Anisotropic MRI Super Resolution using cGANs

Medical Image Computing (CS 736)

UNDER THE GUIDANCE OF

Prof. Suyash P. Awate

Department of CSE

IIT Bombay

SUBMITTED BY:

Mohit Kumar Meena

213070021

Department of Electrical Engineering

Outline

- Timeline of the project
- Introduction to the Problem
- Dataset and Data Preprocessing
- Current approaches
- SRGAN
- Results
- Conclusions and future scopes
- References

Timeline of the project

Project Idea

Looked for 3D SR dataset but due to unavailability of data and limited GPU resources, choosed to work on 2D BraTS registration dataset.

Data Processing

Scans are resized, cropped, downsampled and gaussian noise is added to them.

Model Deployment

SRGAN architecture is implemented and model is trained on google colab GPU.

Model Testing

Model tested on randomly selected testing images and SSIM and PSNR is calculated.

Model Comparison

Result of model is compared with bilinear and bicubic interpolation of same factor.

Introduction

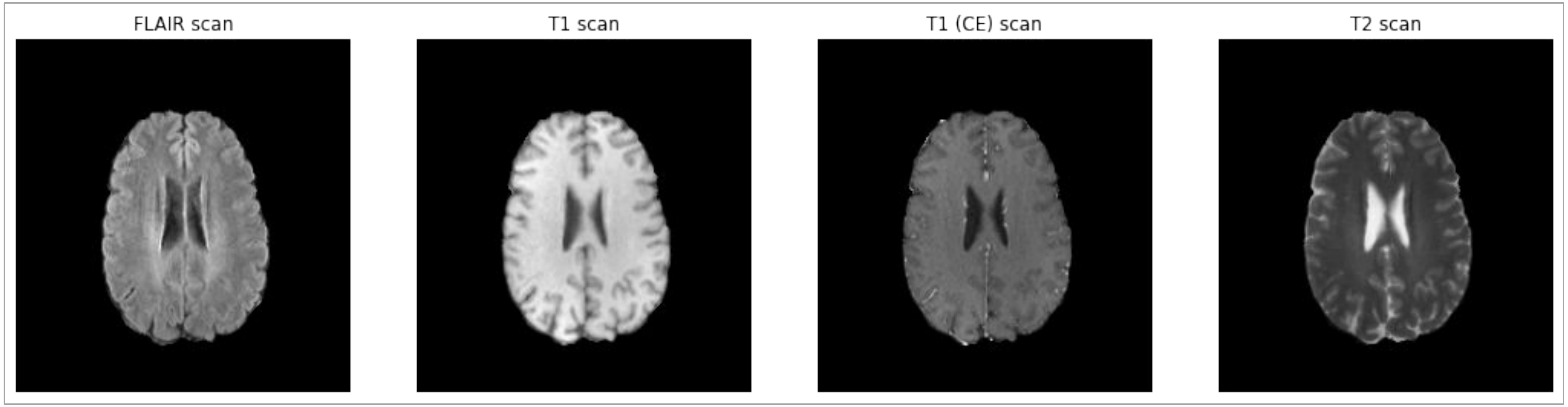
- Capturing high resolution MRI images is time taking process and sometimes not suitable for medical emergencies.
- High resolution images requires huge amount of memory to store.
- Low resolution imaging is faster but it compromises with fine details that might be helpful in diagnosis of a diseases.
- So we need a way that ensures good quality images in less amount of time.
- **Super resolution** aimed to reconstruct detailed High resolution (HR) images from low resolution images.
- Interpolation method fails to recover high frequency information like edges.
- Non learning method has an limitation that they require sound prior knowledge about the data representation.
- Learning based methods learns directly from the set of images provided. Hence no prior information required.

Introduction (Contd.)

- Several state of art methods with deep learning techniques have shown great performance on **natural images**.
- These deep learning techniques approaches includes:
    - SRCNN (Super resolution CNN)
    - Fast SRCNN
    - Efficient Sub-pixel CN (ESPCN)
    - DRCN (deeply recursive CN)
    - SRGAN (Super resolution GAN)
- Among all the above techniques, SRGAN outperformed in terms of PSNR and SSIM.
- In this project, architecture of original SRGAN paper is implemented.

Dataset and Data Preprocessing

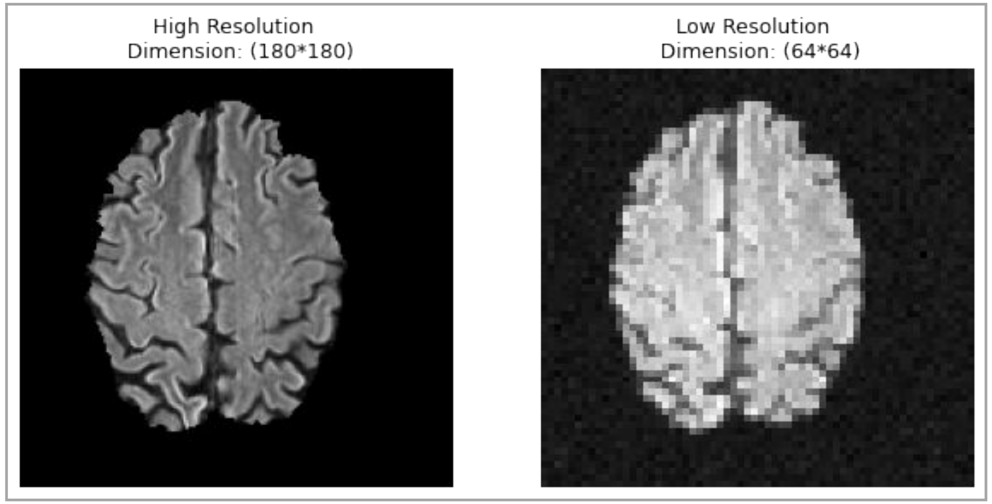
- Dataset used in this project is taken from ISBI 2022 challenge of BraTS registration (Not super resolution task).
- Dataset contains pre-operative and follow-up MRI scans of 160 patients (including testing).
- Dataset contains T1w, T2w, T1 (ce) and FLAIR MRI sequences. FLAIR is taken as a choice.

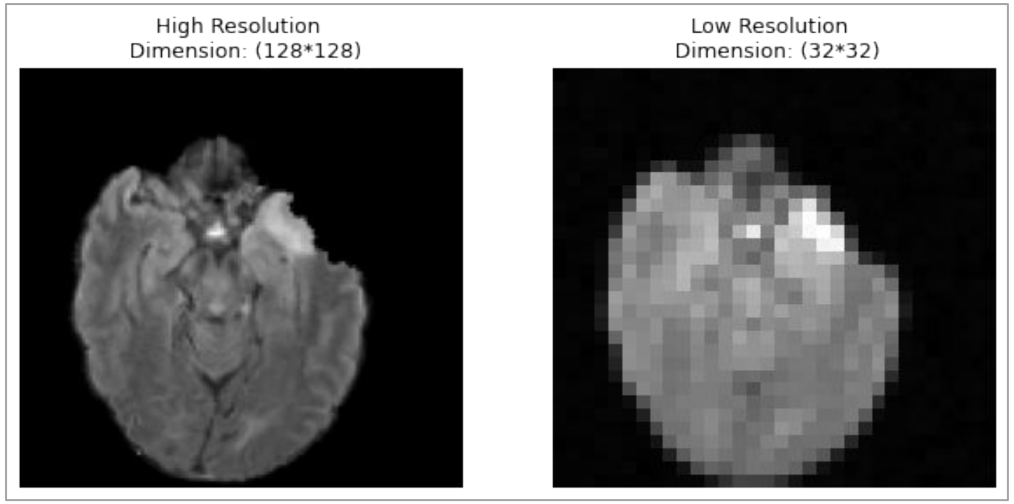


Dataset link: https://www.med.upenn.edu/cbica/brats-reg-challenge/

Dataset and Data Preprocessing (contd.)

- Scans are resized and cropped from (155,240,240) to (70,180,180) and gaussian noise is added.
- Slices of scans are downsampled from (180,180) to produce LR image of size (64,64).
- But since pretrained **VGG19** model is used so they are again downsampled to (128,128) for HR and (32,32) for LR.





Current Approaches

- **SRCNN** directly learns an end to end mapping b/w the low/high resolution images by jointly optimizing all the layers. This was the first deep learning method for SISR.

- **FSRCNN** is the faster version of SRCNN, it does it by introducing a deconv layer at the end of n/w and by shrinking the input feature dimension before mapping and expanding back afterwards.

- **ESPCN** introduces an efficient sub-pixel conv layer which learns an array of upsampling filters. This method is capable of real time SR of 1080p video .

- **DRCN** proposes that increasing recursion path can improve performance w/o introducing new parameters for additional convolutions.

SRGAN

- Generative adversarial network (GAN) for image super resolution.
- Able to recover photo-realistic textures from heavily downsampled images (4x in our case).
- They introduced a perceptual loss function consisting of an adversarial loss and a content loss.

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

- Perceptual loss is the weighted sum of content loss and an adversarial loss.
- Perceptual loss model distinct desirable characteristics of the recovered SR image.

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.

SRGAN (contd.)

Content loss:

Most commonly used **MSE** loss is given by

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

This MSE loss often lacks high frequency content. So the content loss is given by

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

$\phi_{i,j}$ is the feature map obtained by jth conv before the ith mapping layer within VGG19 n/w.

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.

SRGAN (contd.)

Adversarial loss:

Adversarial loss function is given by

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

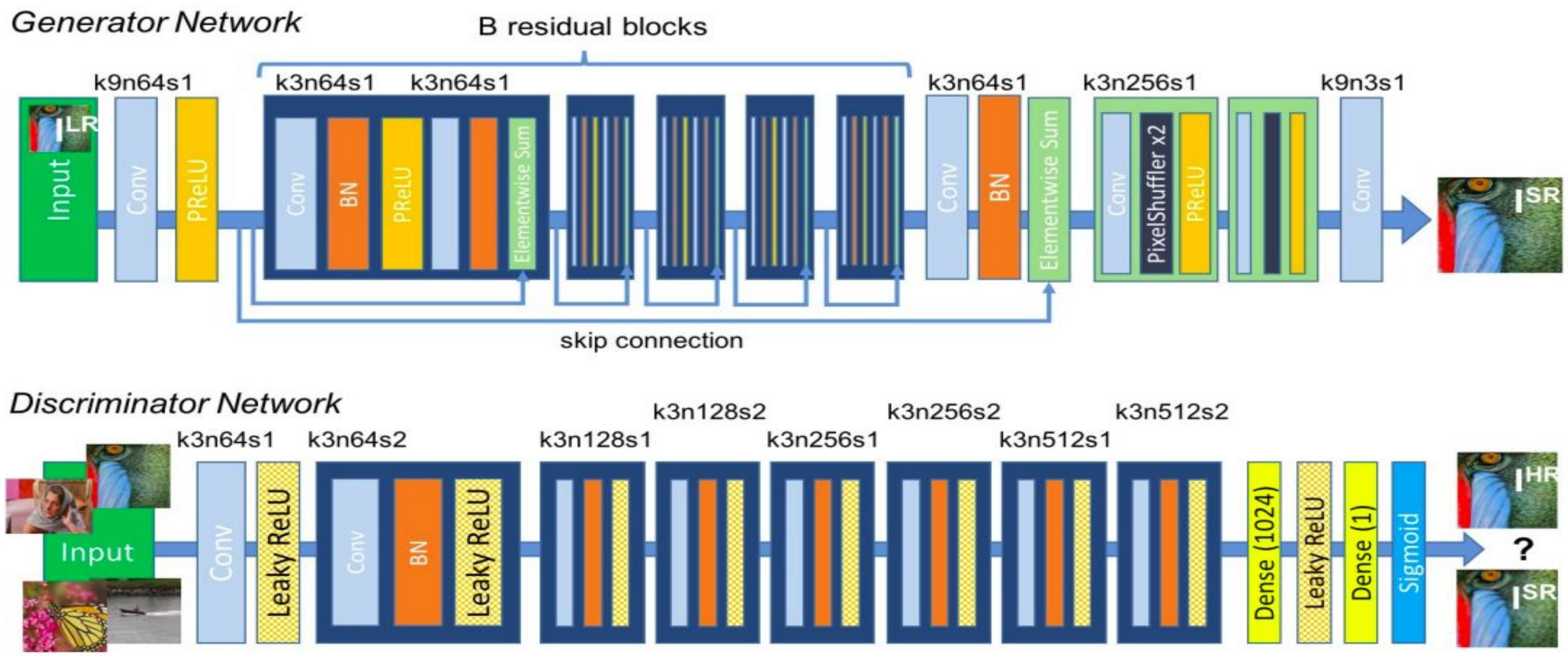
This loss function pushes our solution to the HR image manifold using a discriminator n/w.

Optimization function:

$$\hat{\theta}_G = \arg\min_{\theta_G} \frac{1}{N} \sum_{n=1}^{N} l^{SR}(G_{\theta_G}(I_n^{LR}), I_n^{HR})$$

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.

SRGAN Architecture



Here,

K=kernel size

n=number of feature map

s=stride

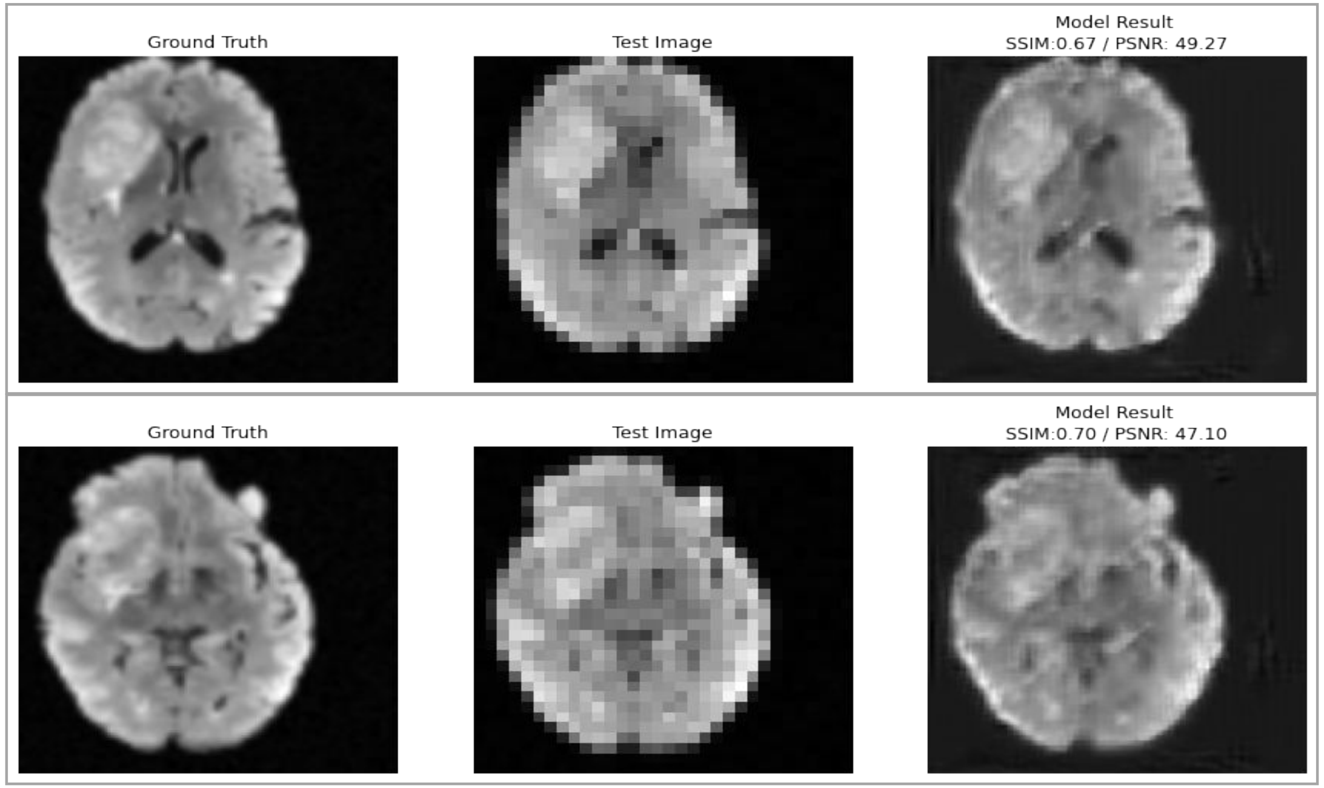
PReLU can learn the slope parameter using backprop w/o increasing the cost of training.

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.

Results

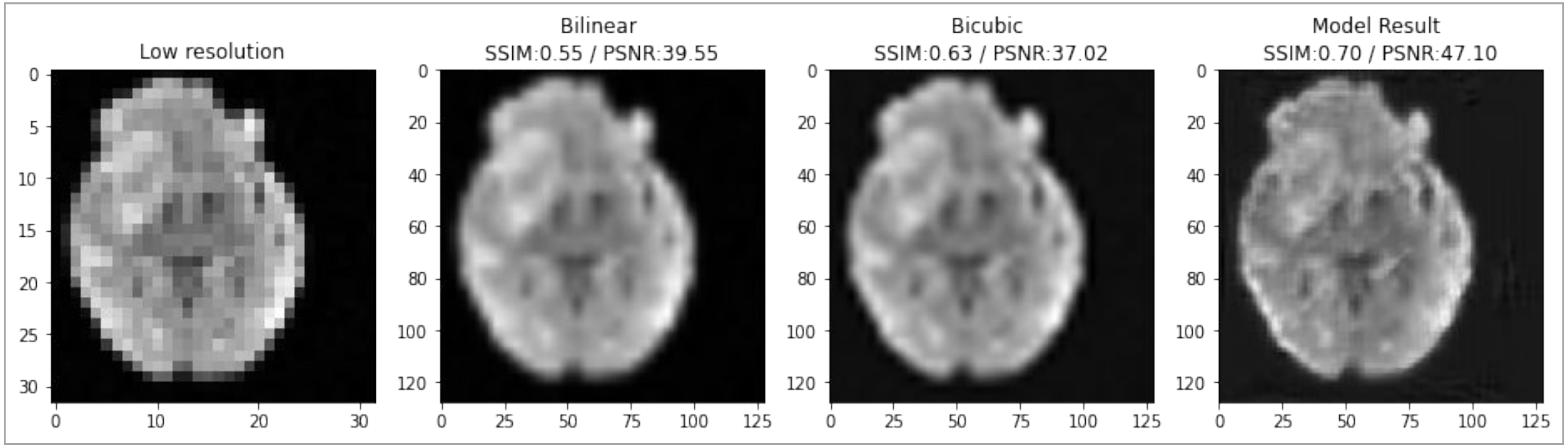
- Results of the model outperforms when compared with bicubic and bilinear interpolation with maximum **SSIM: 0.70** and **PSNR: 47.10**

Model is trained for 20 Epochs which took 462 minutes (~8 hours) on google colab GPU



Results (contd.)

- Bilinear interpolation yields **SSIM: 0.55** and **PSNR: 39.55**
- Bicubic interpolation yields **SSIM: 0.63** and **PSNR: 37.02**



Conclusions and Future scopes

- SRGAN model implemented in this project outperformed the bilinear and bicubic interpolation.
- More training on higher dimensional data needs to done to produce better results for practical usage.
- Model needs to be trained on 3D dataset to exploit the inter-frame dependencies.
- Larger dataset is required to produce promising and authentic results.
- Future work includes training and testing the model on 3D dataset and comparing the results with other recent state -of- art methods.

References

- 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.
- Almahfouz Nasser, Sahar, et al. "Perceptual cGAN for MRI Super-resolution." *arXiv e-prints* (2022): arXiv-2201.
- Ward, Chris M., et al. "Image quality assessment for determining efficacy and limitations of Super-Resolution Convolutional Neural Network (SRCNN)." *Applications of Digital Image Processing XL*. Vol. 10396. International Society for Optics and Photonics, 2017.
- Yang, Wenming, et al. "Deep learning for single image super-resolution: A brief review." *IEEE Transactions on Multimedia* 21.12 (2019): 3106-3121.

Thankyou!