Stability Analysis of Data and Image Domain Learning-based Reconstruction Approaches

M. Usman Ghani and W. Clem Karl

{mughani,wckarl}@bu.edu

Department of Electrical and Computer Engineering
Boston University

BOSTON UNIVERSITY

Overview and Outline

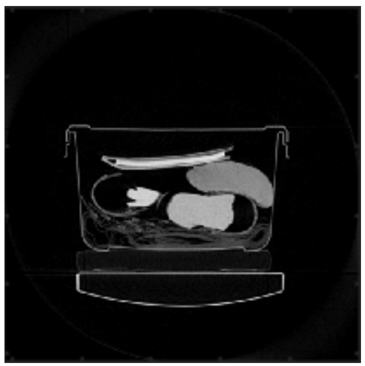
- Goal: Gain insight into robustness of some deep-learning-based reconstruction approaches

- Image Reconstruction and Learning
  - Data-domain Learning (DDL)
  - Image-domain Learning (IDL)
  - Data and Image-domain Learning (DIDL)

- Stability Analysis
  - Adversarial Perturbations
  - Random Perturbations
  - Structural Perturbations

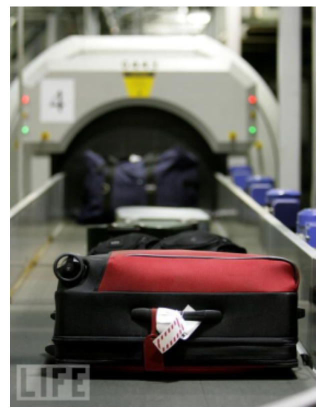
- Summary

Image Reconstruction



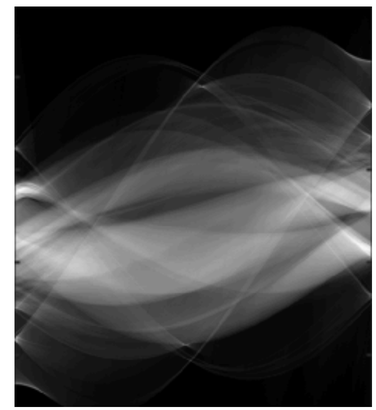
Image

x



Physical System

A



Sinogram

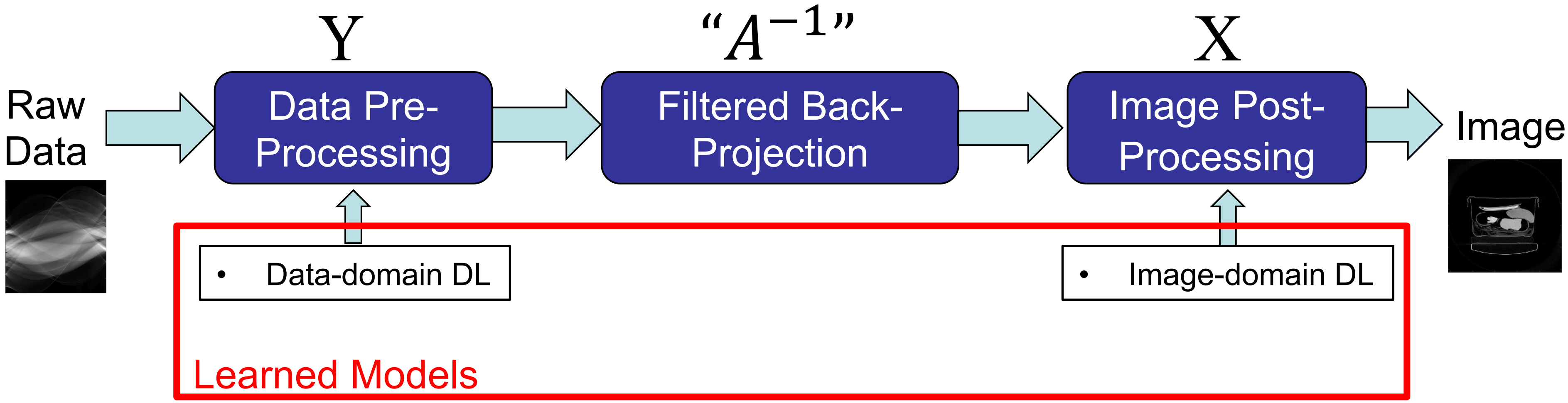
Data

y

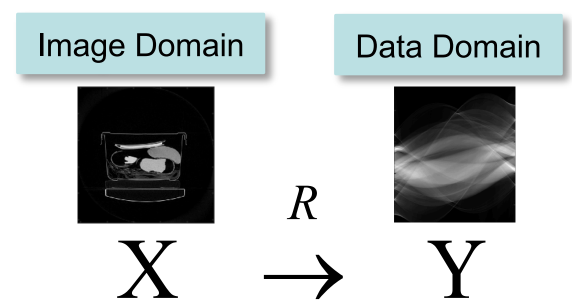
- Model: $y = Ax$ ➔ Goal: make an image x from y

BOSTON UNIVERSITY

Deep Learning for Computational Imaging



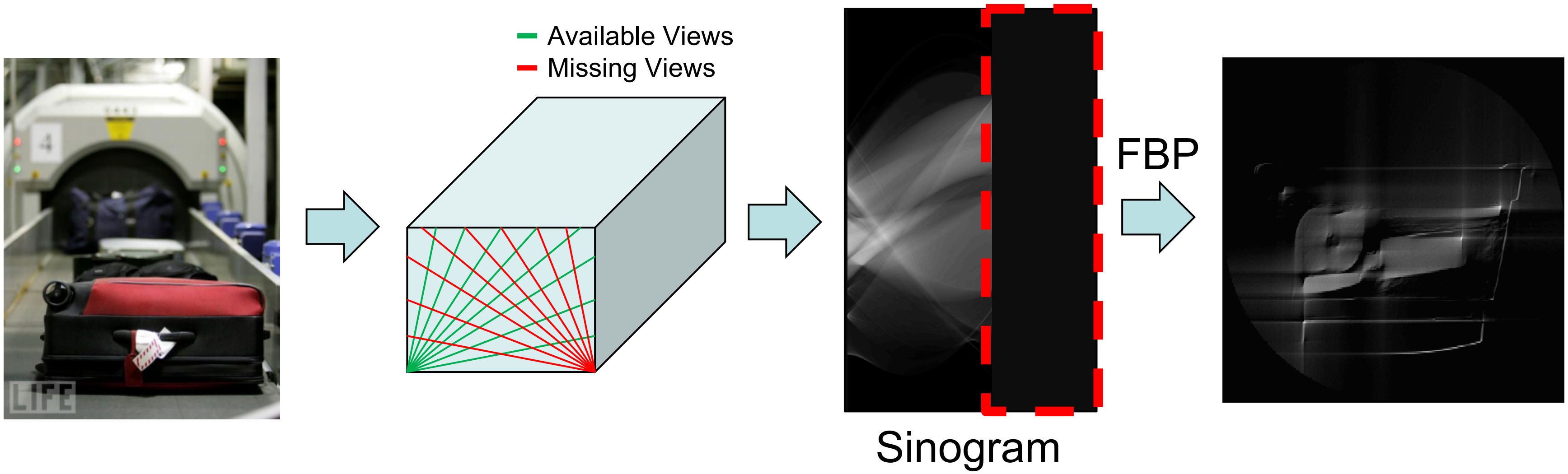
- Image Domain Learning: X→X mappings
- Data Domain Learning: Y→Y mappings



- Approaches considered:
  - Data-domain Learning (DDL)
  - Image-domain Learning (IDL)
  - Data and Image-domain Learning (DIDL)

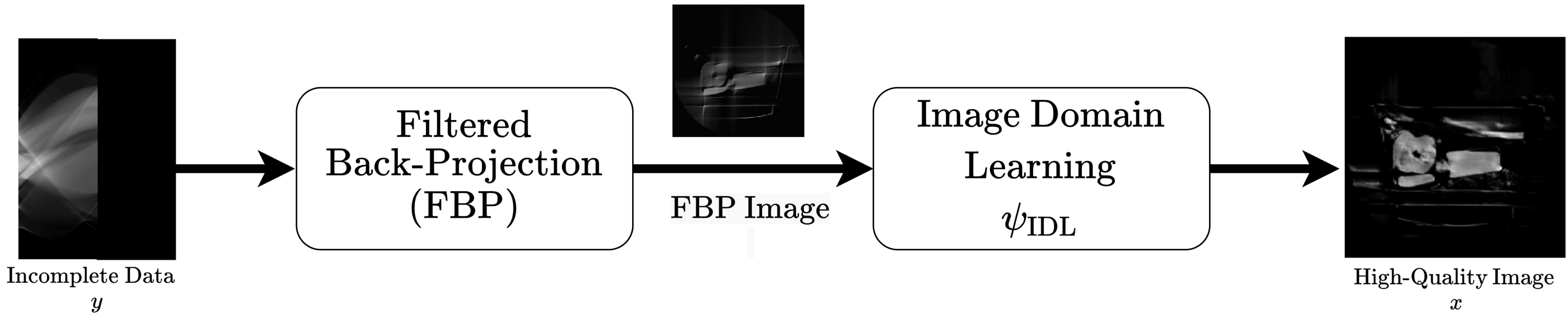
Limited Angle CT Example

- Security Systems with Non-rotational Scanning
- Imaging highly-dynamic scenes



- Available Views: $[0^0, 90^0]$
- $x =$ Reconstructed Image
- $y =$ Incomplete sinogram data

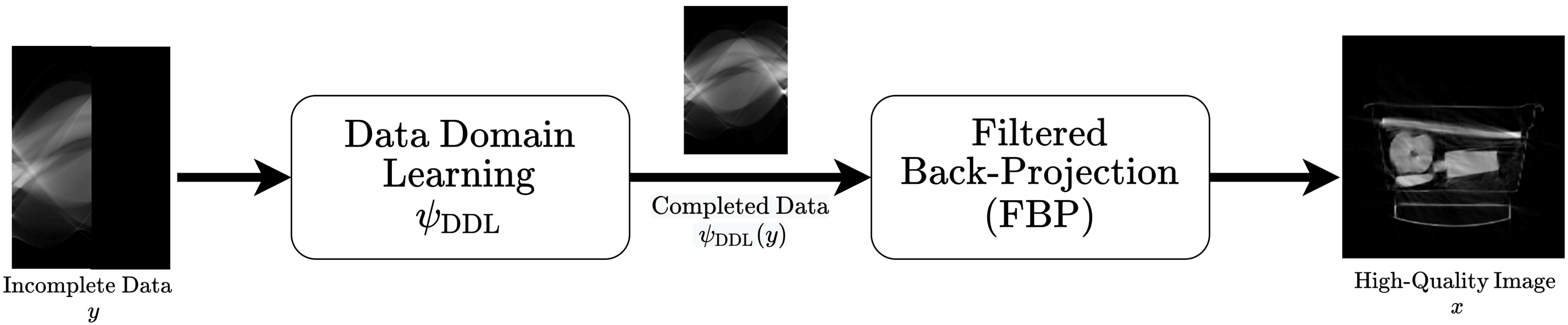
Image-Domain Learning (IDL)



- Conventional reconstruction using incomplete data
- Image-domain post-processing using DL

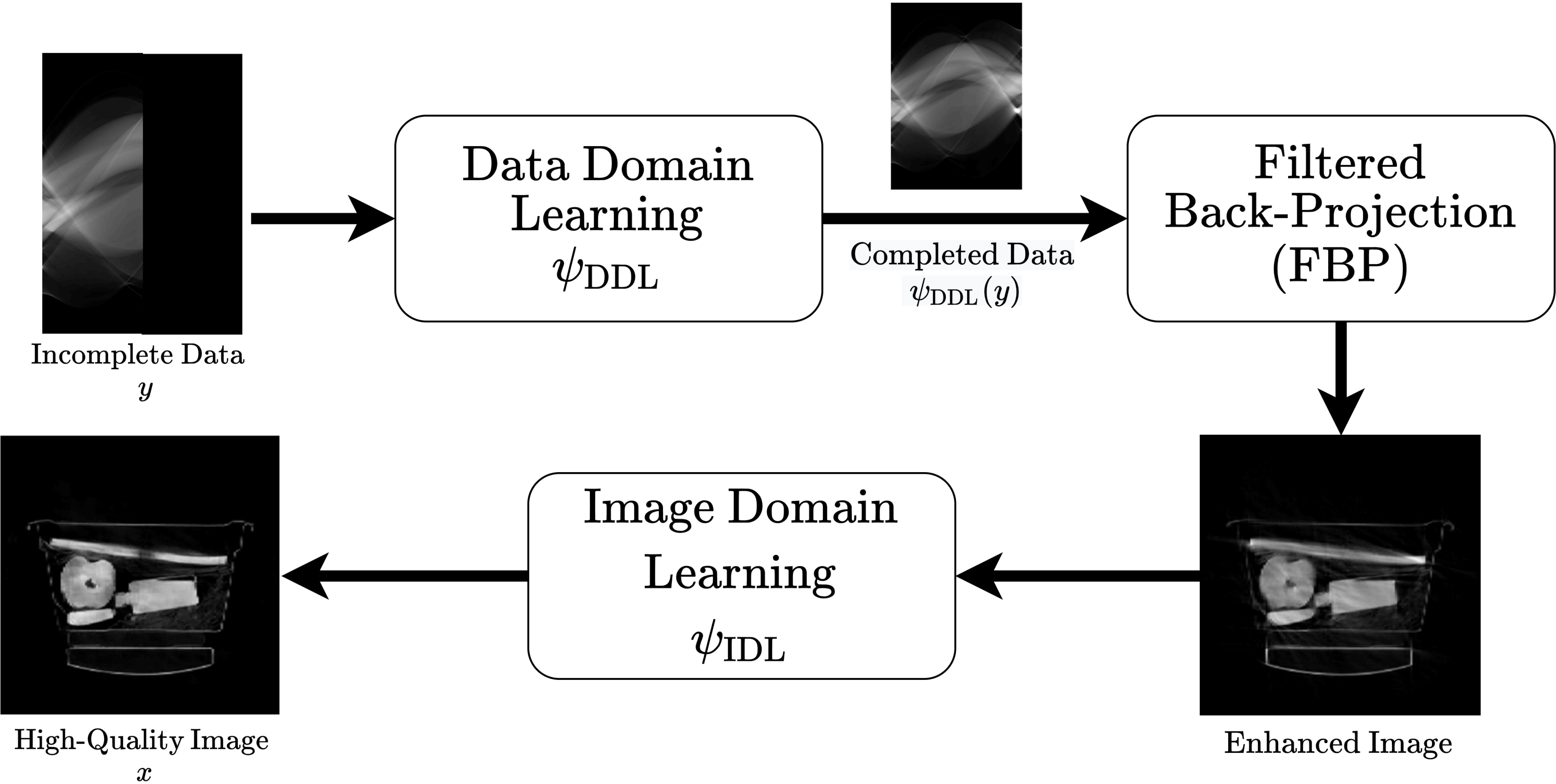
BOSTON UNIVERSITY

Data-Domain Learning (DDL)



- Data-domain pre-processing using DL
- Conventional reconstruction using completed data

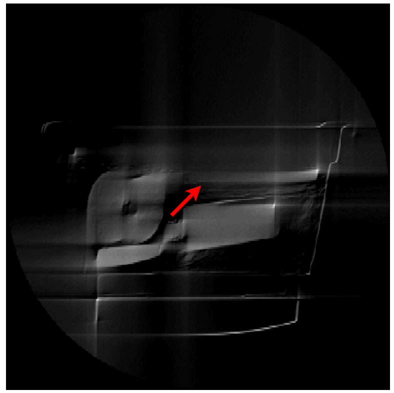
Data and Image-Domain Learning (DIDL)



BOSTON UNIVERSITY

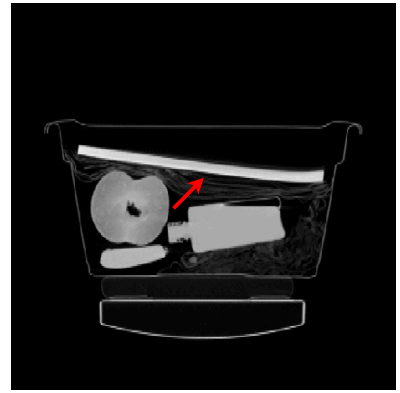
Limited-angle CT Results

FBP



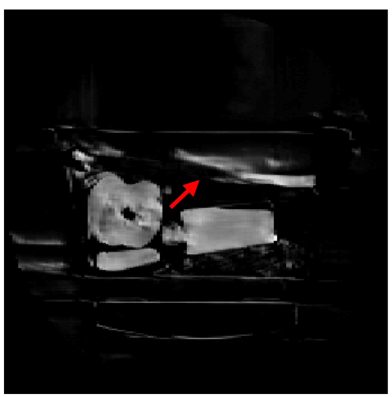
Limited Angle
Conventional Image

Reference



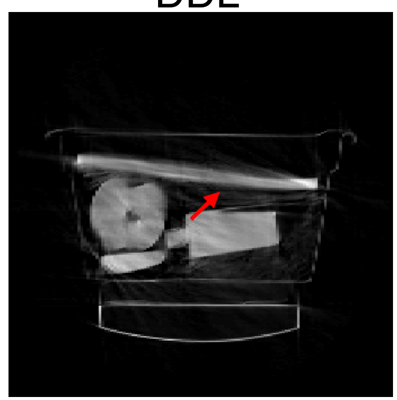
Full-view MBIR
Reconstruction

IDL



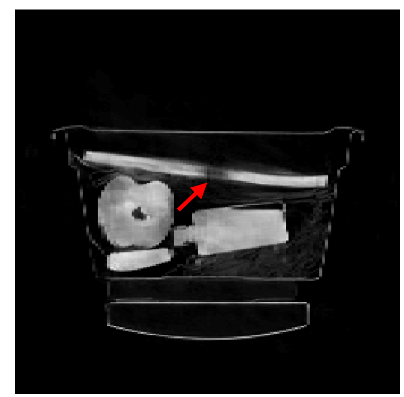
Post-processing Image-domain Learning Only

DDL



Pre-Processing Data-Domain Learning Only

DIDL



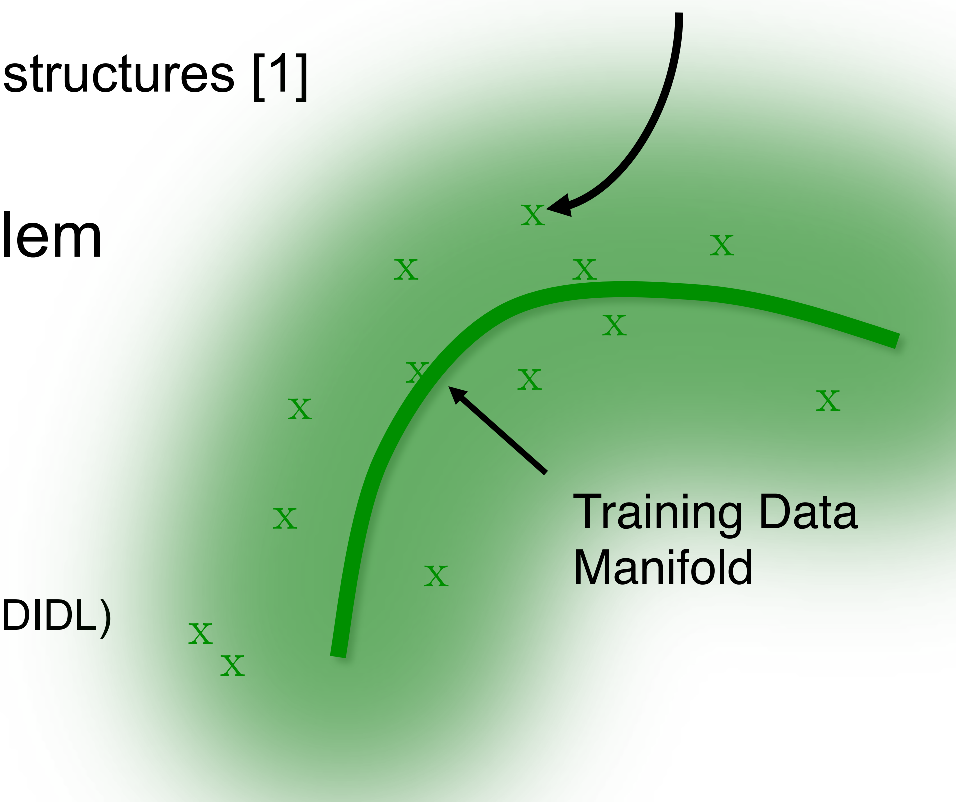
Decoupled Data and Image-Domain Learning

Stability Analysis

Stability Analysis

- 3 Perturbations Considered:
  - Adversarial: worst-case perturbations [1]
  - Random Perturbations [2]
  - Structural: small, significant structures [1]
- 90^0 Limited-angle CT problem
- Analyzed approaches
  - Data-domain Learning (DDL)
  - Image-domain Learning (IDL)
  - Data and Image-domain Learning (DIDL)

Evaluate at points away from training set Manifold



[1] Antun, V., Renna, F., Poon, C., Adcock, B., and Hansen, A. C. (2020). On instabilities of deep learning in image reconstruction and the potential costs of AI. Proceedings of the National Academy of Sciences.
[2] Gottschling, N. M., Antun, V., Adcock, B., and Hansen, A. C. (2020). The troublesome kernel: why deep learning for inverse problems is typically unstable. arXiv preprint arXiv:2001.01258.

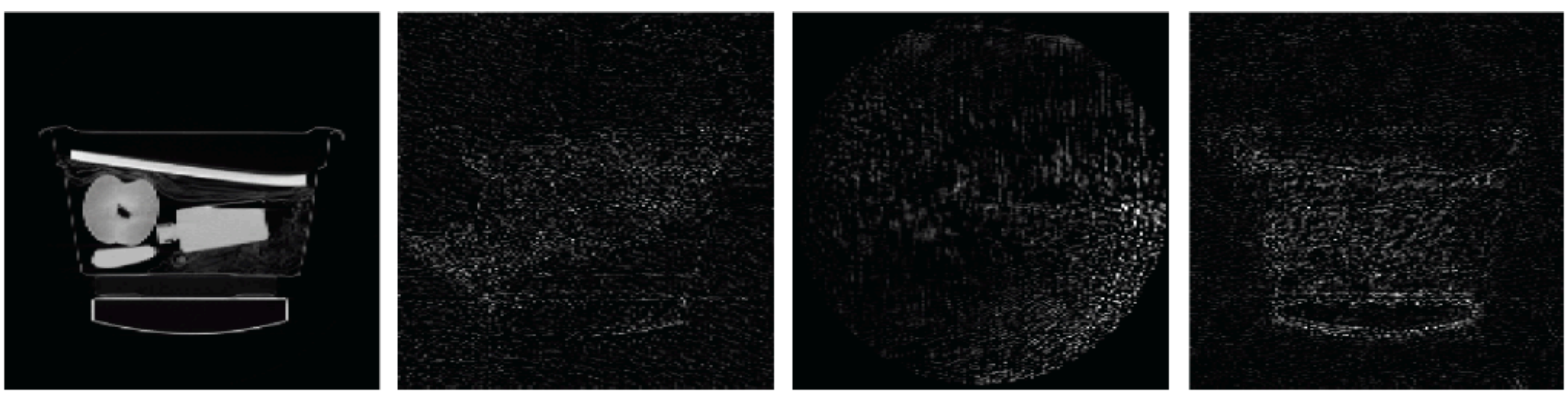
Adversarial Perturbations – Optimization

- Finding the worst-case via optimization:

$$r^* = \underset{r}{\mathrm{argmax}}\ \frac{1}{2}\overbrace{\left\|\underbrace{\psi\left(y+\boldsymbol{A}r\right)}_{\text{Perturbed Input to network}} - x_{\text{ref}}^{(\text{image})}\right\|}^{\text{Error}}{}^{2} - \frac{\lambda}{2}\|r\|^2$$

Reference Image

- Perturbations



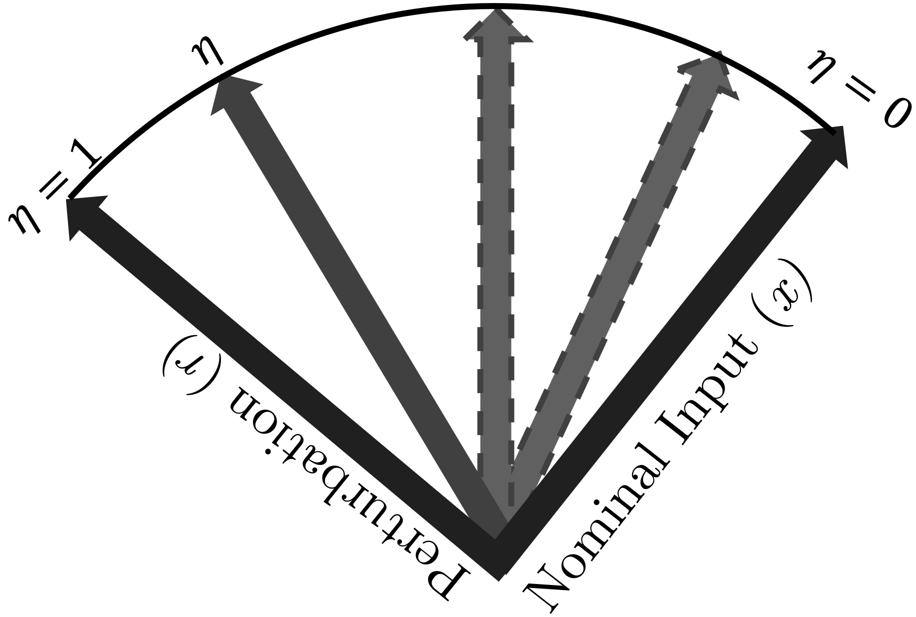
Reference

r_{IDL} Image-Domain Learning

r_{DDL} Data-Domain Learning

r_{DIDL} Data and Image Learning

Interpolate Nominal Input and Perturbation



$$\mathrm{SLERP}(x, r; \eta) = \frac{\sin[(1-\eta)\Omega]}{\sin\Omega} x + \frac{\sin[\eta\Omega]}{\sin\Omega} r$$

- η: relative perturbation contribution

BOSTON UNIVERSITY

Adversarial Perturbations – Application

Data-Domain Learning (DDL)

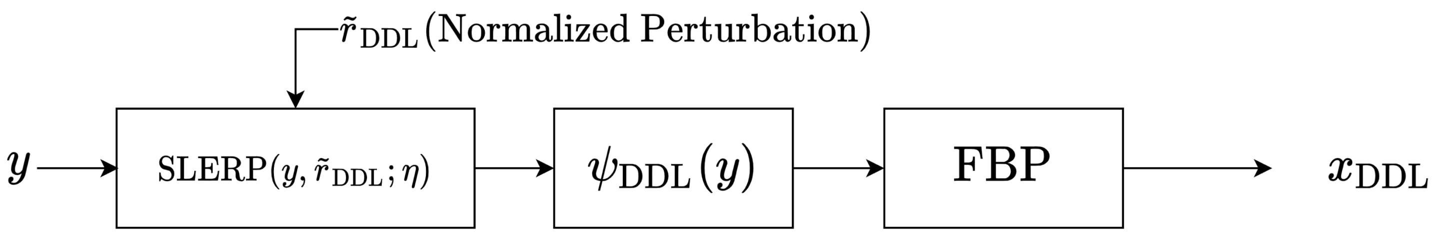
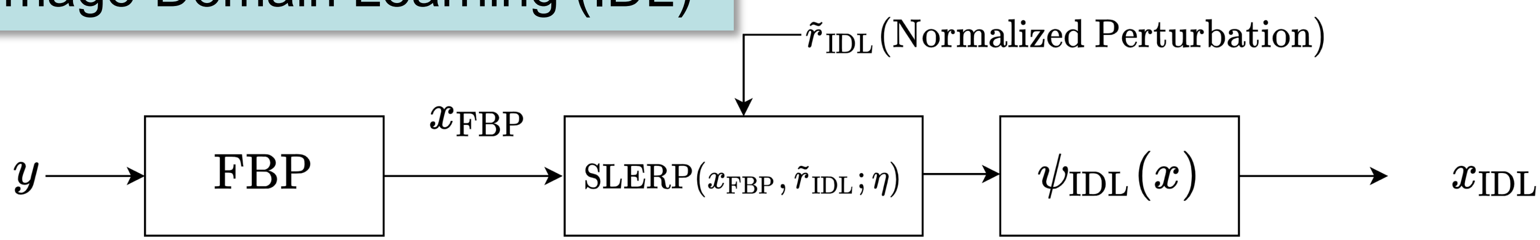
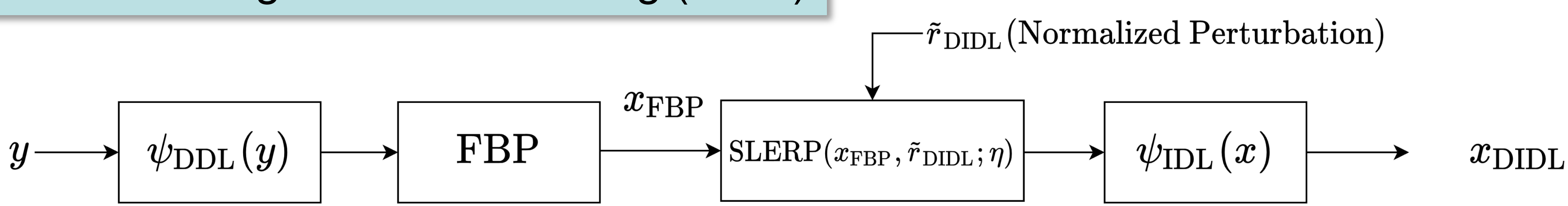


Image-Domain Learning (IDL)



Data and Image Domain Learning (DIDL)



Random and Structural Perturbations

- Random
  - 100 instances of Gaussian Noise $\sim N(0, I)$
  - Fix directions, normalize such that $\|\tilde{r}\| = \|x_{ref}\|$
  - Perturb original image $\tilde{x} = \mathrm{SLERP}(x_{ref}, \tilde{r}; \eta)$
- Structural
  - Study learned biases by using anomalous structures

• Structure from similar dataset	• Shepp-Logan Phantom	• Card suit symbols
• Small text	• Large text	• Add two images

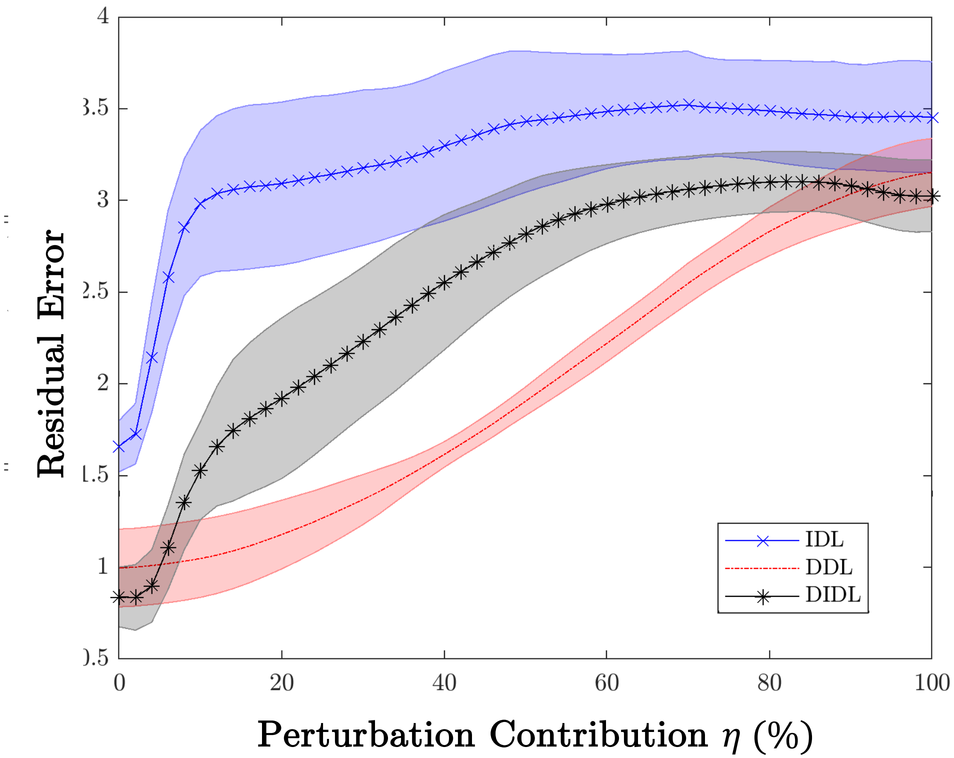
- Generated simulated observations $y = \boldsymbol{A}\,\tilde{x}$
- K-means-based Segmentation with $k = 2$

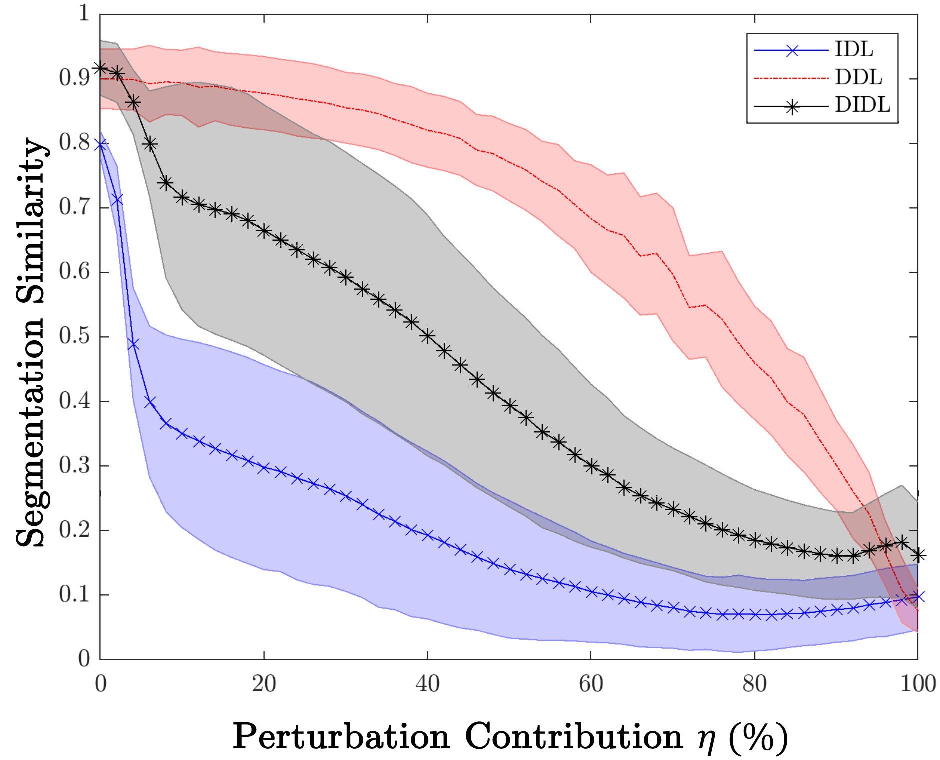
BOSTON UNIVERSITY

Results

Quantitative Results – Adversarial Perturbations

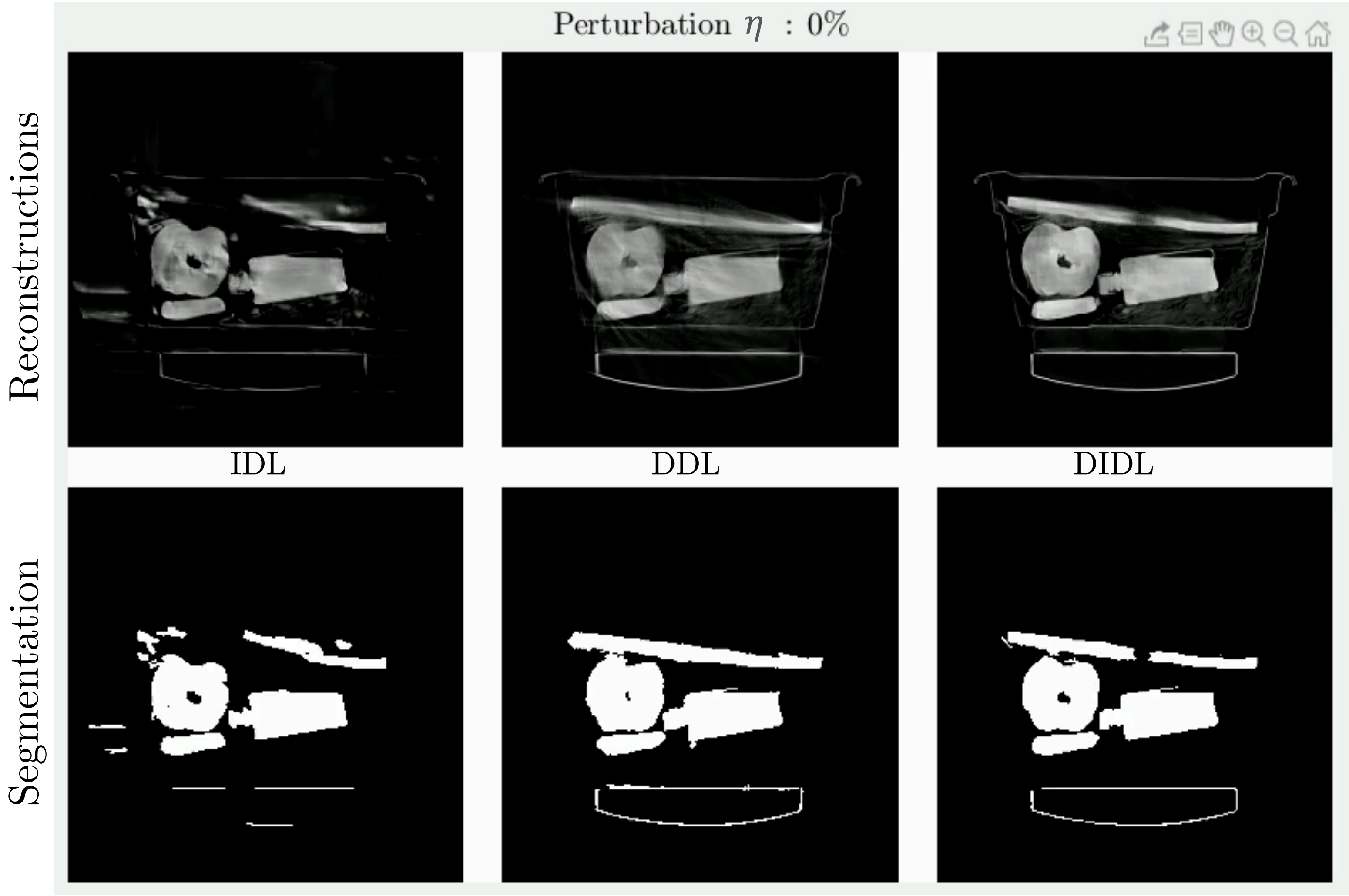
- Mean and STD computed over 10 examples
- **Image-domain DL appears more sensitive**





Reconstruction Error (Lower is better)	Segmentation Similarity Score (Larger is better)

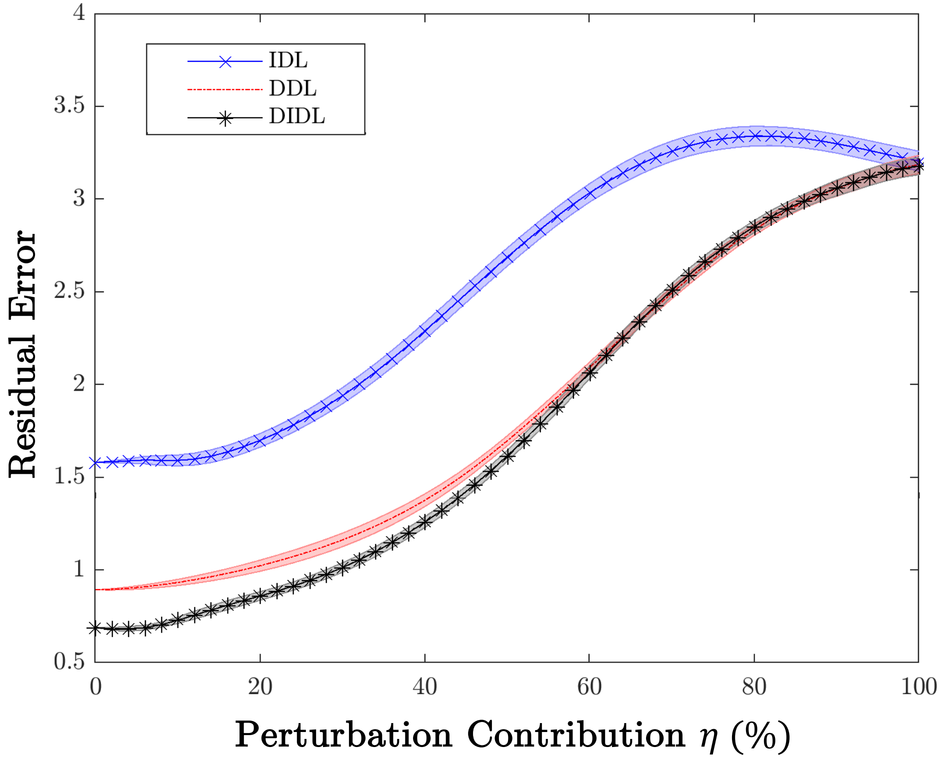
Qualitative Results – Adversarial Perturbations

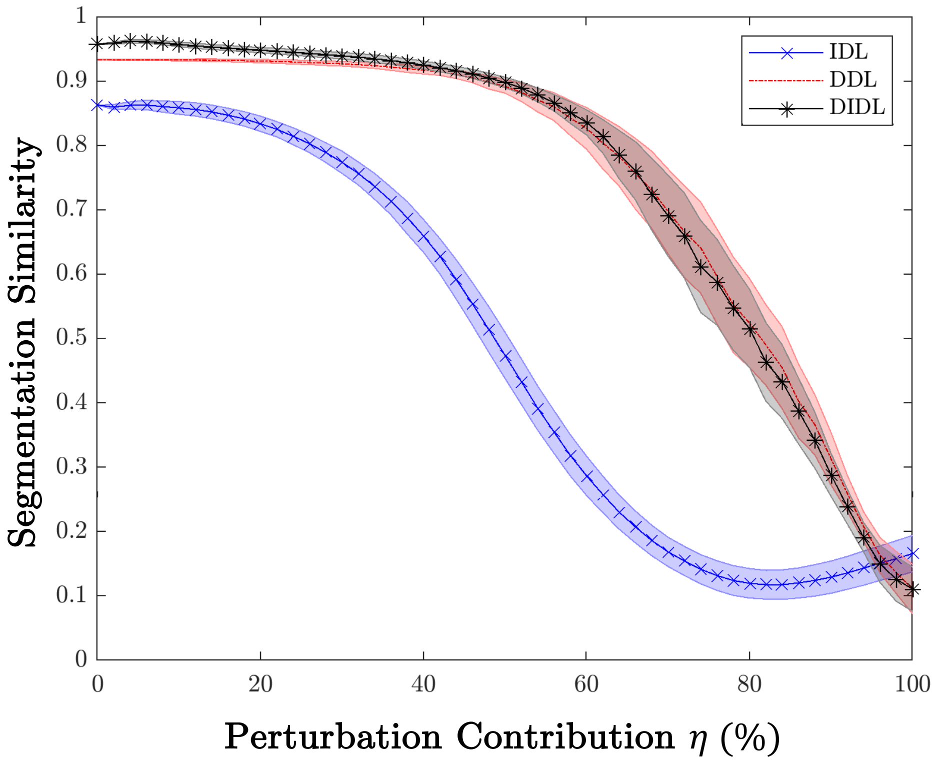


BOSTON UNIVERSITY

Quantitative Results – Random Perturbations

- Mean and STD computed over 100 noise instances
- **Image-domain DL appears more sensitive**

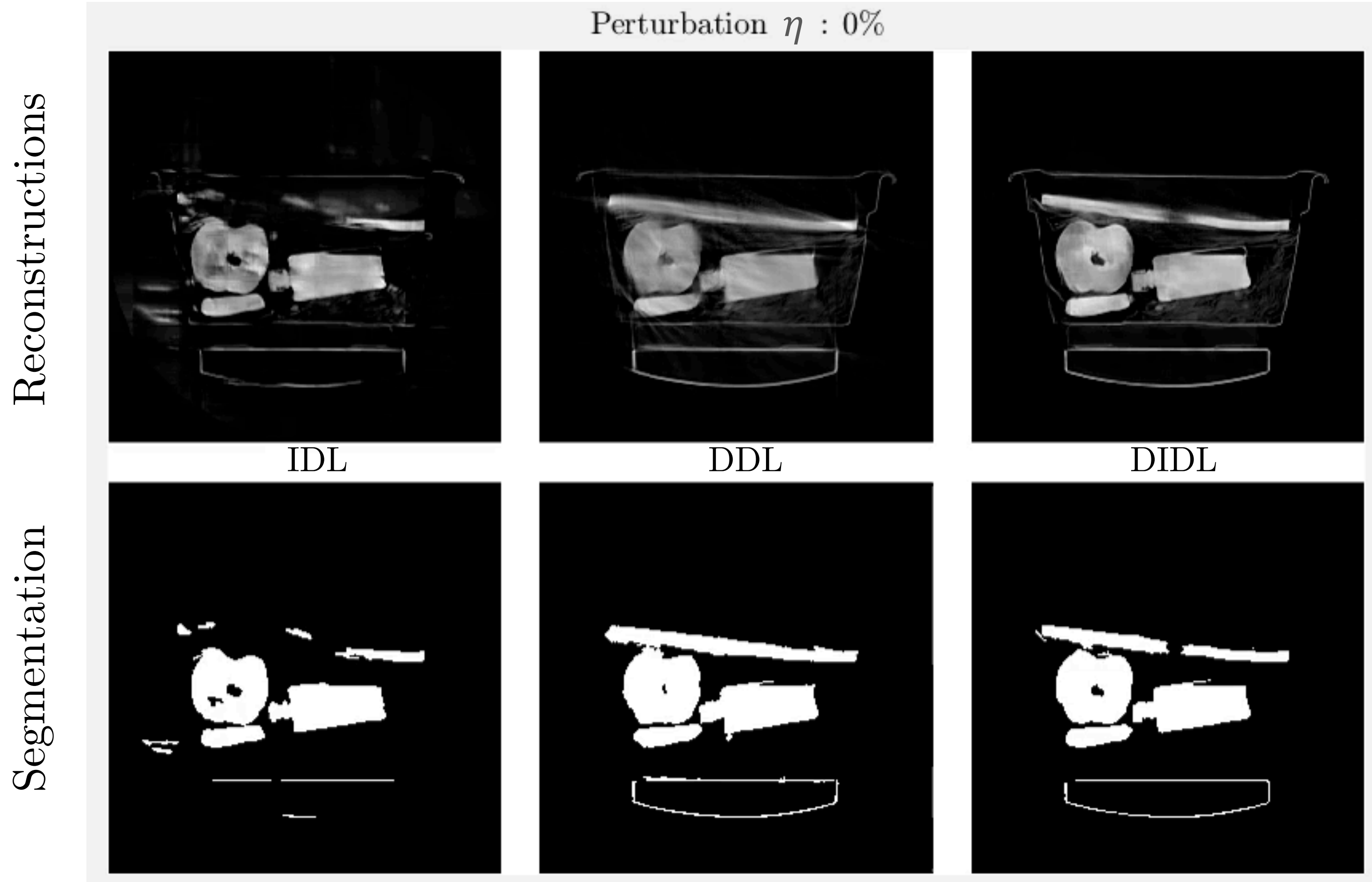




Reconstruction Error (Lower is better)	Segmentation Similarity Score (Larger is better)

Qualitative Results – Random Perturbations

- Worst 1 out of 100 noise instances for each η



BOSTON UNIVERSITY

Structural Perturbations (1)

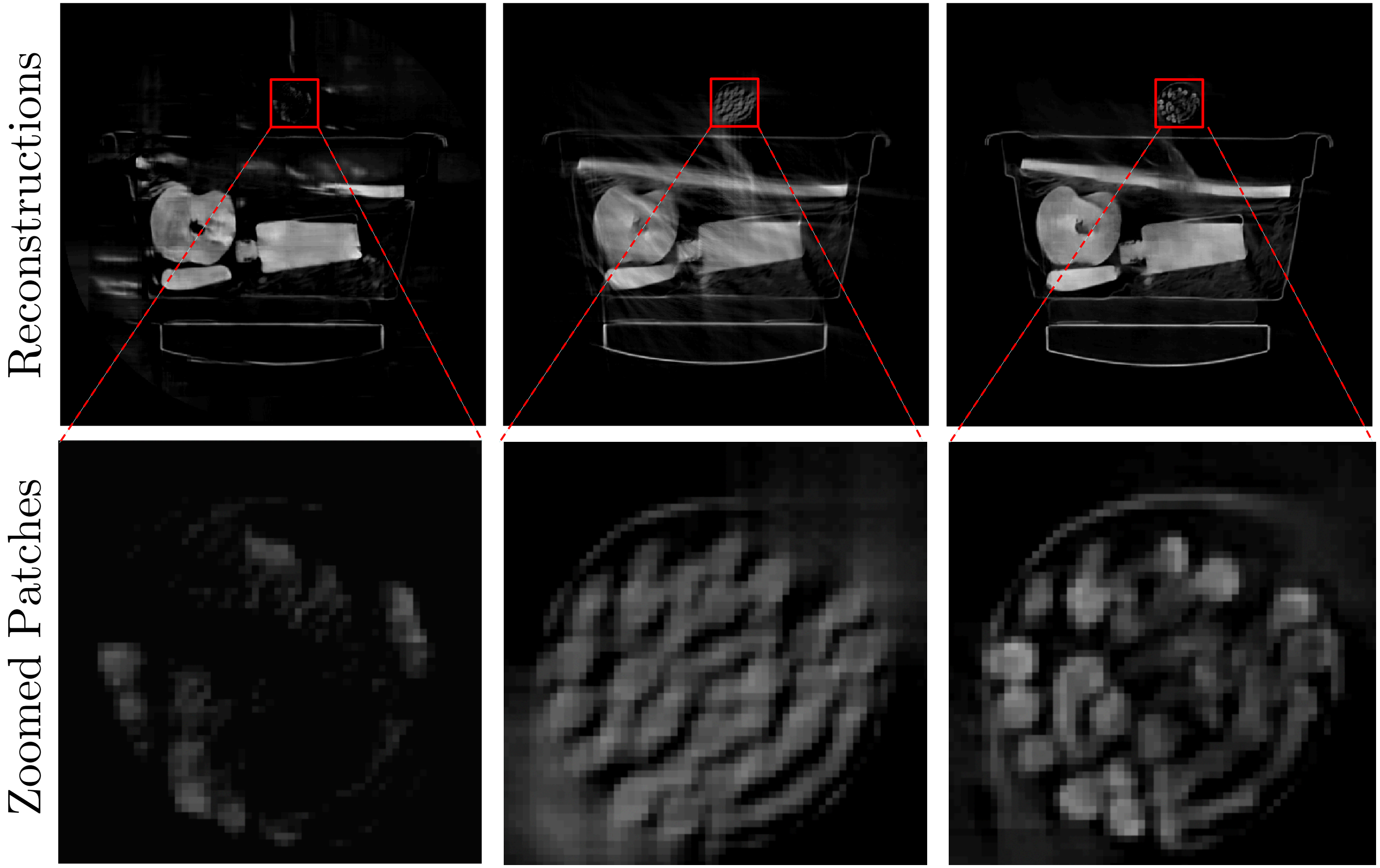


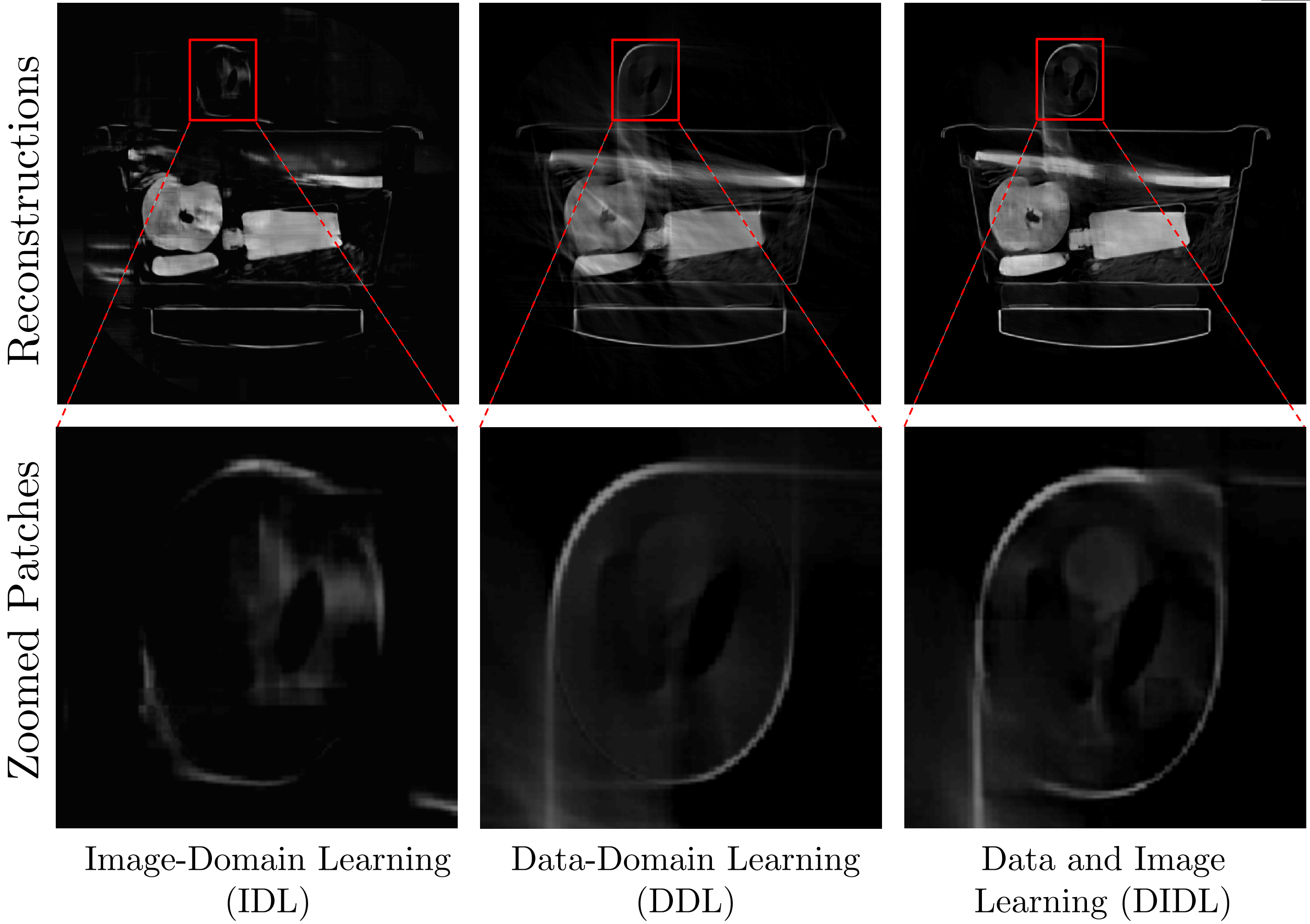
Image-Domain Learning (IDL)

Data-Domain Learning (DDL)

Data and Image Learning (DIDL)

BOSTON UNIVERSITY

Structural Perturbations (2)



BOSTON UNIVERSITY

Summary

- Image post-processing
  - Gets severely damaged by adversarial perturbations
  - Produces ghost features in response to random Gaussian perturbations
  - Fails to reconstruct new structural features

- Performance of data-domain learning method seems to degrade more gracefully in face of all perturbations

- Combined Data and Image domain method have superior performance when perturbations (η) are contained

Thank You!