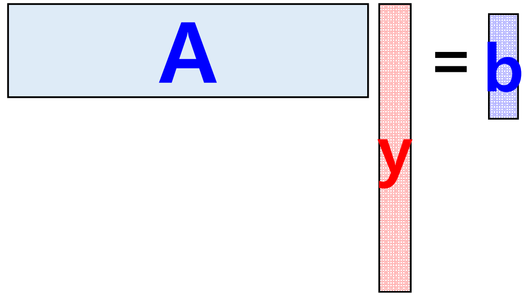
- **Machine-made DL Approach**

Image prior $\{(x^{(n)}, y^{(n)}) : n = 1, \cdots, N\}$

$$f = \underset{f \in \text{Neural Nets}}{argmin} \sum_k ||y_k - f(x_k)||^2$$

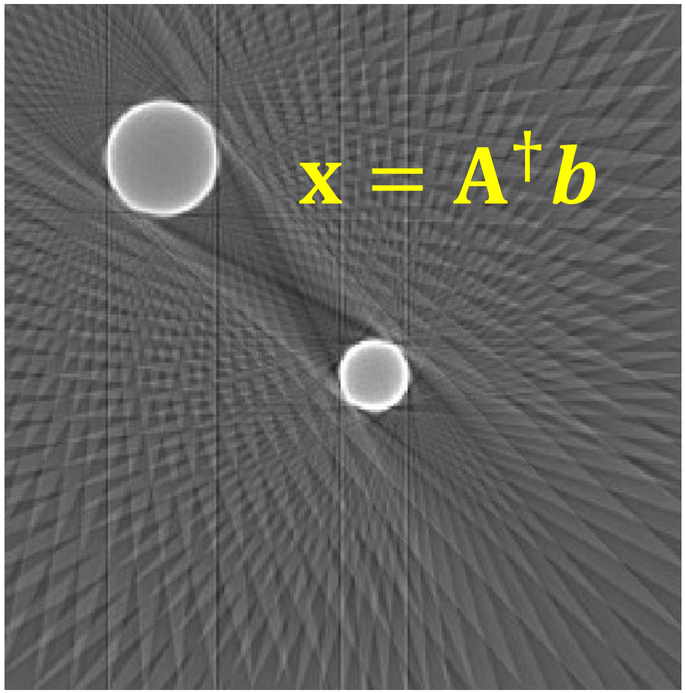
Group data fidelity

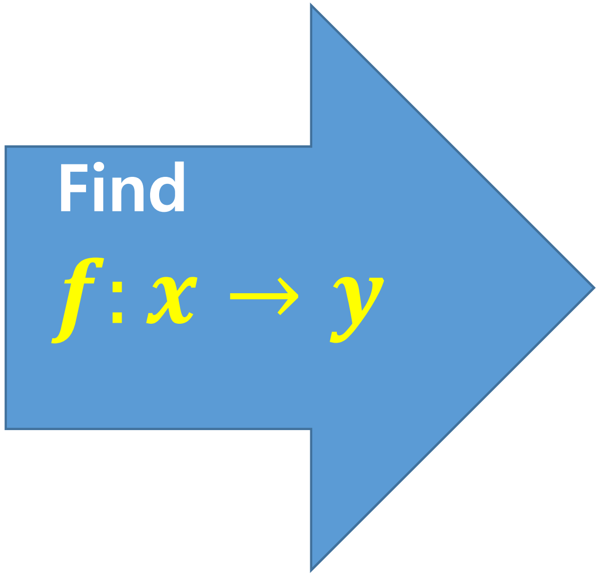
Comparison:
Hand-Made vs Machine-Made

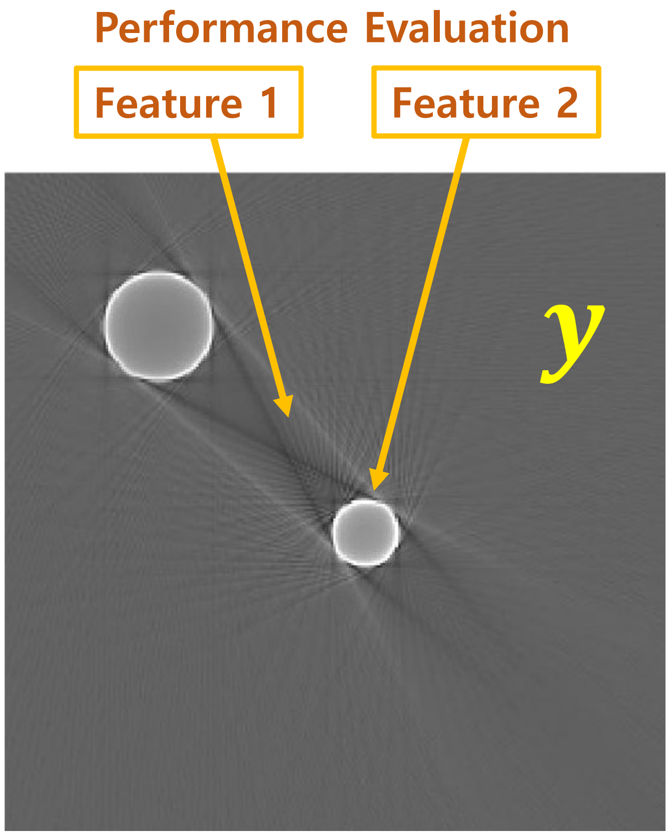


Test problem: Sparse View CT model with specially chosen M_{image}

In this sparse-view CT model, CS methods are known to work well.







Comparison: Man-Made vs Machine-Made

$$\boxed{A}\; y = b$$

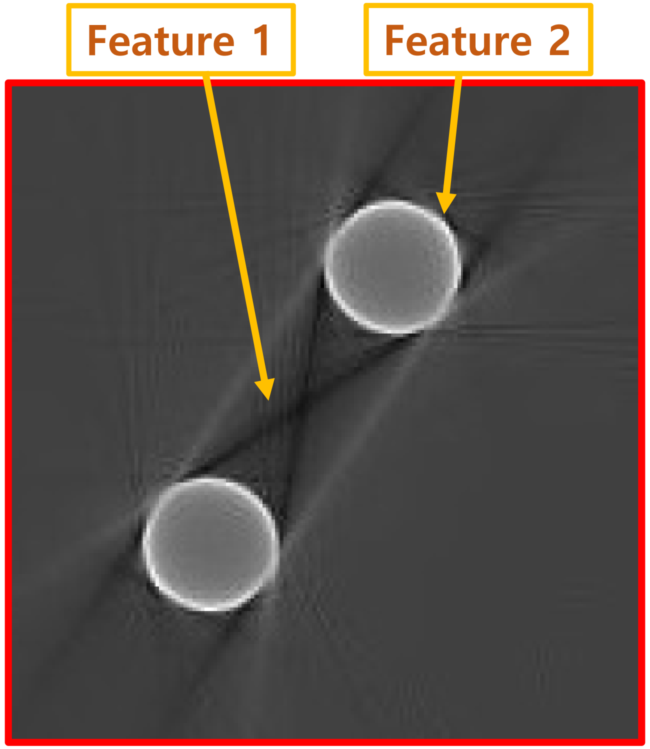
For this sparse-view CT problem, we use
a special solution manifold M_{image} (assumed to be unknown).

Dimension of $M_{image} = 7$

$$\mathcal{M}_{\text{image}} := \{\ \mathfrak{G}(\mathbf{h}) \in \mathbb{R}^{\partial_{\text{image}}}\ :\ \mathbf{h} \in \mathcal{K}\}$$

$$\mathfrak{G}(\mathbf{h}) = -\frac{1}{4\pi}\mathcal{R}^*\mathcal{I}^{-1}\left[\ln\left(\frac{\sinh\left(h_7\mathcal{R}\chi_{D_{\mathbf{h}}}\right)}{h_7\mathcal{R}\chi_{D_{\mathbf{h}}}}\right)\right]$$

Performance Evaluation



Since this solution manifold is only 7 dimension,
$Ay = b$ can be solvable only with 7 equations.

Comparison: Hand-Made vs Machine-Made

- **CS and linear approaches eliminate the feature 1.**

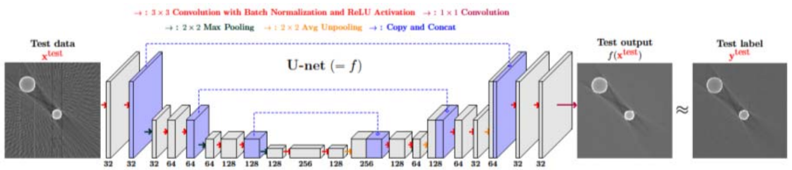
$$f(\mathbf{x}) = \mathbf{D}\mathbf{h}, \quad \mathbf{h} = \underset{\mathbf{h}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{D}\mathbf{h} - \mathbf{A}\mathbf{x}\|_{\ell^2}^2$$

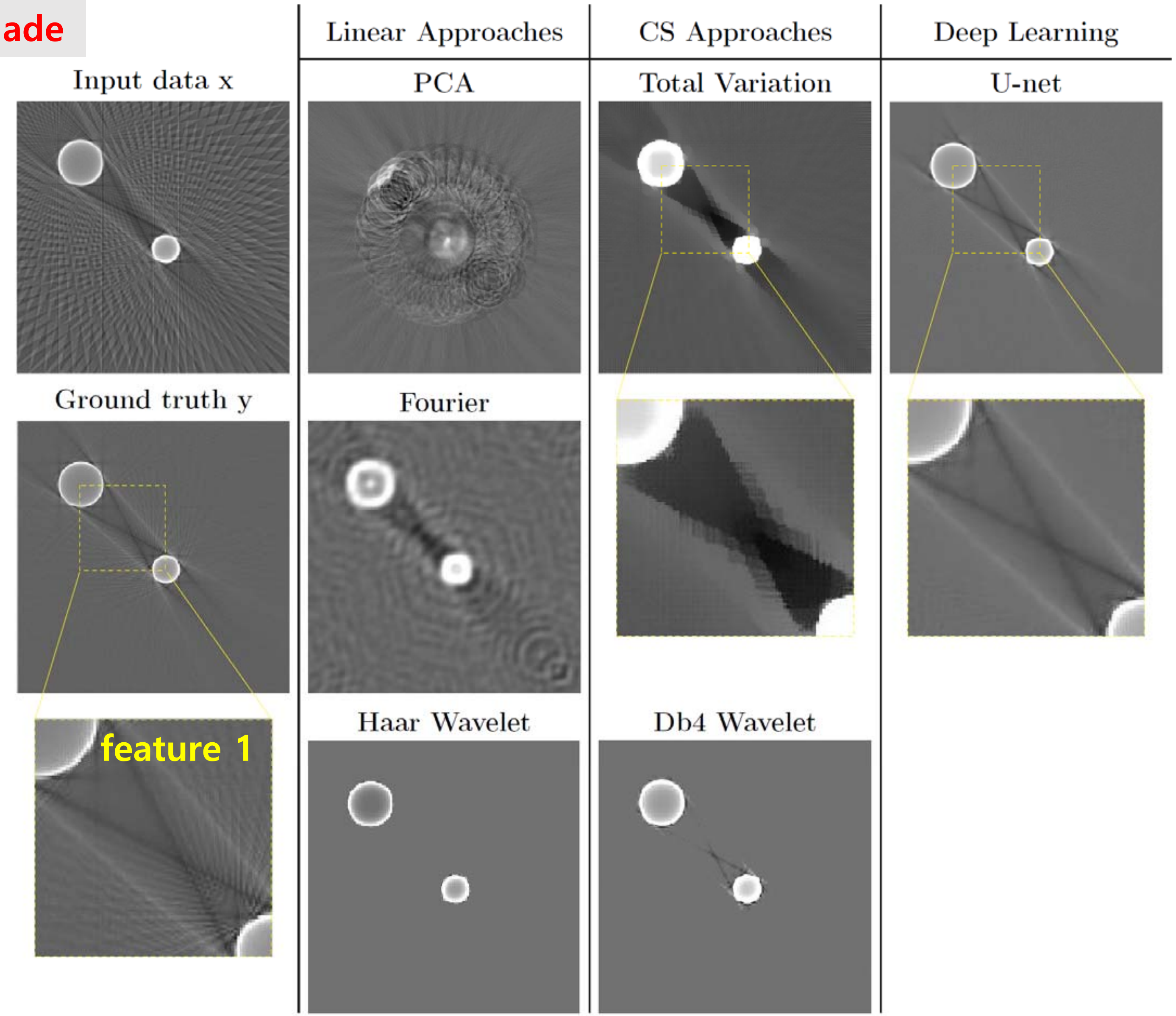
$$f(\mathbf{x}) = \mathbf{D}\mathbf{h}, \quad \mathbf{h} = \underset{\mathbf{h}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{D}\mathbf{h} - \mathbf{A}\mathbf{x}\|_{\ell^2}^2 + \lambda\|\mathbf{h}\|_{\ell^1}$$

$$f(\mathbf{x}) = \underset{\mathbf{y}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{y} - \mathbf{A}\mathbf{x}\|_{\ell^2}^2 + \lambda\|\nabla\mathbf{y}\|_{\ell^1}$$

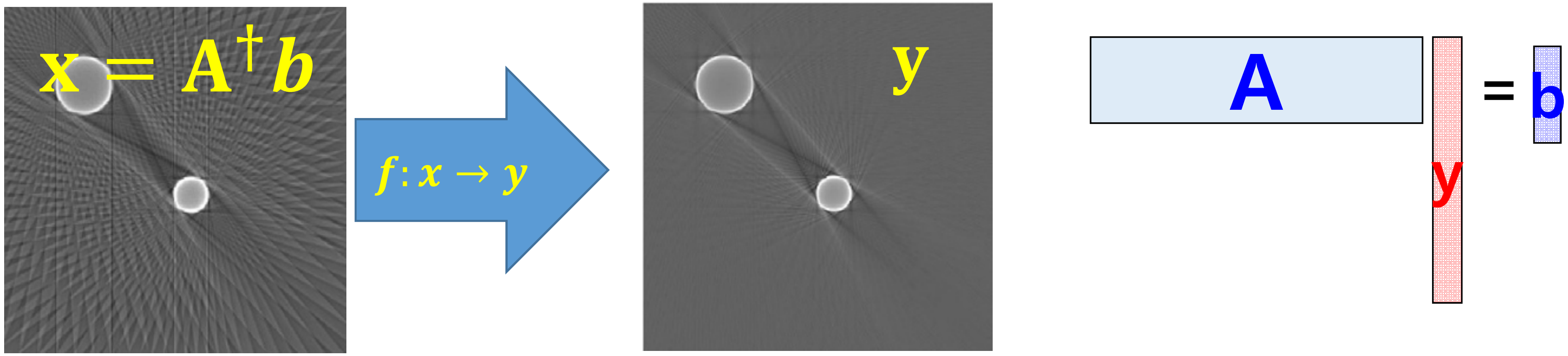
- **Deep learning preserves the feature 1.**

$$f = \underset{f \in \mathbb{NN}}{\operatorname{argmin}} \frac{1}{n_{\text{data}}} \sum_{n=1}^{n_{\text{data}}} \|f(\mathbf{x}^{(n)}) - \mathbf{y}^{(n)}\|_{\ell^2}^2$$

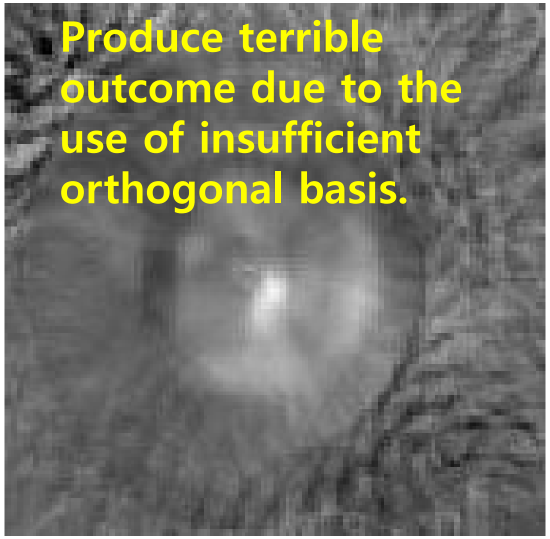




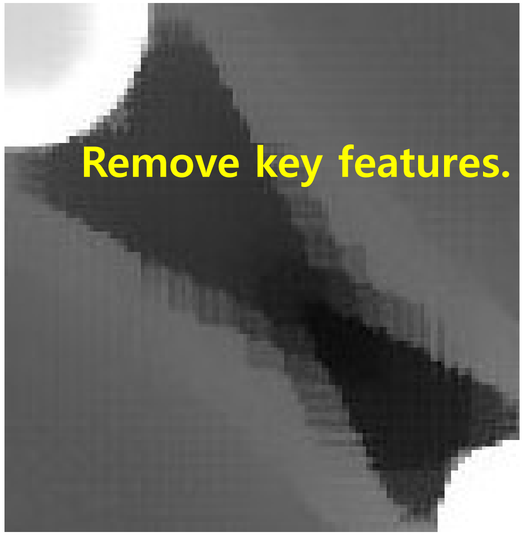
Man-Made vs Machine-Made



PCA

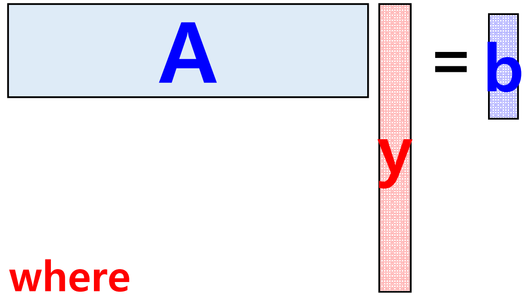


Total Variation



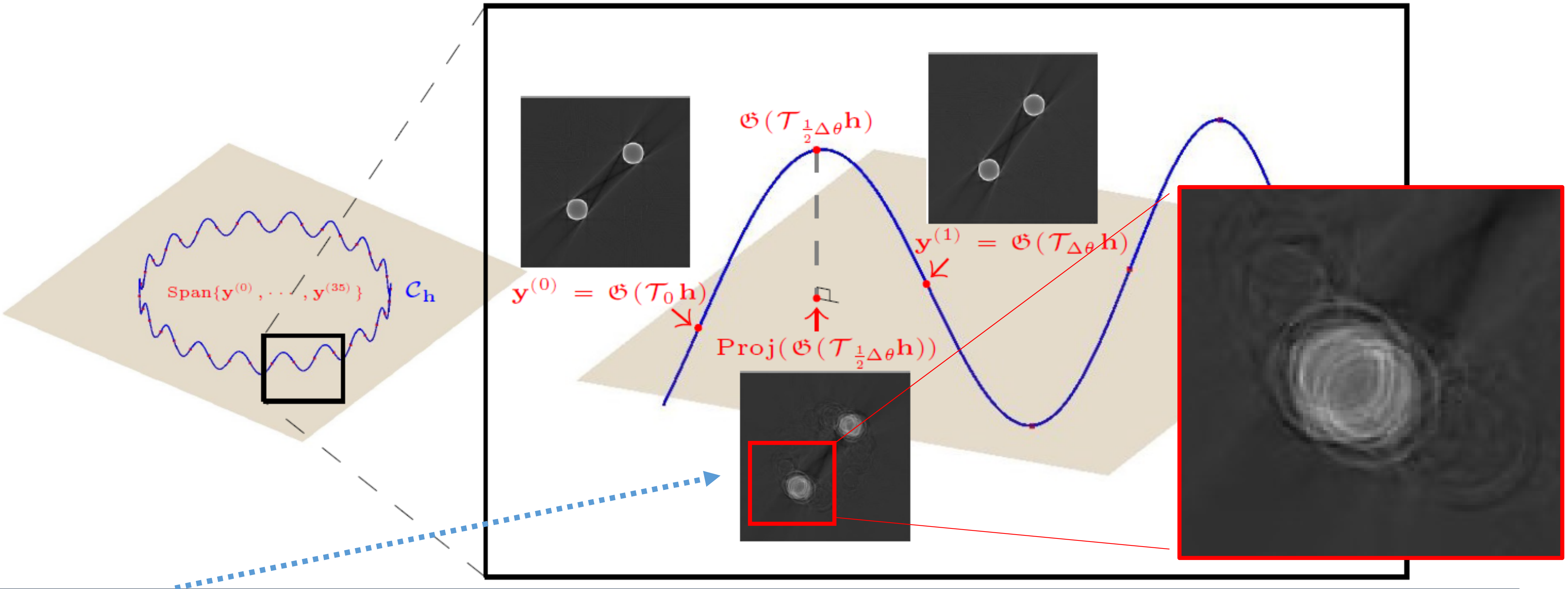
Deep Learning

Man-Made vs Machine-Made



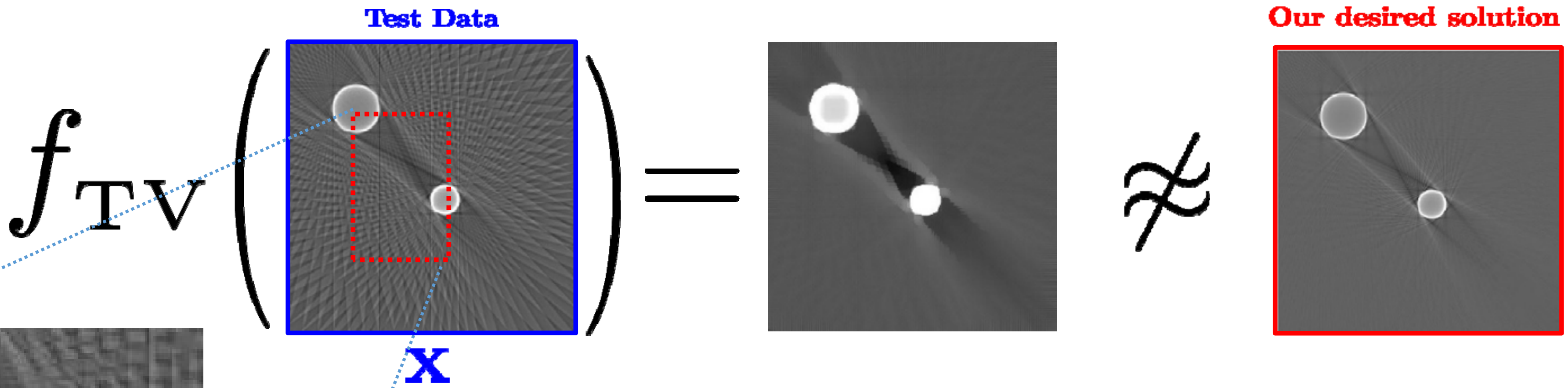
Linear methods (PCA, Wavelet decomposition) may be unable to deal with the highly curved solution manifold.

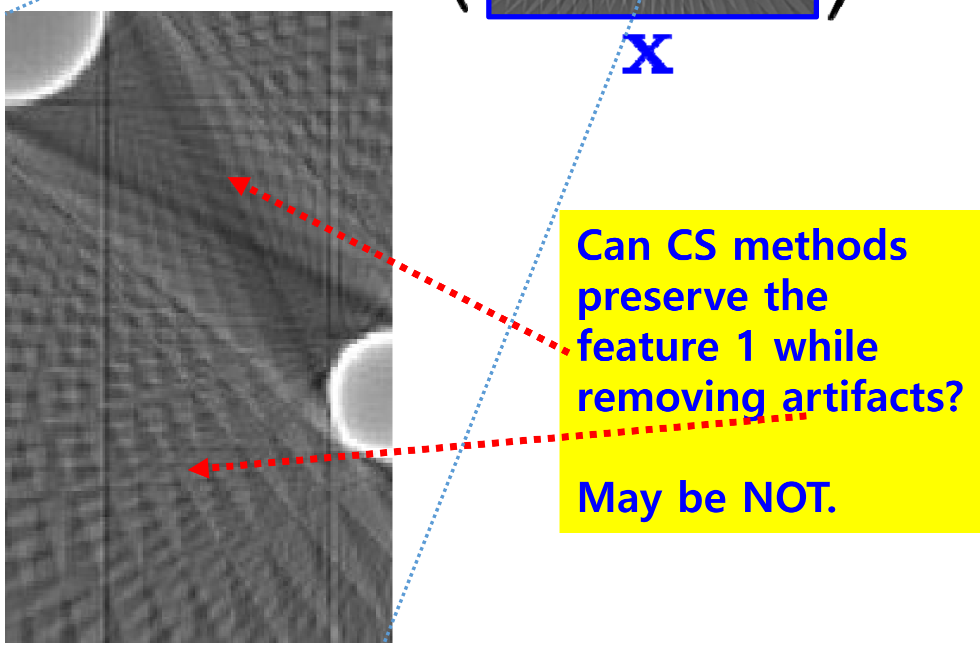
Consider the vector space spanned by images $\{y^{(0)}, y^{(1)}, \cdots, y^{(N)}\}$ where $y^{(k)}$ is $k\pi/N$ rotation of image $y^{(0)}$.



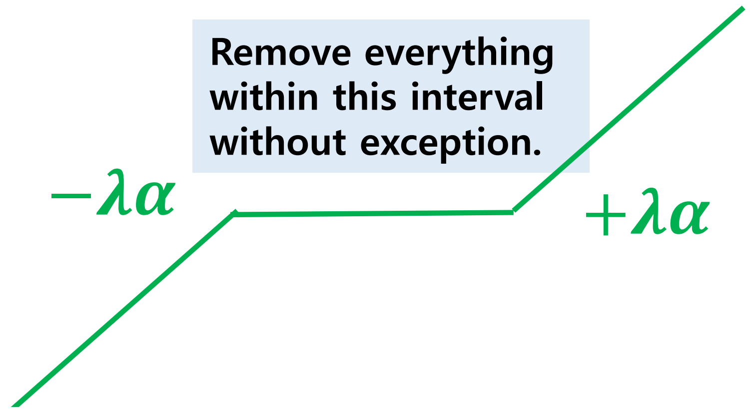
The middle image between $y^{(0)}$ & $y^{(1)}$ cannot be expressed properly by the space spanned by $\{y^{(0)}, y^{(1)}, \cdots, y^{(N)}\}$.

TV approach: $f_{\mathrm{TV}}(x) = \underset{y}{\operatorname{argmin}} ||Ay - x||^2_{\ell^2} + \lambda||\nabla \mathbf{y}||_{\ell^1}$





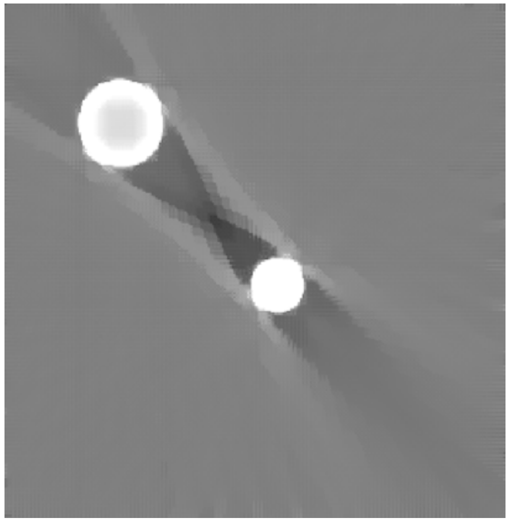
It may not preserve some detailed structures that may contain crucial medical information.



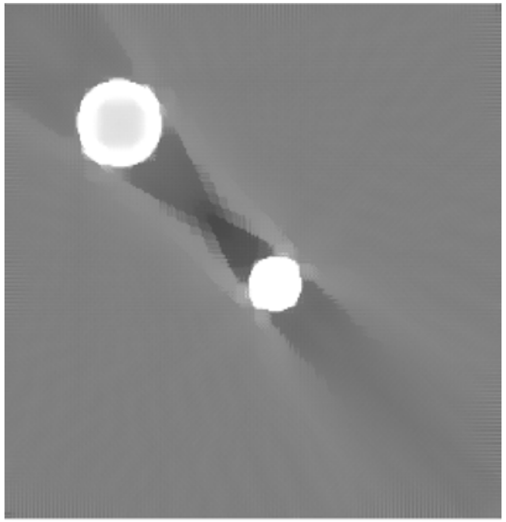
TV approach: $f_{\mathrm{TV}}(x) = \underset{y}{\mathbf{argmin}} ||A\mathbf{y} - x||_{\ell^2}^2 + \lambda||\nabla \mathbf{y}||_{\ell^1}$

- ✓ The performance depends on the regularization parameter.
- ✓ Need several iterations to find a sparse expression.

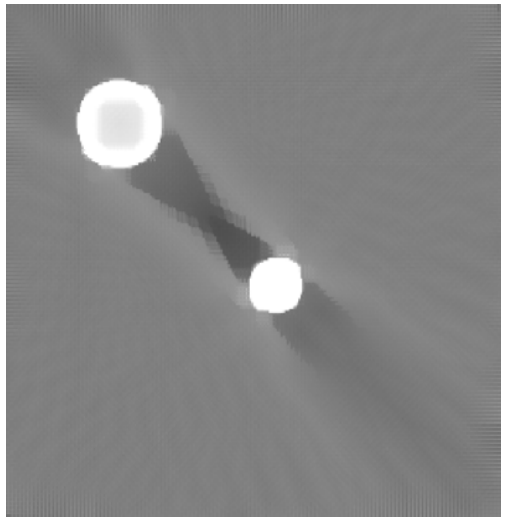
$\lambda = 0.5\,\alpha$



$\lambda = \alpha$

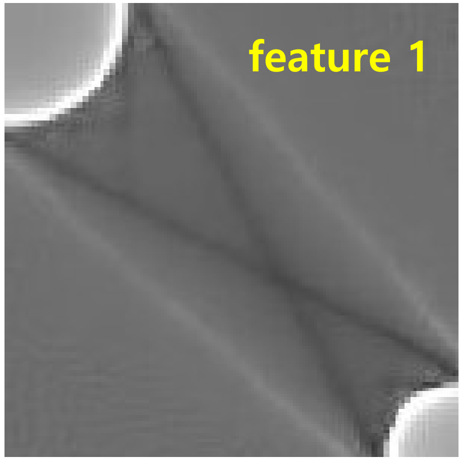


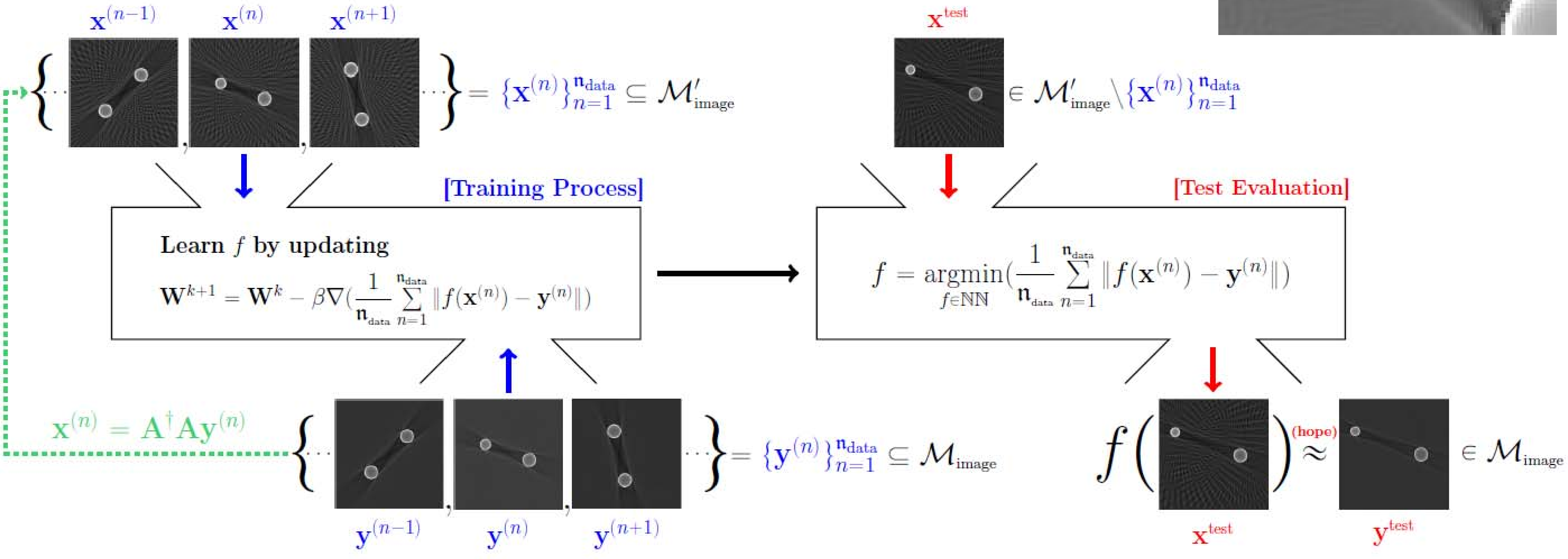
$\lambda = 1.5\,\alpha$



Man-Made vs Machine-Made

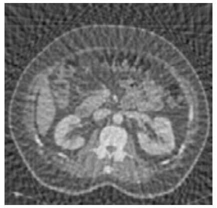
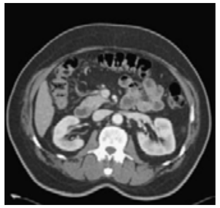
- **DL approach can selectively preserve the feature 1.**

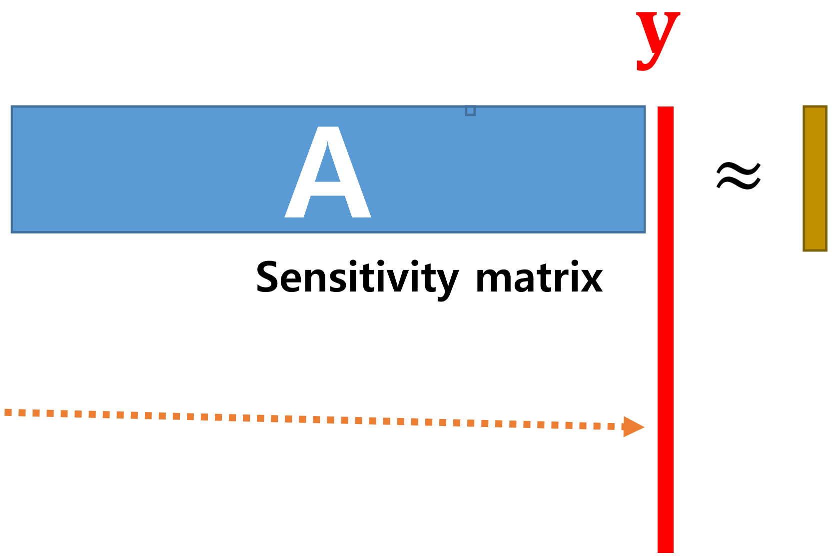




Deep Learning Approach

$\mathbf{A}^{\dagger}\,\mathbf{b}$ $\qquad\qquad$ $\mathbf{y}$

$\boldsymbol{f}\big($  $\big) =$ 



Use training data to learn both f and image manifold such that

$$\boldsymbol{f}(\mathbf{A}^{\dagger}\,\mathbf{A}\mathbf{y}) = \mathbf{y} \quad \forall\, \mathbf{y} \in \textbf{ Image Manifold.}$$

One of DL's most important advantages is to provide non-iterative reconstruction methods for highly non-linear problems.

What is learnable.
What is NOT.

unknown

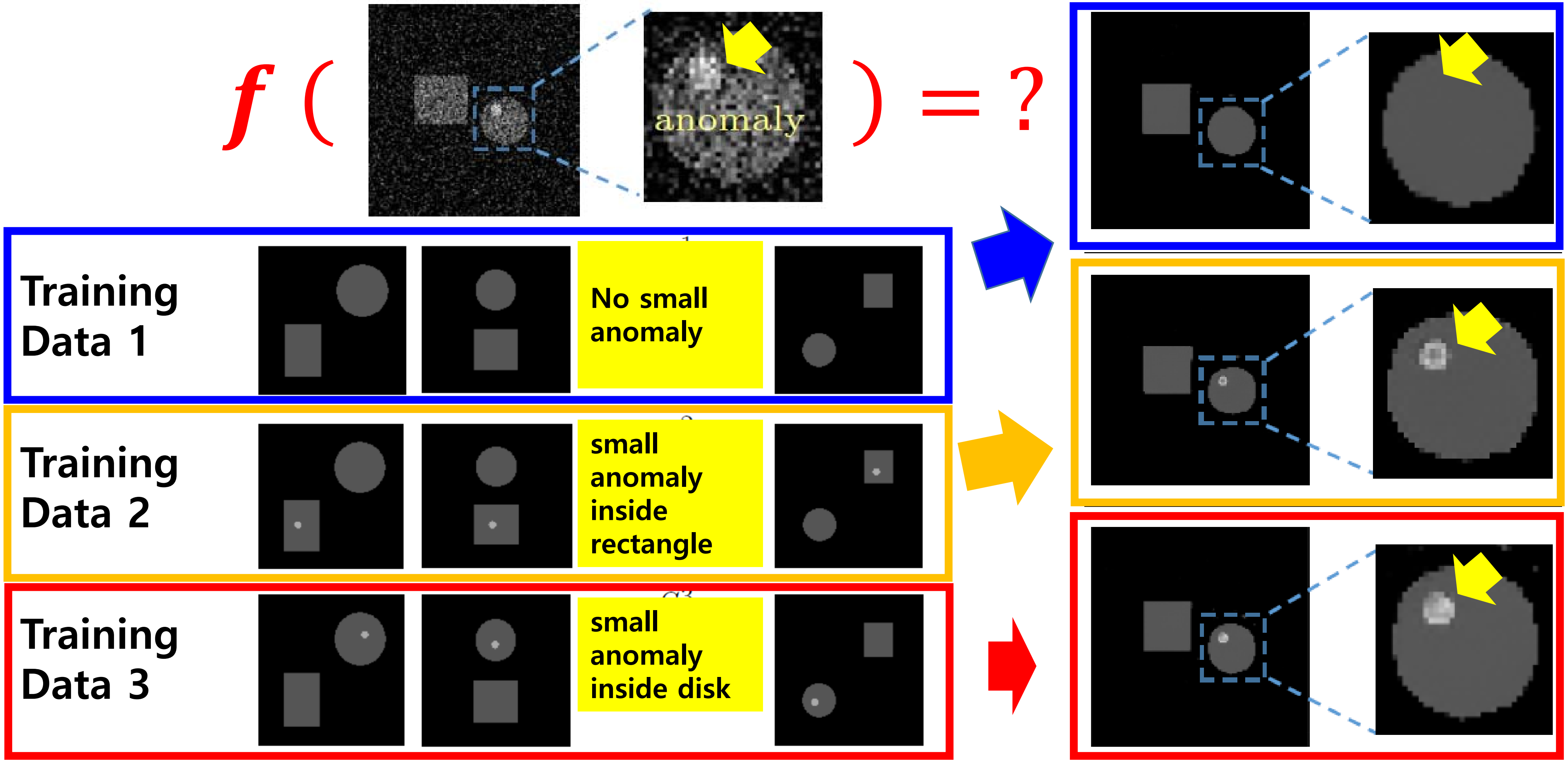
The necessary condition for learning f is that

$$f(\mathrm{A}^\dagger \mathrm{A}\mathbf{y}) = \mathbf{y}, \quad \forall\, \mathbf{y} \in \text{ Image Manifold.}$$

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

Impact of Training Data: It is critical to choose suitable training datasets to reflect the appropriate image priors, in order to preserve detailed features of the images.

Joint work with Hyungsuk Park



Observation: The reconstruction map $f: \mathbf{x} = \mathbf{A}^{\dagger} b \rightarrow y$ is learnable if A satisfies the M-RIP (manifold restricted isometry property) condition.

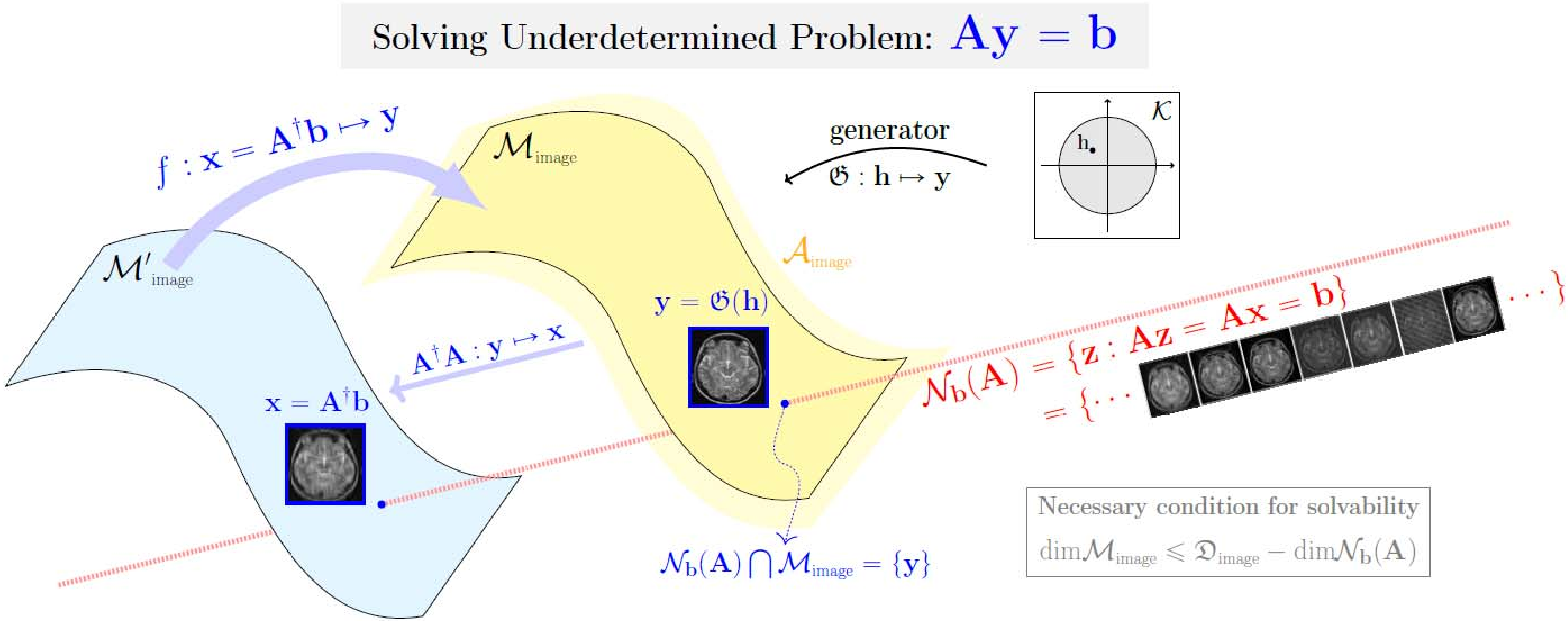
$$\boxed{c\|\mathbf{y}-\mathbf{y}'\| \leq \|\mathbf{A}\mathbf{y}-\mathbf{A}\mathbf{y}'\| \leq \frac{1}{c}\|\mathbf{y}-\mathbf{y}'\| \quad \text{for all} \quad \mathbf{y}, \mathbf{y}' \in \mathcal{M}_{\text{image}}}$$

✓ Necessary condition for learnability is

$$\dim M_{image} \leq Rank\ A$$

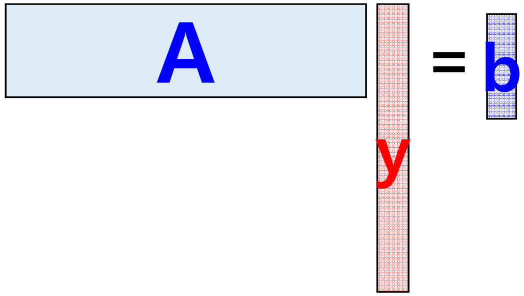
✓ M_{image} indicates the unknown solution Manifold

M-RIP condition: $N_b(A) \cap M_{image} = \{y\}$ (uniqueness & stability).



- M_{image} indicates a solution **manifold** that is assumed to be a good regression of MR head image data distributions $\{y^{(n)} : n = 1, \cdots, N\}$.

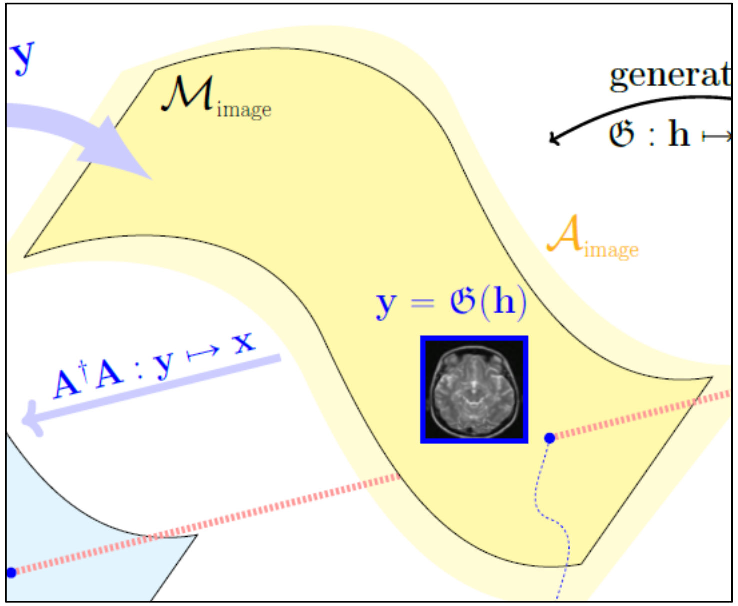
- $\mathrm{x} = \mathrm{A}^\dagger b$ is the minimum-norm solution which will be used to find the true solution y.



is a highly nonlinear problem!

- A is (# equations) $\times$ (#unknowns) matrix
- $M_{image} = \{\, y:\ y = G(h),\ h \in K\}$, where G is a generator & K is a compact subset of R^k.

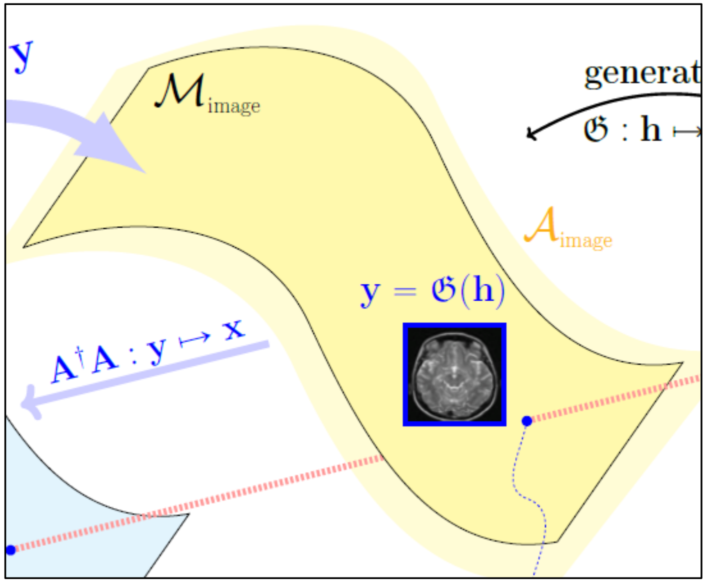
Observation: $f: \mathrm{x} = \mathrm{A}^{\dagger}b \to y$ is nonlinear if dim (span$\{\partial_j G(h): h \in K\}$) $>$ # *equations* .



✓ The map $f: \mathrm{x} = \mathrm{A}^{\dagger}b \to y$ can be viewed as an image restoration function with filling-in missing data in x. Therefore, $\nabla f(x)$ depends on the image structure in x.

✓ The nonlinearity of f is affected by sampling and the degree of bending of the manifold M_{image}

Observation: $f: \mathbf{x} = \mathbf{A}^{\dagger} b \to y$ is nonlinear if dim $(\text{span}\{\partial_j G(h): h \in K\}) > \#\, equations$.



Proof:

- $f(x) = \mathbf{y} \quad \to \quad f(\mathbf{A}^{\dagger}\mathbf{A}\mathbf{y}) = \mathbf{y} \quad \to \quad f(\mathbf{A}^{\dagger}\mathbf{A}\mathbf{G}(\mathbf{h})) = \mathbf{A}\mathbf{G}(\mathbf{h})$
- If $f: x = A^{\dagger} b \to y$ is linear, $B = \nabla f(\cdot)$ is a constant matrix &

$$\mathbf{B}\mathbf{A}^{\dagger}\mathbf{A}\nabla\mathbf{G}(\mathbf{h}) = \nabla\mathbf{G}(\mathbf{h}) \text{ for all } \mathrm{h} \in K.$$

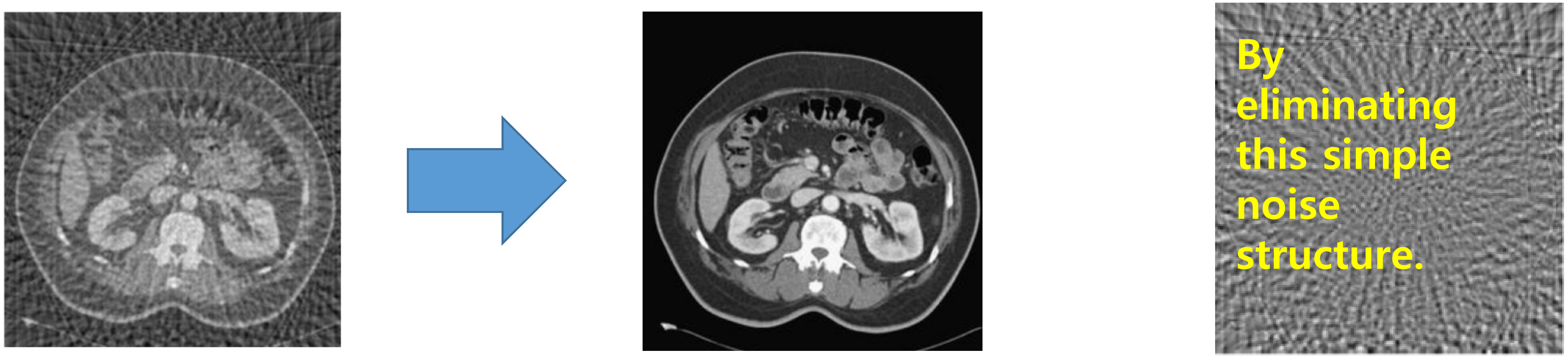
✓ Hence, all $\partial_j G(h) \in \mathbf{Eigen}_1(\mathbf{B}\mathbf{A}^{\dagger}\mathbf{A})$, the eigenspace of $\mathrm{BA}^{\dagger}\mathrm{A}$ corresponding to the eigenvalue 1.

✓ This is not possible if dim $(\text{span}\{\partial_j G(h): h \in K\}) > \#\, equations$.

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

Example 1: Sparse View CT

Both deep learning and compressed sensing work very well for this kind of problems.



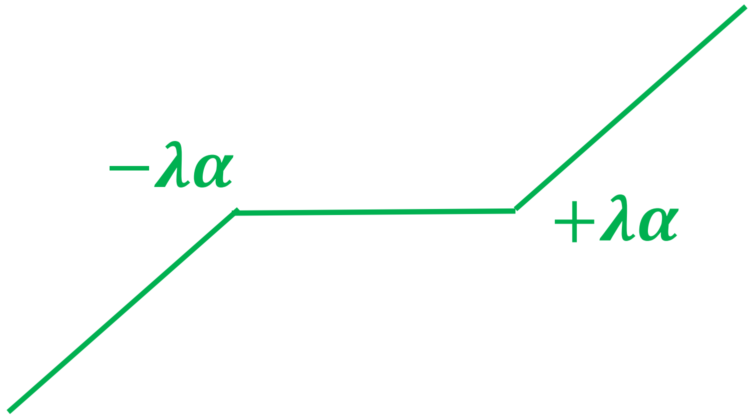
L^1 −Regularized data fitting technique

Assume that there exist W such that $y = Wz$ with z being sparse.

$$min\ E(z) := \ ||g(Wz) - x||^2 + \lambda\ ||\ z||_{\ell^1}$$

$$\mathbf{z} = \mathbf{S}_{\lambda\alpha}(\mathbf{z} - \alpha\nabla||\ \mathbf{g}(\mathbf{Wz}) - \mathbf{x}||^2)$$

$$\mathbf{S}_{\lambda\alpha}(z) = sign(z)\ max\{\ |z| - \lambda\alpha, 0\}$$



Example 2: Local CT

Deep learning works well because of unique continuation of analytic function along the vertical direction.

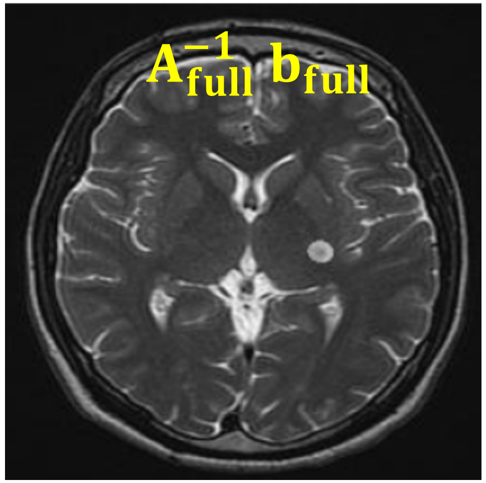
Low dimensional latent representation possible!

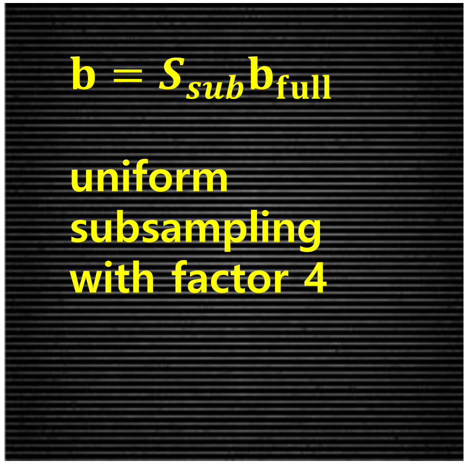
$$\Psi^{out}_{\boldsymbol{\theta}_0}\mathfrak{b}(\boldsymbol{t}_{\boldsymbol{\theta}_0}(a)) = \frac{1}{2\pi^2}\int_{\boldsymbol{t}_{\boldsymbol{\theta}_0}(a')\notin\Omega_{\mathrm{ROI}}} \frac{\mathcal{R}^*_{\boldsymbol{\theta}_0}\frac{\partial}{\partial a'}\mathfrak{b}(\boldsymbol{t}_{\boldsymbol{\theta}_0}(a'))}{a'-a}da'.$$

Example 3: Underdetermined MRI

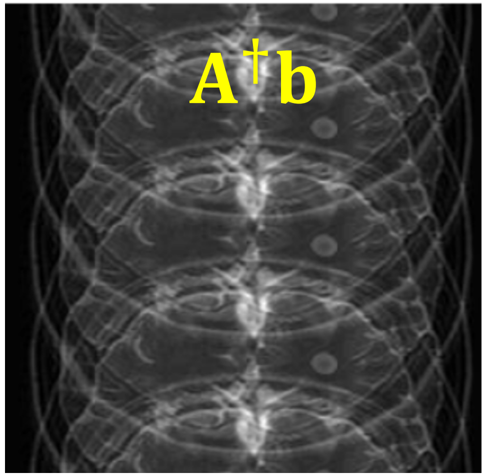
According to the Poisson summation formula, the discrete Fourier transform of $\mathbf{b} = S_{sub}\mathbf{b}_{\text{full}}$ (uniformly subsampled data with factor 4) produces the following four-folded image.

Inverse Fourier Transform





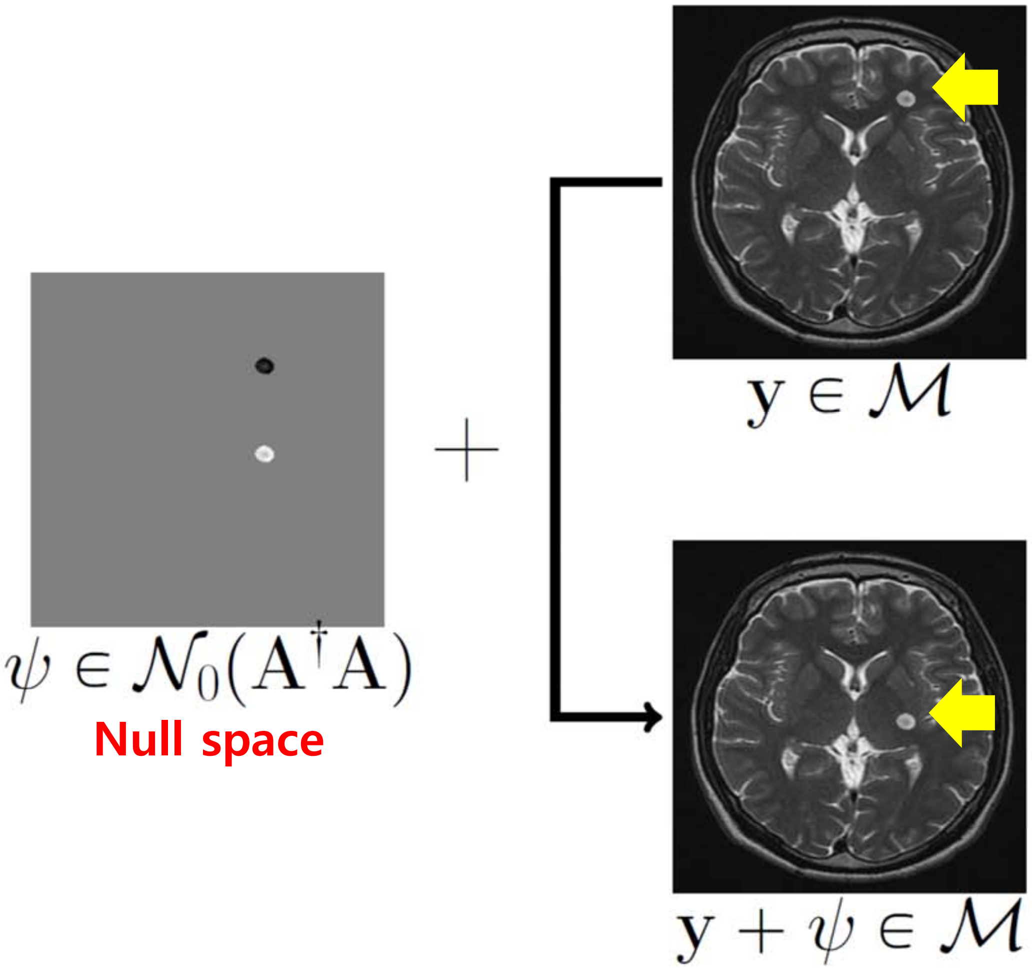
Inverse Fourier Transform



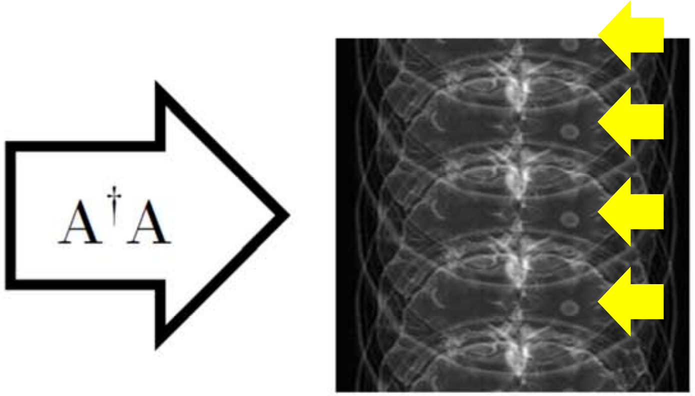
If we use **uniform subsampling** S_{sub} with factor 4,

it is difficult to learn f $s.t.$ $f(\mathrm{A}^\dagger \mathrm{A}\mathrm{y}) = \mathrm{y} \quad \forall \, \mathrm{y} \in$ Image Manifold

- ✓ Fail to satisfy M-RIP condition.
- ✓ DL is NOT a magic.



Null space

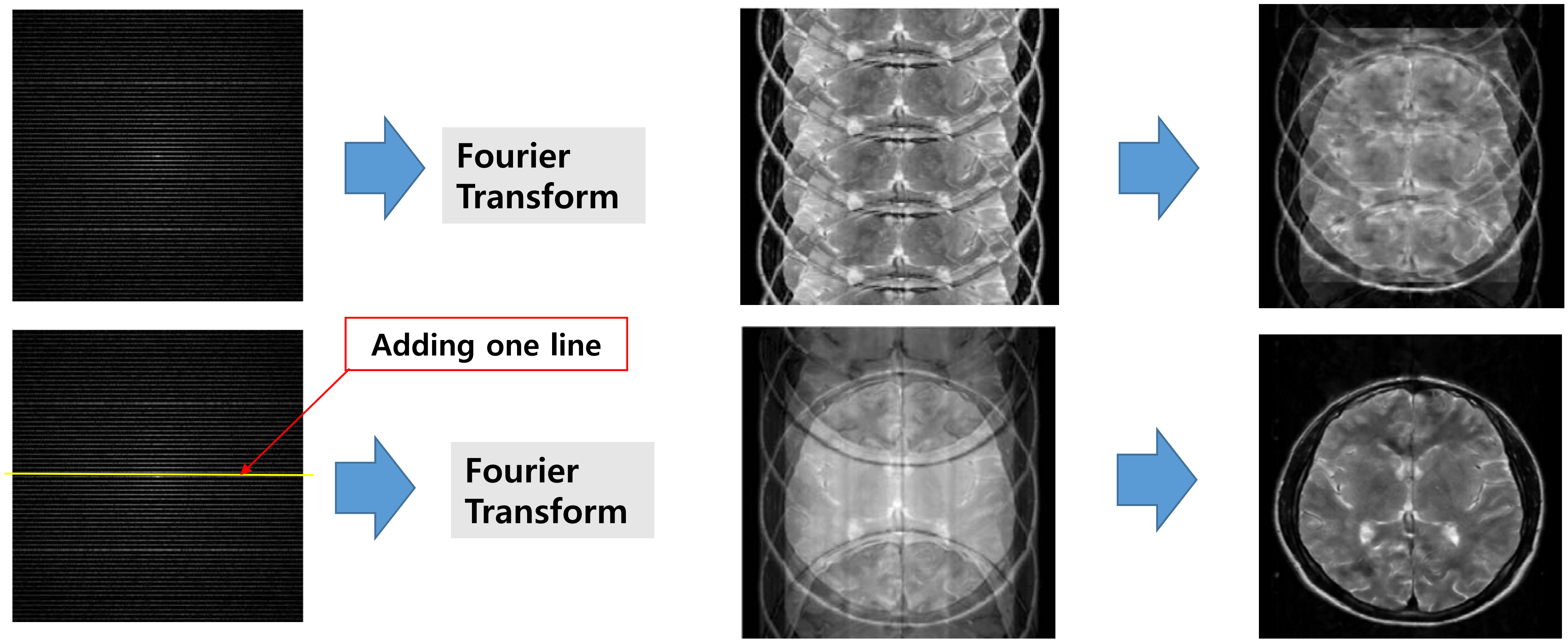


$$\mathrm{x} = \mathrm{A}^\dagger \mathrm{A}\mathrm{y} = \mathrm{A}^\dagger \mathrm{A}(\mathrm{y} + \psi)$$

Poisson summation formula

However, the result changes dramatically by adding only one line in k-space.

$$\mathbf{b} = S_{sub}\mathbf{b}_{\mathbf{full}} \qquad \mathbf{x} = \mathbf{A}^{\dagger}\,\mathbf{b} \qquad \mathbf{f}(\mathbf{x})$$

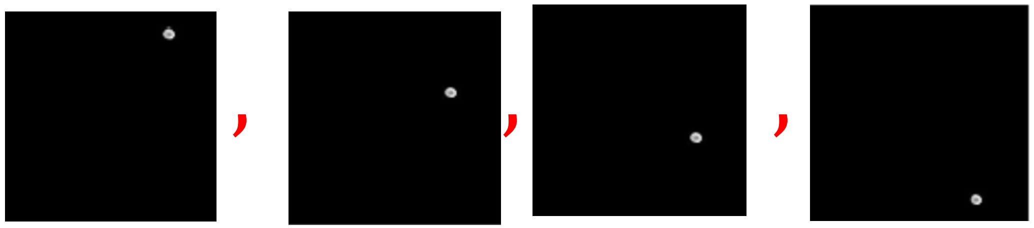


Why does the learning effect dramatically change by adding only one line in k-space?

- Let $\mathbf{A}_1$ be the sensitivity matrix corresponding to uniform sampling with factor 4.
- Let $\mathbf{A}_2$ be the sensitivity matrix corresponding to the row just above the center.

$$\{\mathbf{y}: \mathbf{A}_1^{\dagger}\mathbf{A}_1\mathbf{y} =$$

$$\} = \{$$



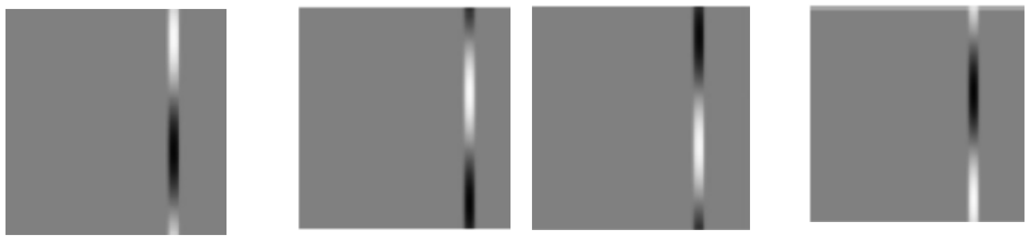
$$\}$$

Indistinguishable

$$\mathbf{A}_2^{\dagger}\mathbf{A}_2$$

Distinguishable

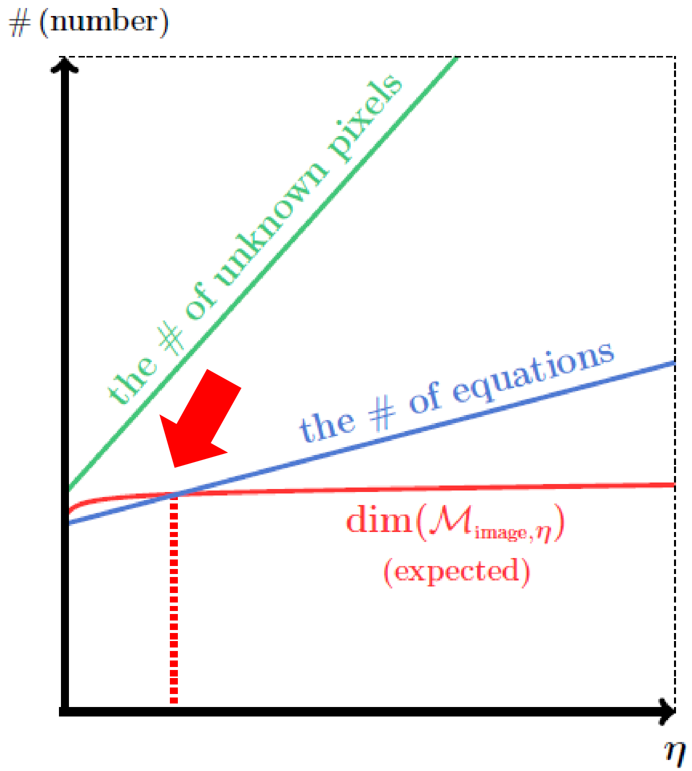
Inverse Fourier transform of the single line in k-space

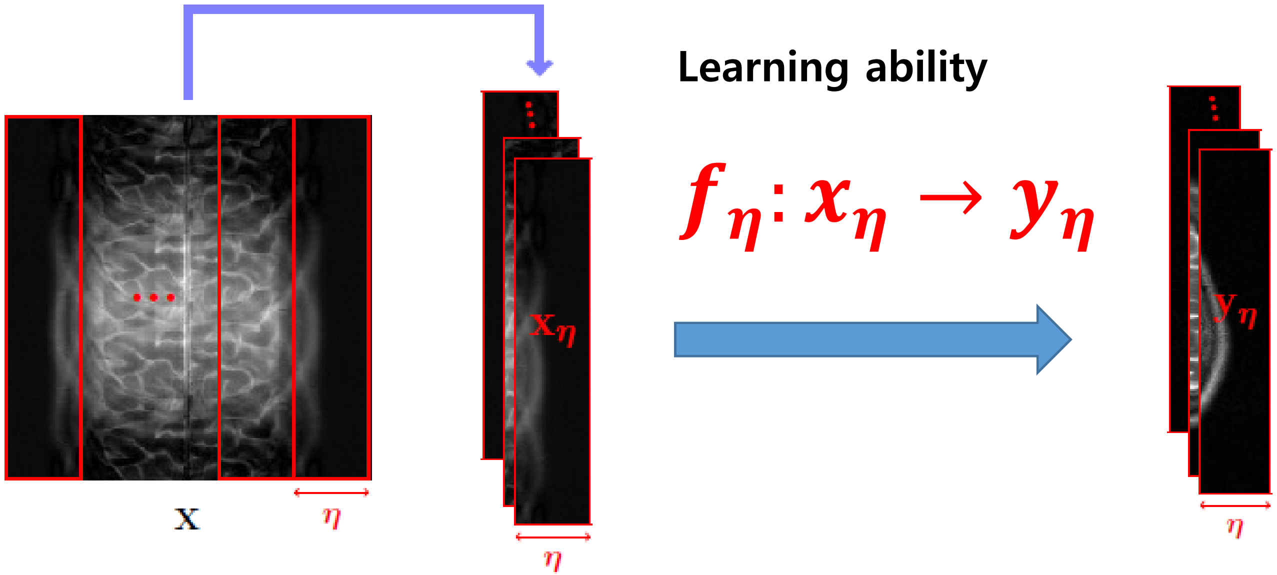


Because of this rough signature, it is capable of learning f $s.t.$
$$f(\mathbf{A}^{\dagger}\mathbf{A}\mathbf{y}) = \mathbf{y} \quad \forall\, \mathbf{y} \in \text{ Image Manifold}$$

Let us consider learning ability issue:

Patch images vs full image





As η increases, the number of unknowns increases more rapidly than the number of equations.

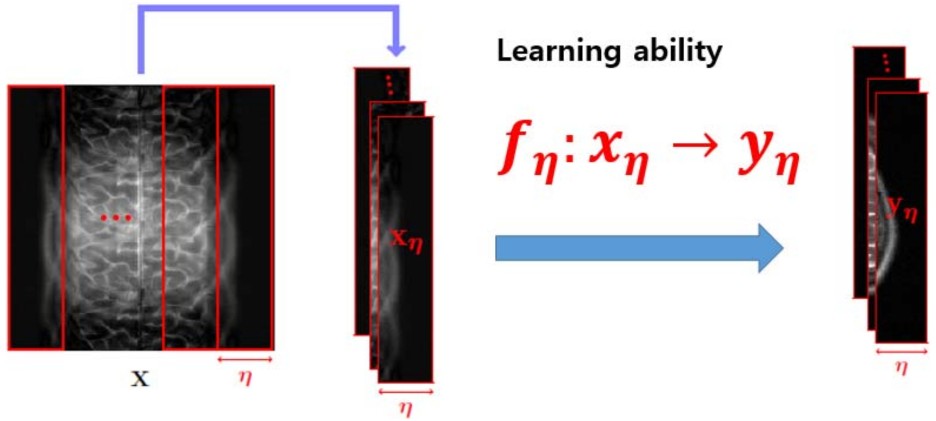
$M_{image}^{\eta} = \{ y_\eta : y_\eta$ is a $256 \times \eta$ image patch extracted from $y \in M_{image} \}$

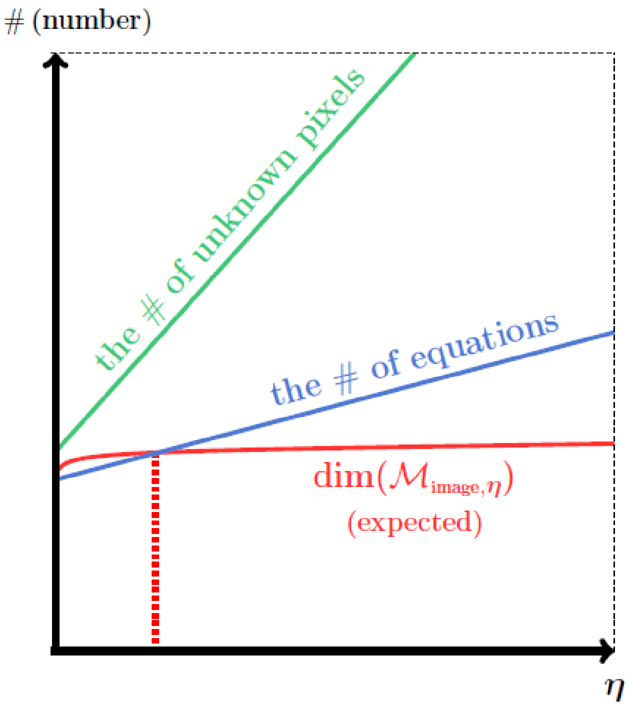
My personal opinion

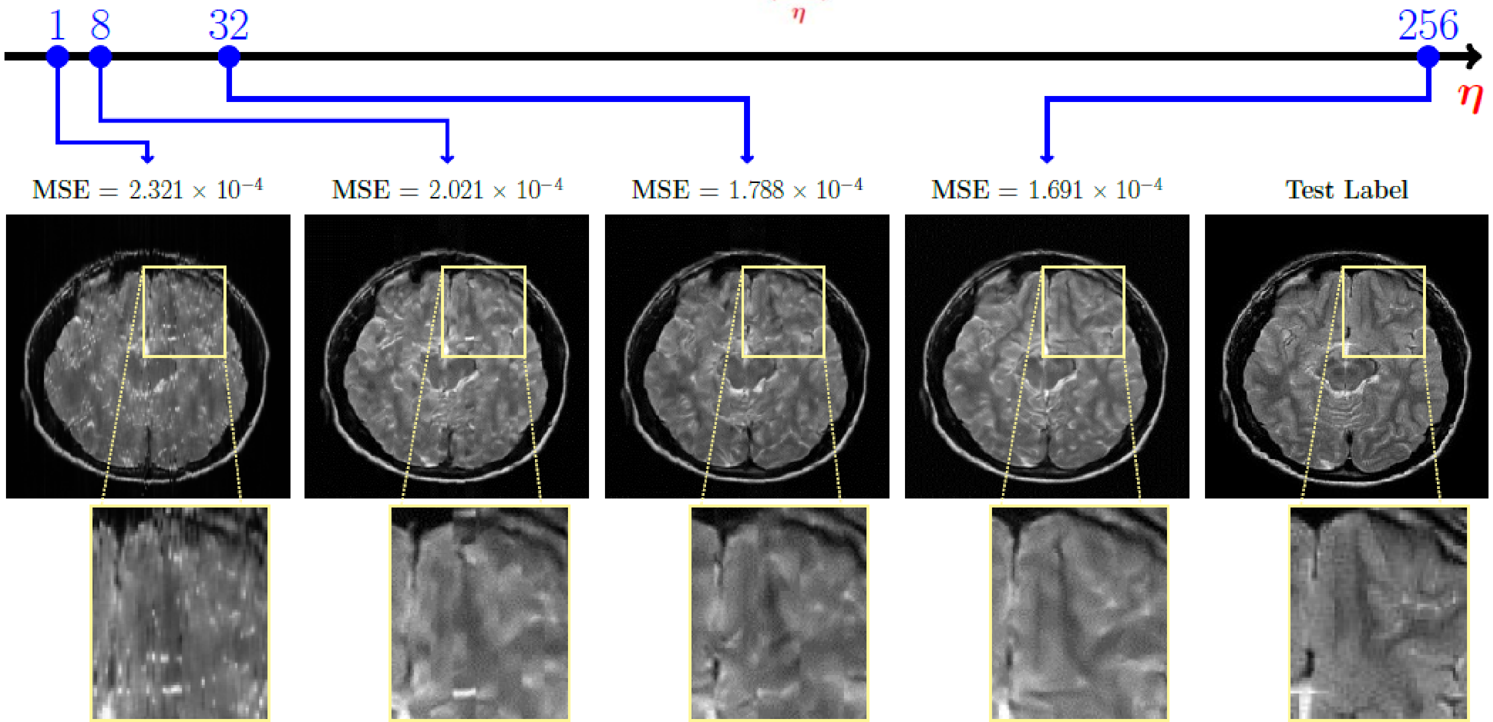
- ✓ Dimension of the manifold M_{image}^{η} does not increase proportionally to η.
- ✓ Hence, the learning ability about $f_\eta : x_\eta \rightarrow y_\eta$ is gradually improved as η increases.

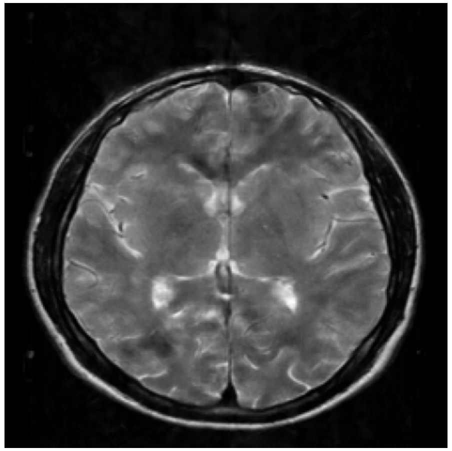
Experimental results demonstrate that the learning ability about $f_\eta : x_\eta \rightarrow y_\eta$ is gradually improved as η increases.

.



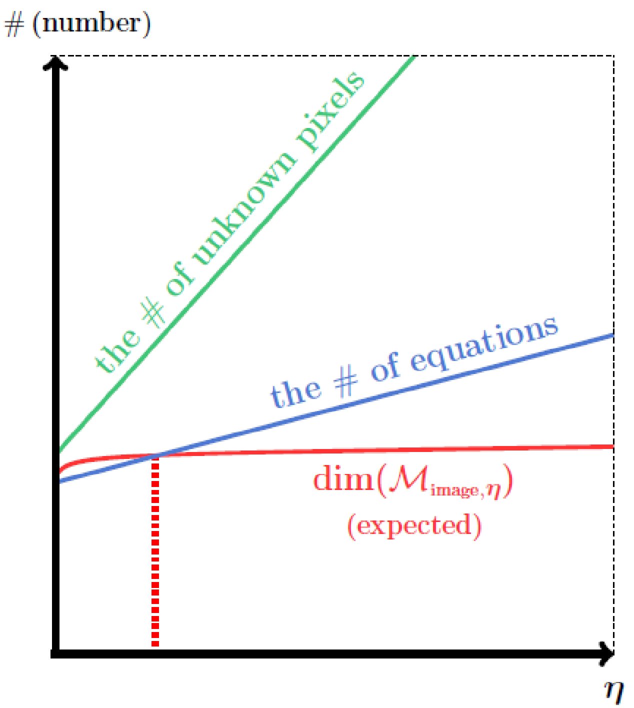






Reasons for expecting dim M^{η}_{image} to grow significantly slowly as η increases.

$M^{\eta}_{image} = \{\, y_\eta : y_\eta$ is a $256 \times \eta$ image patch extracted from $y \in M_{image}\}$



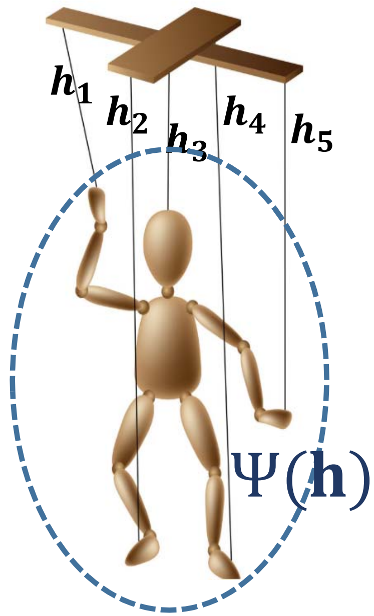
- ✓ Assume that M_{image} is the set of all the human head MR images.
- ✓ Then, all the images in M_{image} possess a similar anatomical structure that consists of skull, gray matter, white matter, cerebellum, among others.
- ✓ In addition, every skull and tissue in the image have distinct features that can be represented nonlinearly by a relatively small number of latent variables, and so does for the entire image.
- ✓ Notably, the skull and tissues of the image are spatially interconnected, and even if a part of the image is missing, the missing part can be recovered with the help of the surrounding image information.

Challenging Issue: Low-dimensional representation of MR and CT images (high dimensional data: $512 \times 512 \times 400$ voxels)

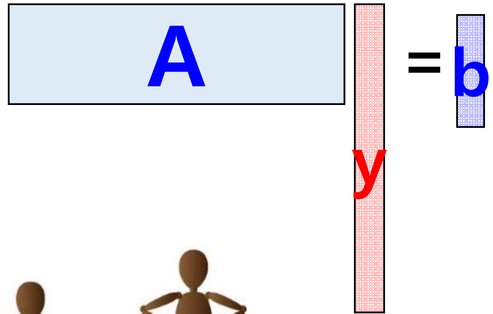
GAN (Generative Adversarial Network)

VAE (Variational Autoencoder)

Given data distributions $\{y^{(n)}: n = 1, \cdots, N\}$ in medical images (e.g. dental CBCT data), can we find a low dimensional latent generator (decoder) $\Psi: \mathrm{h} \to y$ and an encoder $\Phi : y \to h$ such that $\Psi \circ \Phi(y) \approx y$ for all $y \in M_{image}$.



One of challenging issues for solving an ill-posed problem is to find a low-dimensional representation.



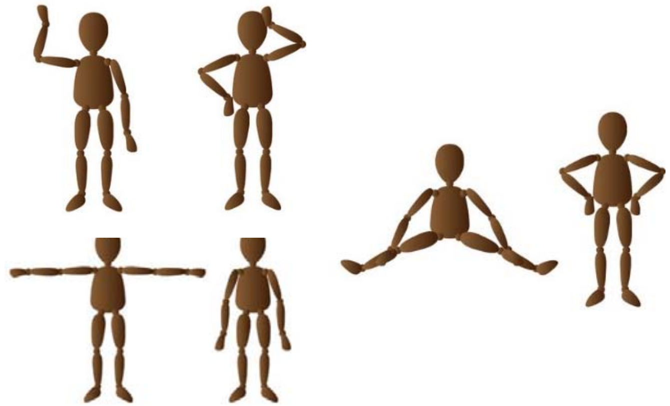
5 Latent variables

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

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

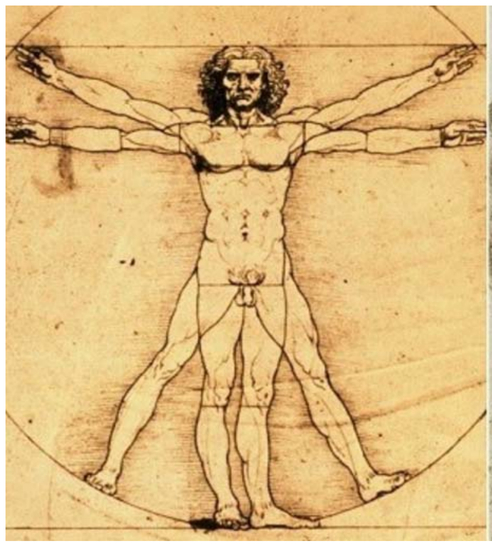
Generator /decoder

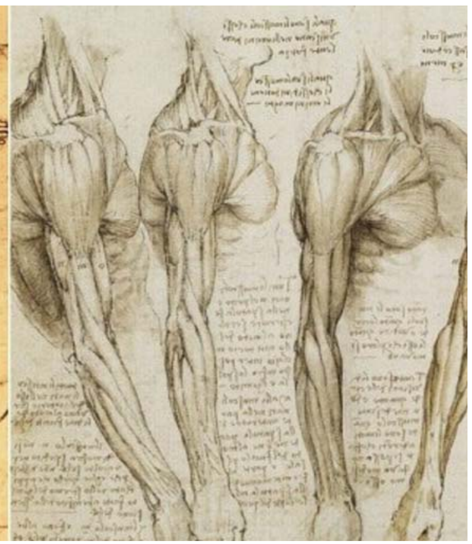
Latent variable



??

Disentangled expression with extracting the underlying explanatory axis





Example

Electrical Impedance Tomography is known to be a highly ill-posed problem.

16384

208

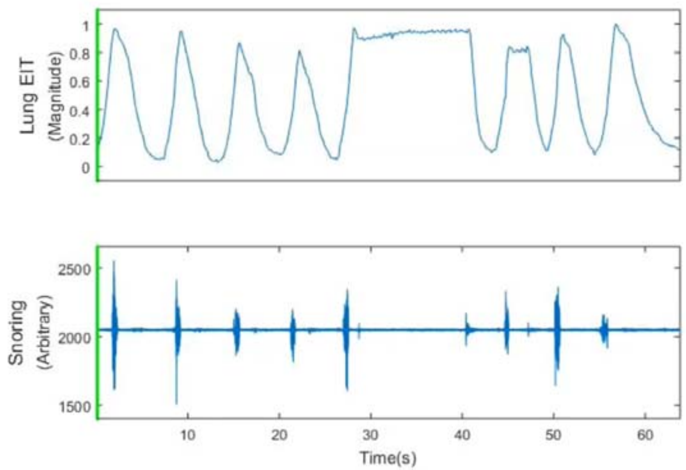
$$A = \frac{db}{dy} \quad \text{(sensitivity matrix)}$$

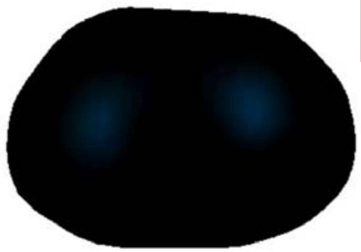
y $\approx$ b

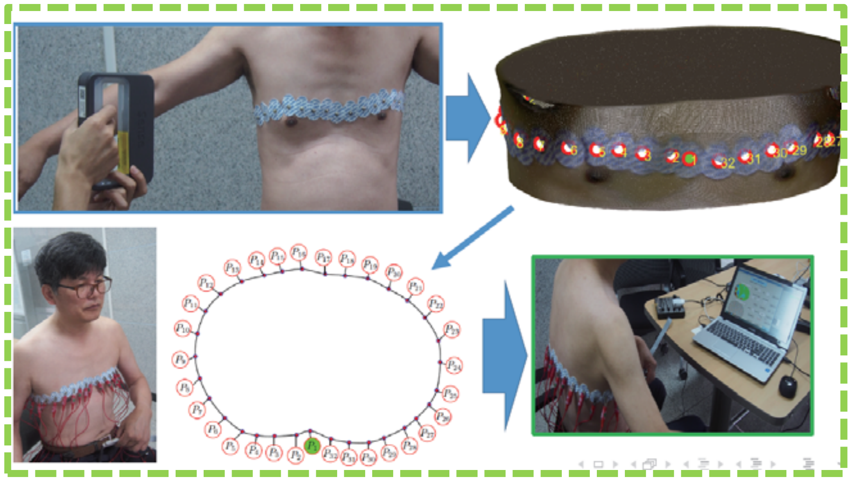
$$dim\{y \in \mathbb{R}^{16384} : Ay = 0\} > 16384 - 208$$

However, it can be well-posed if we give up excessive ambition or find a way to make a low dimensional expression.

Data acquisition





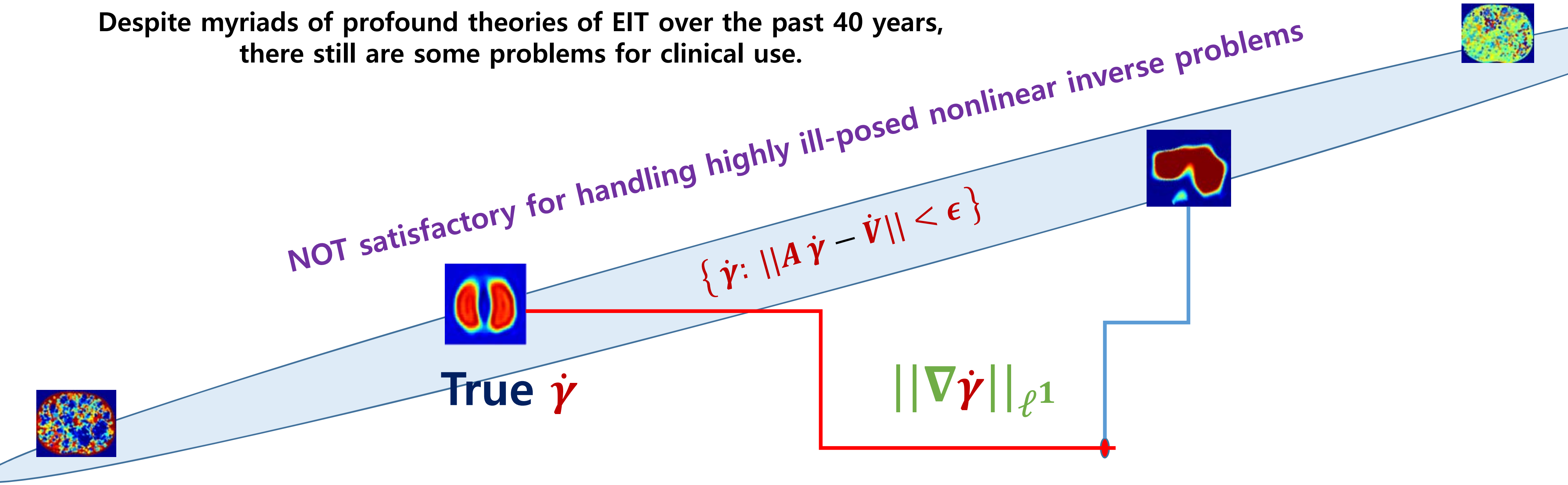


Hand-made regularization techniques may not be effective for EIT imaging.

$$\dot{\gamma} = \underset{\dot{\gamma}}{\operatorname{argmin}} || \boldsymbol{A}\,\dot{\gamma} - \dot{\boldsymbol{V}} || \; + \lambda\, \boldsymbol{Reg}(\dot{\gamma})$$

(L^2, L^1, TV regularization)

Despite myriads of profound theories of EIT over the past 40 years,
there still are some problems for clinical use.



16384

208

$$A = \frac{db}{dy} \quad \text{(sensitivity matrix)}$$

$y \approx b$

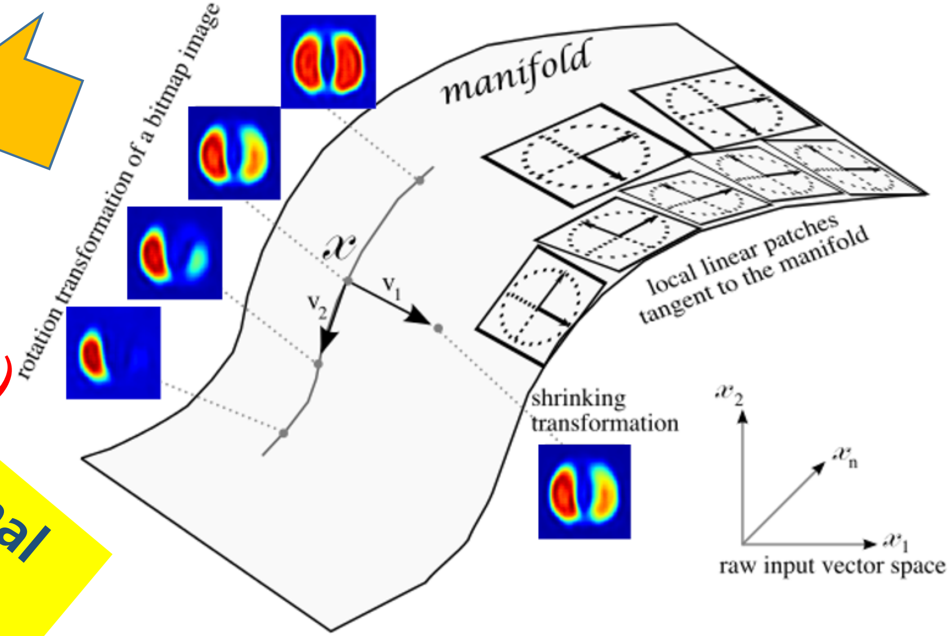
$$dim\{y \in \mathbb{R}^{16384} : Ay = 0\} > 16384 - 208$$

208=# of equations (data) ≪
16384=# of unknowns (pixels of image).

This can be well-posed if we can find a low dimensional representation of soltuions.

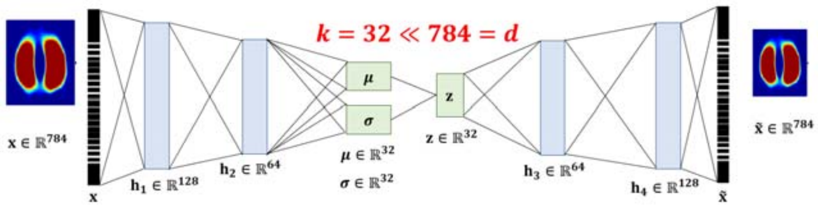
$y = \Psi(h)$

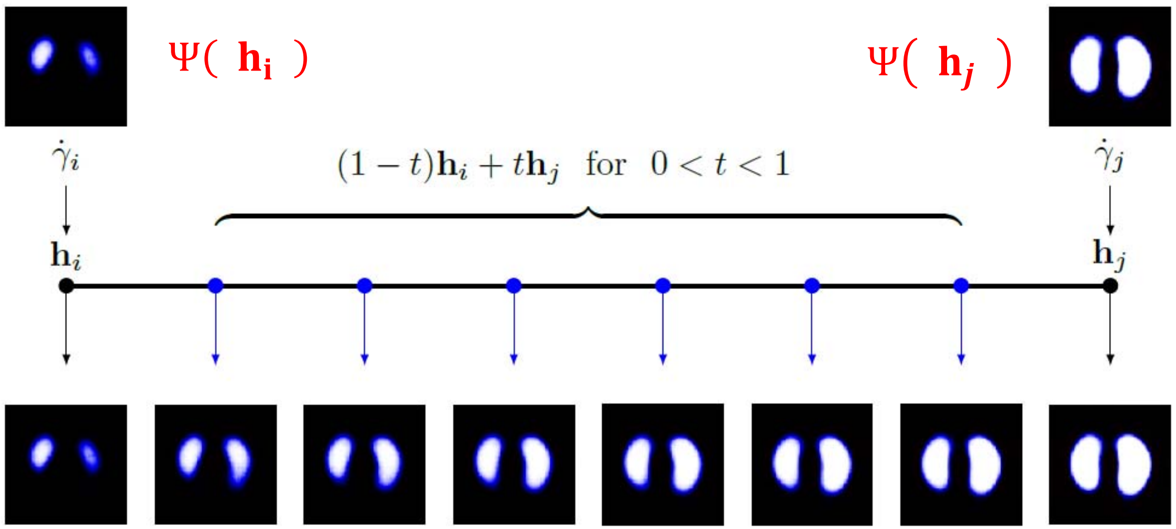
Learn low-dimensional latent representation



Deep learning framework may provide a nonlinear regression on training data which acts as learning complex prior knowledge on the output.

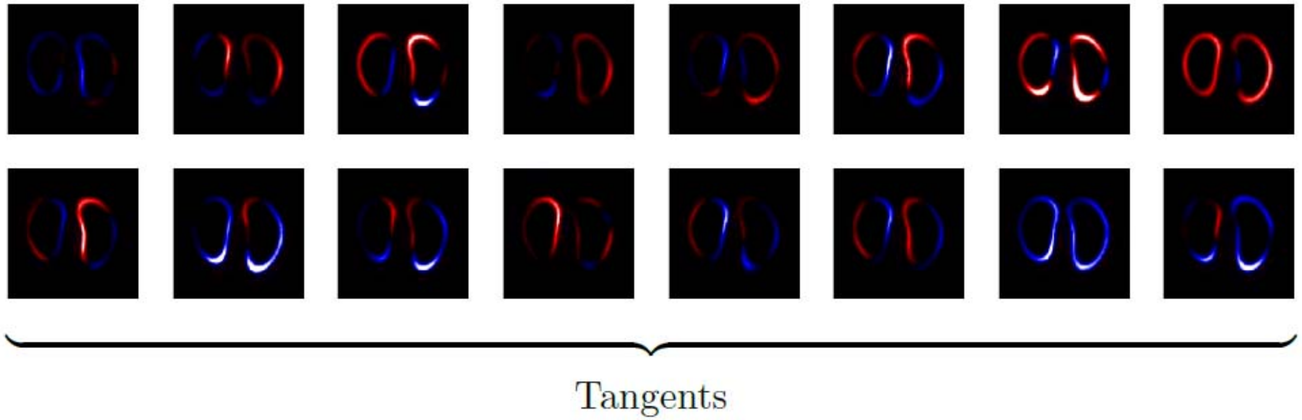
Low-dimensional latent representation produces $\mathcal{M}$anifold.



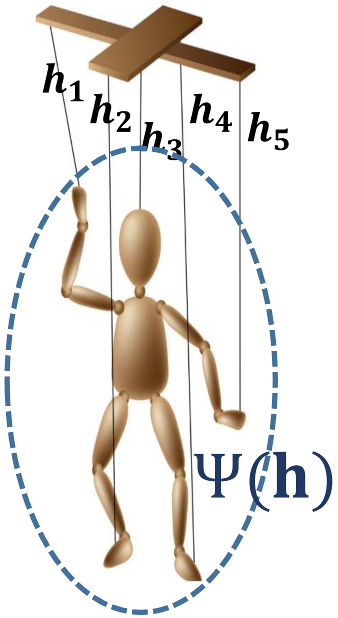


- Interpolation between two points $\mathbf{h}_i$ and $\mathbf{h}_j$ in the latent space. Between the two given images, VAE can generate the interpolated image.

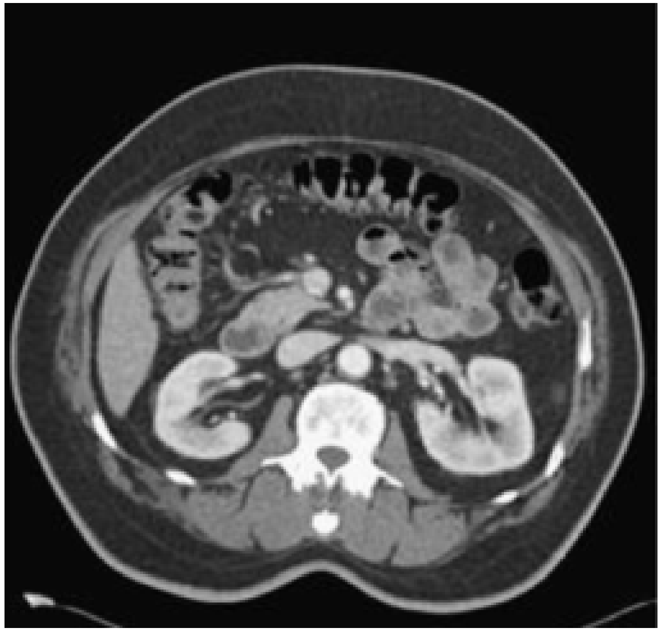
- Tangent vectors on manifold $\mathcal{M}$



What about low-dimensional representation of high dimensional images such as MR and CT images.

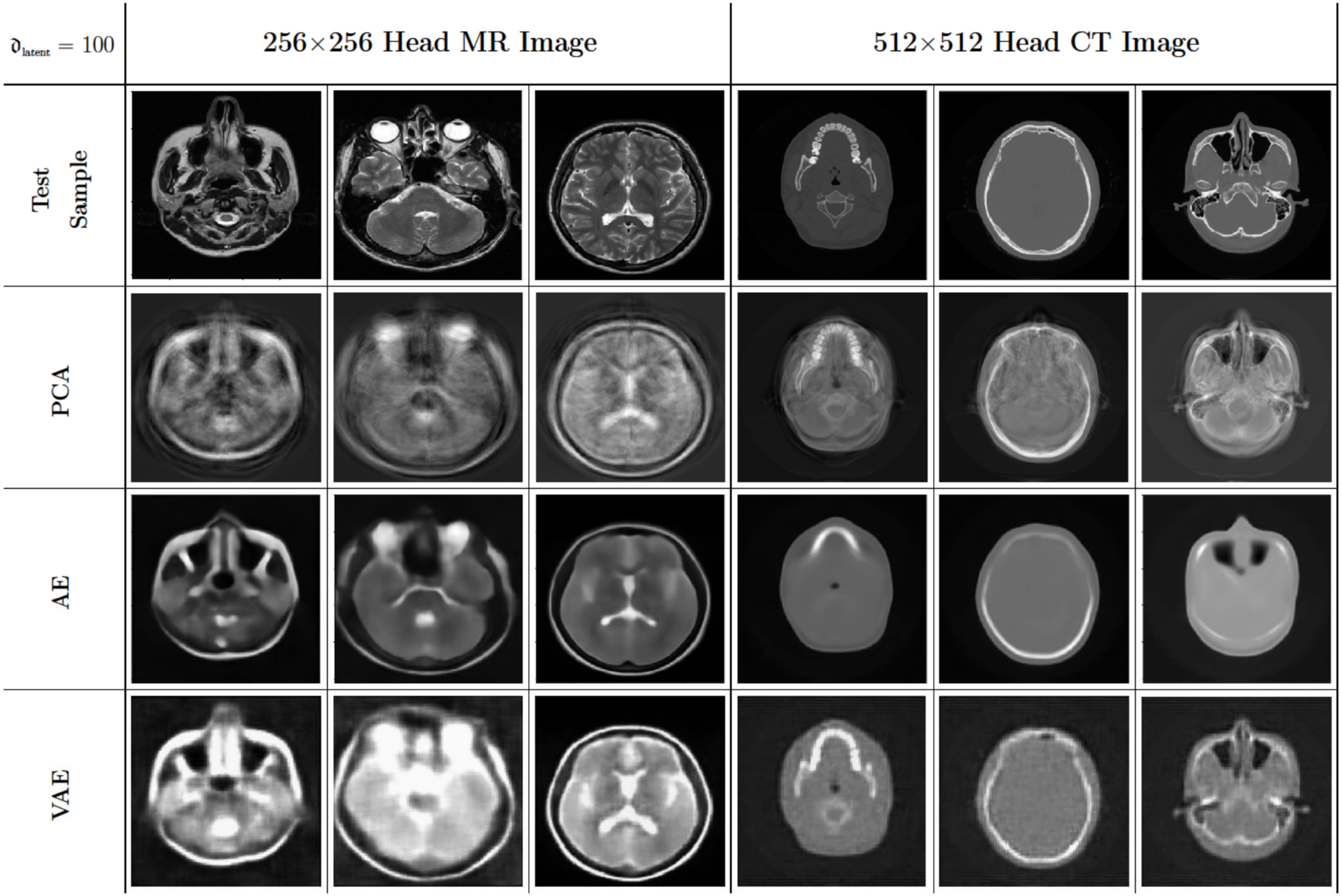


$$\Psi(\mathbf{h}) =$$



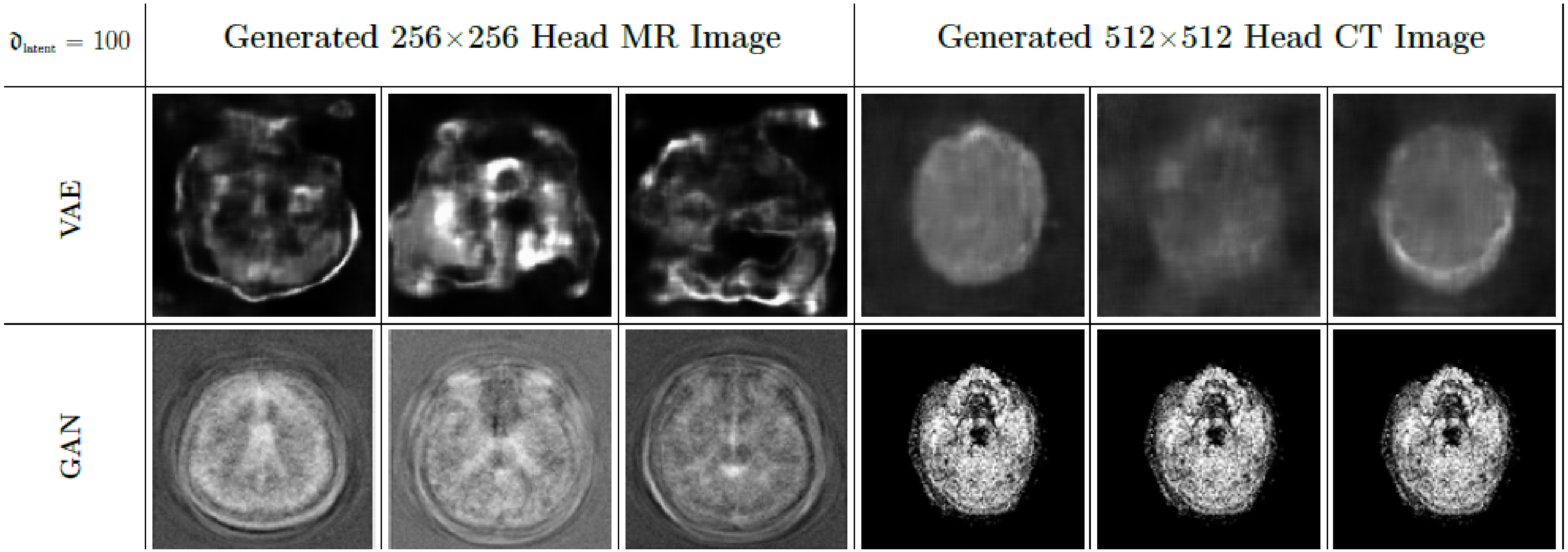
So far, my team has tried several kinds of GANs and VAE, but has not succeeded.

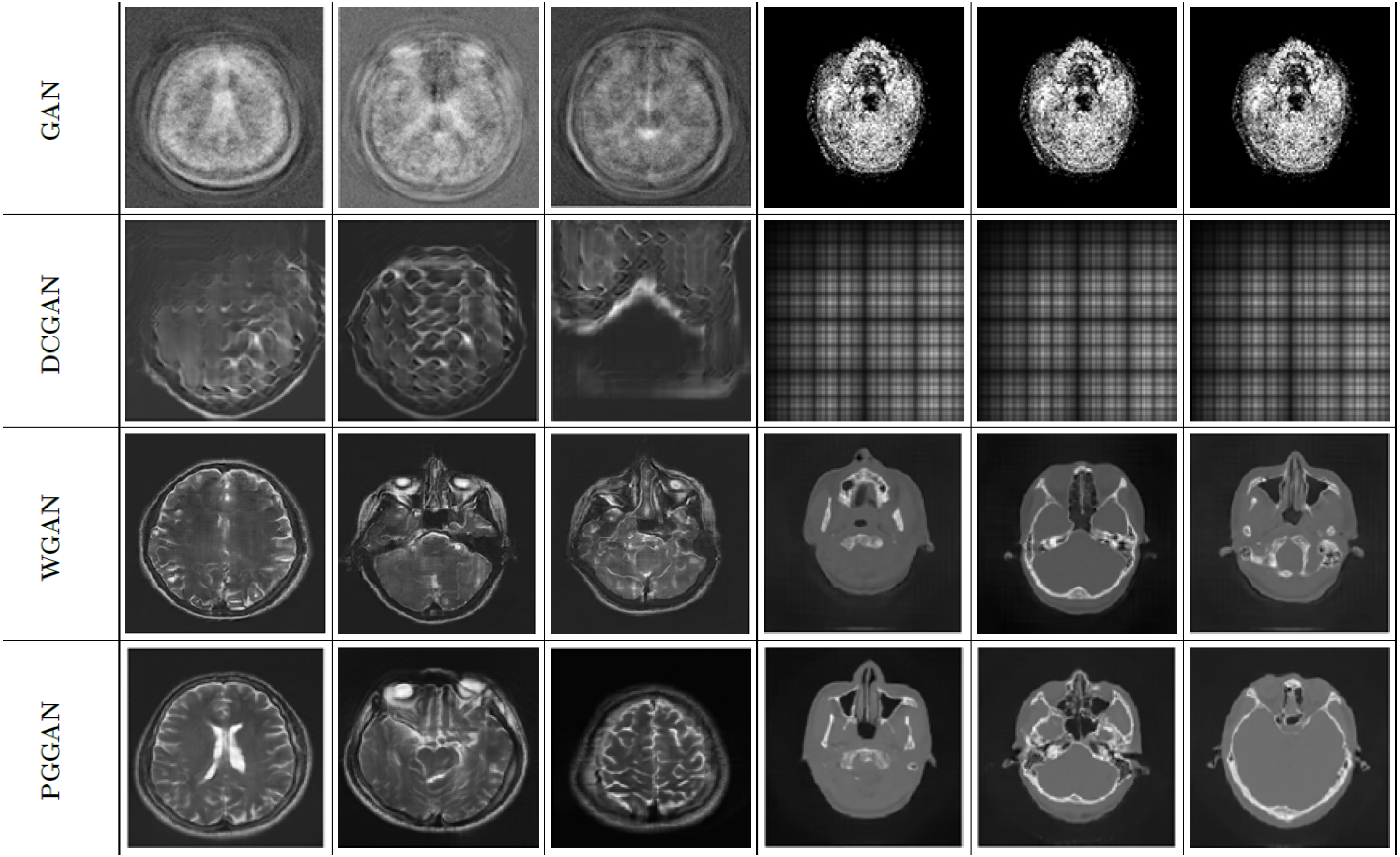
For high dimensional data, AEs suffer from image blurring and loss of small details.



Generative Adversarial Network

- ✓ GANs have shown remarkable success in generation of various realistic images. However, there exist some limitations in synthesizing high resolution medical data.
- ✓ The GAN's approach makes it difficult to deal with high-dimensional data because the generated image can be easily distinguished from the training data, which can lead to collapse or instability during training process.




GAN
DCGAN
WGAN
PGGAN

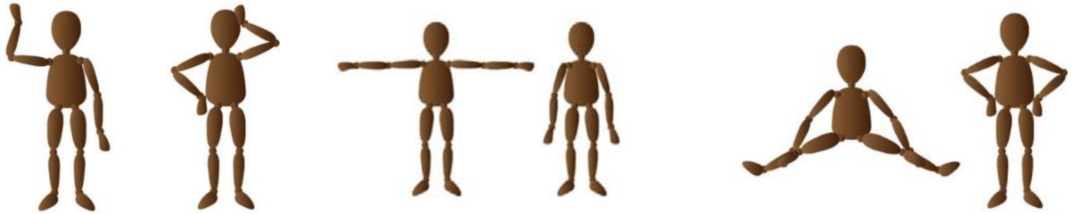
My personal opinion

AEs learns a bidirectional mapping(encoder and decoder), while GANs learn only the unidirectional mapping (decoding) in high dimensional medical images.

AE can control this latent variables

However, for high dimensional data, AEs suffer from image blurring and loss of small details.

GANs have a remarkable ability to generate these images.



GANs have difficulties in encoding high dimensional images.

Challenging Issue: Generalization

Training error ≈ 0

$$\sum_k ||y_k - f(x_k)||^2 \approx 0$$

Memorize learning materials well

Hope

Recognize and generalize features

Test error ≈ 0

$$y_{test} - f(x_{test}) \approx 0$$

Problems that do not appear in the tutorial also find the correct answer.

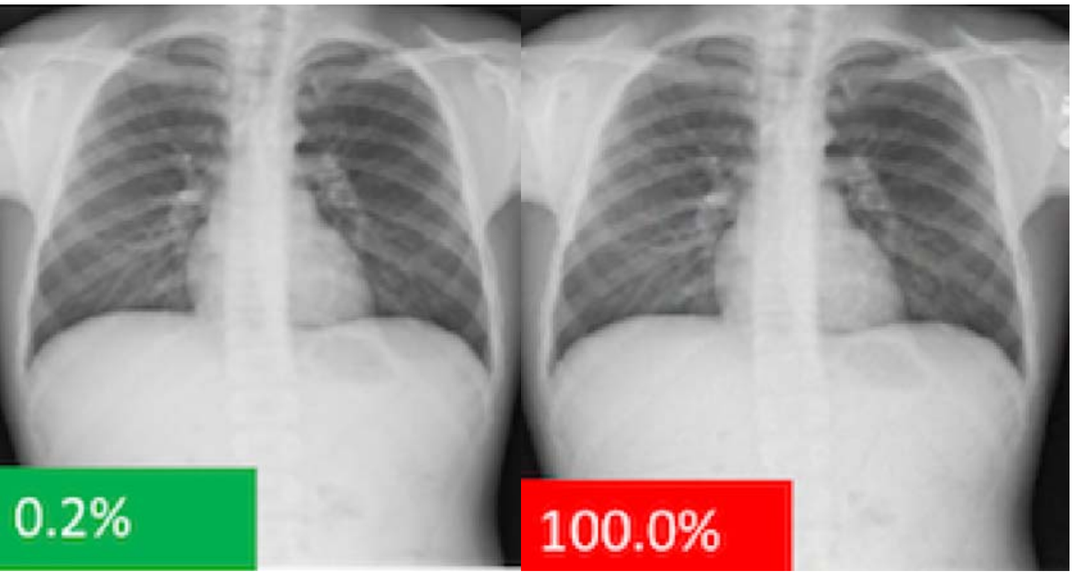
Example of Memorization without Generalization

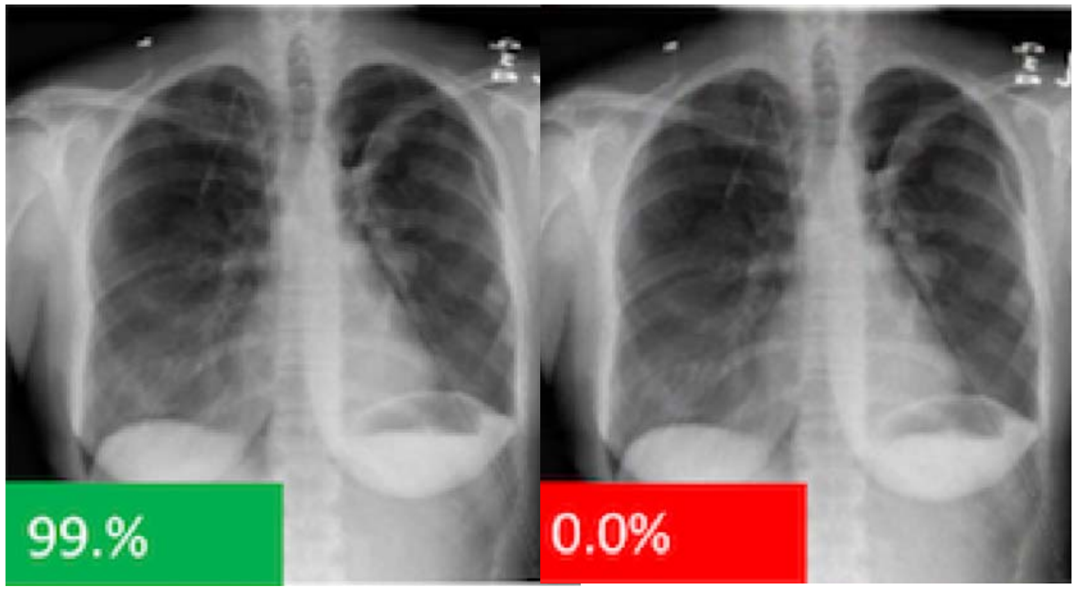
Recently, several experiments regarding adversarial classifications (false positive output of cancer) have shown that deep neural networks (obtained via gradient descent-based error minimization procedure) are vulnerable to various noisy-like perturbations, resulting in incorrect output (that can be critical in medical environments).

Adversarial attacks against medical deep learning systems

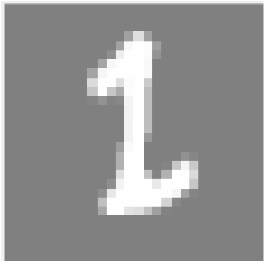
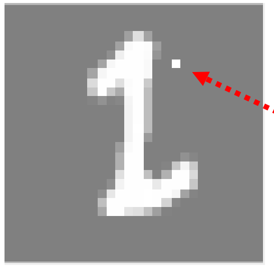
by Samuel G. Finlayson et al (2018)

The percentage represents the probability of Pneumothorax.



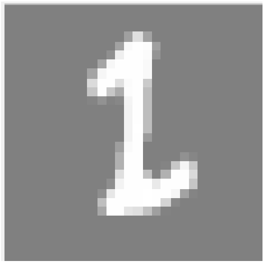
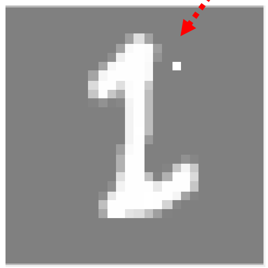


MNIST example of Memorization without Generalization

$dist_{human}($  ,  $) = 0$

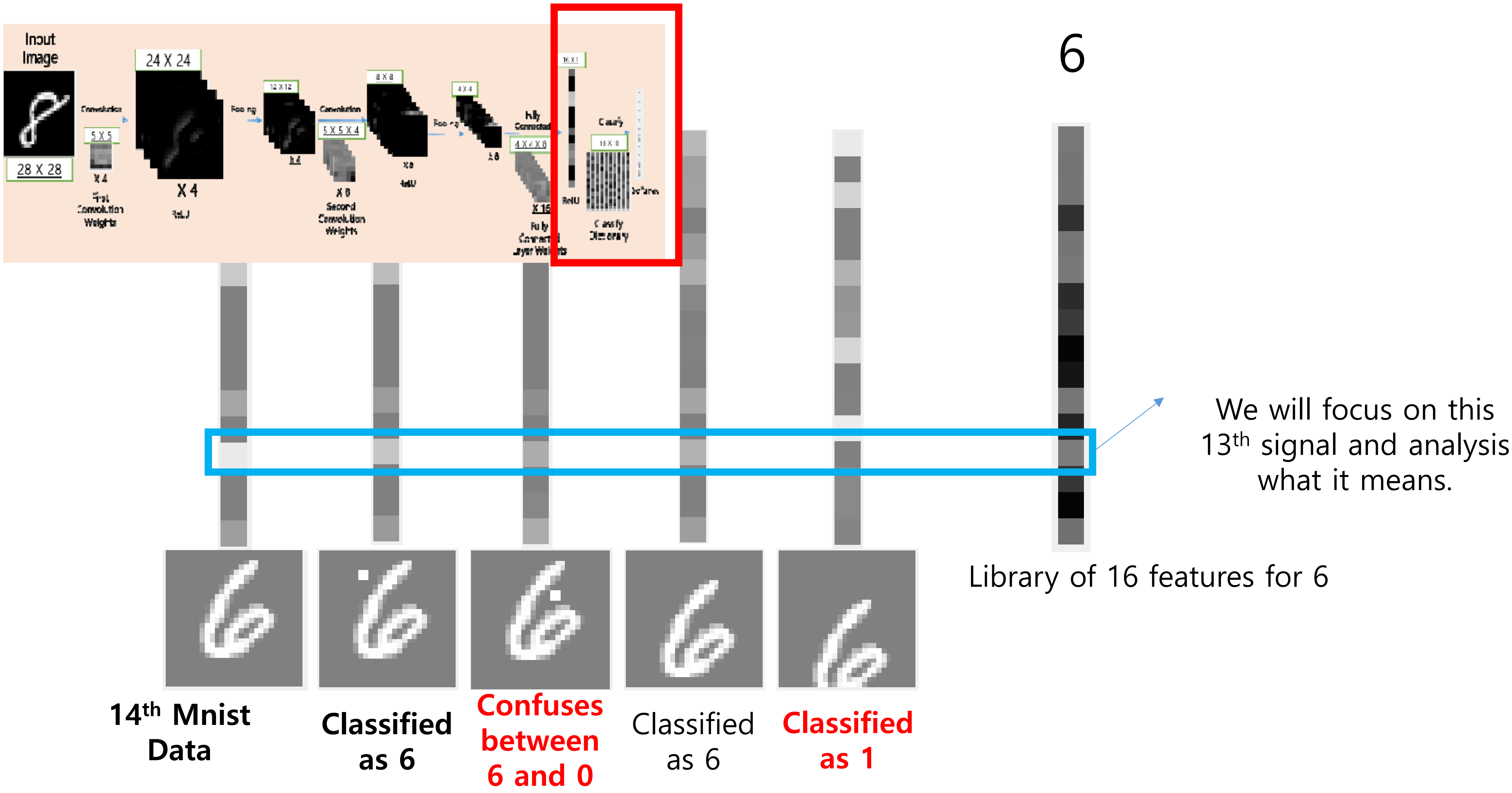
One pixel attack

However, deep learning may provide

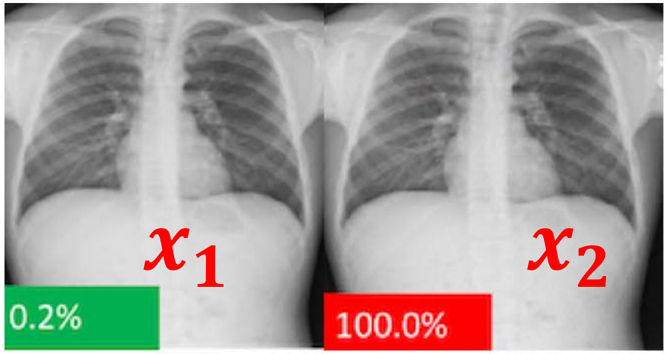
$f($  $) = 1 \neq 8 = f($  $)$

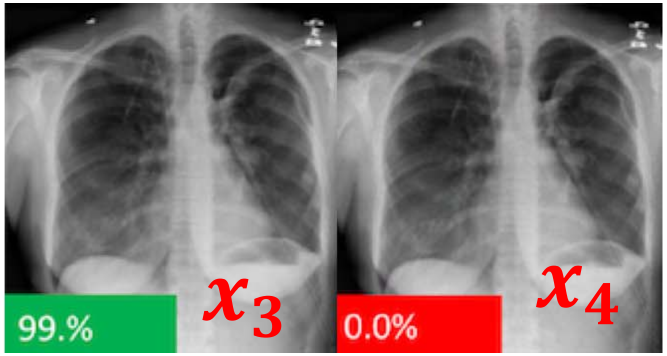
$$\mathbf{f}(x) = \boldsymbol{\sigma}\left(\boldsymbol{W}^{L} \otimes \left(\boldsymbol{\sigma} \circ \boldsymbol{P} \circ \boldsymbol{\sigma}\left(\boldsymbol{W}^{L-1} \otimes \left(\cdots\cdots \boldsymbol{\sigma} \circ \boldsymbol{P} \circ \boldsymbol{\sigma}\left(\boldsymbol{W}^{1} \otimes \boldsymbol{x} + \boldsymbol{b}_1\right)\cdots\cdots\right) + \boldsymbol{b}_{L-1}\right) + \boldsymbol{b}_L\right)$$

Adversarial attacks against MNIST handwritten classification



Challenging issue: Normalization of input data





Adversarial attacks against medical deep learning systems
by Samuel G. Finlayson et al (2018)

$$dist_{radiologist}(x_1, x_2) = 0 \;\;\&\;\; dist_{radiologist}(x_3, x_4) = 0$$

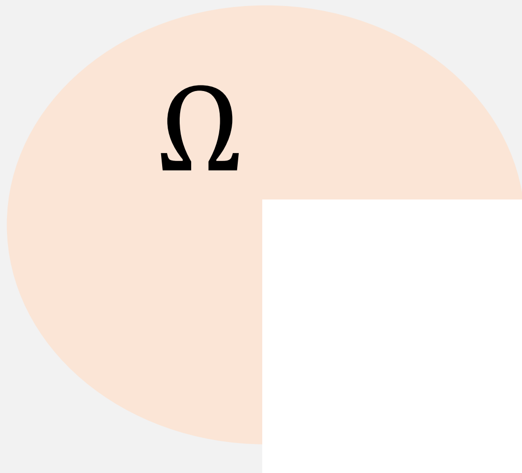
- ✓ These adversarial examples show that a well-trained function $f: x \rightarrow y$ works only in the immediate vicinity of a manifold, whereas producing incorrect results if the input deviates even slightly from the training data manifold.
- ✓ In practice, the measured data is exposed to various noise sources such as machine dependent noise; therefore, the developed algorithm must be stable against the perturbations due to noise sources.
- ✓ Hence, normalization of the input data is essential for improving robustness and generalizability of the deep learning network against adversarial attacks.

Final Remark: Historically, our mathematicians have tried to find well-posed model by imposing appropriate constraints to solution spaces. In the simple Dirichlet problem, it took decades to find the appropriate space $W^{1,2}(\Omega)$. It can take decades to solve the challenging problems in DL .

The Dirichlet problem may not be well-posed without the constraint $W^{1,2}(\Omega)$.

$$Au = b \iff \begin{cases} \nabla \cdot \nabla u = 0 \quad \text{in } \Omega \\ u\big|_{\partial\Omega} = b \end{cases}$$

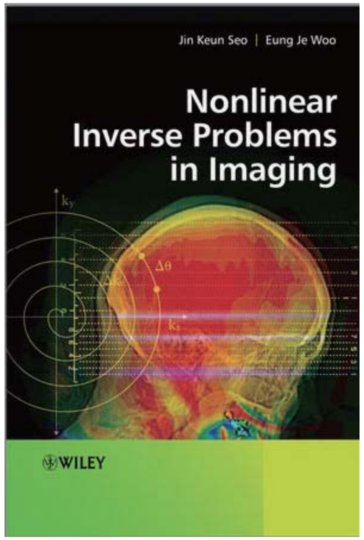
$$\Omega = \left\{ (r, \theta): 0 < r < 1, 0 < \theta < \frac{3}{2}\pi \right\}$$

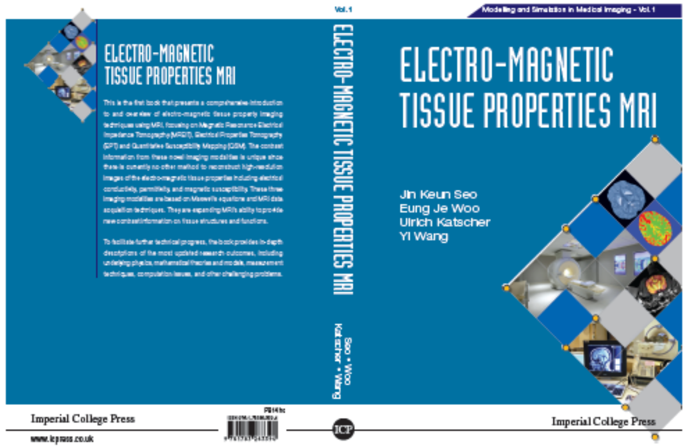


- ✓ **Without the constraint $M = W^{1,2}(\Omega)$, $Au = 0$ has infinitely many solutions in $C^\infty(\Omega)$: $u(r,\theta) = \left(r^{\frac{2n}{3}} - r^{-\frac{2n}{3}}\right)\sin\frac{2n}{3}\theta$, $n = 0, 1, 2, \cdots$**
- ✓ **Wit the constraint $M = W^{1,2}(\Omega)$, $Au = 0$ has the unique solution $u = 0$.**

In terms of M-RPI, note that $\|Au - Au'\|_{H^{\frac{1}{2}}(\partial\Omega)} \approx \|u - u'\|_{H^1(\Omega)}$ for all $u, u' \in M = \{u \in W^{1,2}(\Omega): \nabla \cdot \nabla u = 0 \text{ in } \Omega\}$.

I hope that we will discuss various challenging issues during this meeting. Thank you!





SAMSUNG 삼성미래기술육성재단 Samsung Science and Technology Foundation

BrainKorea21 PLUS

CMAC 응용해석 및 계산센터

MEDIVIEWSOFT