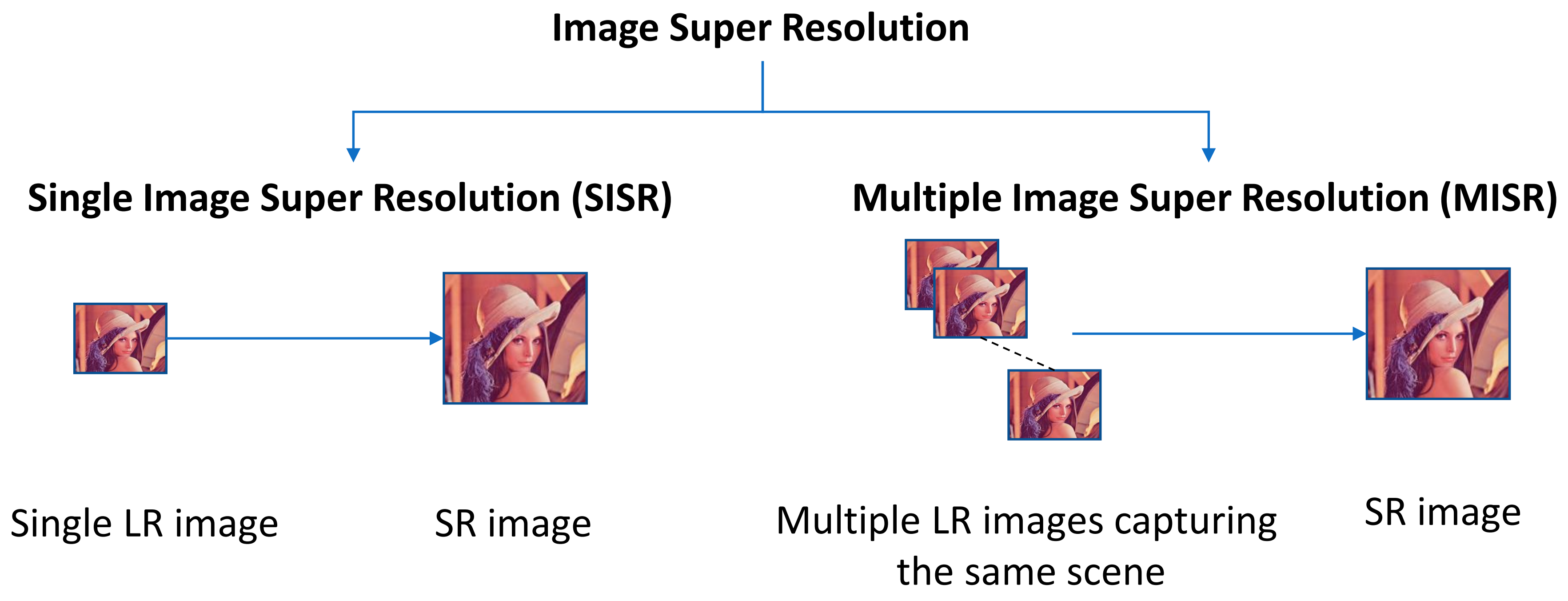
Self-supervised Super Resolution of Ultrasound images

Yukta Thapliyal

Dept. of Electrical & Computer Engineering, McGill University, Montreal

Table of Contents

- Introduction
- Motivation
- Related Work
- Background
- Methodology
- Experimental results
- Conclusion and future work



Key Features:

- Upscales the image to higher resolution
- Improves details within the image
- Part of computer vision and artificial intelligence

Limitations:

- Dependent upon input LR image
- Can create details within SR image that were not originally present

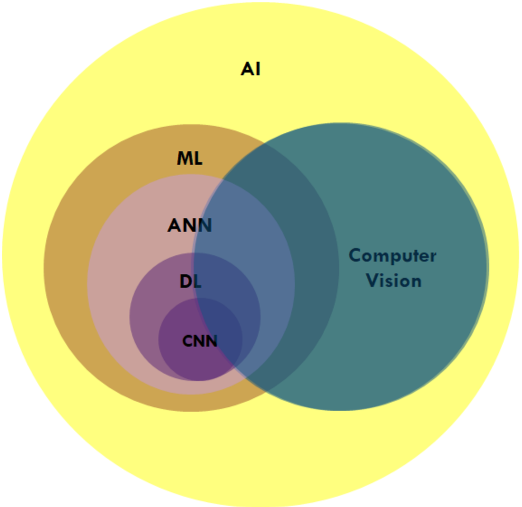
Why deep learning (DL) based SISR for US images?

Explore applications of DL algorithm in US:

- Current clinical post-processing of US scans uses proprietary methods and are based on signal processing techniques
- Attempt to see if DL can be used as well

Learn US image enhancement methods:

- Attempt to learn US image enhancement methods that can be use for post-processing of US images.



Contribution of this project:

- Review and summarise CNN based works in both natural image SR and US image SR
- Perform SISR of US images using self-supervised methods with no assumption on source of LR images
- Combine DL networks and wavelets decomposition

Taxonomy of ML

Supervised Learning:
- Learning from labelled datasets to known output

Unsupervised Learning:
- Explore patters in unlabelled datasets and predict output

Self-supervision Learning:
- Create supervisory signal from unlabelled dataset itself
- No external labelled datasets

Reinforcement Learning:
- Learning in a reward and action systems

Natural Image SISR

Traditional CNN based:
- General method: first upsample the image e.g. by bicubic interpolation, and then feed into conventional CNN
- Supervised
- Eg: SRCNN [1], VDSR [2]

Non- traditional CNN based:
- Add modifications to traditional CNN based models like residuals, recursions, sub-pixel convolution, wavelets
- Supervised
- Eg: MWCNN [3], ESPCN [4]

GAN based:
- Generate synthetic data samples based on provided data
- Unsupervised
- Eg: SRGAN [5], ESRGAN [6]

Self supervised learning:
- Takes only one test image and no other human annotated data as input
- Generates samples from input image for training
- Unsupervised
- Eg: ZSSR [7], KernelGAN [8]

US image SISR

GAN based:

- Given input LR US image, tries to generate synthetic HR images based on HR image representations
- Semi-supervised or unsupervised
- Eg: DRNN [9]

Deep CNN based:

- Use deep CNN
- Eg: DECUSR [10]

U-Net based:

- Uses popular biomedical image segmentation network
- Eg: deepULM [11]

Challenges in US-SISR:

- No DL model accepted as standard of US-SISR
- Limited dataset
- No existence labelled dataset, i.e., tagged pairs (LR-HR pairs) for training
- According to our literature review, all models make assumption about the source of LR images
- SR is difficult in case of unknown source of degradation

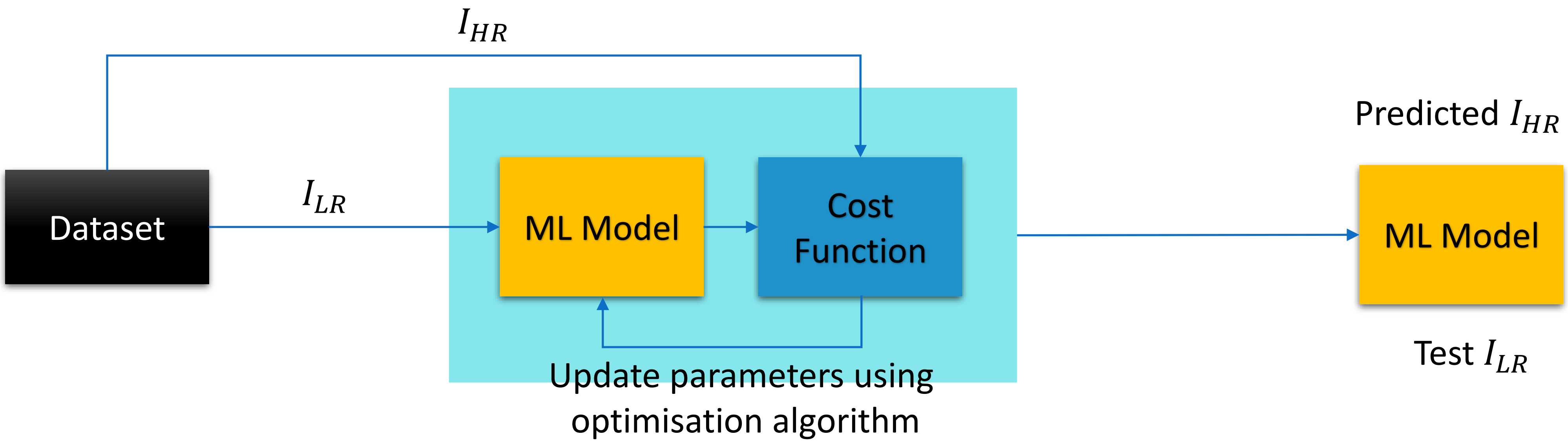
ML Algorithm

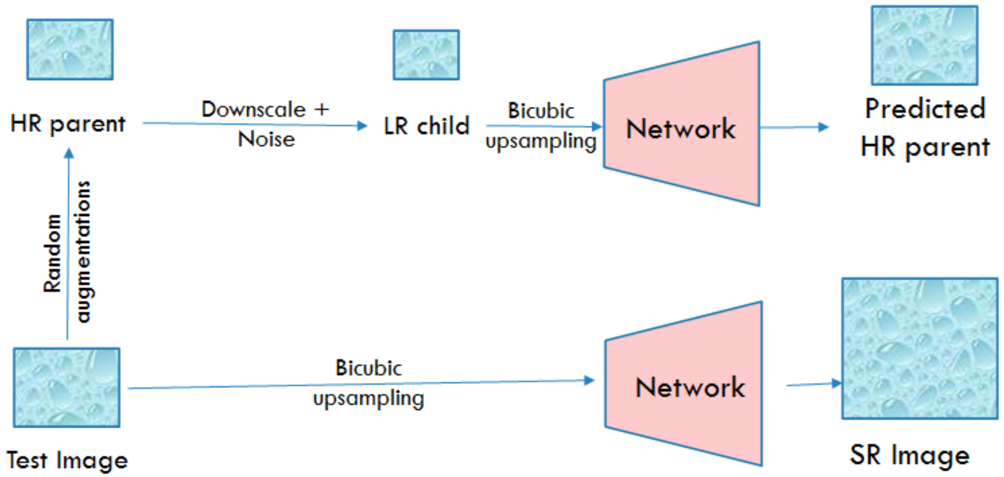
Model:
- Core of the ML algorithm that performs the ML task

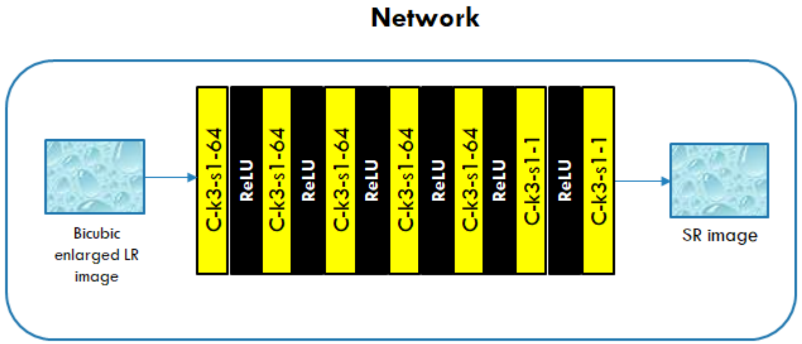
Cost Function and Optimisation algorithm :
- Cost function formulates mathematical relationship between learned output and target and is minimises through an optimisation algorithm

Dataset:
- A collection of instances

Performance Analysis:
- Metrics for analysis of test results



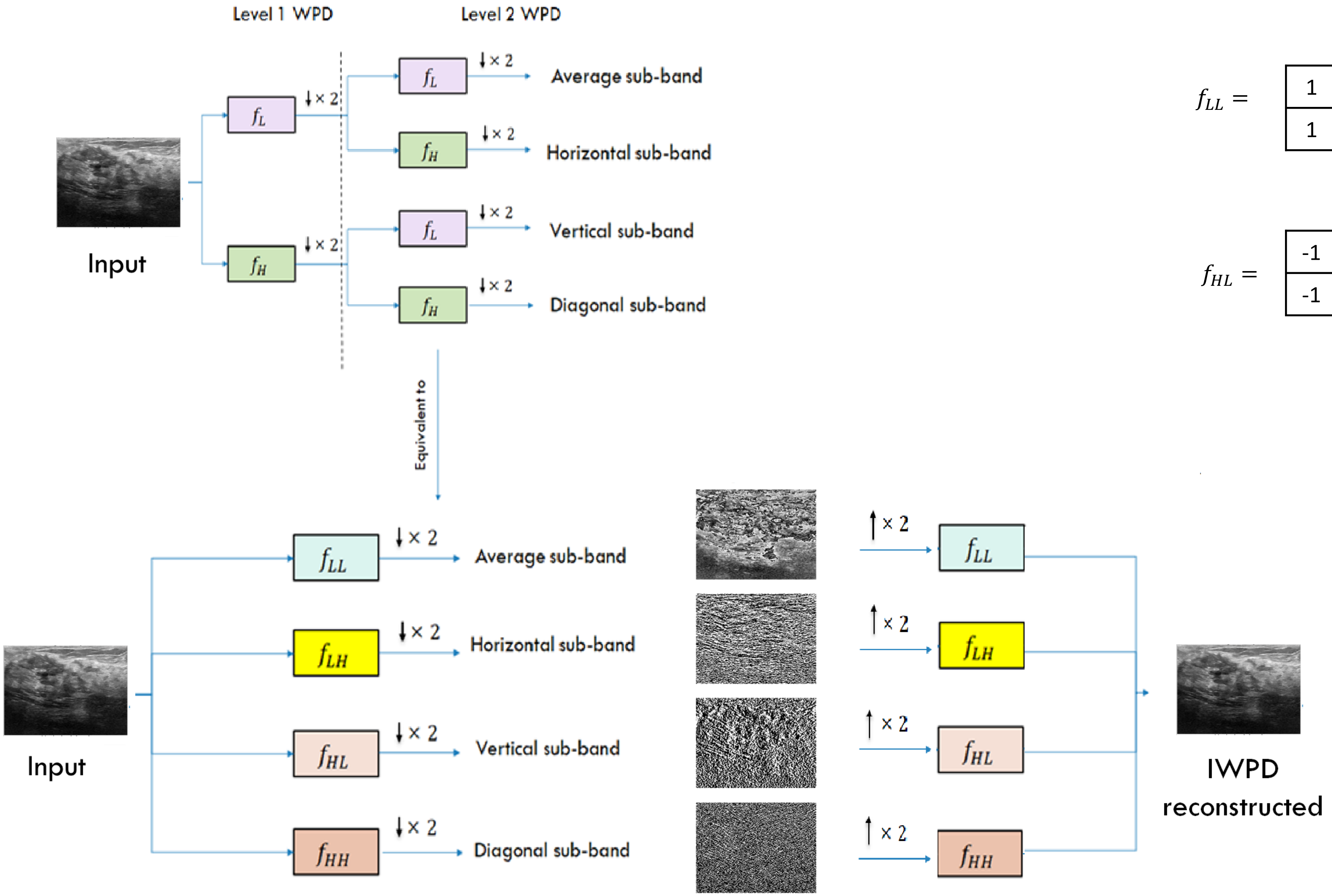




Reconstruction error/ Loss function = $|HR\ parent\ - ZSSR(LR\ child)|_1$

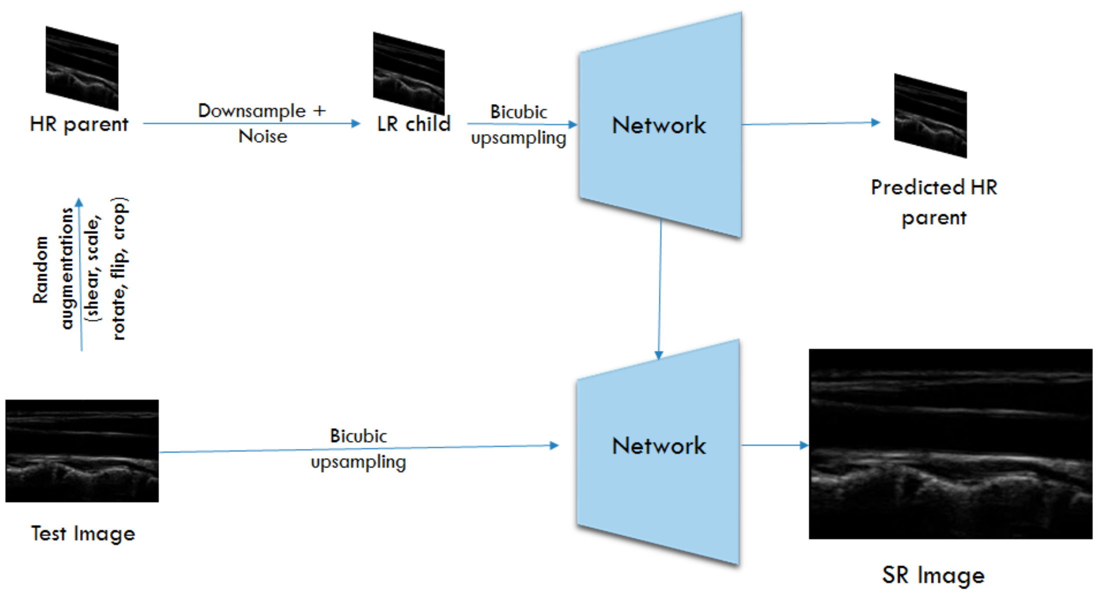
Image from [7]

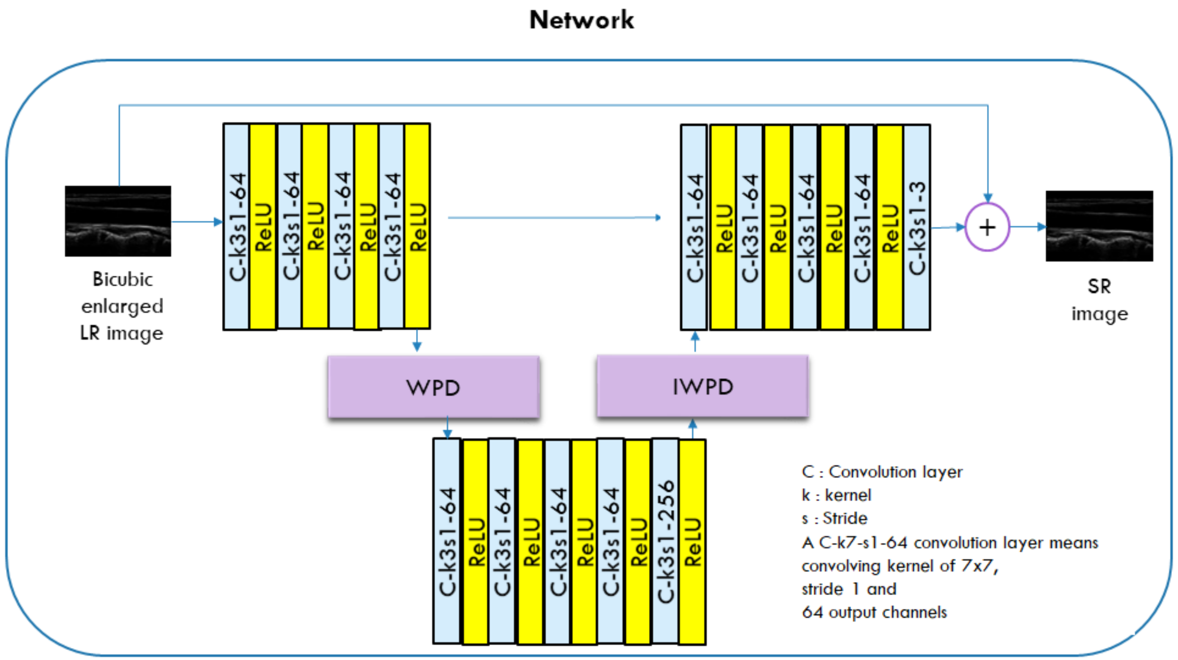
Background: WPD and IWPD



$$f_{LL} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \qquad f_{LH} = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}$$

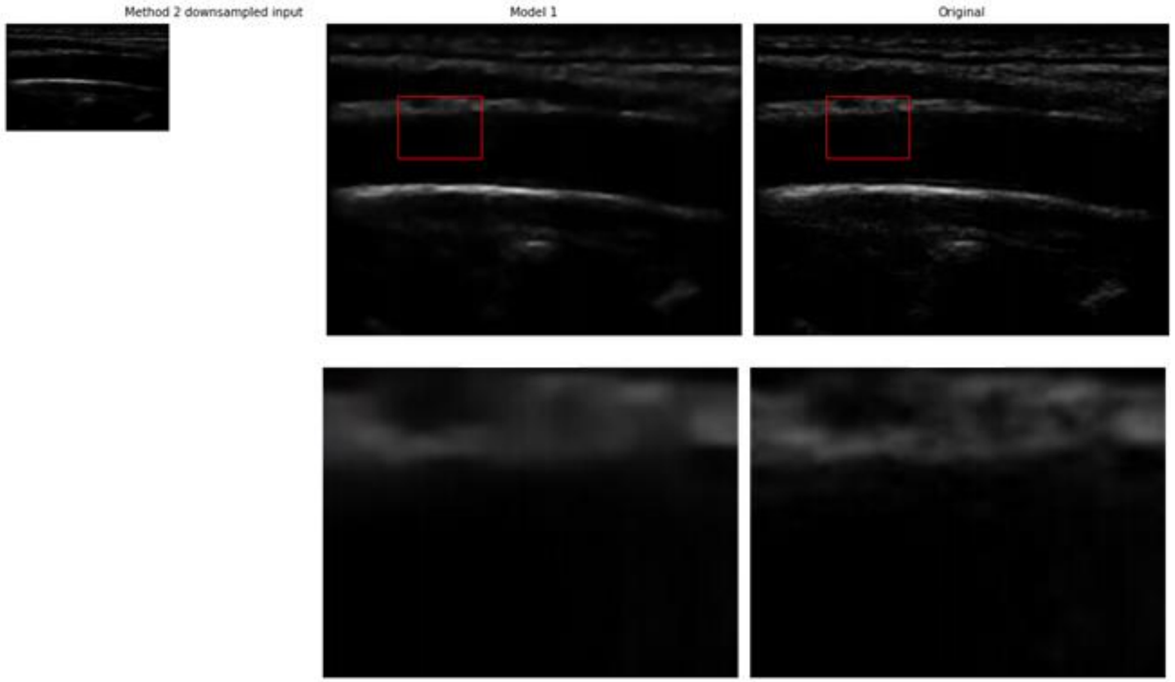
$$f_{HL} = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \qquad f_{HH} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$



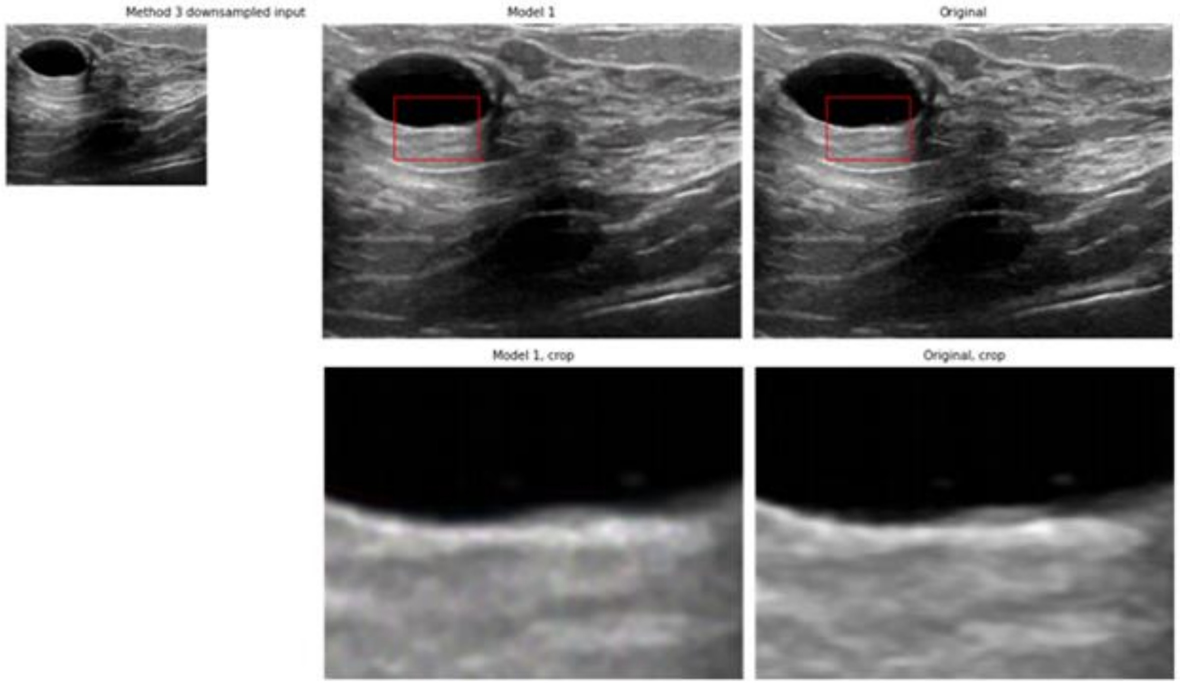


$$Loss = \,|\, I_{HR\ parent} - M_1(I_{LR\ child})|_1 + \lambda\left[1 - \frac{1}{1 + VGG_{loss}}\right], \qquad \lambda = 10$$

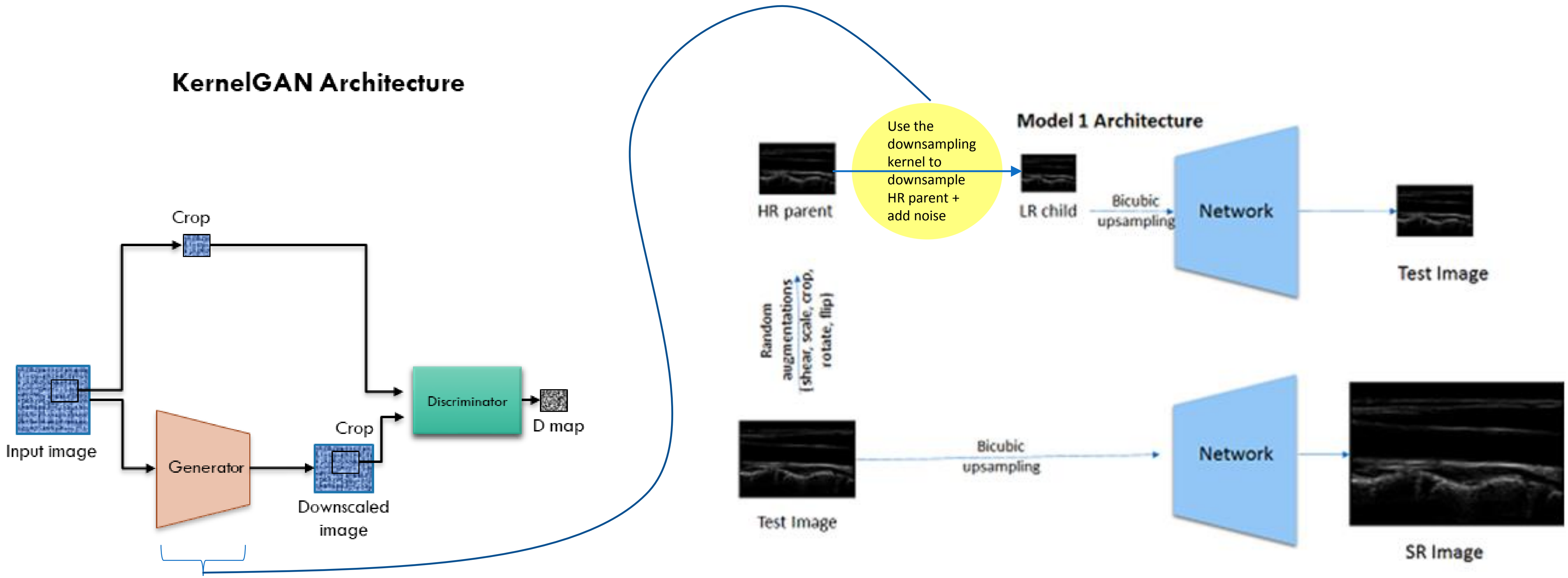
$$VGG_{loss(j,i)} = \frac{1}{W_{j,i} H_{j,i}} \sum_{x=1}^{W_{j,i}} \sum_{y=1}^{H_{j,i}} \left(\phi_{j,i}(I_{HR})_{x,y} - \phi_{j,i}\,(M_1(I_{LR}))_{x,y}\right)^2, \; j = 1, i = 2$$



$\times 2$ upscaling on an image from CCA – US dataset taking input from method 2



$\times 2$ upscaling on an image from Kaggle dataset taking input from method 3



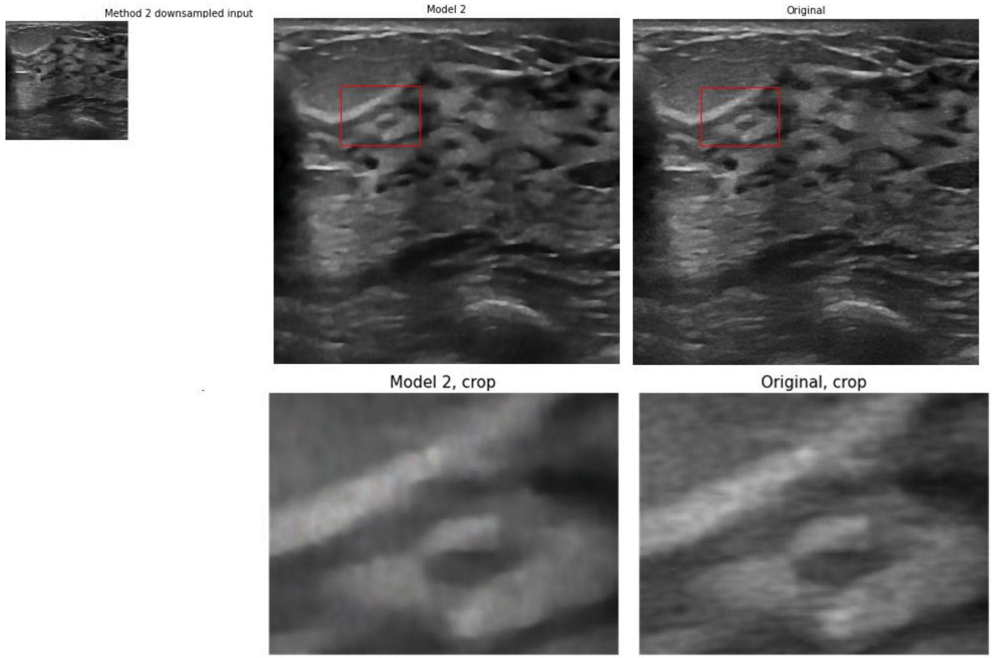
$$\min_G \max_D V(D,G) = E_{x\sim patches(I_{HR\ parent})}\left[\log\left(D(x)\right)\right] + E_{y\sim patches(I_{LR\ child})}\left[\log\left(1 - D(y)\right)\right]$$

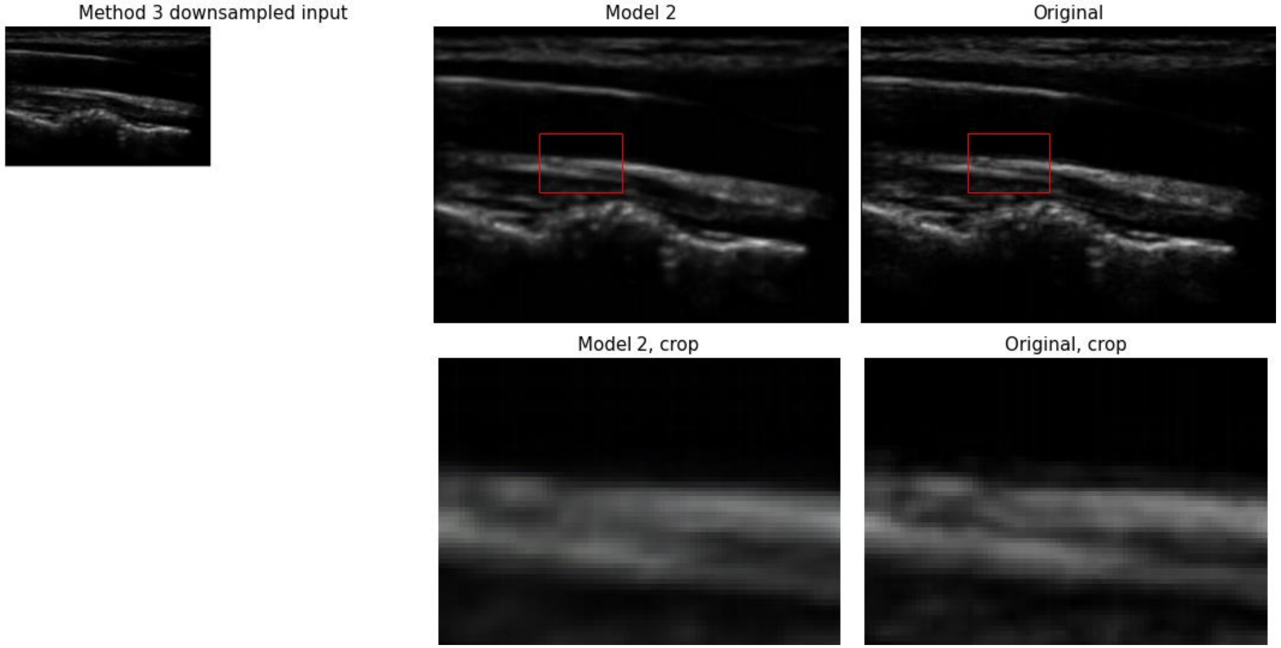
$$G_{loss} = -\log(D(G(patches\ from\ I_{LR} child)))$$

$$D_{loss} = -[\log\left(D\left(patches\ from\ I_{HR\ parent}\right)\right) + \log(1 - D(G(patches\ from\ I_{HR\ child})))]$$

How to obtain downsampling kernel?:

Convolve all the layers of the generator





$\times 2$ upscaling on an image from Kaggle dataset taking input from method 2

$\times 2$ upscaling on an image from CCA – US taking input from method 3

Dataset Generation

Non-ideal downsampling

Ideal downsampling/
Bicubic interpolation/
Method 1

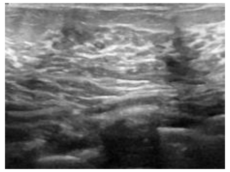
$$I_{LR} = I_{HR} \downarrow_{bic} \times \text{sf}$$

Lanczos kernel +
anisotropic diffusion filter +
speckle noise / Method 2

$$I_{LR} = ((I_{HR} \downarrow_{lan3} \times \text{ sf}) * f_{ad}) + n_r \times ((I_{HR} \downarrow_{lan3} \times \text{ sf}) * f_{ad})$$

Lanczos kernel +
gaussian filter +
speckle noise/ Method 3

$$I_{LR} = \left((I_{HR} \downarrow_{lan3} \times \text{ sf}) * f_g\right) + n_r \times ((I_{HR} \downarrow_{lan3} \times \text{ sf}) * f_g)$$



Original image

$\times\frac{1}{2}$ $\times\frac{1}{4}$

LR images

$\times\frac{1}{2}$ $\times\frac{1}{4}$

LR images

$\times\frac{1}{2}$ $\times\frac{1}{4}$

LR images

- Benchmarking against 3 natural image SISR models: ZSSR, KernelGAN, MWCNN and one US image SISR (named CycleGAN-US-SISR for convenience)
- Generated 6 type of datasets for testing
- Created **2 models** that perform self-supervised SISR of US images.

I_{ref}: Reference image

I_{der}: Derived image

1. PSNR (Peak Signal to Noise Ratio):

$$MSE = \frac{1}{w \times h} \sum_{i=0}^{w-1} \sum_{i=0}^{h-1} \left[I_{ref} - I_{der}\right]^2, PSNR = 20log \frac{Max\ pixel\ value}{\sqrt{MSE}}$$

2. (M)SSIM:

$$SSIM\left(I_{ref}, I_{der}\right) = \frac{(2\mu\left(I_{ref}\right)\mu(I_{der})+C_1)(2\sigma\left(I_{ref}I_{der}\right)+C_2)}{(\mu\left(I_{ref}\right)^2+\mu(I_{der})^2+C_1)(\sigma\left(I_{ref}\right)^2+\sigma(I_{der})^2+C_2)}$$

$$MSSIM\ \left(I_{ref}, I_{der}\right) = \frac{1}{m} \sum_{n=1}^{n} \left(SSIM(I_{ref,i}, I_{der,i})\right)$$

3. Alex Loss :

$$\text{Alex Loss} = \frac{1}{W_j H_j} \sum_{x=1}^{W_j} \sum_{y=1}^{H_j} \left[\phi_j(I_{HR})_{x,y} - \phi_j\left(\text{Model}(I_{LR})\right)_{x,y}\right]^2, j = 5$$

Experimental Results: $\times$ 2, $\times$ 4 upscaling

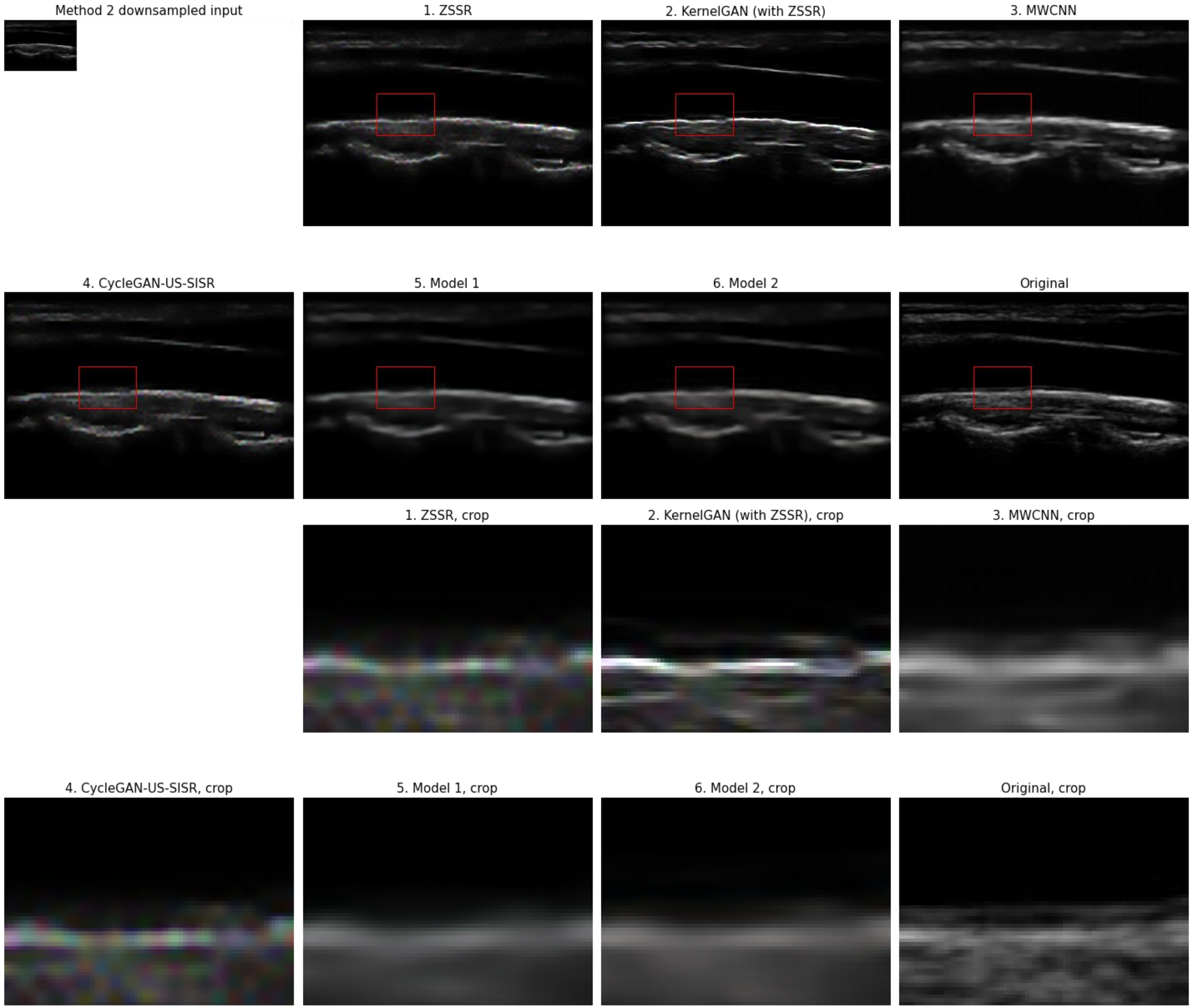
CCA-US dataset (7 images): $\times$2 upscaling				
Model	Method 1	Method 2	Method 3	Evaluation metric
ZSSR [28]	**45.3611** ± 2.1210	34.0592 ± 1.2132	35.5616 ± 1.8275	PSNR
	0.9893 ± 0.0000	0.9188 ± 0.0001	0.9566 ± 0.0000	SSIM
	0.0338 ± 0.0000	0.1522 ± 0.0003	0.1045 ± 0.0002	Alex Loss
KernelGAN [29] + ZSSR [28]	35.8389 ± 2.5026	34.3130 ± 1.8194	37.7225 ± 2.2358	PSNR
	0.9570 ± 0.0001	0.9341 ± 0.0020	0.9657 ± 0.0000	SSIM
	0.0647 ± 0.0001	0.1556 ± 0.0049	**0.0657** ± 0.0001	Alex Loss
MWCNN [12]	29.1289 ± 2.5026	31.2508 ± 2.1547	31.3971 ± 5.8312	PSNR
	0.9570 ± 0.0001	0.8561 ± 0.0020	0.8054 ± 0.0059	SSIM
	0.1856 ± 0.0002	0.1556 ± 0.0049	0.1463 ± 0.0037	Alex Loss
CycleGAN-US-SISR [42]	44.9243 ± 2.3342	34.2508 ± 1.1446	35.7947 ± 1.8226	PSNR
	0.9886 ± 0.0000	0.9184 ± 0.0001	0.9558 ± 0.0001	SSIM
	0.0289 ±n 0.0001	0.1469 ± 0.0003	0.0995 ± 0.0000	Alex Loss
Model1	42.6696 ± 2.2986	**36.9016** ± 0.9783	**38.9662** ± 2.3373	PSNR
	0.9823 ± 0.0000	**0.9397** ± 0.0001	**0.9703** ± 0.0000	SSIM
	0.0781 ± 0.0001	**0.1105**± 0.0002	0.0884± 0.0002	Alex Loss
Model2	40.8408 ± 1.2525	36.3347 ± 1.0041	37.4261 ± 2.1760	PSNR
	0.9360 ± 0.0004	0.9228 ± 0.0003	0.8949 ± 0.0005	SSIM
	0.1301 ± 0.0010	0.1582 ± 0.0030	0.1657 ± 0.0014	Alex Loss

Kaggle dataset (33 images): $\times$2 upscaling				
Model	Method 1	Method 2	Method 3	Evaluation metric
ZSSR [28]	**41.4393** ± 5.3378	25.6254 ± 0.0056	26.0542 ±3.4005	PSNR
	0.9788 ± 0.0000	0.5838 ± 0.0006	0.6542 ± 0.0062	SSIM
	0.0243 ± 0.0000	0.1374 ± 0.0006	0.1125 ± 0.0001	Alex Loss
KernelGAN [29]	34.0756 ± 9.3306	27.7581 ± 4.1716	28.4265 ± 4.7513	PSNR
	0.9272 ± 0.0008	0.6886 ± 0.0049	0.7611 ± 0.0047	SSIM
	0.0497 ± 0.0004	0.0911 ± 0.0005	0.0824 ± 0.0002	Alex Loss
MWCNN [12]	31.4682 ± 18.0504	29.9679 ± 11.1570	30.7567 ± 13.5010	PSNR
	0.9109 ± 0.0008	**0.8254** ± 0.0011	**0.8810** ± 0.0006	SSIM
	0.0558 ± 0.0006	0.0796 ± 0.0008	0.0543 ± 0.0003	Alex Loss
CycleGAN-US-SISR [42]	40.9962 ±4.8578	25.4931 ± 3.0656	25.8789 ± 3.5966	PSNR
	0.9773 ± 0.0000	0.5763 ± 0.0061	0.6439 ± 0.0068	SSIM
	0.0208 ± 0.0004	0.1398 ± 0.0005	0.1062 ± 0.0001	Alex Loss
Model1	38.5234 ± 3.9060	**31.3743** ± 1.5939	**32.9226** ± 2.8843	PSNR
	0.9563 ± 0.0002	0.7911 ± 0.0011	0.8727 ± 0.0008	SSIM
	0.0460 ± 0.0003	**0.0687** ± 0.0007	**0.0536** ± 0.0003	Alex Loss
Model2	33.3628 ± 7.8995	30.0942 ± 1.8006	30.9415 ± 1.7527	PSNR
	0.8784 ± 0.0058	0.7495 ± 0.0034	0.8115 ± 0.0032	SSIM
	0.0821 ± 0.0015	0.1045 ± 0.0036	0.0778 ± 0.0015	Alex Loss

CCA-US dataset (7 images): $\times$4 upscaling				
Model	Method 1	Method 2	Method 3	Evaluation metric
ZSSR [28]	35.6793 ± 2.4680	31.2926 ± 0.8435	**32.2217** ± 1.7551	PSNR
	0.9467 ± 0.0001	**0.8506** ± 0.0004	**0.8979** ± 0.0003	SSIM
	0.0808 ± 0.0002	0.1869 ± 0.0006	**0.1621** ± 0.0003	Alex Loss
KernelGAN [29]	26.1786 ± 2.3820	26.0638± 2.2.3605	29.1976 ±3.4894	PSNR
	0.8321 ± 0.0010	0.8199 ± 0.0009	0.8694 ± 0.0006	SSIM
	0.2074 ± 0.0009	0.2379 ± 0.0004	0.1678 ± 0.0005	Alex Loss
MWCNN [12]	28.2801 ± 7.4760	26.9909 ± 1.9565	26.5557 ± 5.8312	PSNR
	0.7838 ± 0.0090	0.5567 ± 0.0042	0.7227 ± 0.0058	SSIM
	0.1896 ± 0.0059	0.2372 ± 0.0020	0.2249 ± 0.0057	Alex Loss
CycleGAN-US-SISR [42]	**36.5823** ± 2.8956	31.3216± 0.9152	31.9597 ± 2.1805	PSNR
	0.9503 ± 0.0001	0.8468 ± 0.0004	0.8943 ± 0.0004	SSIM
	0.0975 ±0.0002	0.1913 ± 0.0010	0.1685 ± 0.0004	Alex Loss
Model1	33.8589 ± 2.3945	**32.0137** ± 1.4641	31.7447 ± 2.0199	PSNR
	0.8827 ± 0.0001	0.8425 ± 0.0003	0.8593 ± 0.0003	SSIM
	0.2556 ± 0.0011	0.2769 ± 0.0016	0.2736 ± 0.0011	Alex Loss
Model2	33.4878 ± 1.8064	28.5800 ± 2.4044	31.1756 ± 0.6893	PSNR
	0.8795 ± 0.0008	0.6984 ± 0.0017	0.8353 ± 0.0006	SSIM
	0.2313 ± 0.0014	**0.1226** ± 0.0027	0.2755 ± 0.0009	Alex Loss

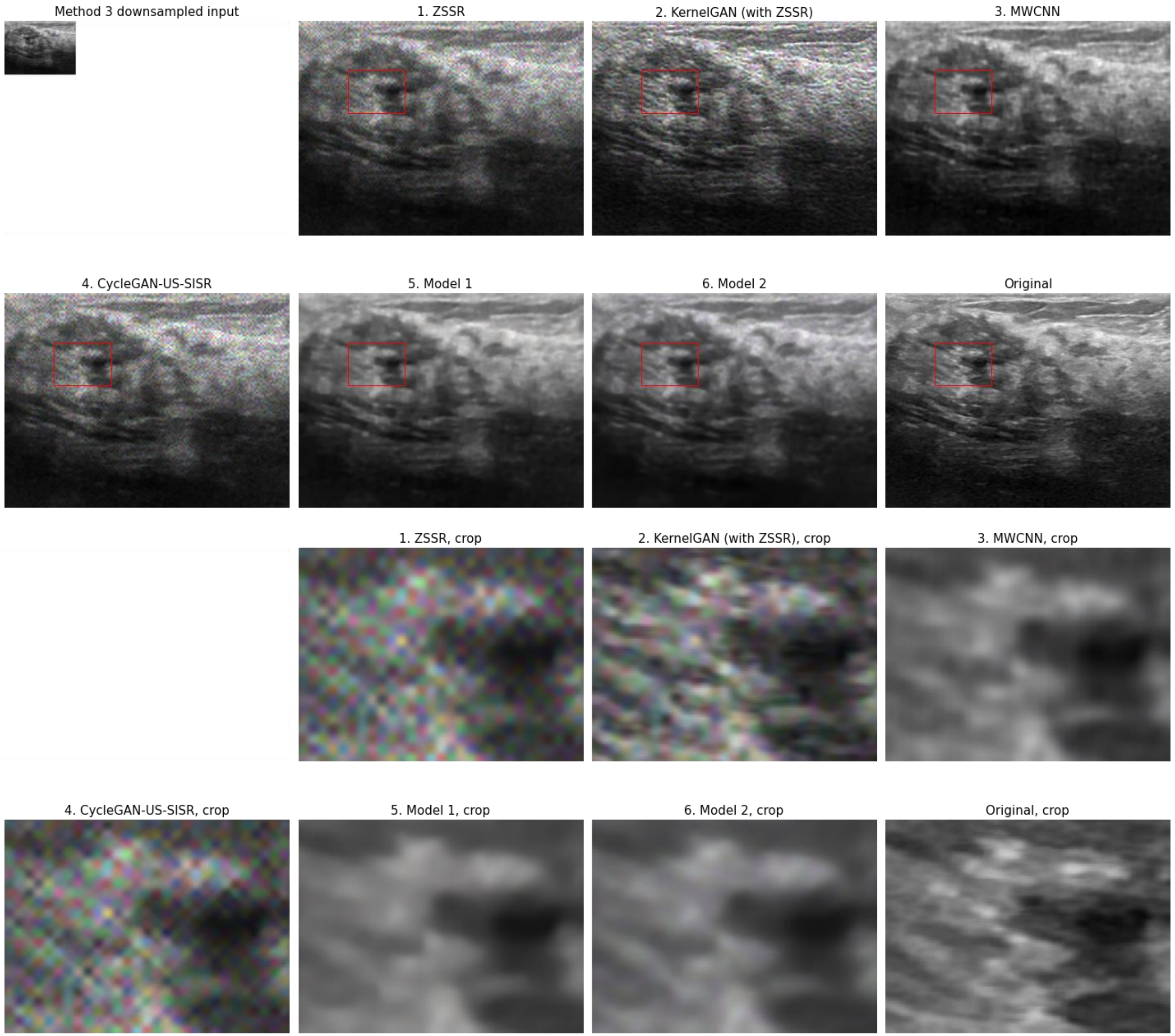
Kaggle dataset (33 images): $\times$4 upscaling				
Model	Method 1	Method 2	Method 3	Evaluation metric
ZSSR [28]	**33.2300** ± 5.2050	24.8055 ± 2.5436	25.1098 ± 3.0107	PSNR
	0.8809 ± 0.0007	0.5298 ± 0.0049	0.5817 ± 0.0054	SSIM
	0.0677 ± 0.0010	0.2001 ± 0.0005	0.1958 ± 0.0002	Alex Loss
KernelGAN [29]	25.9043 ± 10.3544	24.1170± 2.5436	25.5240 ±4.6654	PSNR
	0.7110 ± 0.0080	0.5528 ± 0.0067	0.6358 ± 0.0048	SSIM
	0.0912 ± 0.0008	0.1431 ± 0.0008	0.1310 ± 0.0008	Alex Loss
MWCNN [12]	29.2626 ± 14.4879	27.1243 ± 8.3364	26.5236 ± 10.6172	PSNR
	0.8482 ± 0.0011	**0.7089**± 0.0018	0.7277 ± 0.0058	SSIM
	0.0791 ± 0.0009	**0.1204** ± 0.0013	0.1216 ± 0.0012	Alex Loss
CycleGAN-US-SISR [42]	32.9072 ± 4.7455	24.7451± 2.5508	25.0184 ± 3.0784	PSNR
	0.8744 ± 0.0007	0.5257± 0.0050	0.5759 ± 0.0058	SSIM
	0.0713 ±0.0011	0.1997 ± 0.0005	0.1904 ± 0.0003	Alex Loss
Model1	31.4986 ± 3.9234	**28.7633** ± 2.1375	**29.0328** ± 3.3374	PSNR
	0.8171 ± 0.0013	0.7046 ± 0.0017	**0.7439** ± 0.0020	SSIM
	0.1082 ± 0.0010	0.1206 ± 0.0020	**0.1109** ± 0.0014	Alex Loss
Model2	30.8769 ± 3.3308	28.5800 ± 2.4044	28.6963 ± 3.1207	PSNR
	0.7863 ± 0.0015	0.6984 ± 0.0017	0.7201 ± 0.0019	SSIM
	0.1265 ± 0.0025	0.1226 ± 0.0027	0.1417 ± 0.0037	Alex Loss

Experimental Results: Picture comparison



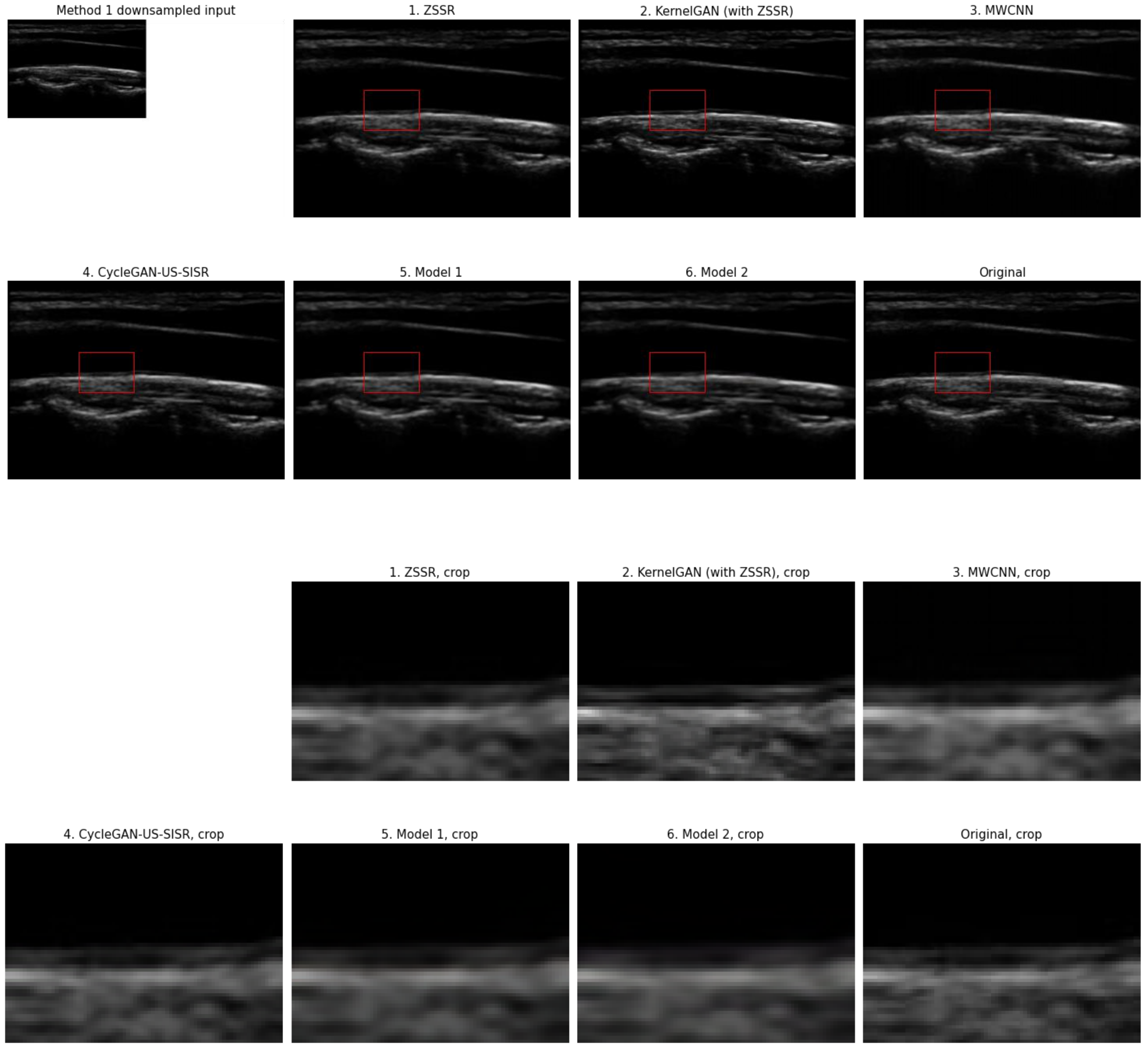
$\times 4$ upscaling on an image from CCA – US dataset taking input from method 2

Experimental Results: Picture comparison

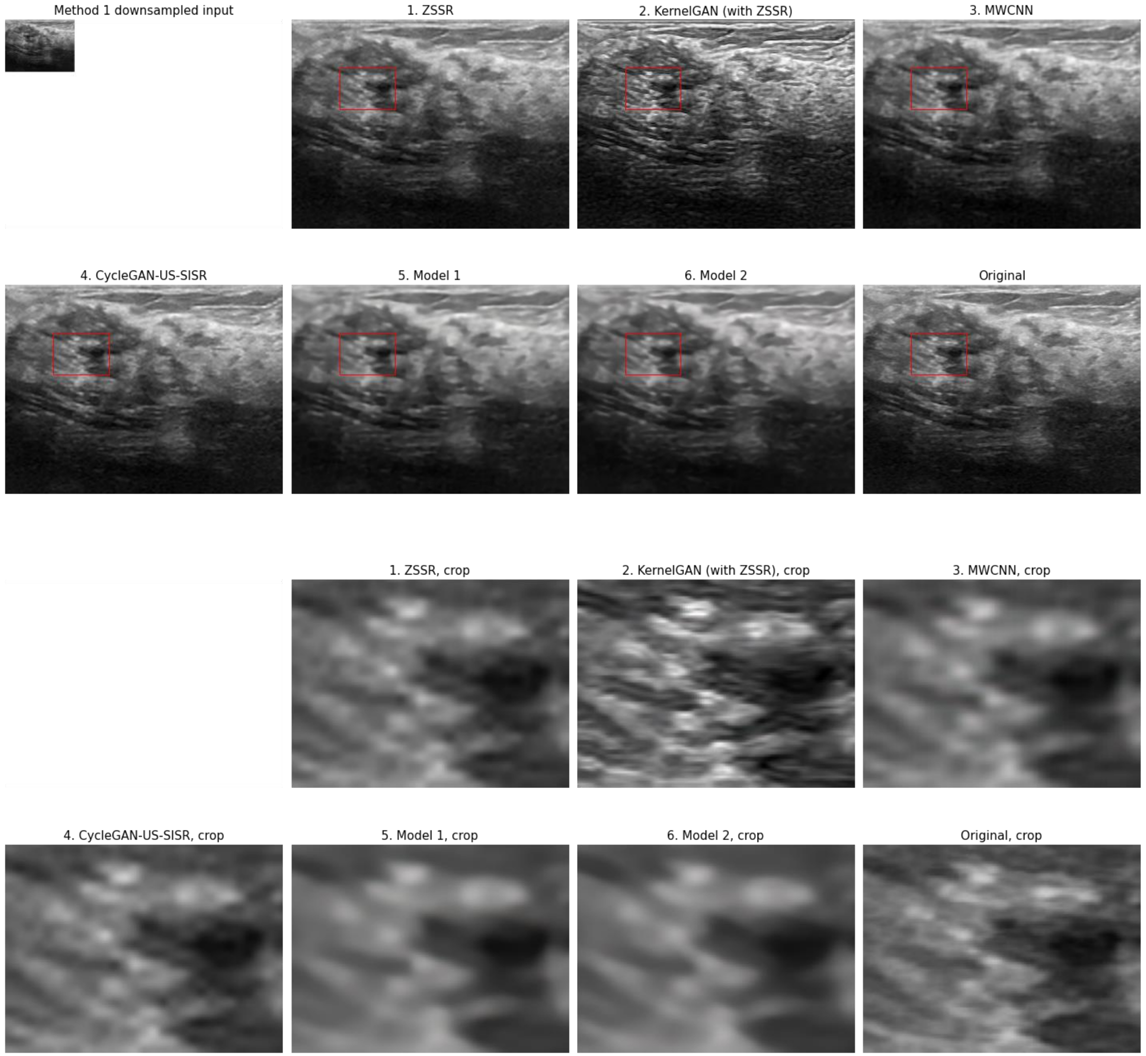


$\times 4$ upscaling on an image from kaggle dataset taking input from Method 3

Experimental Results: Picture comparison



$\times 2$ upscaling on an image from CCA − US dataset taking Method 1 input

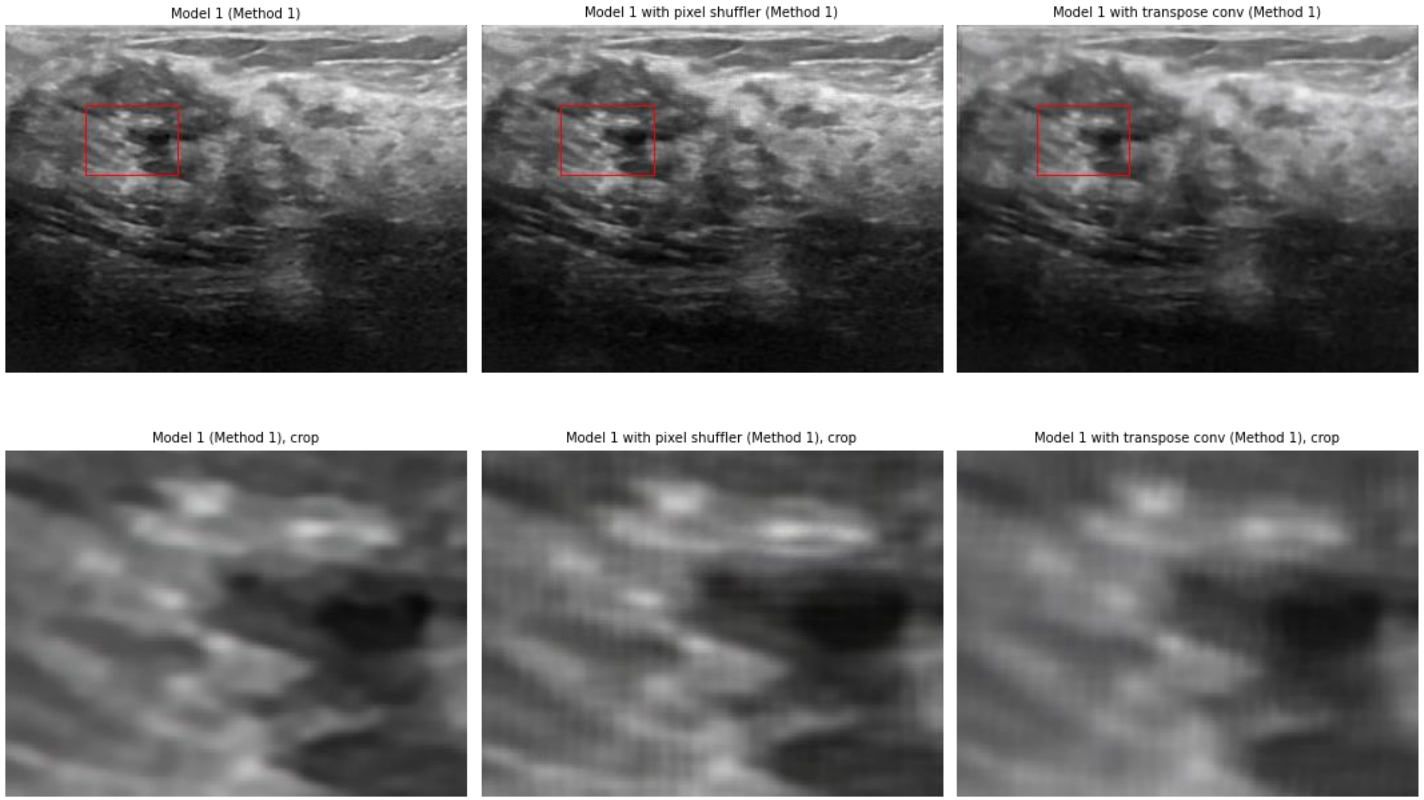


$\times 4$ upscaling on an image from kaggle dataset taking Method 1 input

Conclusion

- Proposed **self-supervised models for SISR of US images** without forming any assumption about the source of input images.
- Based on evaluation metrics, the models give a good performance in comparison to benchmarking models in more than half of the test cases.
- Gives lower but comparable performance in ideal downsampling cases.
- The proposed models do observe **loss of data** in SR results, but it is not only a characteristics if the model itself but SR task as well.
- Based on experiments here, a **simple model** with simple losses may be well suited for US SISR.

Future work



- Repeated bicubic upsampling and downsampling causes loss of data. There could be merit in exploring other **upscaling techniques**. However, transposed convolution and sub-pixel convolutional may not necessarily be a better option as they introduce checkerboard effects.
- Using **image pre-processing** such as wavelets denoising to for HR parent could be of merit.
- Exploring other **blind SR techniques**.

References

[1] C. Dong, C. C. Loy, K. He and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," *in IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295-307, 1 Feb. 2016.

[2] H. Rampal, A. Das, S. Mahajan and S. Singh, "Enhancement of Corn image quality using Very-Deep Super Resolution (VDSR) neural network," *2018 3rd International Conference on Inventive Computation Technologies (ICICT),* 2018, pp. 1-4.

[3] P. Liu, H. Zhang, W. Lian and W. Zuo, "Multi-Level Wavelet Convolutional Neural Networks," in *IEEE Access*, vol. 7, pp. 74973-74985.

[4] M. A. Talab, S. Awang and S. A. M. Najim, "Super-Low Resolution Face Recognition using Integrated Efficient Sub-Pixel Convolutional Neural Network (ESPCN) and Convolutional Neural Network (CNN)," *2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS),* 2019, pp. 331-335.

[5] C. Ledig et al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 105-114.

[6] X. Wang et al. , "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks," in *Computer Vision – ECCV 2018 Workshops. Cham: Springer International Publishing, 2019, pp. 63-79.*

[7] A. Shocher, N. Cohen and M. Irani, "Zero-Shot Super-Resolution Using Deep Internal Learning," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3118-3126.

References

[8] S. B. Kligler, A. Shocher and Michal Irani, "Blind Super-Resolution Kernel Estimation using an Internal-GAN," *NeurIPS*, 2019, pp. 284-293.

[9] D. Mishra, S. Chaudhary, M. Sarkar, and A. S. Soin, "Ultrasound image enhancement using structure oriented adversarial network," *IEEE Signal Processing Letters, vol. 25, no. 9, pp, 1349-1352, 2018.*

[10] H. Temiz and H. S. Bilge, "Super resolution of B-mode ultrasound images with deep learning," *IEEE Access, vol. 8, pp. 78808-78820, 2020.*

[11] R. J. G. van Sloun, O. Solomon, M. Bruce, Z. Z. Khaing, Y. C. Eldar and M. Mischi, "Deep Learning for Super-resolution Vascular Ultrasound Imaging," *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 1055-1059.

[12] I. J. Goodfellow, J. P. Abadie, M. Mirza, B. Xu, D. W.Farley, S. Ozair, A. Courville, and Y. Bengio. 2014, "Generative adversarial nets," in *Proceedings of the 27th International Conference on Neural Information Processing Systems,* Vol. 2 (NIPS'14). MIT Press, Cambridge, MA, USA, 2014, pp. 2672–2680.

[13] J. Y. Zhu, T. park, P. Isola, and A.A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *2017 IEEE International Conference on Computer Vision (ICCV),* 2017, pp. 2242-2251.

[14] H. Liu, J. Liu, T. Tao, S. Hou, and J. Han, "Perception consistency ultrasound image super-resolution via self-supervised CycleGAN," *Neural Computing and Applications,* Jan. 2021.

Special thanks to:

- **Prof. Ioannis Psaromiligkos,** Dept. of Electrical and Comp. Eng., McGill University for his continuous support and guidance throughout the project.
- **Prof. Fabrice Labeau,** Dept. of Electrical and Comp. Eng., McGill University for his valuable time.