

# **Recurrent Feature Reasoning in Medical Image Inpainting**

CS736 - Medical Image Computing  
Prof. Suyash Awate

Ashwin Goyal (210050024), Ravi Kumar (210010052)

# What is Inpainting?

- Object removal - power lines or tourists in a scenic photo.
- Image completion - completing an image that has missing portions
- Creative Editing - Artistic effects by filling in areas with imaginative content.



*Masked Image*

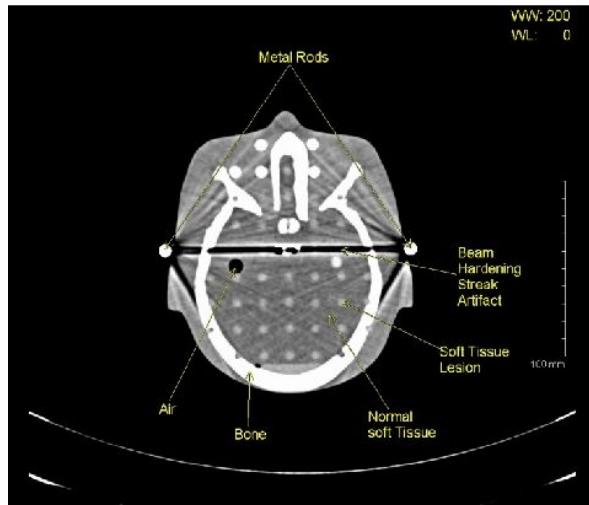


*Inpainted Image*

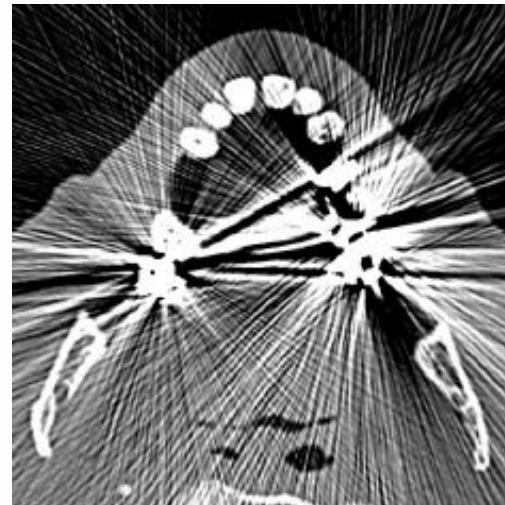
(Source: <https://arxiv.org/abs/2008.03737>)

# Inpainting for Medical Images

- Motion artifacts - distortions - MRI
- Beam hardening - Motion artifacts - Metallic implants - dark/light streaks - CT
- Loss of data



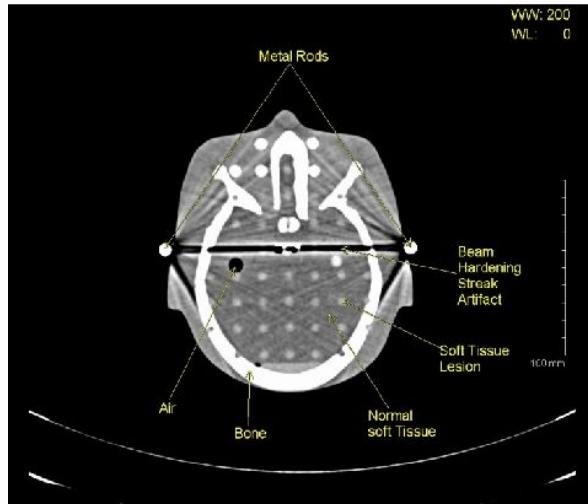
Beam hardening



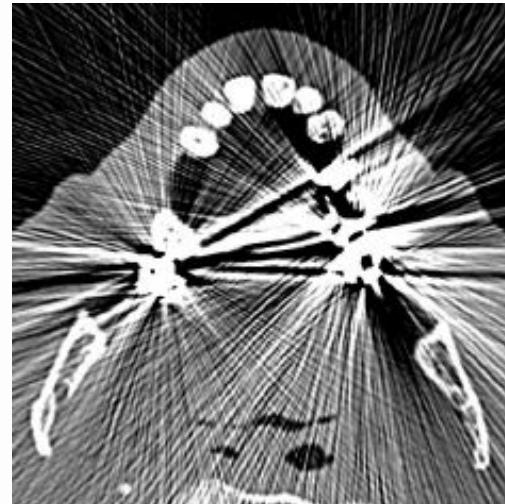
Streaks due to dental fillings

# Inpainting for Medical Images

- Reduces image quality and accuracy of diagnosis.
- Affects image segmentation, classification & attenuation correction in PET/MRI
- Can be used to analyze multi-modal data

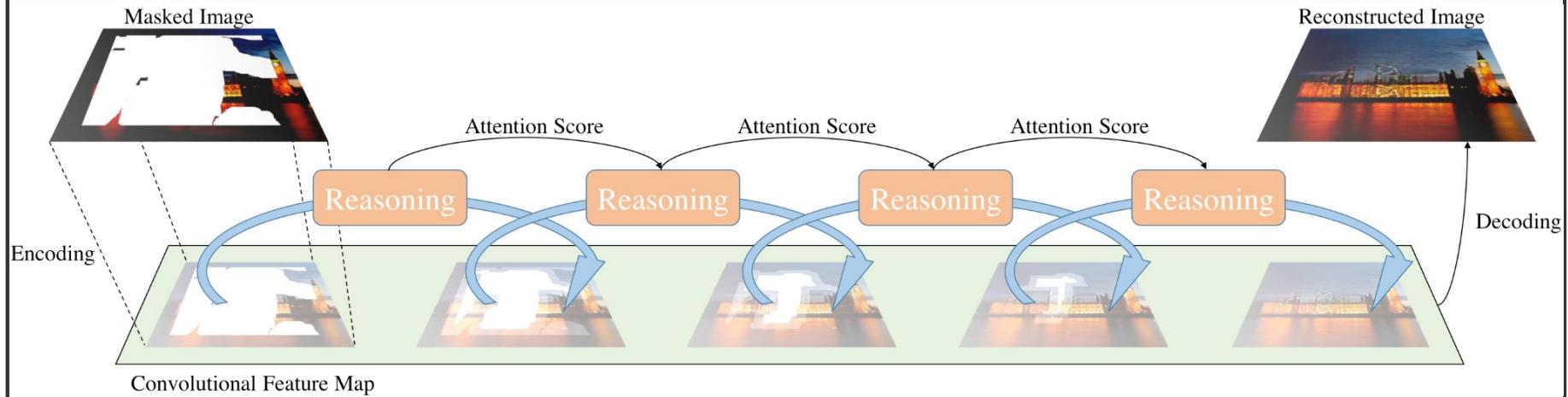


Beam hardening



Streaks due to dental fillings

# Recurrent Feature Reasoning



*RFR inpainting scheme*

[https://openaccess.thecvf.com/content\\_CVPR\\_2020/papers/Li\\_Recurrent\\_Feature\\_Reasoning\\_for\\_Image\\_Inpainting\\_CVPR\\_2020\\_paper.pdf](https://openaccess.thecvf.com/content_CVPR_2020/papers/Li_Recurrent_Feature_Reasoning_for_Image_Inpainting_CVPR_2020_paper.pdf)

# Recurrent Feature Reasoning

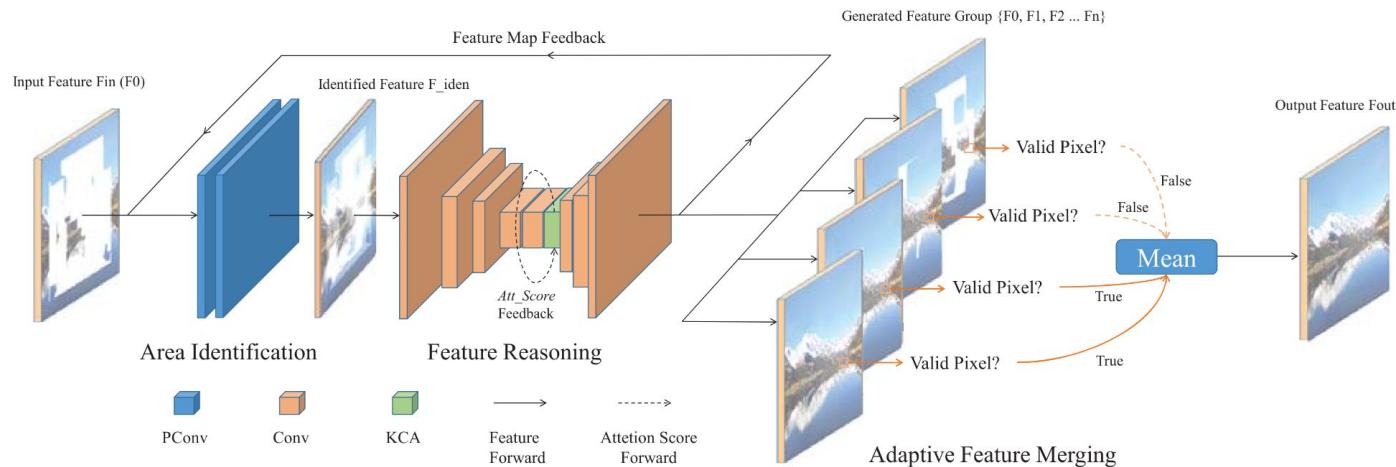


Figure 2. Illustration of the Recurrent Feature Reasoning module. The area identification process and the feature reasoning process are performed continuously. After several times of reasoning, the feature maps are merged in an adaptive fashion and an output feature map of a fixed channel numbers are generated. The module is Plug-In-and-Play and can be placed in any layer of an existing network.

# Recurrent Feature Reasoning

$$L_{perceptual} = \sum_{i=1}^N \frac{1}{H_i W_i C_i} |\phi_{pool_i}^{gt} - \phi_{pool_i}^{pred}|_1 \quad (10)$$

Similarly, the computation of the style loss is as follows:

$$\phi_{pool_i}^{style} = \phi_{pool_i} \phi_{pool_i}^T \quad (11)$$

$$L_{style} = \sum_{i=1}^N \frac{1}{C_i \times C_i} \left| \frac{1}{H_i W_i C_i} (\phi_{pool_i}^{style_{gt}} - \phi_{pool_i}^{style_{pred}}) \right|_1 \quad (12)$$

Further,  $L_{valid}$  and  $L_{hole}$  which calculate L1 differences in the unmasked area and masked area respectively are also used in our model. In summary, our total loss function is:

$$L_{total} = \lambda_{hole} L_{hole} + \lambda_{valid} L_{valid} + \lambda_{style} L_{style} + \lambda_{perceptual} L_{perceptual} \quad (13)$$

# Comparison Metrics

- All metrics are calculated in the masked region only
- **Structural Similarity Index Measure (SSIM)**
  - Compares the luminance, contrast and structure of the reference image with the distorted image
  - Ranges from -1 (worst) to 1 (best/identical)

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

SSIM between two same sized windows x and y

# Comparison Metrics

- Peak Signal-to-Noise Ratio (PSNR)
  - Computes SNR between images, higher is better

$$\text{PSNR} = 10 \log_{10} \left( \frac{\text{MAX}^2}{\text{MSE}} \right)$$

*PSNR between two images. MAX is the maximum possible pixel value and MSE is the mean squared error between the two images*

# Comparison Metrics

- Root Mean Squared Error (RMSE)
  - Error between two images, lower is better

$$\text{RMSE} = \frac{1}{N} \sqrt{\sum_{i=1}^N \sum_{j=1}^N (I_{\text{gt}}^{ij} - I_{\text{rec}}^{ij})^2}$$

*RMSE between two images*

# Experiments

- Dataset Variability
  - Presence of disease affecting output?
- Hole size
  - Different sized masks
- Types of masks
  - Square or strip
- Number of holes/strips
  - 1, 2, 4, etc. per mask
- Orientation of masks
  - Strip mask - horizontal vs vertical
- Out-of-Distribution tests
  - Mix and match models/masks/datasets

# **Dataset**

**Chest-xray-pneumonia**

# Standard results, Hole size = $\frac{1}{8}$ image size

SSIM - 0.992294

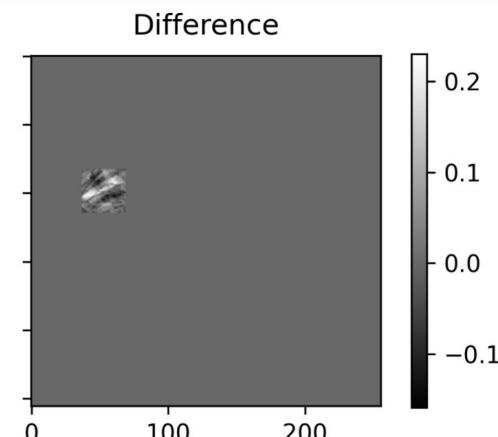
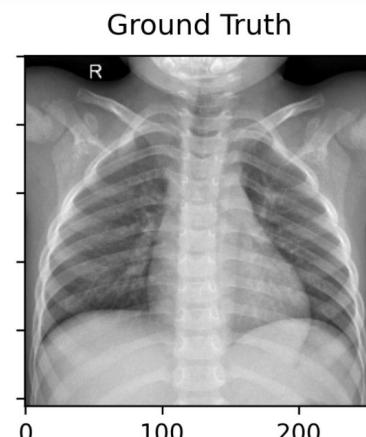
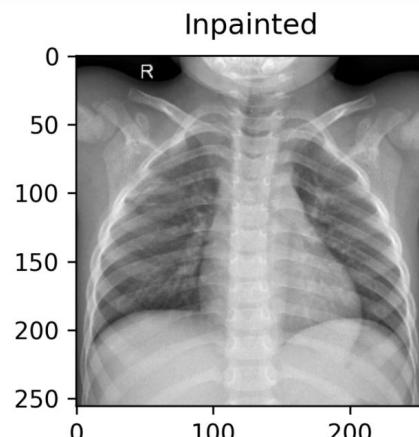
PSNR - 41.2384

RMSE - 0.00845

SSIM - 0.6841

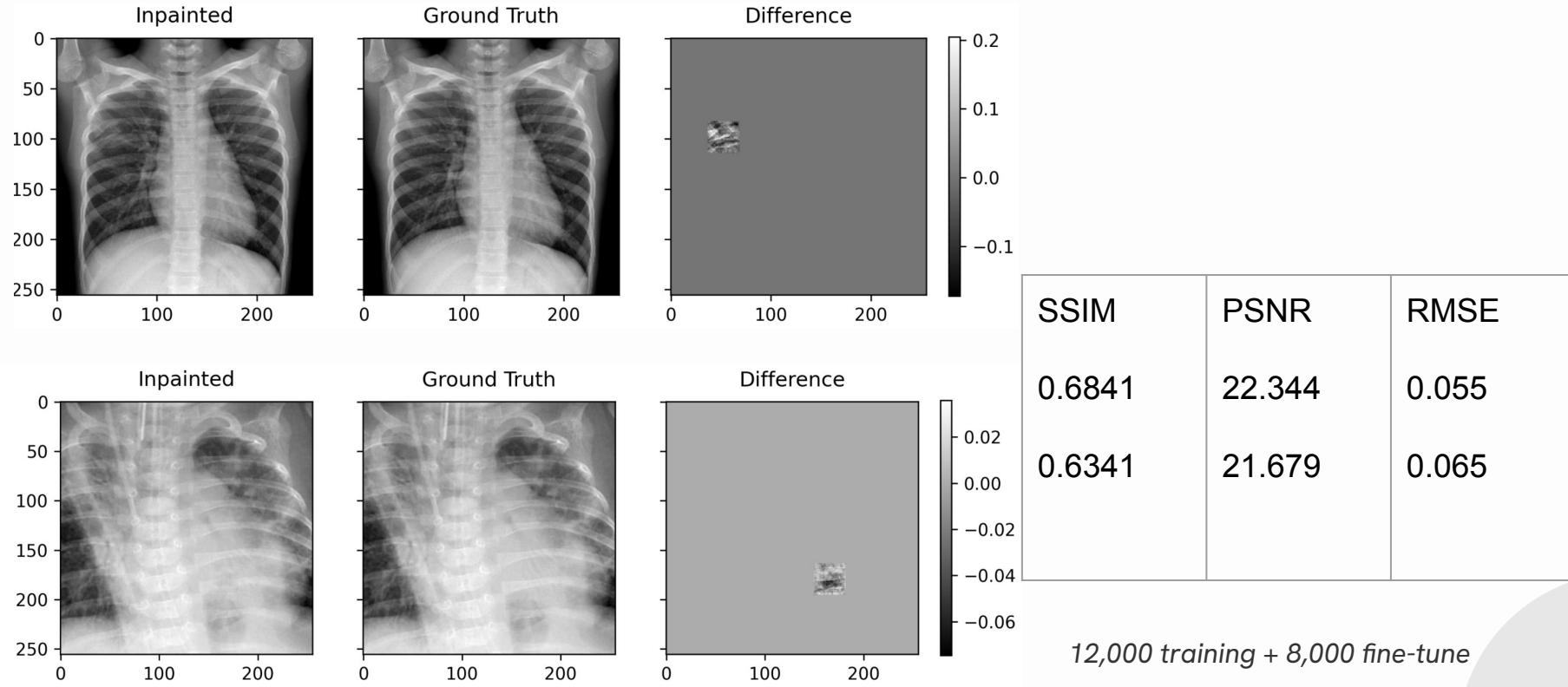
PSNR - 22.344

RMSE - 0.055



12,000 training + 8,000 fine-tune

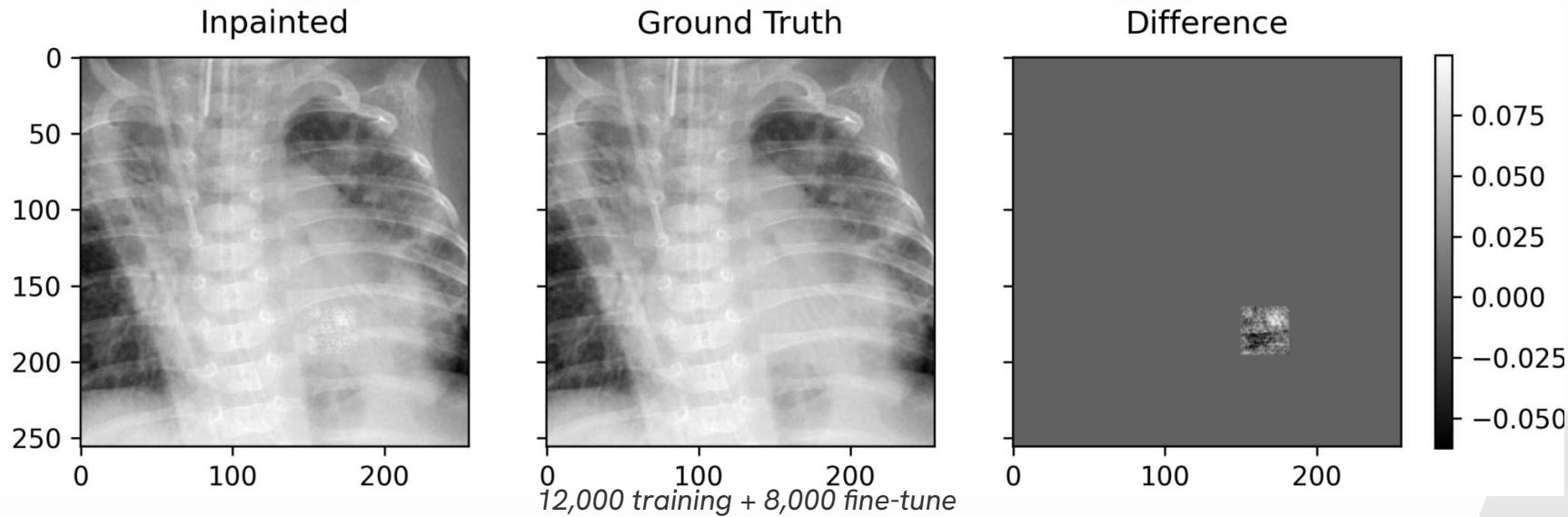
# Only normal vs normal+pneumonia



# Out of Distribution Results

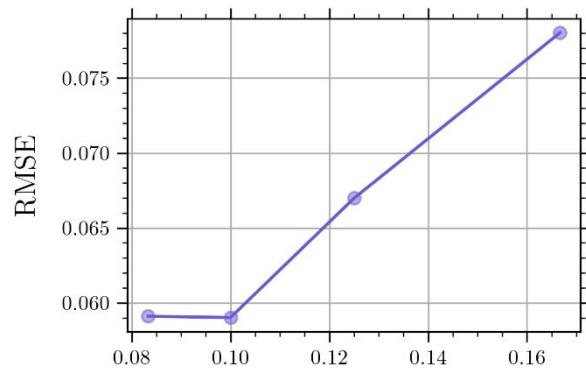
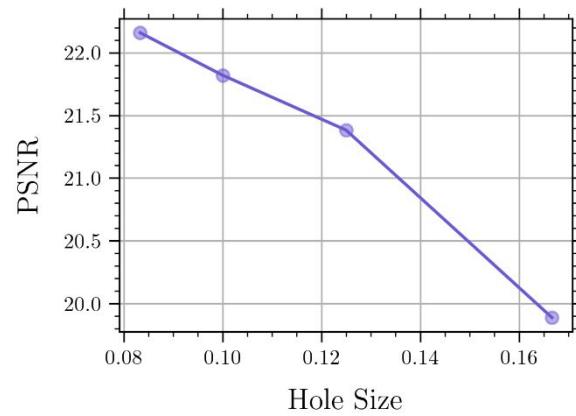
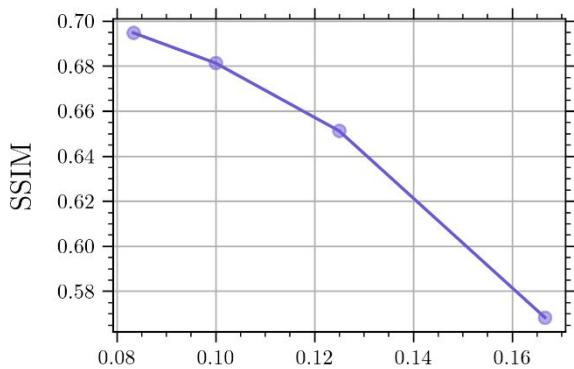
- Trained on only normal, test on pneumonia+normal

	SSIM	PSNR	RMSE
ID	0.684	22.34	0.055
	0.639	22.10	0.057



# Hole size

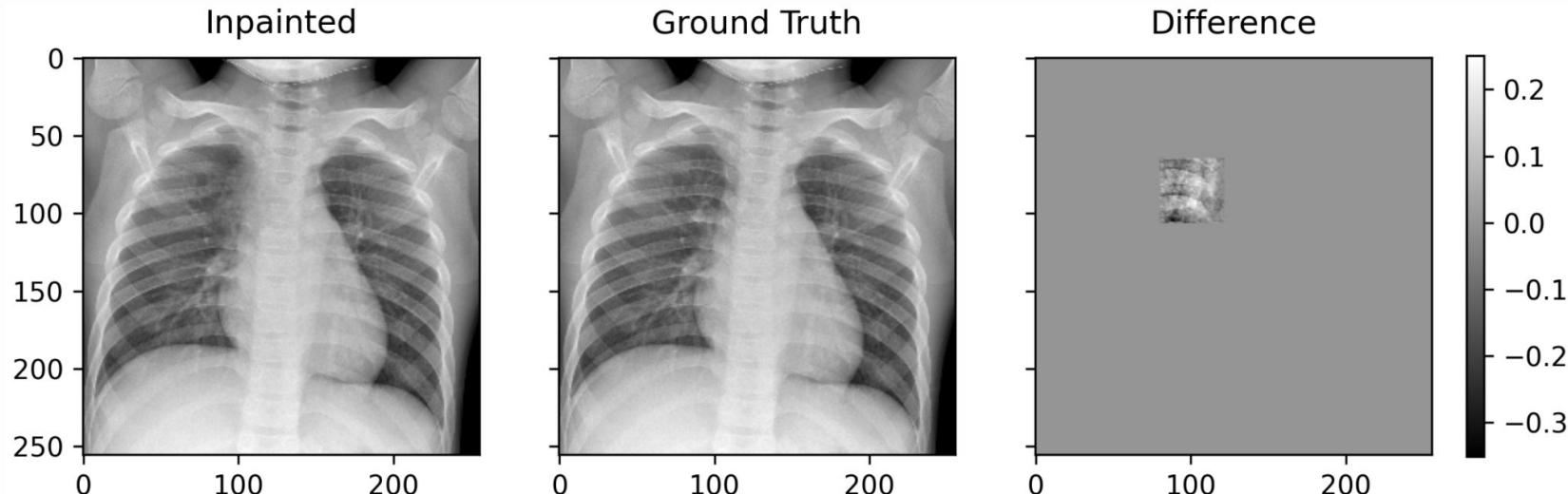
- Hole ratio is defined as the ratio of the image being masked ( $256/\text{ratio}$ )



12,000 training + 8,000 fine-tune

# Out of Distribution Results

- Hole size OOD performs similar to hole size ID



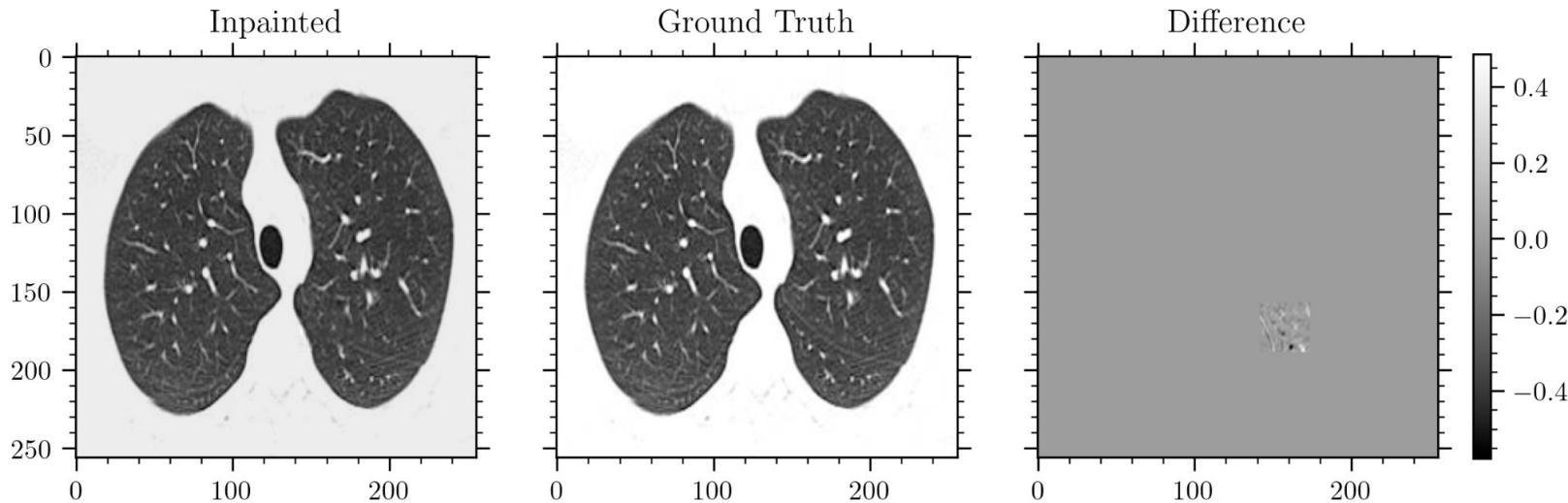
12,000 training + 8,000 fine-tune

# **Dataset**

## **SARS-COV-2 Ct-Scan**

# Types of masks

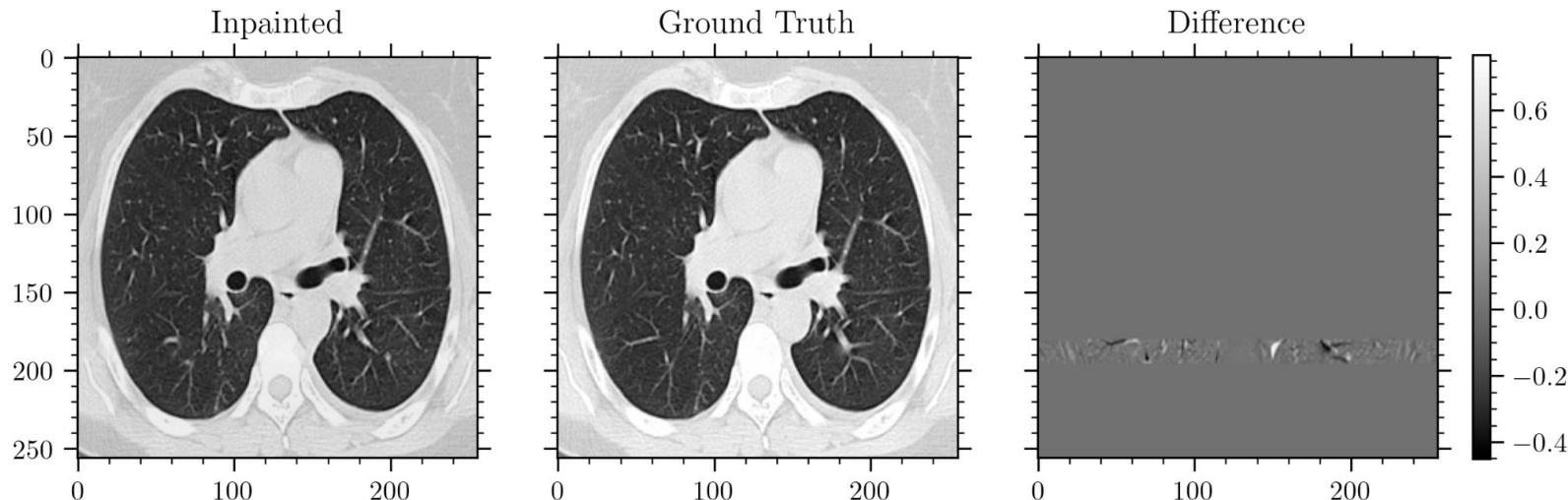
- Square holes - 32x32 placed randomly
  - Simulate regional artifacts that need inpainting



*16,000 training + 8,000 fine-tune*

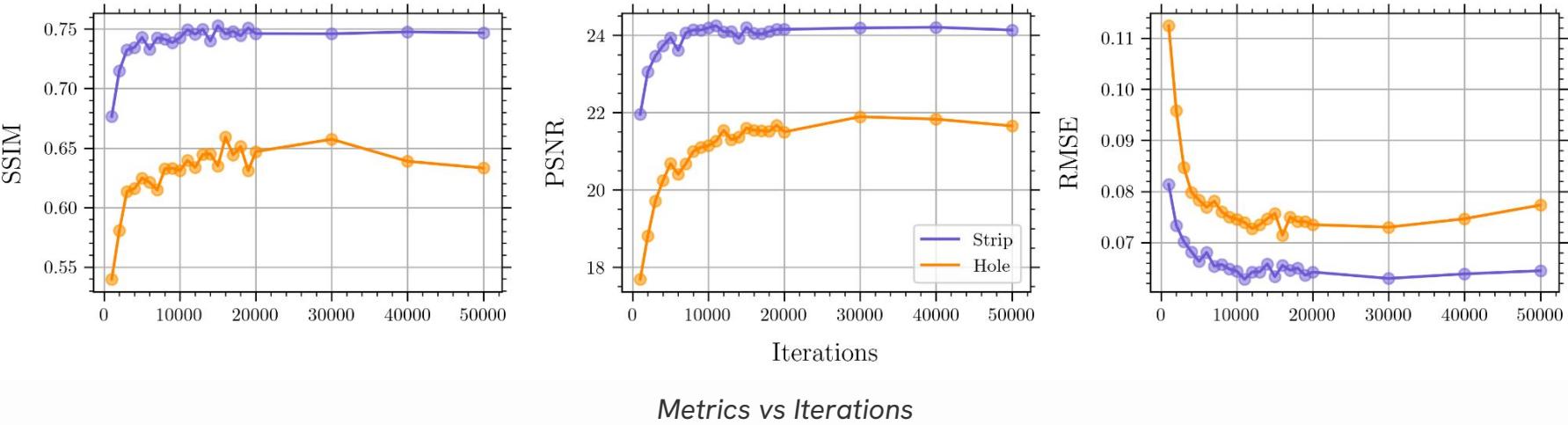
# Types of masks

- Strips - 16x256 placed randomly
  - Simulate beaming artifacts that need inpainting

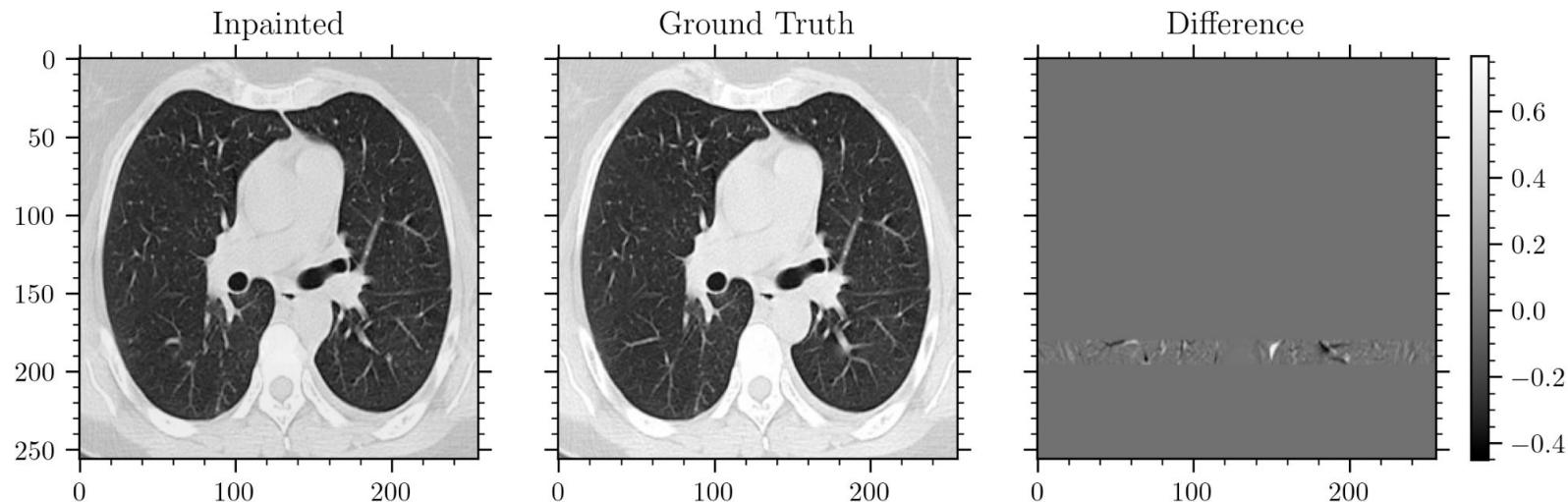


*15,000 training + 7,500 fine-tune*

# Types of Masks

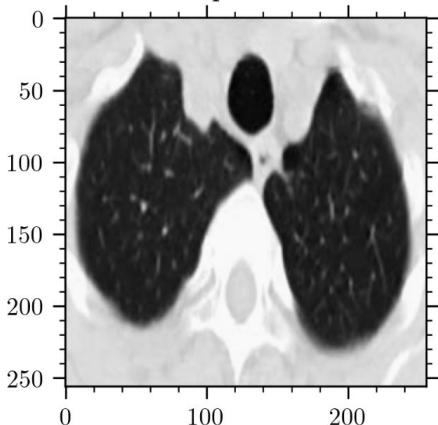


# Number of Holes/Strips

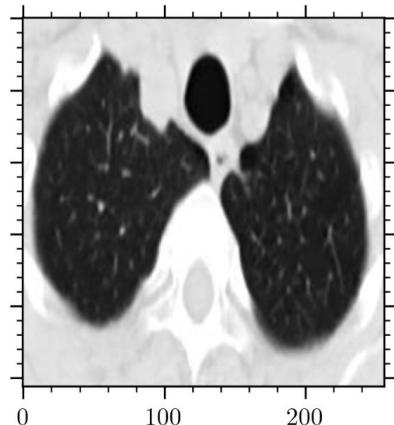


1 Strip: 15,000 training + 7,500 fine-tune

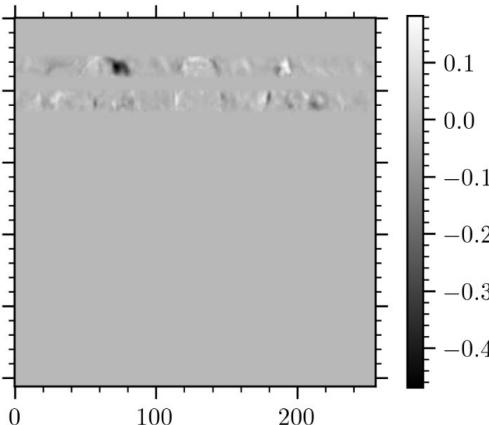
Inpainted



Ground Truth

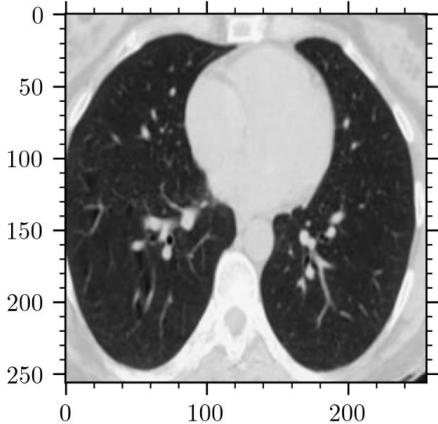


Difference

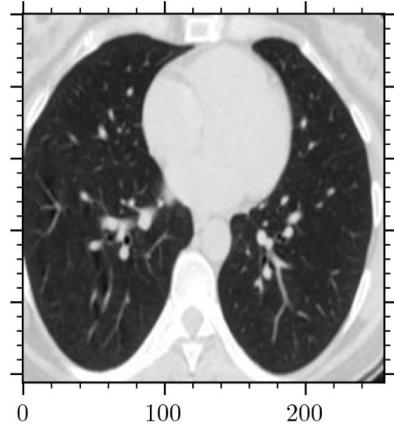


2 Strips: 18,000 training  
+ 9,000 fine-tune

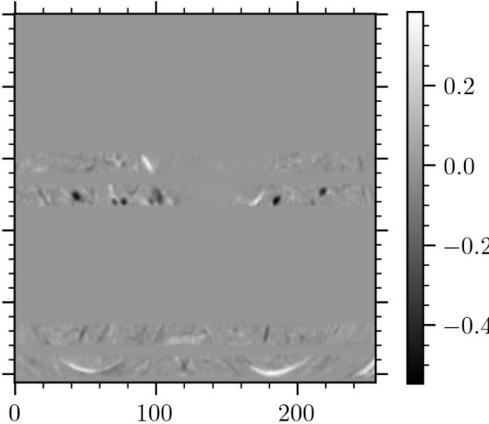
Inpainted



Ground Truth

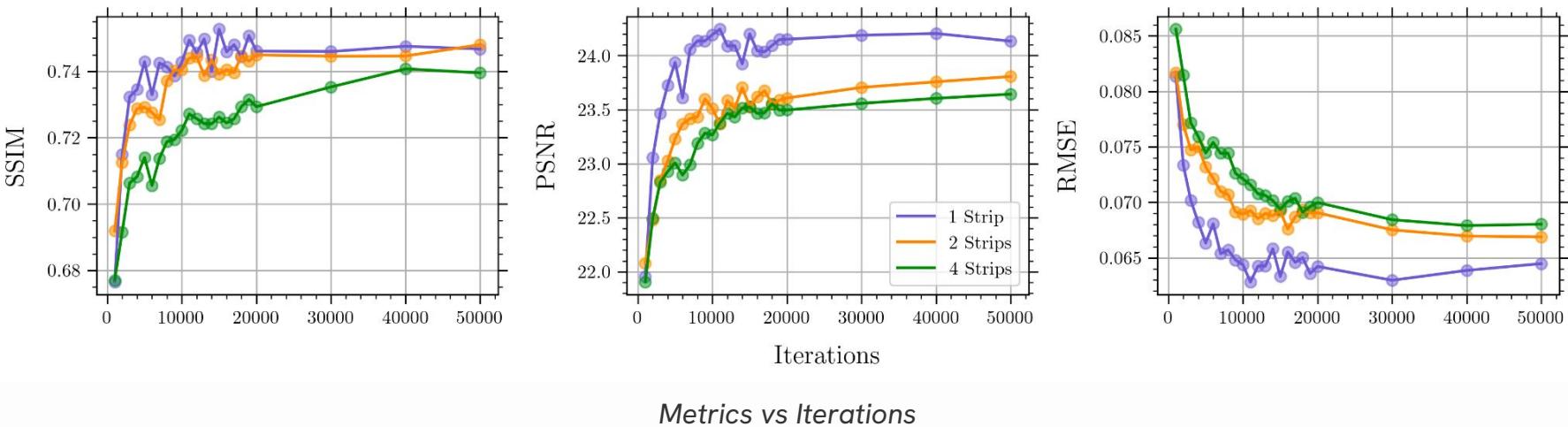


Difference

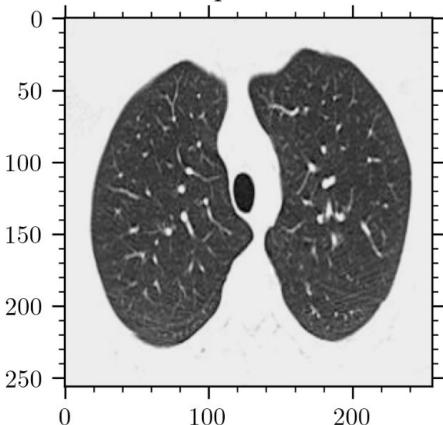


4 Strips: 40,000 training  
+ 20,000 fine-tune

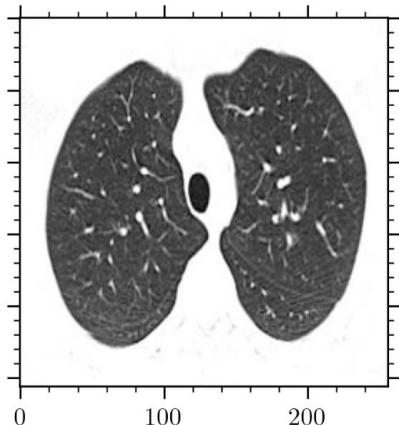
# Number of Holes/Strips



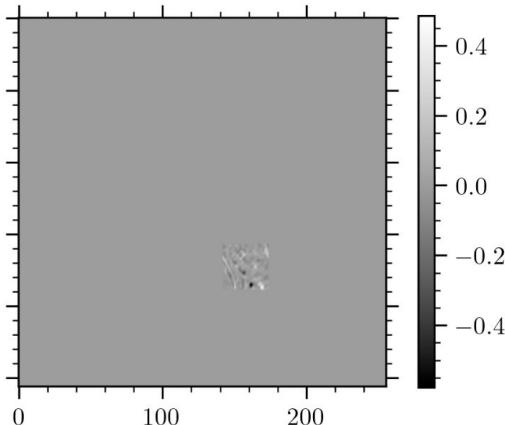
Inpainted



Ground Truth

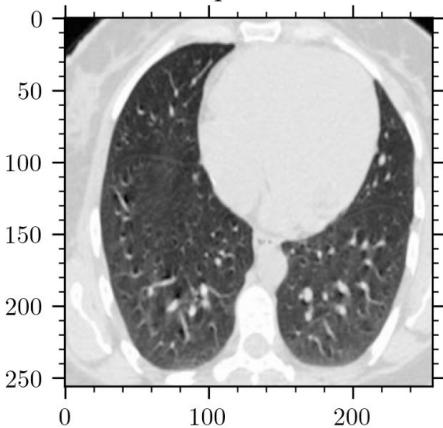


Difference

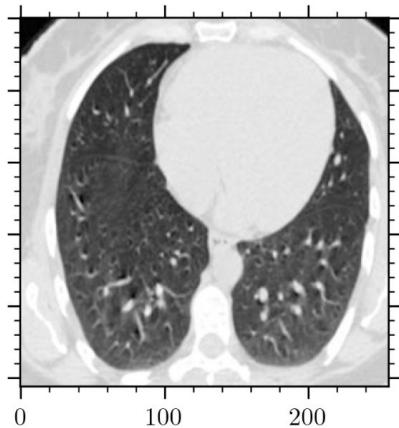


1 Hole: 16,000 training  
+ 8,000 fine-tune

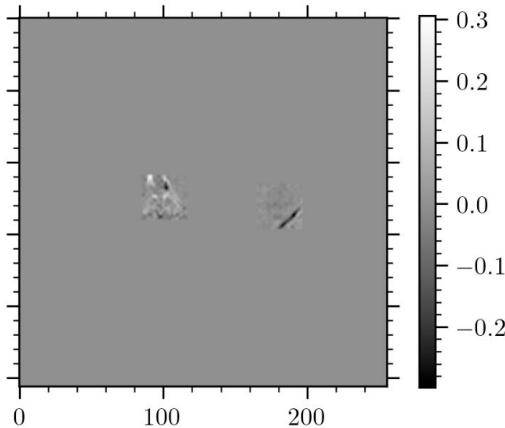
Inpainted



Ground Truth

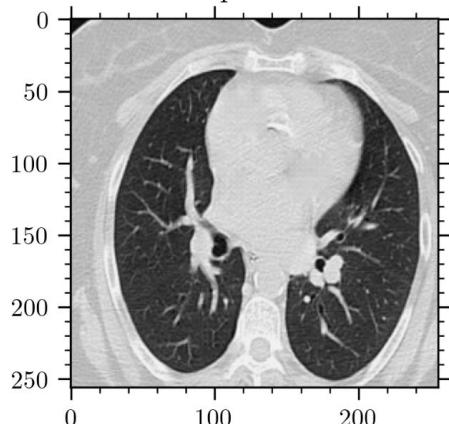


Difference

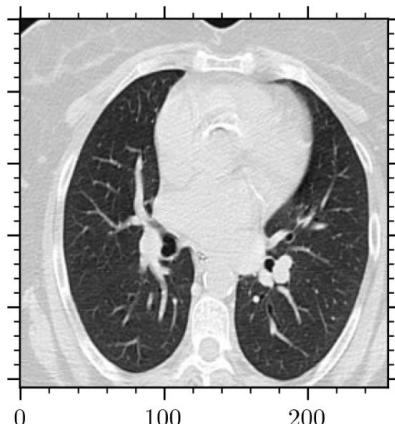


2 Holes: 13,000 training  
+ 6,500 fine-tune

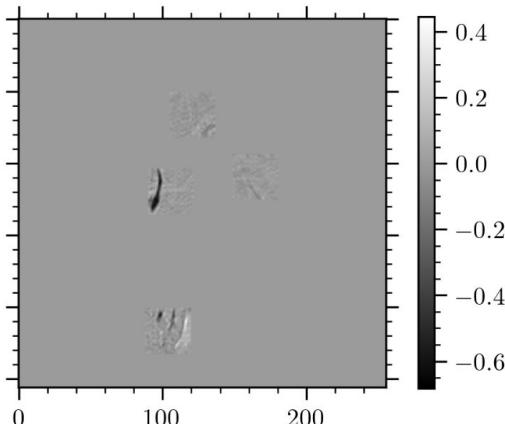
Inpainted



Ground Truth

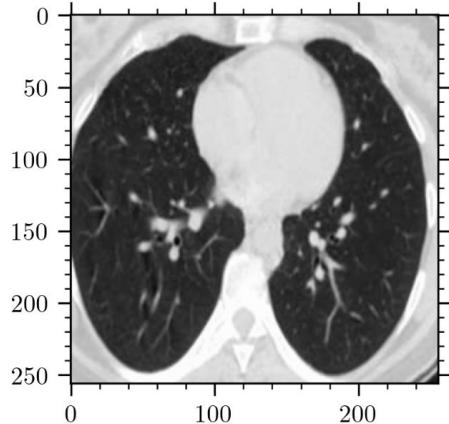


Difference

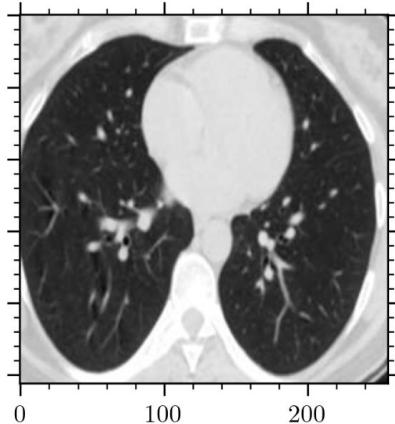


4 Holes: 20,000 training  
+ 10,000 fine-tune

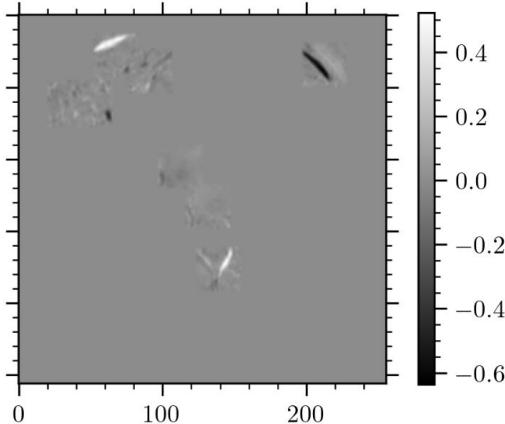
Inpainted



Ground Truth

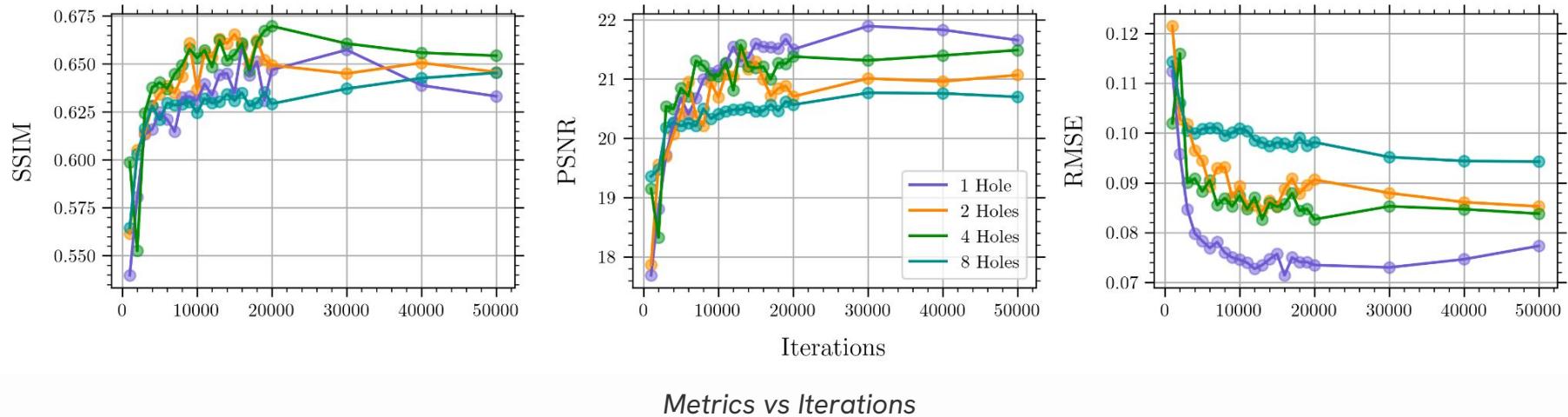


Difference



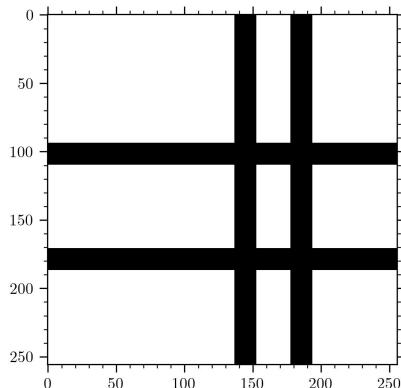
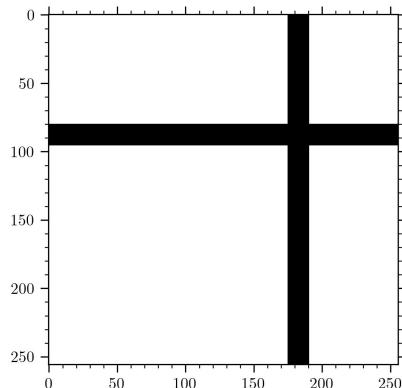
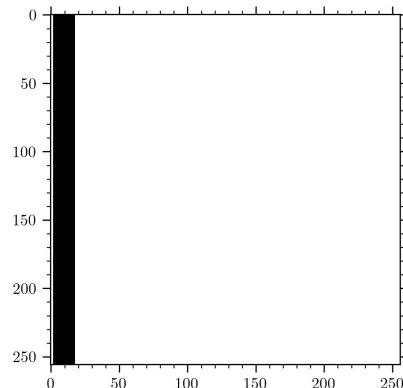
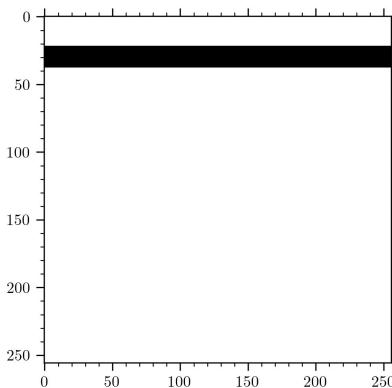
8 Holes: 50,000 training  
+ 25,000 fine-tune

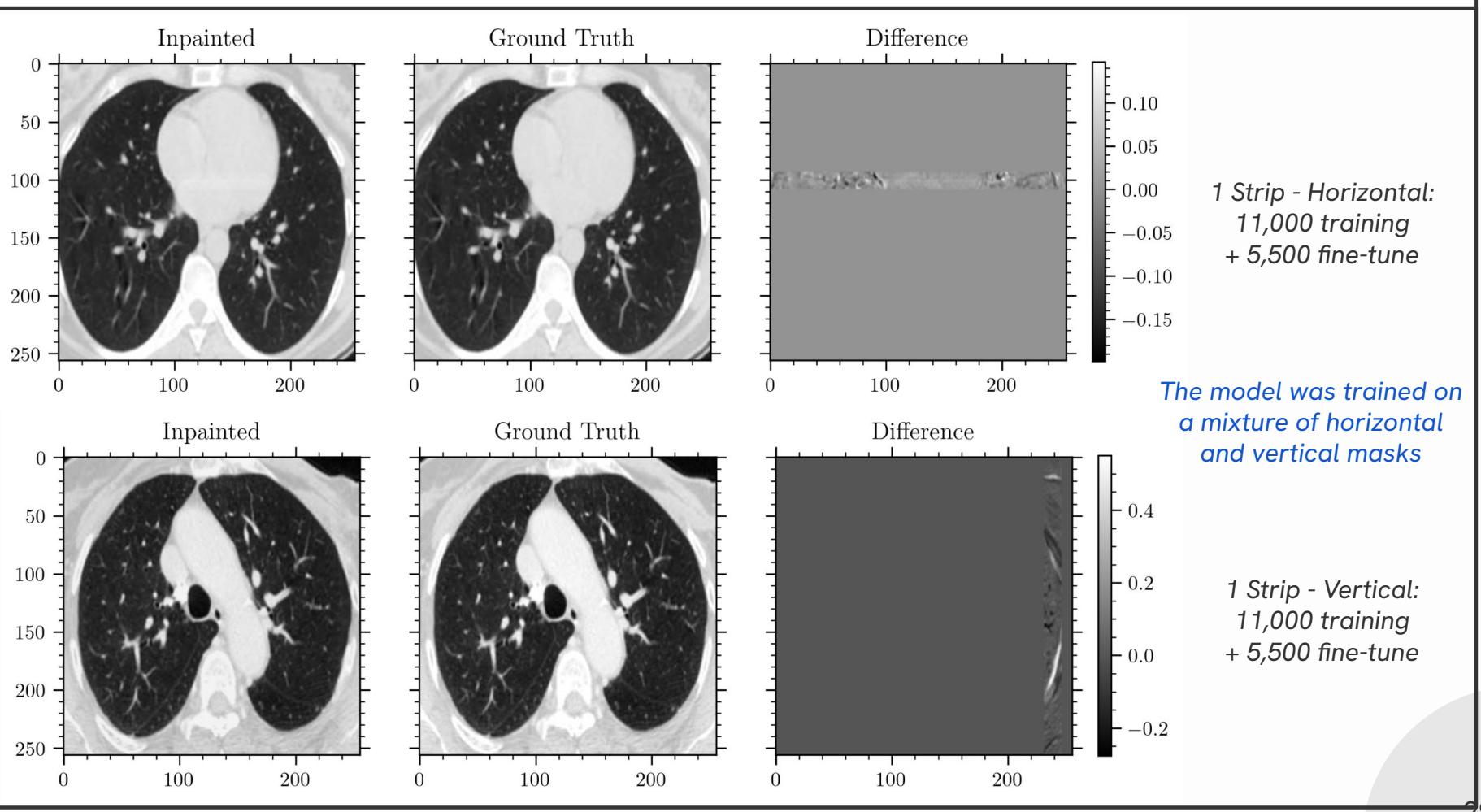
# Number of Holes/Strips



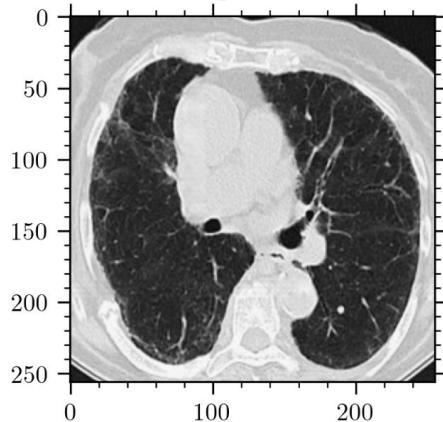
# Orientation of Masks

- Only applicable to strip masks
- Combinations of vertical and horizontal strip masks were used
  - 1 strip per mask - randomly choose orientation
  - 2 strips per mask - ensure 1 vertical and 1 horizontal
  - 4 strips per mask - ensure 2 vertical and 1 horizontal

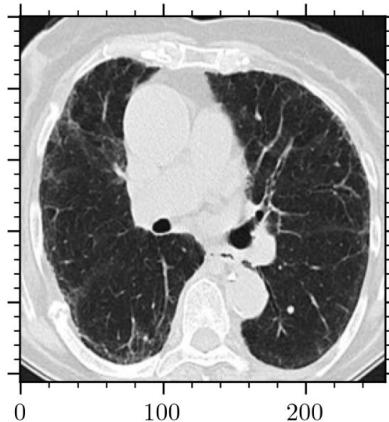




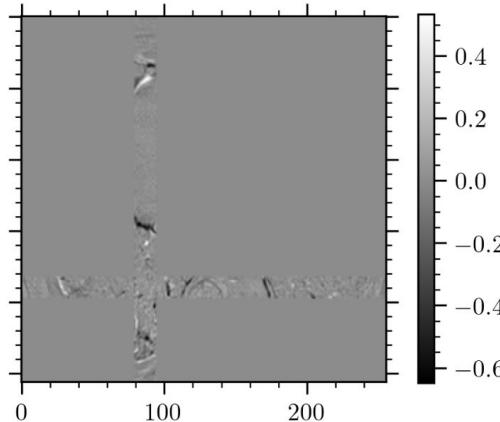
Inpainted



Ground Truth



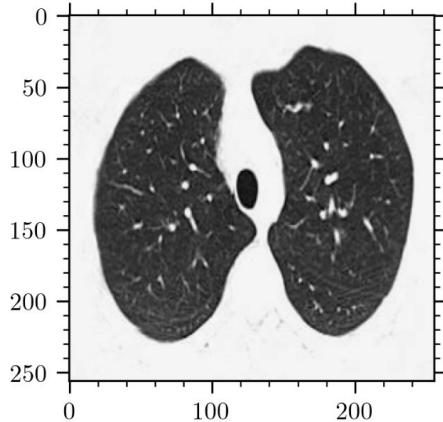
Difference



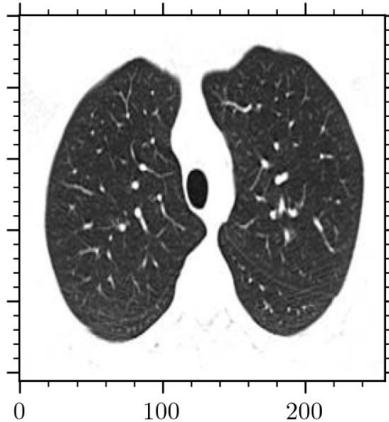
2 Strips

Horizontal+Vertical  
40,000 training  
+ 20,000 fine-tune

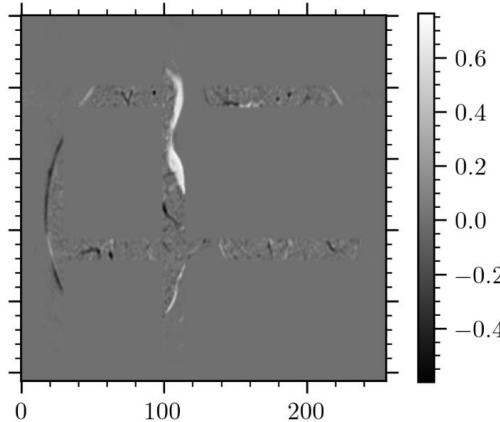
Inpainted



Ground Truth

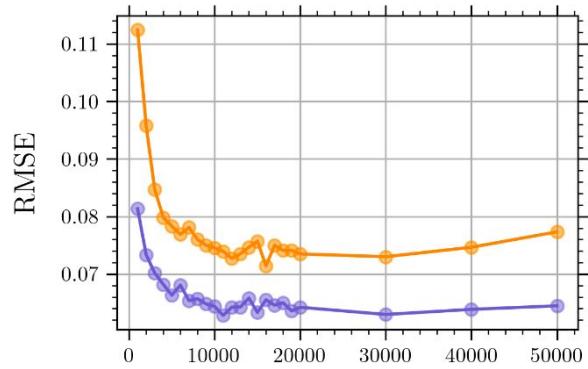
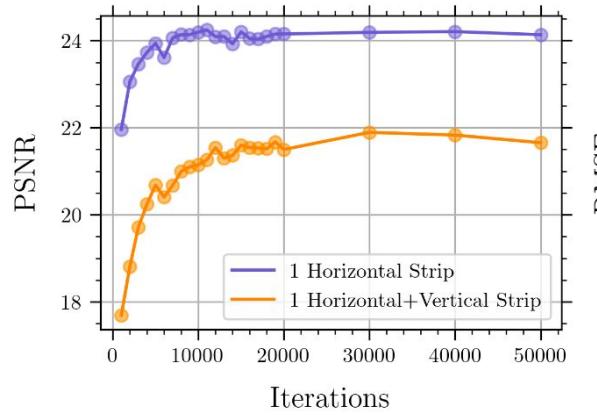
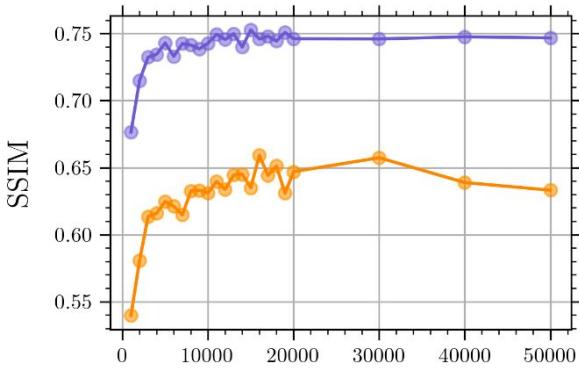


Difference

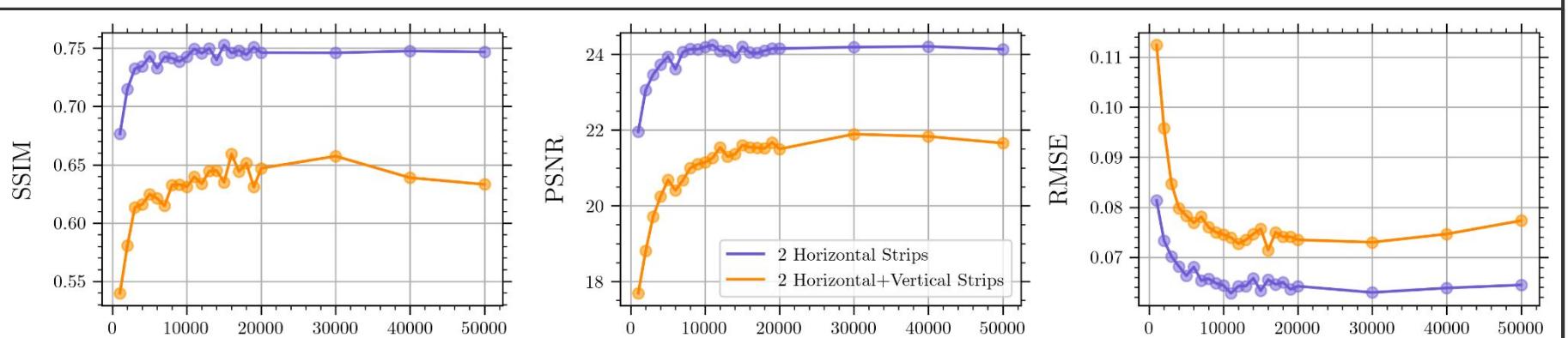


4 Strips

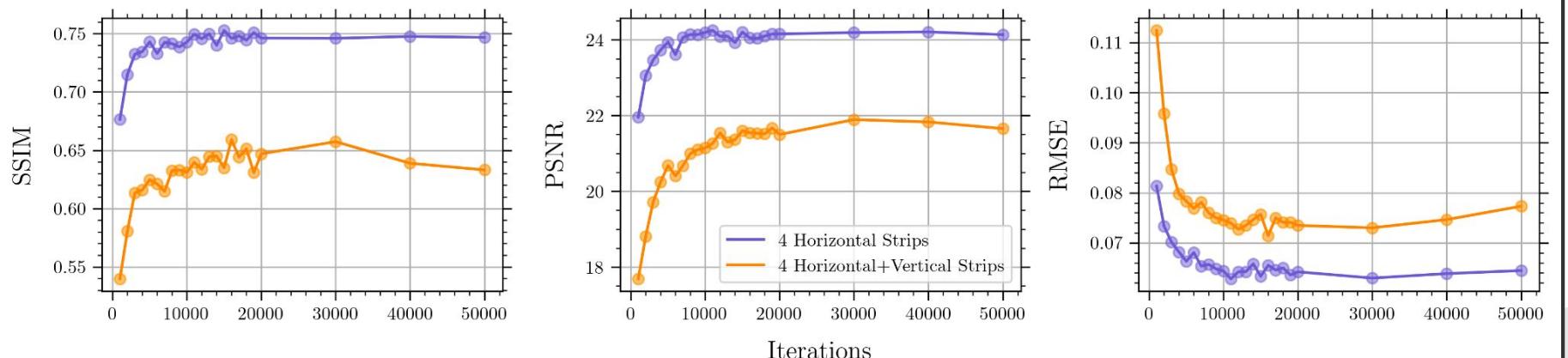
Horizontal+Vertical  
40,000 training  
+ 20,000 fine-tune



*Metrics vs Iterations: Comparison of 1 Horizontal strip with 1 Horizontal+Vertical*



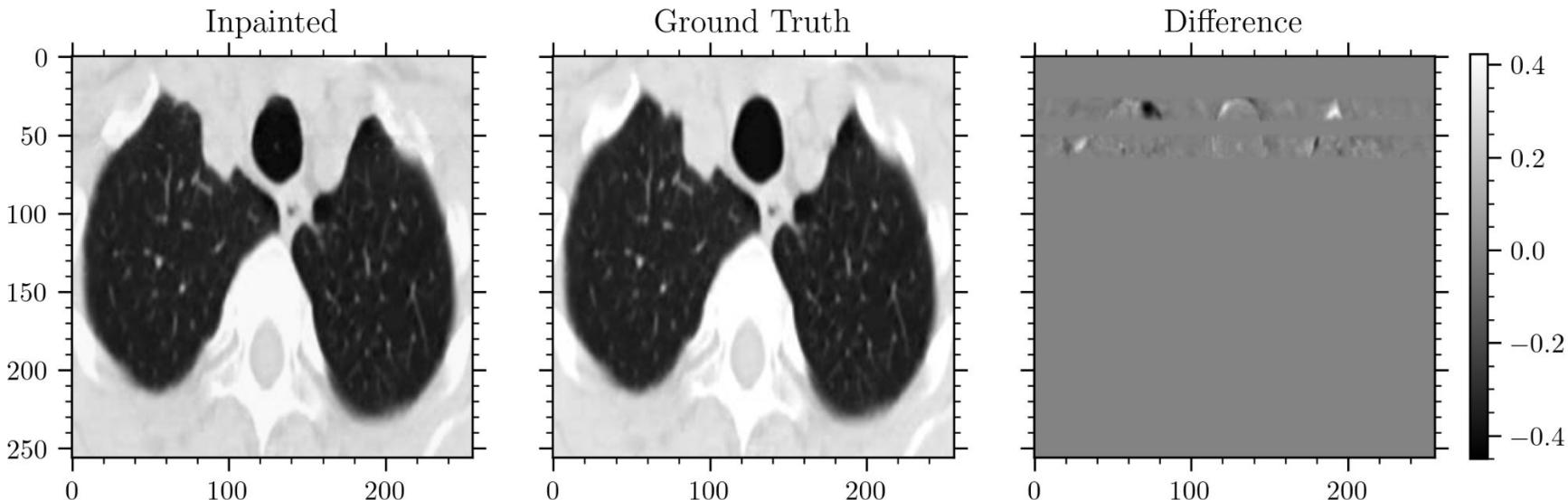
*Metrics vs Iterations: Comparison of 2 Horizontal strips with 2 Horizontal+Vertical*



*Metrics vs Iterations: Comparison of 4 Horizontal strips with 4 Horizontal+Vertical*

# Out of Distribution Results

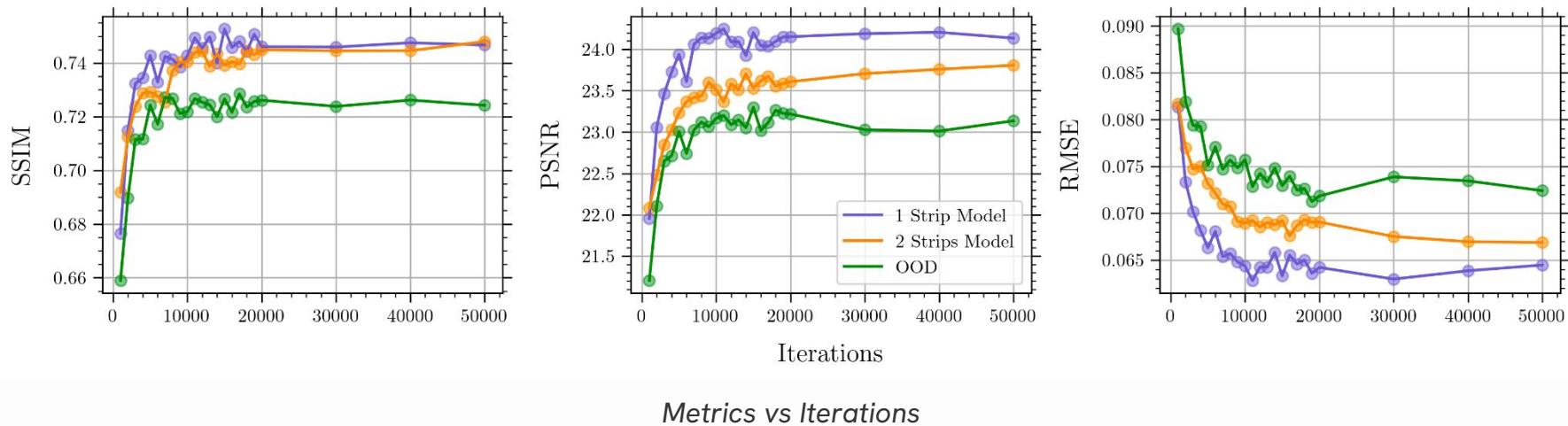
- Applying 1 strip model to 2 strip masks



1 Strip model + 2 Strips masks: 17,000 training + 8,500 fine-tune

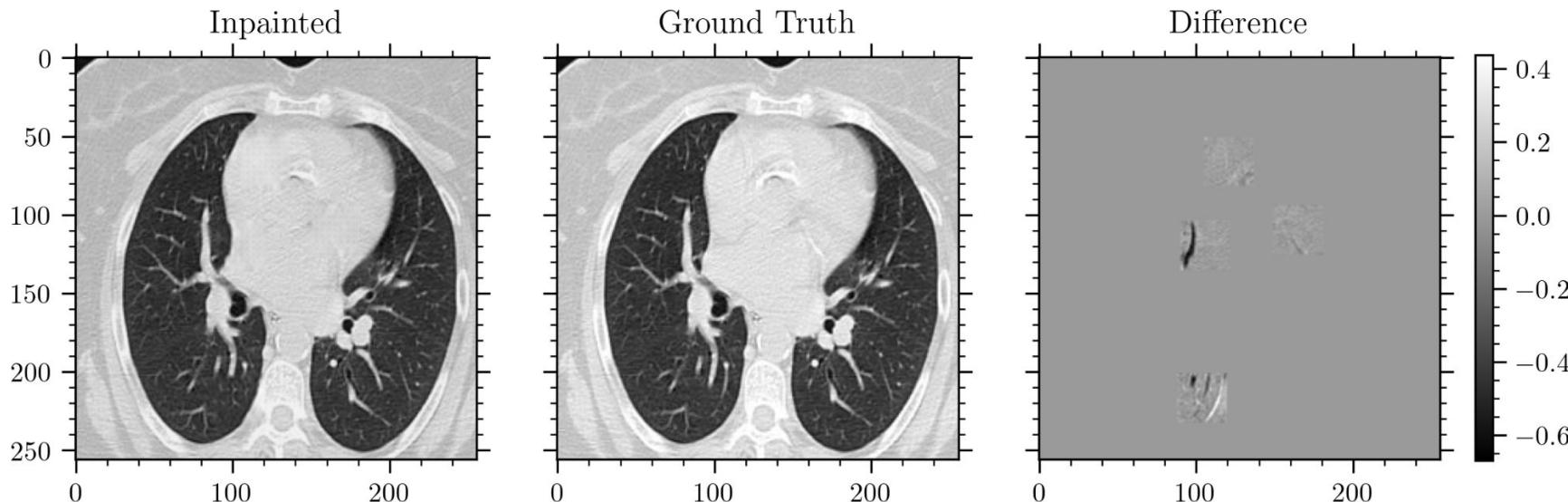
# Out of Distribution Results

- Applying 1 strip model to 2 strip masks



# Out of Distribution Results

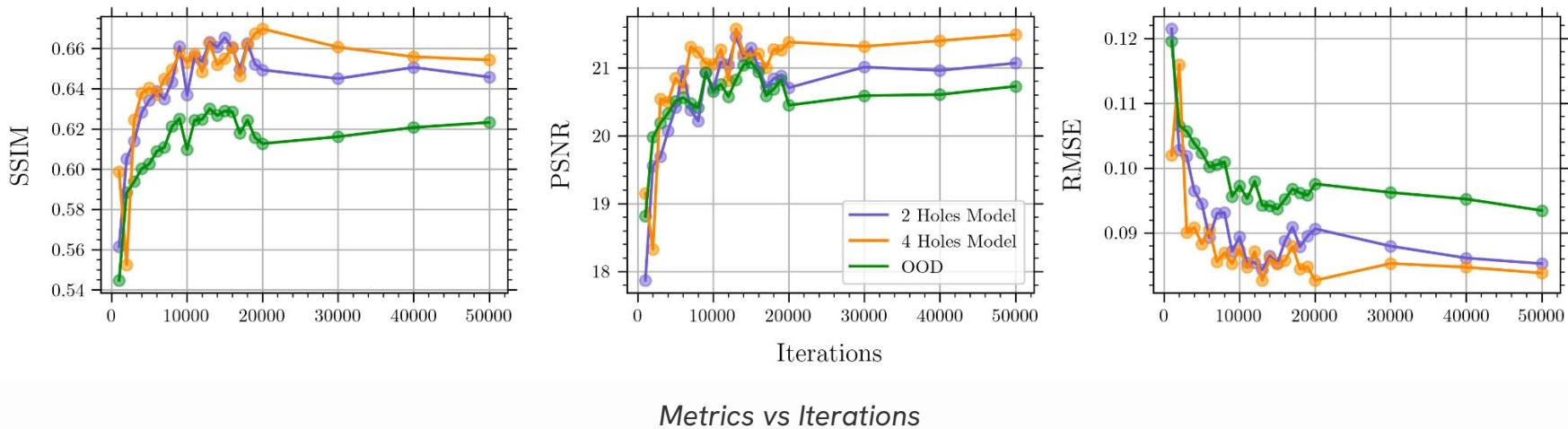
- Applying 2 holes model to 4 hole masks



2 Holes model + 4 Holes masks: 15,000 training + 7,500 fine-tune

# Out of Distribution Results

- Applying 2 holes model to 4 hole masks



# Conclusion

- RFR seems to fit CT and X-ray well visually, and changing hole shape or number doesn't affect the ability of a model to predict by much.
- RFR, leads to visually appealing inpainting, however it lacks accuracy.
- X-ray OOD results suggest that performance will drop significantly if unexpected features are present in the images - bad predictive ability
- Accuracy is crucial in medical image inpainting, hence other methods such as 'Medical image inpainting with edge and structure priors' or 'Multi-Task Learning for Medical Image Inpainting Based on Organ Boundary Awareness' fair better.

# Acknowledgements

- Recurrent Feature Reasoning for Image Inpainting: <https://arxiv.org/pdf/2008.03737>
- Code implementation by authors: <https://github.com/jingyuanli001/RFR-Inpainting/>
- Our code: <https://github.com/ravioli1369/CS736-Project>

# **Thank You**