IMAGE SUPER-RESOLUTION USING GENERATIVE ADVERSARIAL NETWORK

Adithya Chowdary Boppana aboppana@iu.edu

Satyaraja Dasara sdasara@iu.edu

Siva Charan Mangavalli simang@iu.edu

Abstract

- Image super-resolution (SR) can be defined as the process of recovering a high resolution(HR) image from a low resolution(LR) image
- Traditionally methods include upscaling and interpolation techniques, however, recent advances in the field of deep learning have led to novel approaches
- Traditional techniques lack finer texture details due to excessive smoothing of the image
- GANS combined with appropriate loss functions restore more details and output naturally realistic looking images
- Since many HR images can be mapped to one LR image, this problem is a tricky one.
- We reimplemented the paper to replicate the results

Previous Work

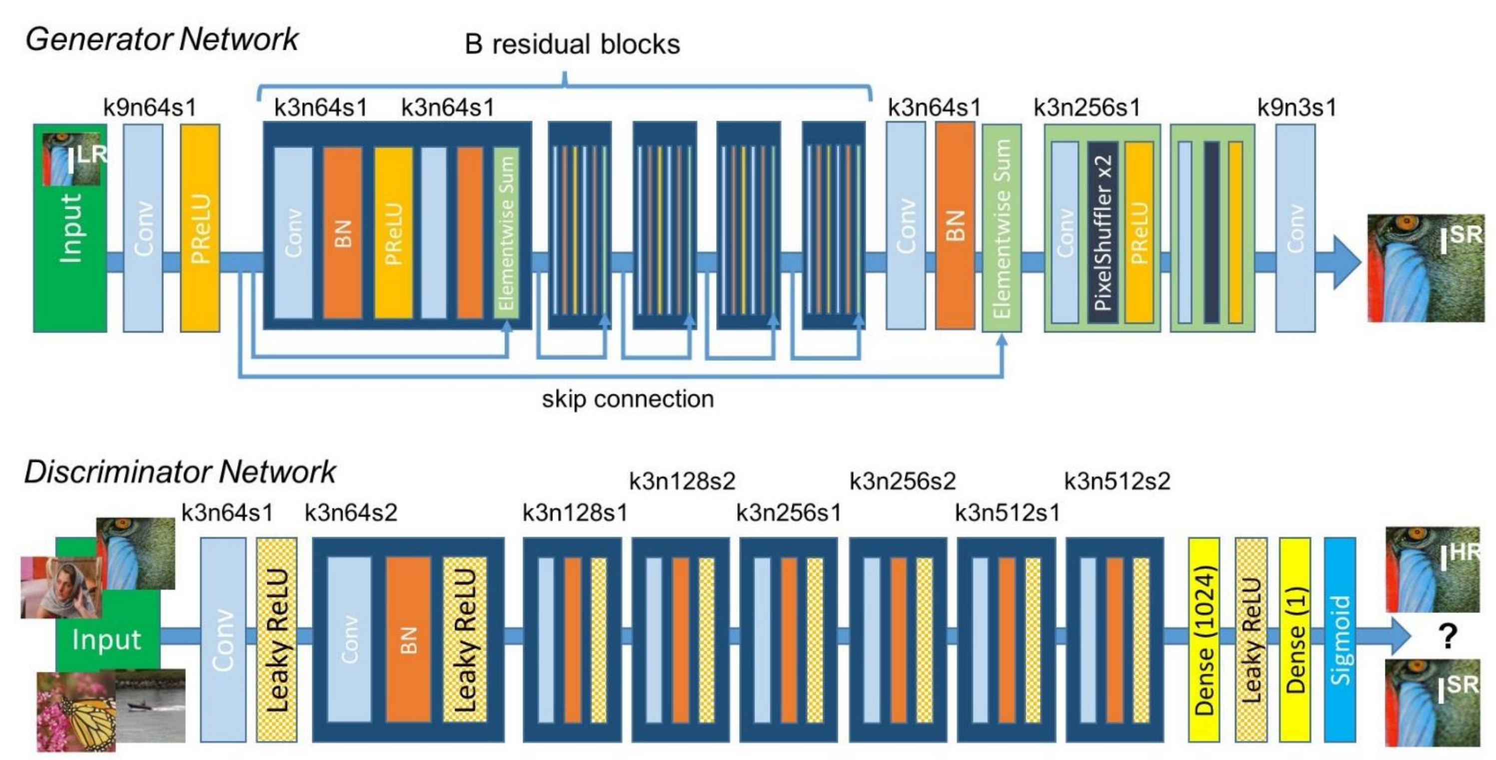
- The most well-known implementation of GAN for image super-resolution is by Christian Ledig[1]. A custom perceptual loss function was used, which was a weighted sum of content loss and adversarial loss to train the generator network
- The loss functions for super-resolution applications have been explored by Justin Johnson[3]. The two loss functions explained in detail are Feature Reconstruction loss and Style Reconstruction loss
- The SR-CNN method presented in Jiwon Kim[2] used a CNN for end to end learning for the image reconstruction at multiple scales

Data & Augmentation

- We have used Div2K dataset which has 800, 100, 100 low resolution and corresponding high resolution images for train, validation and test sets
- Also we have augmented the data by randomly
    - Cropping
    - Flipping
    - Adding Gaussian noise
    - Changing brightness
- The images have been normalized to [-1, 1]
- A total of 4400 train , 1000 validation and 1000 test images have been created

Architecture

We have used the same architecture as proposed by Leidig[1]



Network used in Leidig[1]

Loss Function

Perpetual loss function has been used which is a combination of content loss and adversarial loss[1]

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

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

Loss used in Leidig[1]

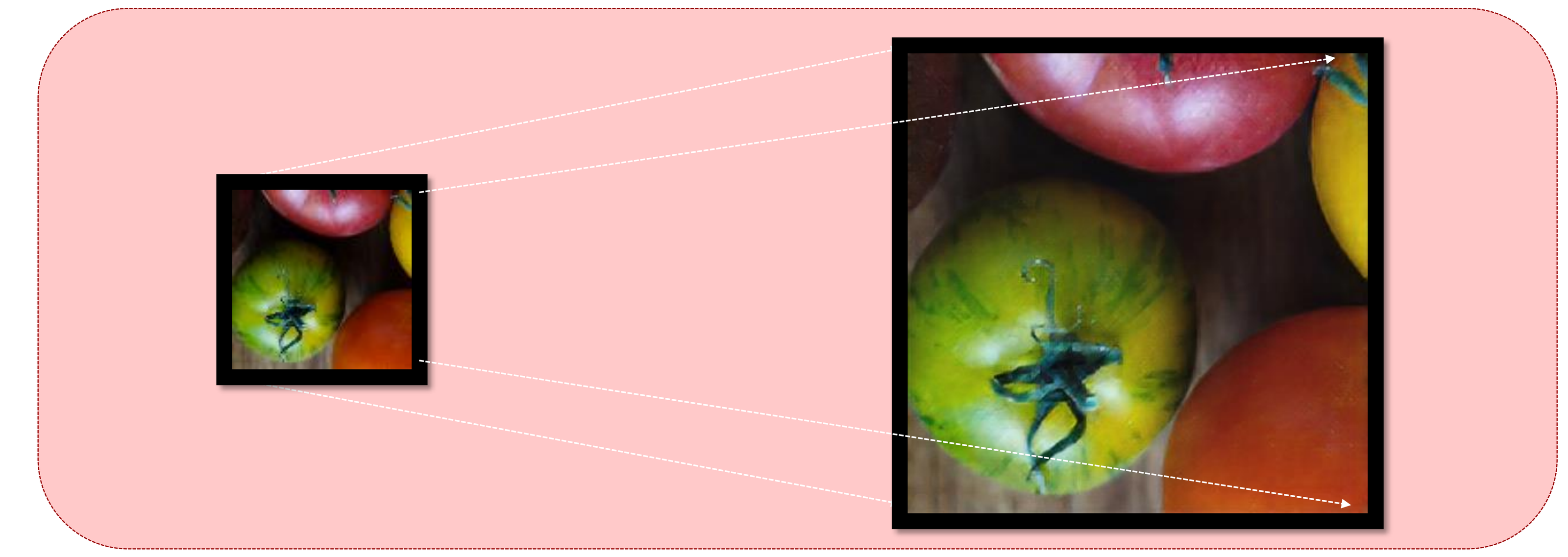
Implementation details

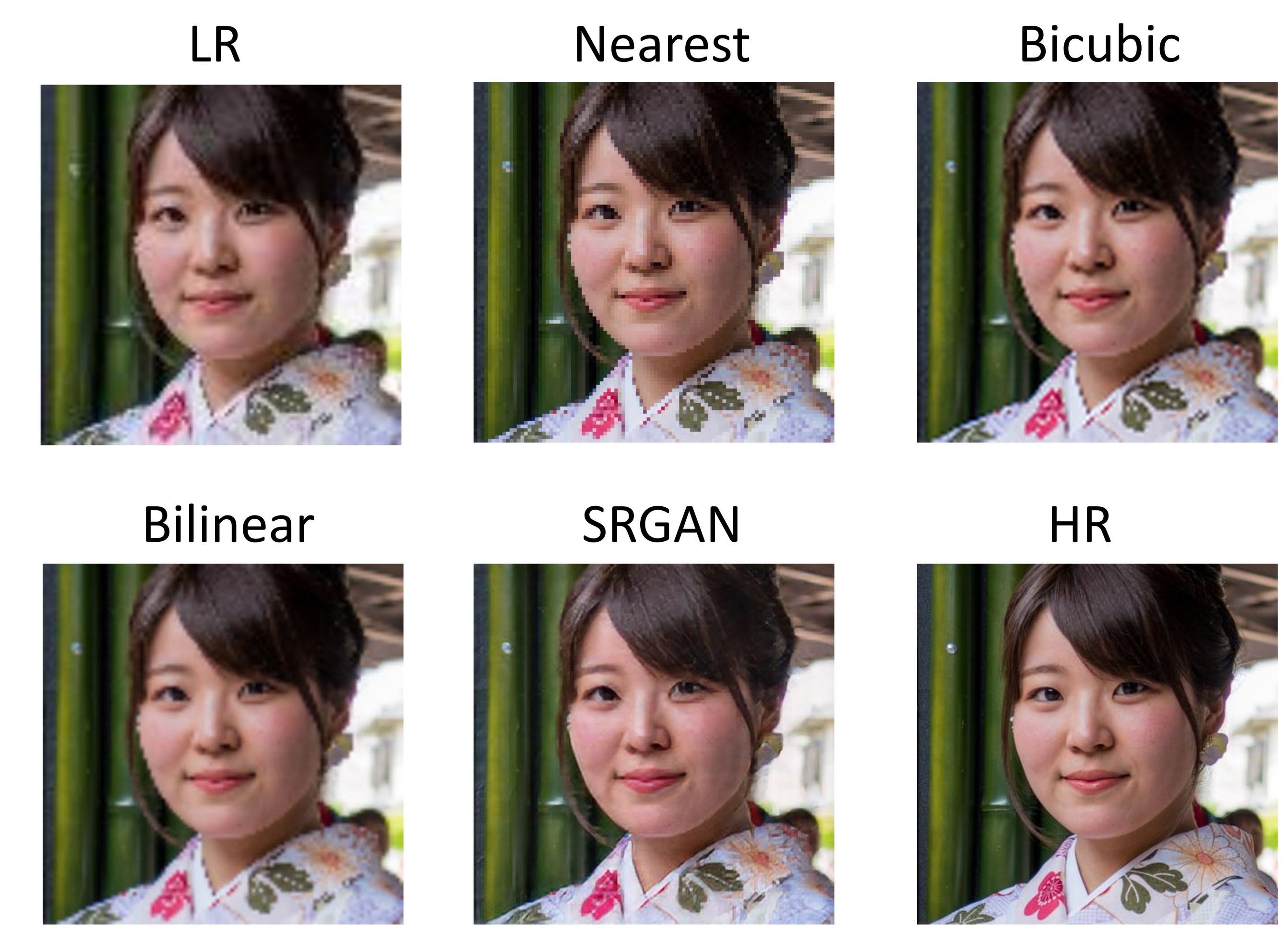
- The model is built and trained using Keras framework
- Used Google Cloud Instance with NVIDIA TESLA P100 GPU and 32gb RAM configuration
- A batch size of 8 images is used to prevent GPU memory allocation error
- Trained 55 epochs for 19 hours

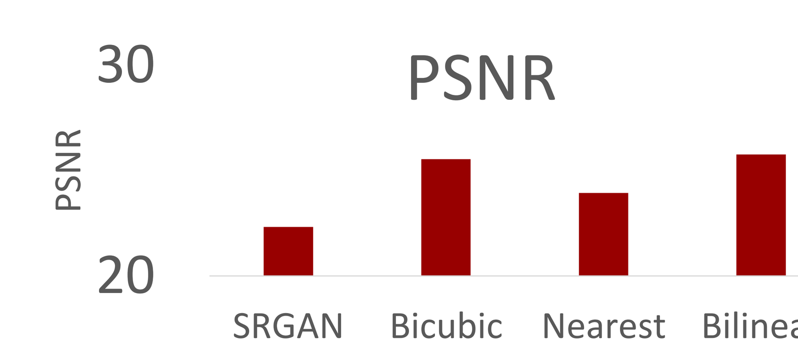
Conclusion and Future work

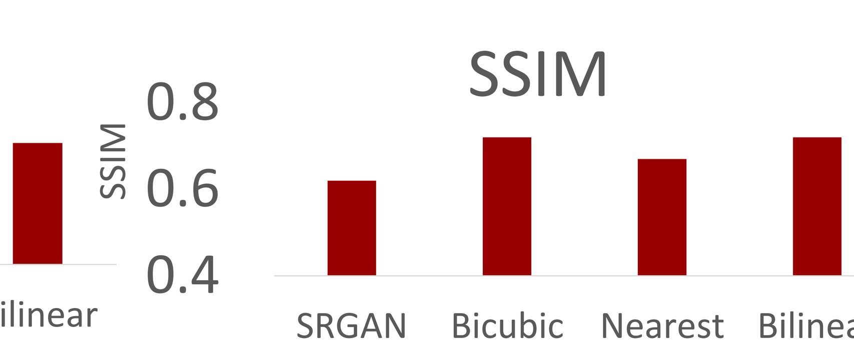
- SRGAN offers superior perceptual performance compared to traditional interpolation techniques
- PSNR and SSIM are not good indicators of SR performance
- Network optimization for speed and memory for future work

Results









References

- C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham,´ A. Acosta, A. P. Aitken, A. Tejani, J. Totz, Z. Wang et al., “Photorealistic single image super-resolution using a generative adversarial network,” in CVPR, 2017.
- J. Kim, J. Kwon Lee, and K. Mu Lee, “Accurate image super resolution using very deep convolutional networks,” in CVPR, CVPR, 2016.
- Justin Johnson, Alexandre Alahi, Li Fei-Fei, “Perceptual Losses Losses for Real-Time Style Transfer and Super-Resolution”