Intelligent Critical Care Center
UNIVERSITY of FLORIDA

Image Super-Resolution

EEL 6935: Deep Learning in Medical Image Analysis

Fall 2025

Wei Sha, PhD

What is Image Super-Resolution?

- Restore high-resolution (HR) images from low-resolution (LR) images.
- Single-image super-resolution (SISR): one input LR image.
- Multi-image super-resolution (MISR): multiple input LR images.

High-Resolution Image

Low-Resolution Image

Output from Diffusion Model

SISR

Single-Image Super-Resolution

Interpolation-based Methods

- Bilinear or bicubic (based on cubic b-spline) interpolation.
- Faster, but low quality.

Question 1: Fundamentals of Image Super-Resolution

- Which of the following best describes the goal of Single-Image Super-Resolution (SISR)?
  A. To enhance multiple low-resolution images into one high-resolution image.
  B. To generate a high-resolution image from a single low-resolution input.
  C. To align multiple 3D MRI volumes using affine transformation.
  D. To reduce the computational cost of convolutional neural networks.

Question 1: Fundamentals of Image Super-Resolution

- Which of the following best describes the goal of Single-Image Super-Resolution (SISR)?
  A. To enhance multiple low-resolution images into one high-resolution image.
  B. To generate a high-resolution image from a single low-resolution input.
  C. To align multiple 3D MRI volumes using affine transformation.
  D. To reduce the computational cost of convolutional neural networks.

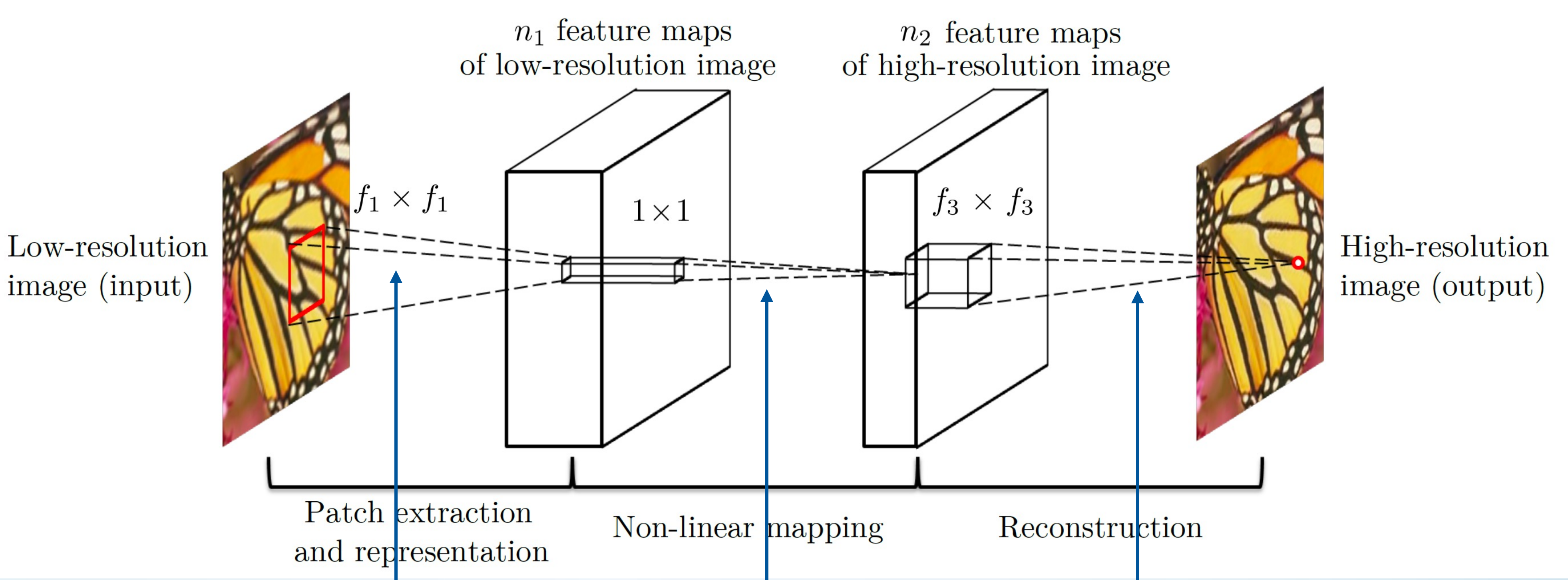
Image Super-Resolution Using Deep Convolutional Networks (SRCNN)

- Proposed by Dong et al. in 2014.
- End-to-end mapping.
- A lightweight structure yet demonstrates SOTA performance.

Original / PSNR   Bicubic / 24.04 dB   SRCNN / 27.58 dB

Dong, Chao, et al. "Learning a deep convolutional network for image super-resolution." *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13*. Springer International Publishing, 2014.

SRCNN Architecture



$F_1(\mathbf{Y}) = \max(0, W_1 * \mathbf{Y} + B_1)$

$W_1 : c \times f_1 \times f_1 \times n_1$

Y : low-resolution image.

$F_2(\mathbf{Y}) = \max(0, W_2 * F_1(\mathbf{Y}) + B_2)$

$W_2 : n_1 \times 1 \times 1 \times n_2$

$F(\mathbf{Y}) = W_3 * F_2(\mathbf{Y}) + B_3$

$W_3 : n_2 \times f_3 \times f_3 \times c$

Dong, Chao, et al. "Learning a deep convolutional network for image super-resolution." *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13*. Springer International Publishing, 2014.

Question 2: Model Architecture

- In SRCNN, what is the role of the first convolutional layer?
  A. To perform pixel-wise classification of image patches.
  B. To upsample the image using bicubic interpolation.
  C. To extract feature representations from the low-resolution input.
  D. To compute the perceptual loss from a pre-trained VGG network.

Question 2: Model Architecture

- In SRCNN, what is the role of the first convolutional layer?
  A. To perform pixel-wise classification of image patches.
  B. To upsample the image using bicubic interpolation.
  C. To extract feature representations from the low-resolution input.
  D. To compute the perceptual loss from a pre-trained VGG network.

Loss Function

- The Mean Squared Error (MSE) is used as the low function.

$$L(\Theta) = \frac{1}{n} \sum_{i=1}^{n} ||F(\mathbf{Y}_i; \Theta) - \mathbf{X}_i||^2$$

$\{\mathbf{X}_i\}$: a set of high-resolution images.

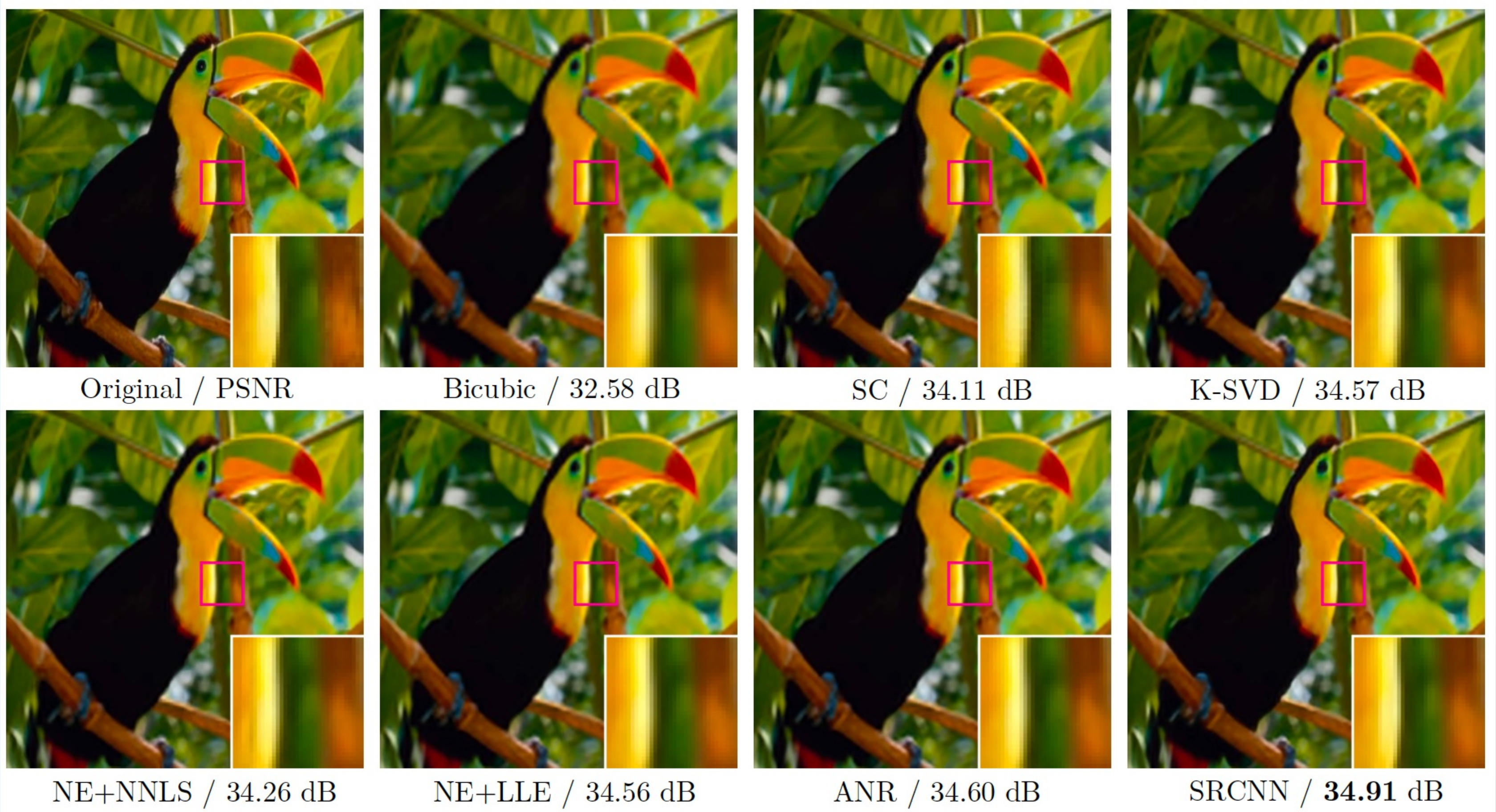
$\{\mathbf{Y}_i\}$: corresponding set of low-resolution images.

Dong, Chao, et al. "Learning a deep convolutional network for image super-resolution." *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13*. Springer International Publishing, 2014.

Evaluation Metrics

- Peak-signal-to-noise ratio (PSNR)
  - Formula: $PSNR = 20 \cdot \log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right).$
  - MAX_I is the maximum pixel value in the image. For 8-bit grayscale images, it is 255.
  - MSE is the Mean Squared Error between the original and the super-resolved image.
  - Higher PSNR values indicate better quality.
- Structural similarity index (SSIM)
  - Measure perceived quality by comparing local patterns of pixel intensities in images.
  - SSIM ranges from -1 to 1, with 1 being the best value.

Qualitative Result



Dong, Chao, et al. "Learning a deep convolutional network for image super-resolution." *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13*. Springer International Publishing, 2014.

Quantitative Result

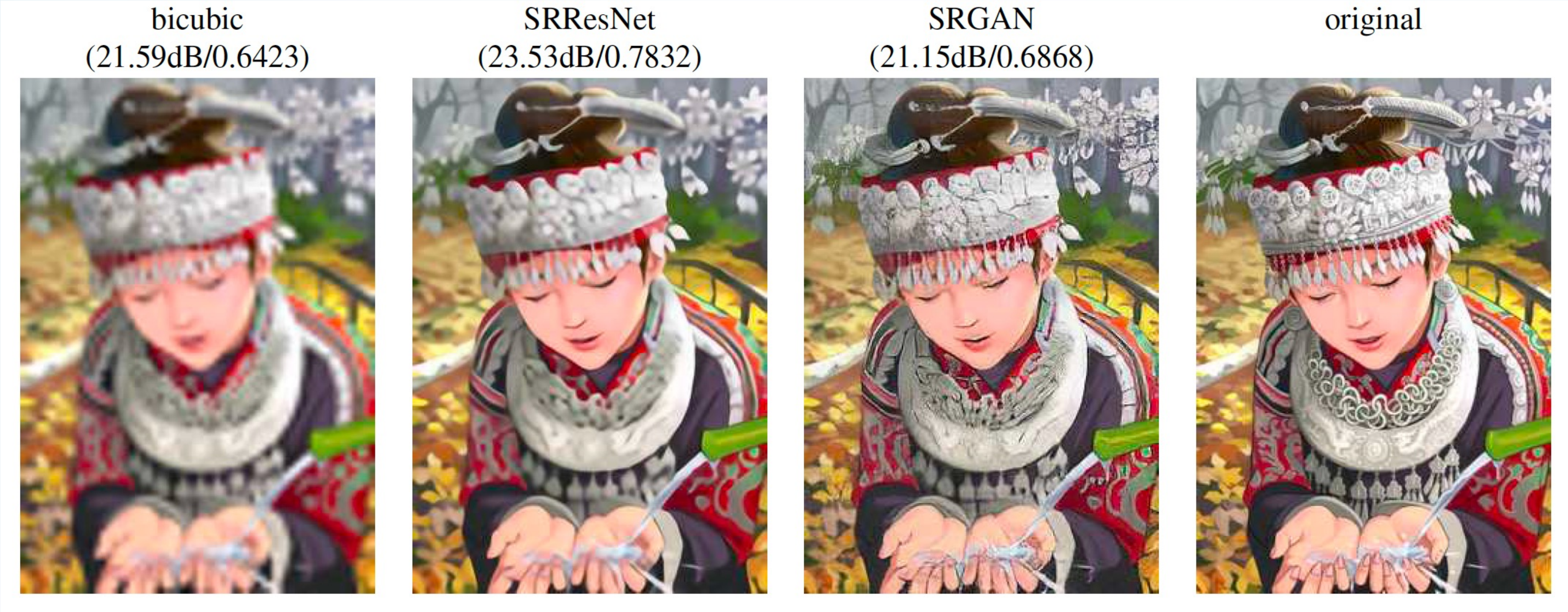
Set5 [2] images	$n_1 = 128,\ n_2 = 64$		$n_1 = 64,\ n_2 = 32$		$n_1 = 32,\ n_2 = 16$	
	PSNR	Time	PSNR	Time	PSNR	Time
	32.60	0.60	32.52	0.18	32.26	0.05

n_1: number of $f_1 x f_1$ convolutional filters in the first convolutional layer.
n_2: number of 1x1 convolutional filters in the second convolutional layer.

Dong, Chao, et al. "Learning a deep convolutional network for image super-resolution." *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13*. Springer International Publishing, 2014.

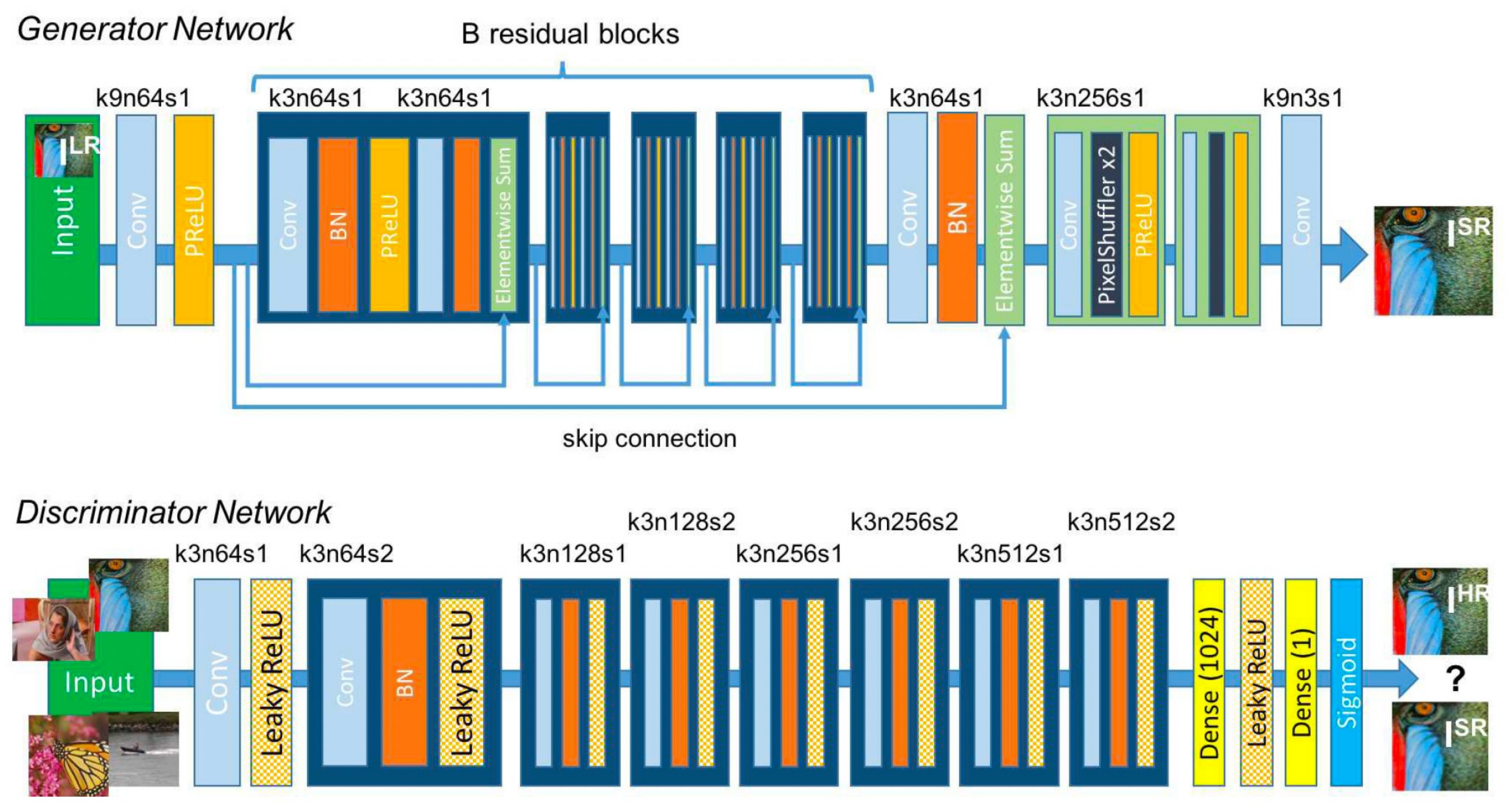
Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network (SRGAN)

- Proposed by Ledig et al. in 2017.
- Used a perceptual loss for generating realistic images.



Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

Architecture of SRGAN

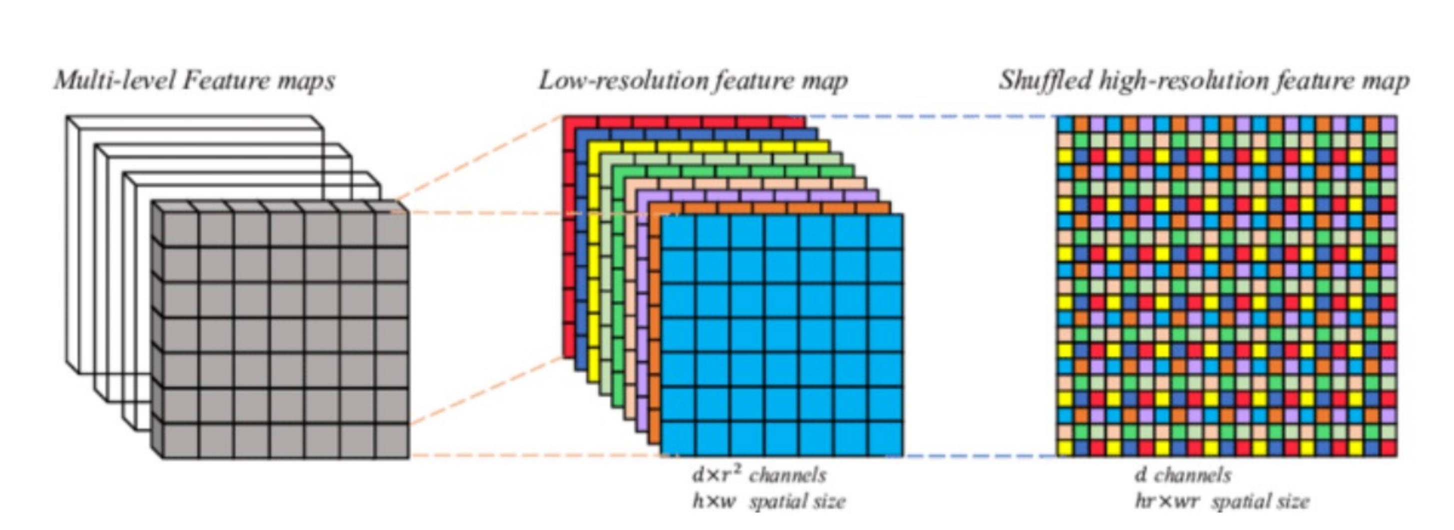


$$\min_{\theta_G} \max_{\theta_D} \mathbb{E}_{I^{HR} \sim p_{\text{train}}(I^{HR})}[\log D_{\theta_D}(I^{HR})] + \mathbb{E}_{I^{LR} \sim p_G(I^{LR})}[\log(1 - D_{\theta_D}(G_{\theta_G}(I^{LR}))]$$

Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

Pixel Shuffle

- Pixel Shuffle (also known as Sub-Pixel Convolution) is a learnable upsampling method.
- Rearranges feature map channels into higher spatial dimensions to increase image resolution.



Loss for the Generator

- Weighted sum of a content loss and an adversarial loss:

$$l^{SR} = \underbrace{\underbrace{l^{SR}_{X}}_{\text{content loss}} + \underbrace{10^{-3} l^{SR}_{Gen}}_{\text{adversarial loss}}}_{\text{perceptual loss (for VGG based content losses)}}$$

- Content loss
  - MSE loss:

$$l^{SR}_{MSE} = \frac{1}{r^2 WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I^{HR}_{x,y} - G_{\theta_G}(I^{LR})_{x,y})^2$$

  - VGG loss:

$$l^{SR}_{VGG/i.j} = \frac{1}{W_{i,j} H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2$$

Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

Adversarial Loss

- Adversarial loss encourages the network to favor solutions that reside on the manifold of natural images.

$$l_{Gen}^{SR} = \sum_{n=1}^{N} -\log D_{\theta_D}(G_{\theta_G}(I^{LR}))$$

Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

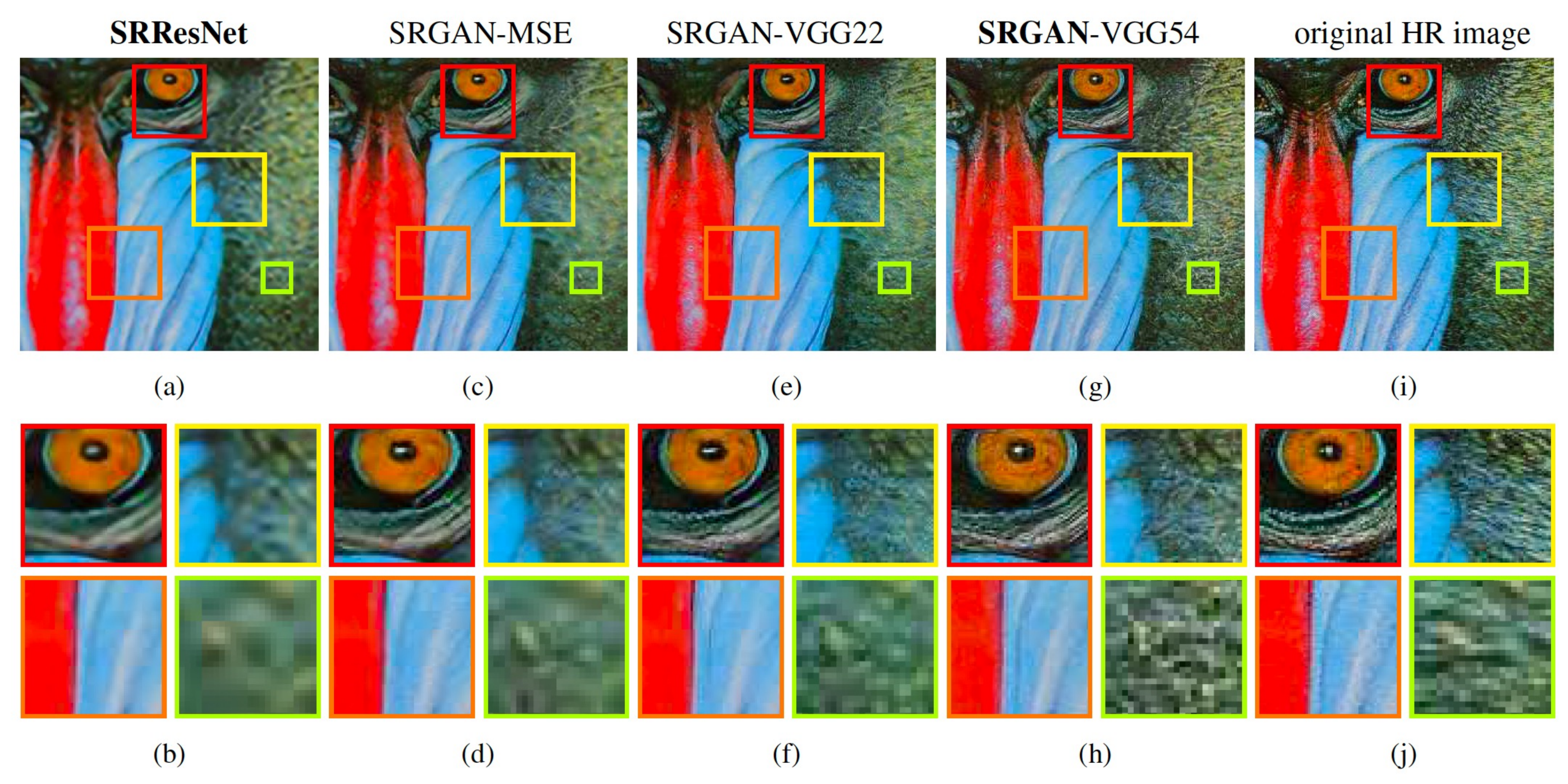
Discriminator Loss

- Maximize the following loss for training the discriminator:

$$\max_{\theta_D} \mathbb{E}_{I^{HR} \sim p_{\text{train}}(I^{HR})}[\log D_{\theta_D}(I^{HR})] + \mathbb{E}_{I^{LR} \sim p_G(I^{LR})}[\log(1 - D_{\theta_D}(G_{\theta_G}(I^{LR})))]$$

Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

Qualitative Result



Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

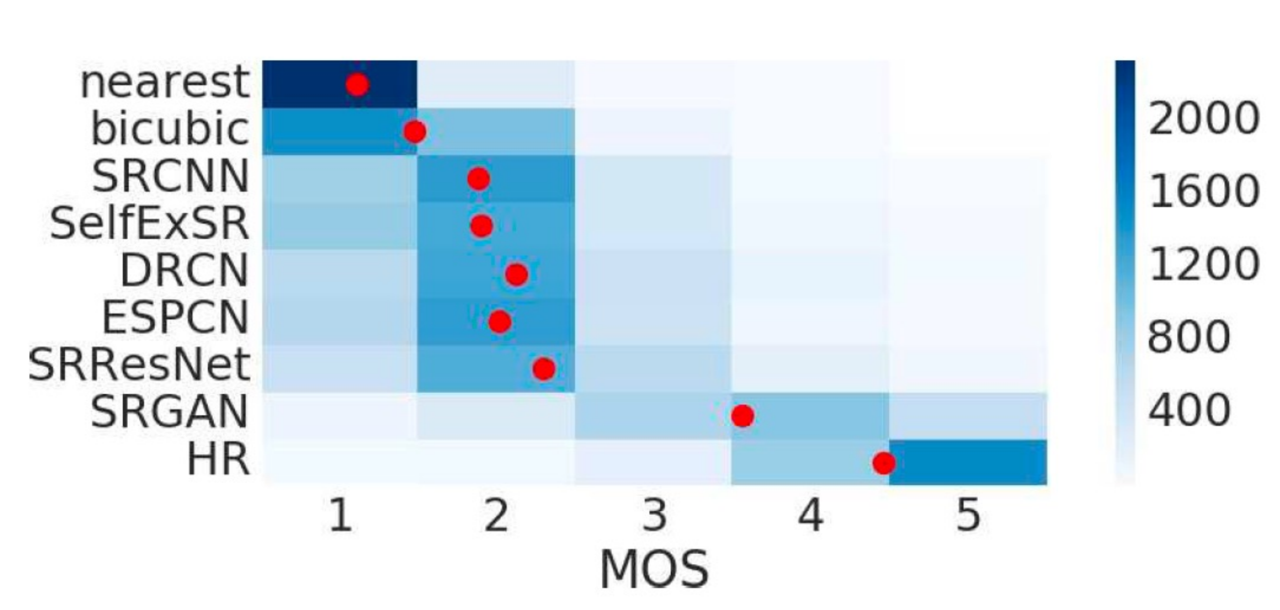
Quantitative Result

Set5	nearest	bicubic	SRCNN	SelfExSR	DRCN	ESPCN	**SRResNet**	**SRGAN**	HR
PSNR	26.26	28.43	30.07	30.33	31.52	30.76	**32.05**	29.40	∞
SSIM	0.7552	0.8211	0.8627	0.872	0.8938	0.8784	**0.9019**	0.8472	1
MOS	1.28	1.97	2.57	2.65	3.26	2.89	3.37	**3.58**	4.32
Set14									
PSNR	24.64	25.99	27.18	27.45	28.02	27.66	**28.49**	26.02	∞
SSIM	0.7100	0.7486	0.7861	0.7972	0.8074	0.8004	**0.8184**	0.7397	1
MOS	1.20	1.80	2.26	2.34	2.84	2.52	2.98	**3.72**	4.32
BSD100									
PSNR	25.02	25.94	26.68	26.83	27.21	27.02	**27.58**	25.16	∞
SSIM	0.6606	0.6935	0.7291	0.7387	0.7493	0.7442	**0.7620**	0.6688	1
MOS	1.11	1.47	1.87	1.89	2.12	2.01	2.29	**3.56**	4.46

Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

Quantitative Result

- Mean opinion score (MOS): evaluate super-resolved images with multiple raters on a scale of 1 (worst) to 5 (best) to assess perceptual quality of various methods.



Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

Question 3: Evaluation Metrics and Loss Functions

- A. SRGAN relies solely on Mean Squared Error (MSE) loss for reconstruction.
  B. The discriminator in SRGAN minimizes a perceptual loss to generate sharper images.
  C. The generator loss combines a content loss (e.g., VGG-based) and an adversarial loss.
  D. SRGAN achieves high PSNR by suppressing texture details.

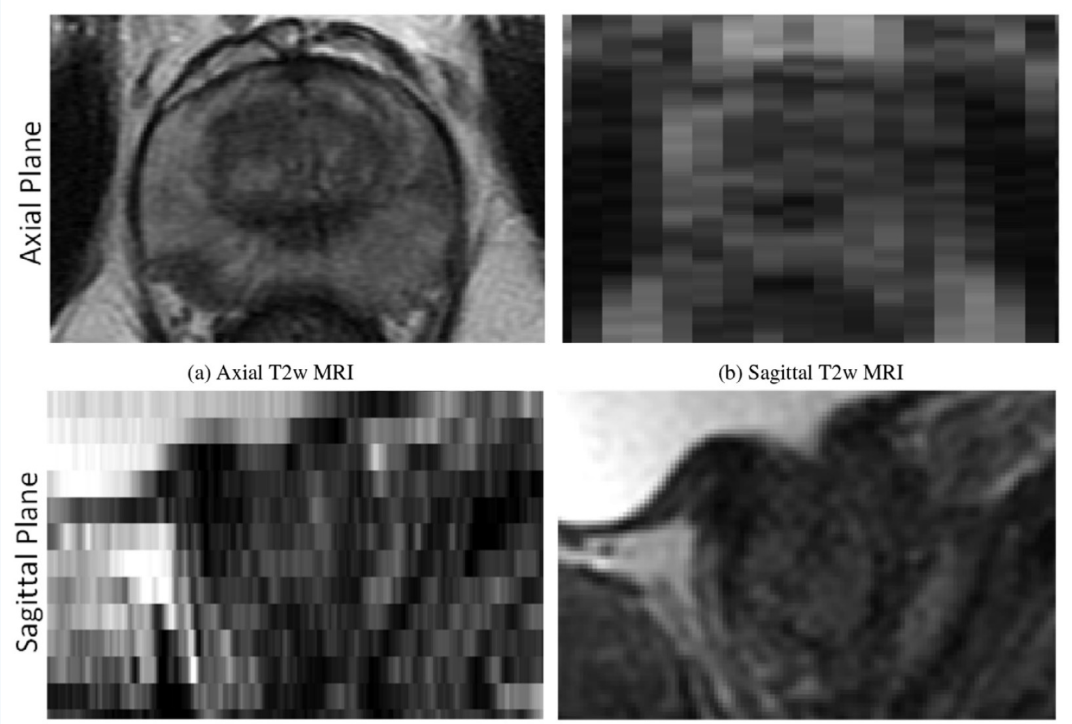
Question 3: Evaluation Metrics and Loss Functions

- A. SRGAN relies solely on Mean Squared Error (MSE) loss for reconstruction.
  B. The discriminator in SRGAN minimizes a perceptual loss to generate sharper images.
  C. The generator loss combines a content loss (e.g., VGG-based) and an adversarial loss.
  D. SRGAN achieves high PSNR by suppressing texture details.

Multi-Image Super-Resolution

Motivation

- Medical images often exhibit high in-plane resolution but have limited through-plane resolution, making them not true 3D volumes.

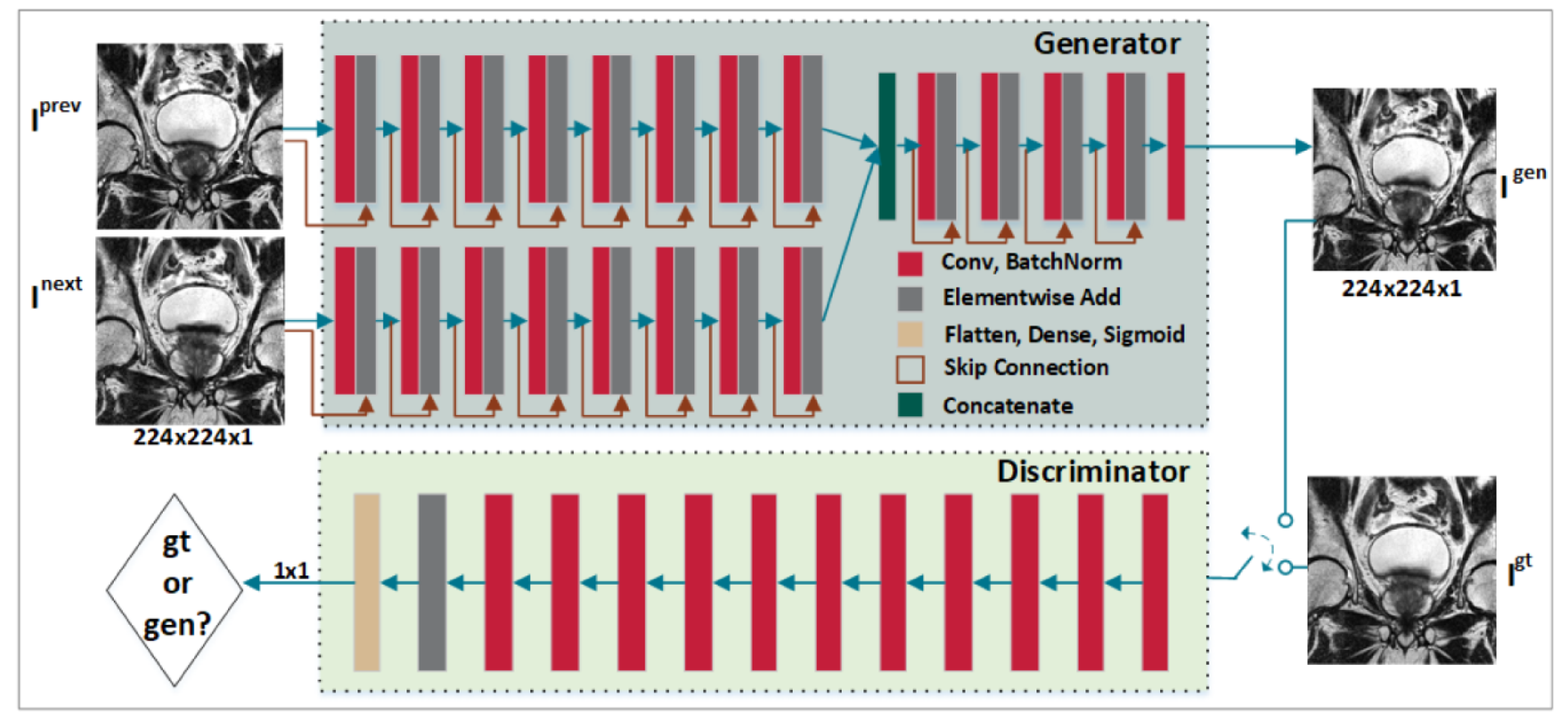


(a) Axial T2w MRI (b) Sagittal T2w MRI

(d) Axial T2w MRI (e) Sagittal T2w MRI

Rewa Sood* and Wei Shao*, et al. "3D Registration of pre-surgical prostate MRI and histopathology images via super-resolution volume reconstruction." *Medical Image Analysis* 69 (2021): 101957. * Equal contribution.

Multi-Image Super-Resolution Generative Adversarial Network



Rewa Sood* and Wei Shao*, et al. "3D Registration of pre-surgical prostate MRI and histopathology images via super-resolution volume reconstruction." *Medical Image Analysis* 69 (2021): 101957. * Equal contribution.

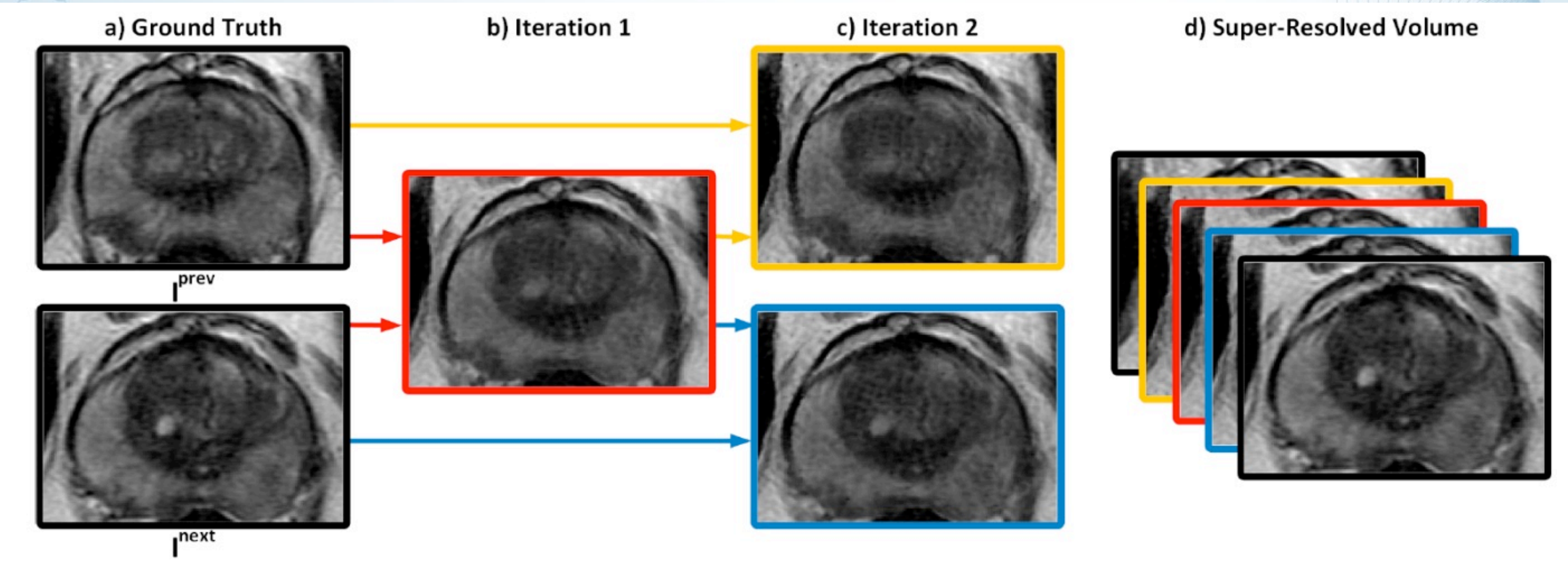
Loss Function

- Loss for the generator: $l_G = l_{MSE} + \beta l_{percep} + \gamma l_{GAN}$

where $l_{MSE} = \frac{1}{N}||I^{gt} - I^{gen}||_2^2$ $l_{percep} = \frac{1}{N}||VGG(I^{gt}) - VGG(I^{gen})||_2^2$ $l_{GAN} = \sum_{n=1}^{N}(-\log D(I_n^{gen}))$

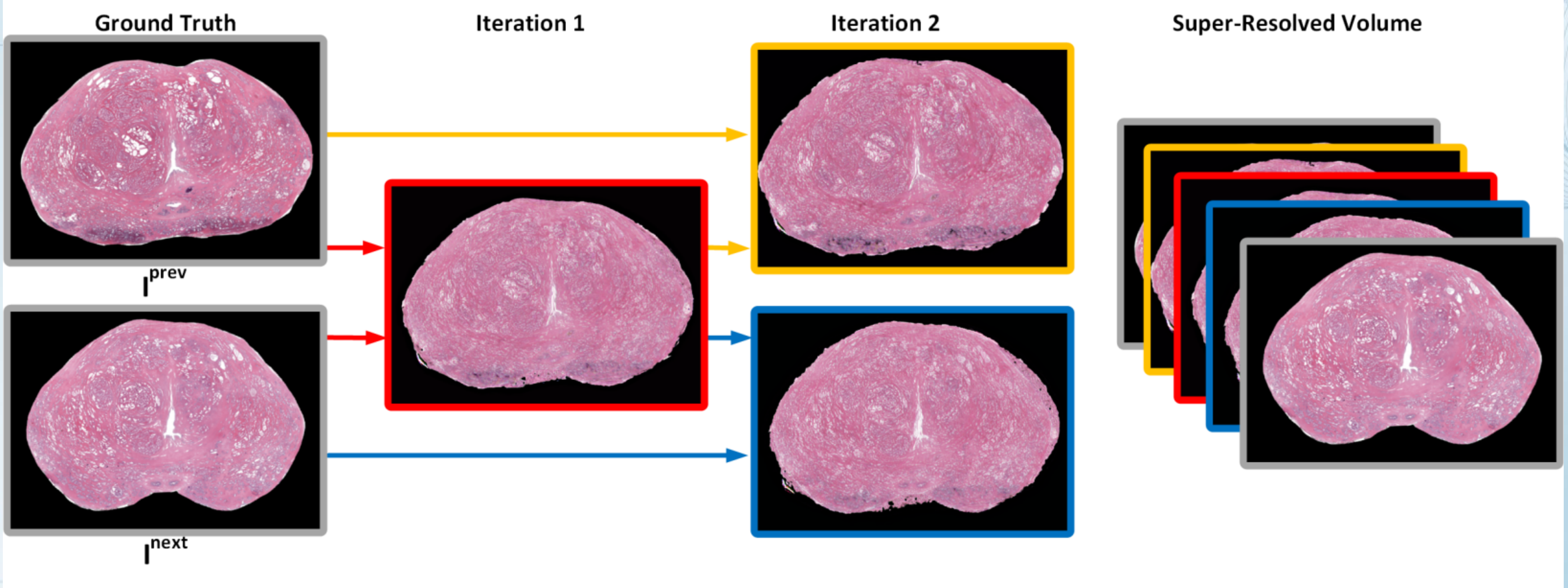
- Loss for the discriminator: $l_D = \sum_{n=1}^{N}\left(\log D(I_n^{gt}) + \log(1 - D(I_n^{gen}))\right)$

MRI Super-Resolution Strategy



Rewa Sood* and Wei Shao*, et al. "3D Registration of pre-surgical prostate MRI and histopathology images via super-resolution volume reconstruction." *Medical Image Analysis* 69 (2021): 101957. * Equal contribution.

Histopathology Super-Resolution Strategy



Rewa Sood* and Wei Shao*, et al. "3D Registration of pre-surgical prostate MRI and histopathology images via super-resolution volume reconstruction." *Medical Image Analysis* 69 (2021): 101957. * Equal contribution.

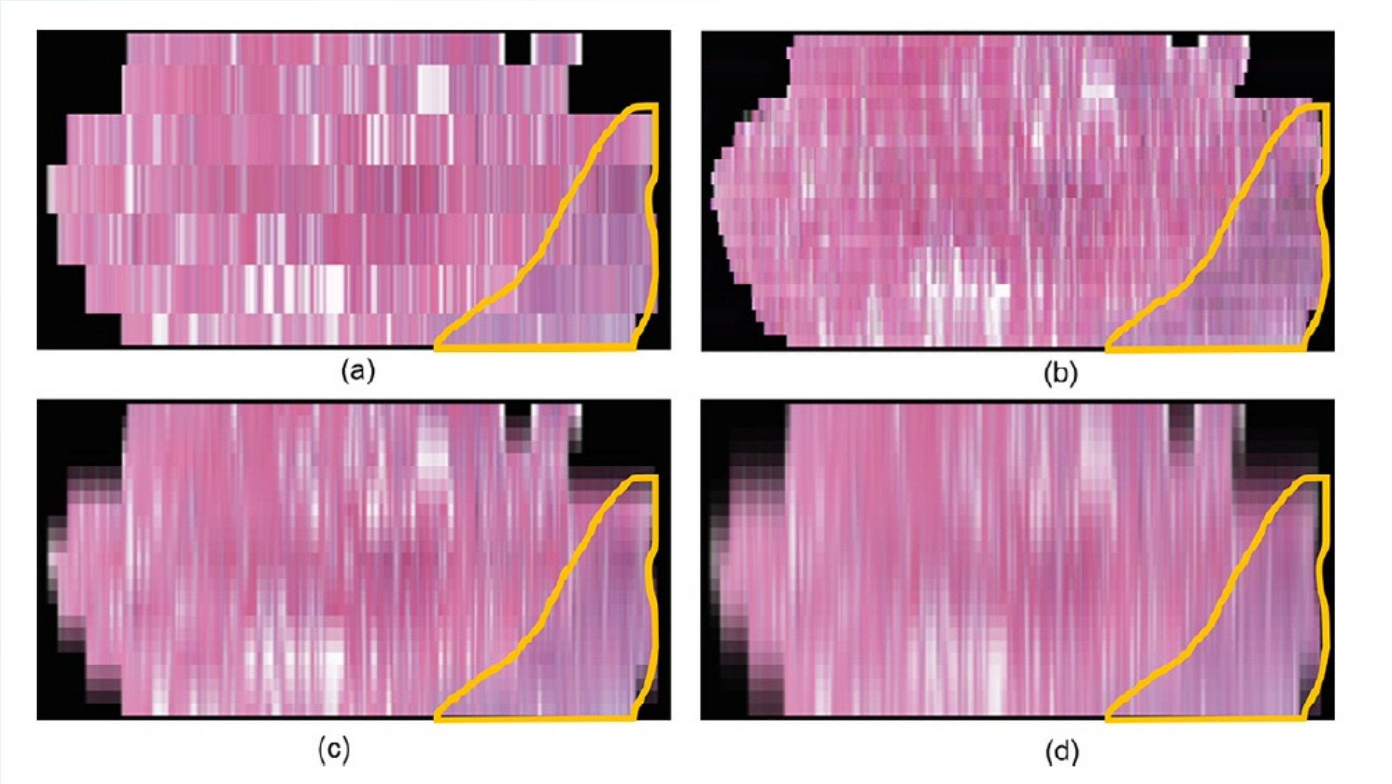
Super-Resolved MRI

(a) Input: 2.5D T2w Axial MRI (b) Super-Resolved MRI (ours) (c) Ground Truth: 2.5D T2w Sagittal MRI

Sagittal view of (a) 2.5D T2w Axial MRI, (b) 3D T2w MRI volumes, and (c) super-resolution

Rewa Sood* and Wei Shao*, et al. "3D Registration of pre-surgical prostate MRI and histopathology images via super-resolution volume reconstruction." *Medical Image Analysis* 69 (2021): 101957. * Equal contribution.

Super-Resolved Histopathology



Sagittal view of the (a) input histopathology, (b) super-resolved histopathology volume, (c) histopathology upsampled by linear interpolation, and (d) histopathology upsampled by third-order B-spline interpolation.

Rewa Sood* and Wei Shao*, et al. "3D Registration of pre-surgical prostate MRI and histopathology images via super-resolution volume reconstruction." *Medical Image Analysis* 69 (2021): 101957. * Equal contribution.