SUPER-RESOLUTION USING GAN

COURSE : CSE 241

NEERUDU NIKHIL
AND
KUTHADI JASWANTH SAI PREM

Generative Adversarial Network

Generative Adversarial Network (GAN) is a type of deep learning model that is used for generating new data that is similar to a training dataset. To understand the term GAN, let’s break it into three separate parts -

- Generative - To learn a generative model, which describes how data is generated in terms of a probabilistic model. In simple words, it explains how fake data is generated.
- Adversarial - The training of the model is done in an adversarial setting.
- Networks - Uses deep neural networks for training purposes.

GANs have gained popularity in recent years due to their ability to generate realistic images, videos, and audio, which has important implications in fields such as art, entertainment, and computer vision.

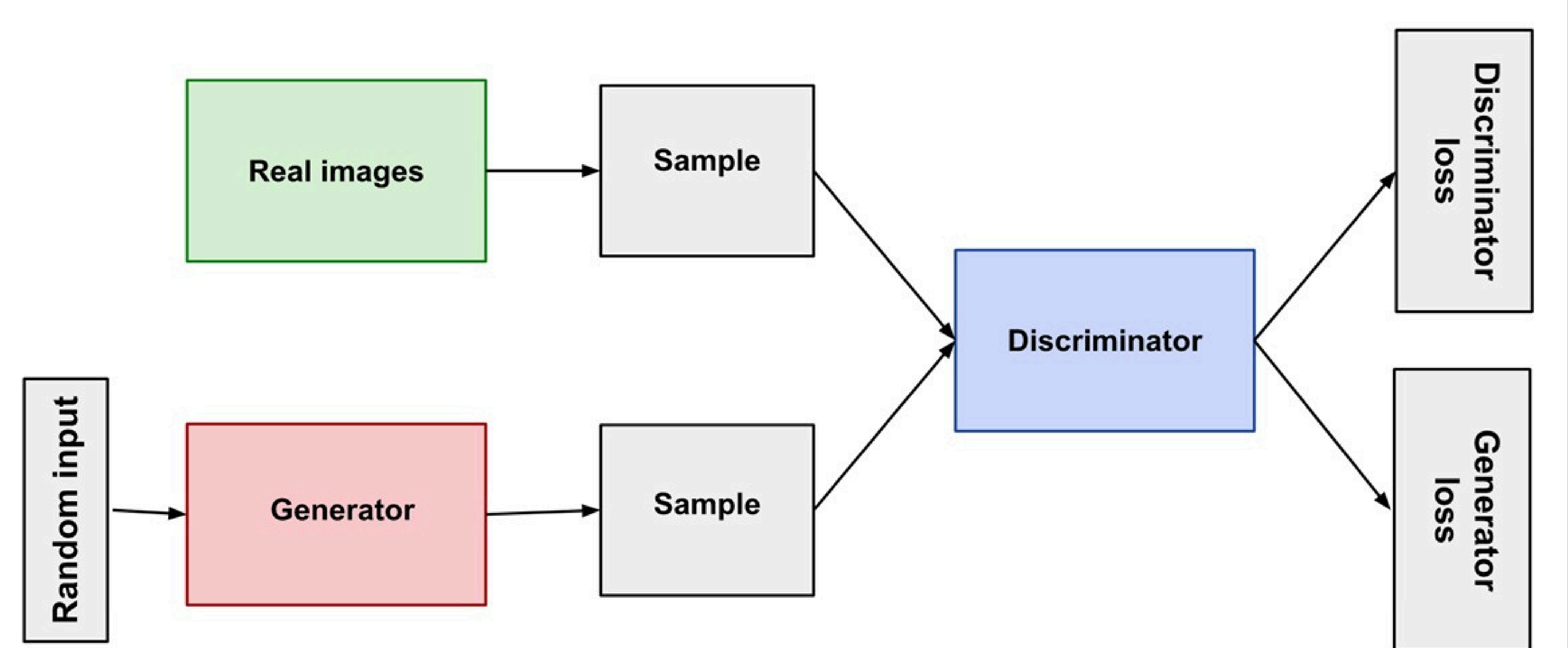
GENERATIVE ADVERSARIAL NETWORK

The GAN model architecture includes two sub-models :

- Generator - Generates data that is fake data based on original(real) data.
- Discriminator - Predicts if the data is fake or real.

The generator generates a batch of samples, and these, along with real examples from the domain are provided to the discriminator, which classifies them as real or fake.

- The discriminator is then updated to get better at discriminating between real and fake samples in the next round
- The generator is updated based on how well or not, the generated samples fooled the discriminator.



SUPER RESOLUTION GAN (SRGAN)

Super-resolution is a technique used to enhance the resolution or quality of an image beyond its original resolution. It involves increasing the pixel density or resolution of an image while maintaining its visual quality as much as possible.
Out of the many applications of GAN, generating super-resolution images from low-resolution images is one of the most important applications.

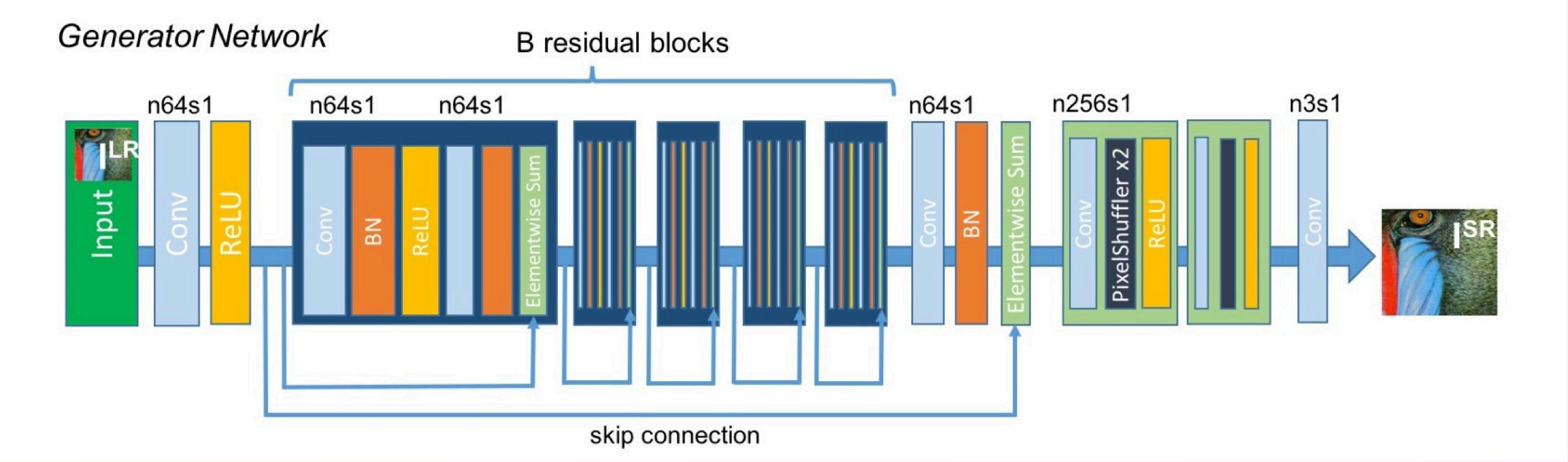
The GAN models used for super-resolution(SRGAN) also include a perceptual loss function that measures the similarity between the generated image and the original high-resolution image based on features extracted from a pre-trained deep neural network. This loss function encourages the generator to produce images that are not only visually similar to the original image but also have similar high-level features.

The perceptual loss function is written as a combination of 2 loss functions-

- Content loss -that measures the similarity between generated image and the actual image
- Adversarial loss- that ables the generator to produce images that minimize the difference between distributions of generated images and real images.

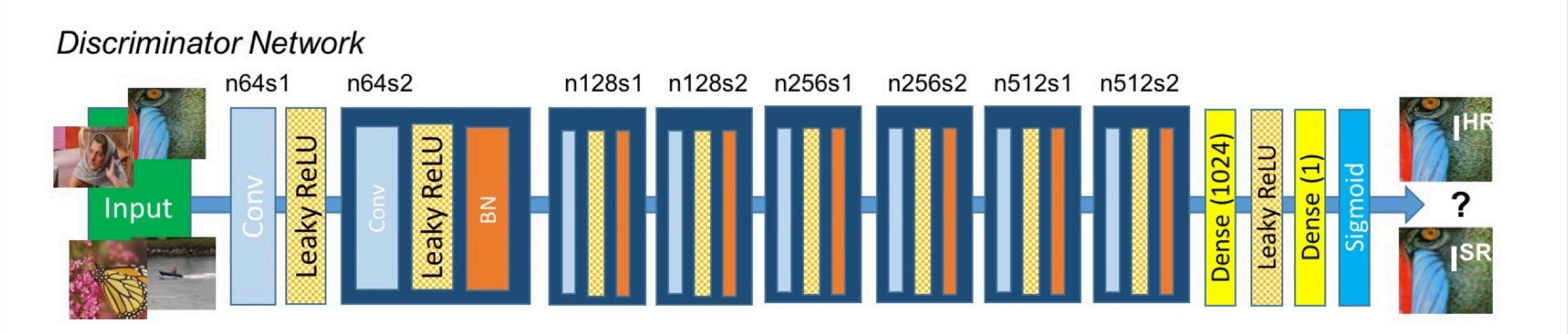
AS AN OPTIMISATION OF SR GAN, MEAN SQUARED ERROR AND PEAK SIGNAL TO NOISE RATIO(PSNR) ARE MONITORED AS THEY CAPTURE IMAGE DIFFERENCES PIXEL-WISE.

Model Implementation

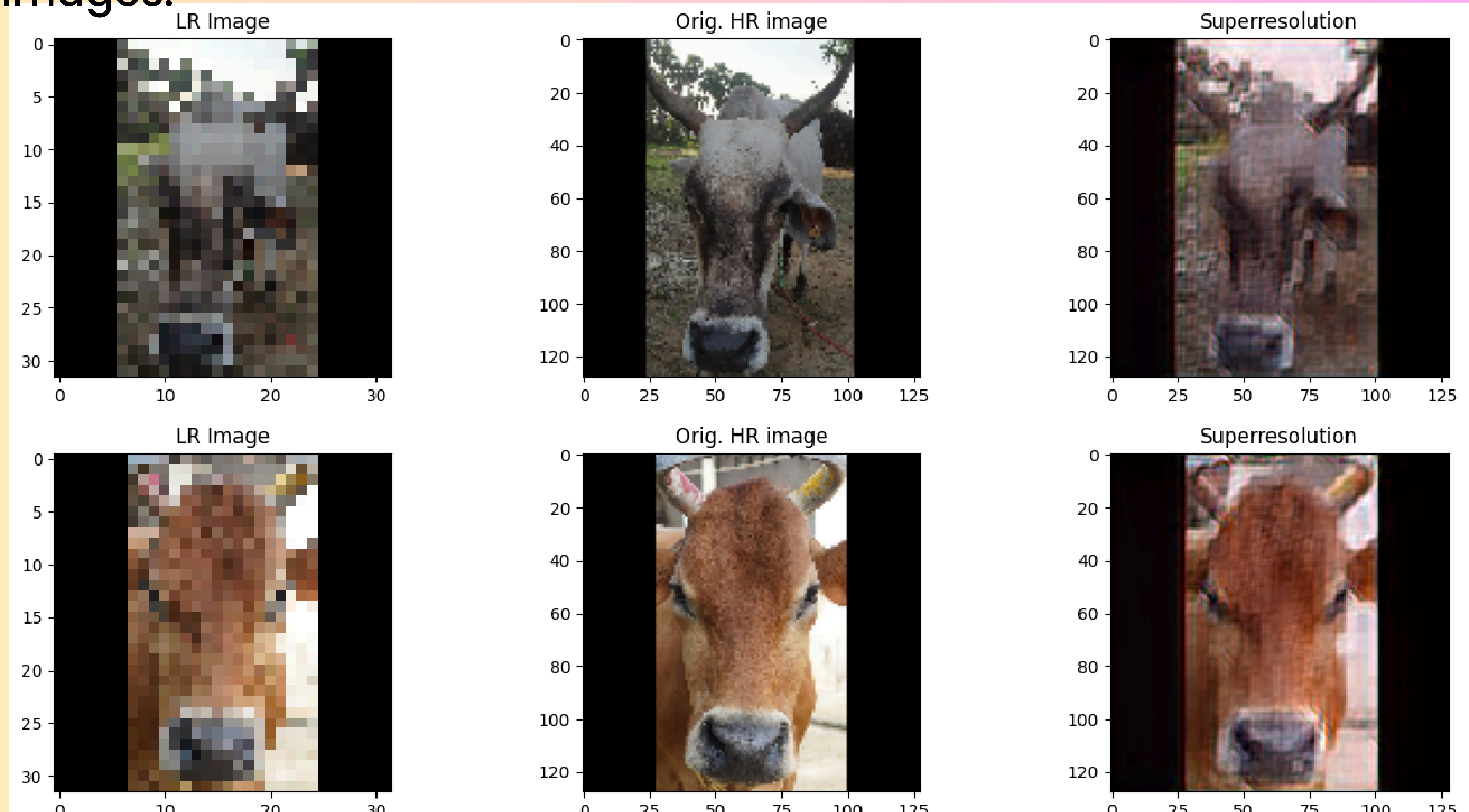


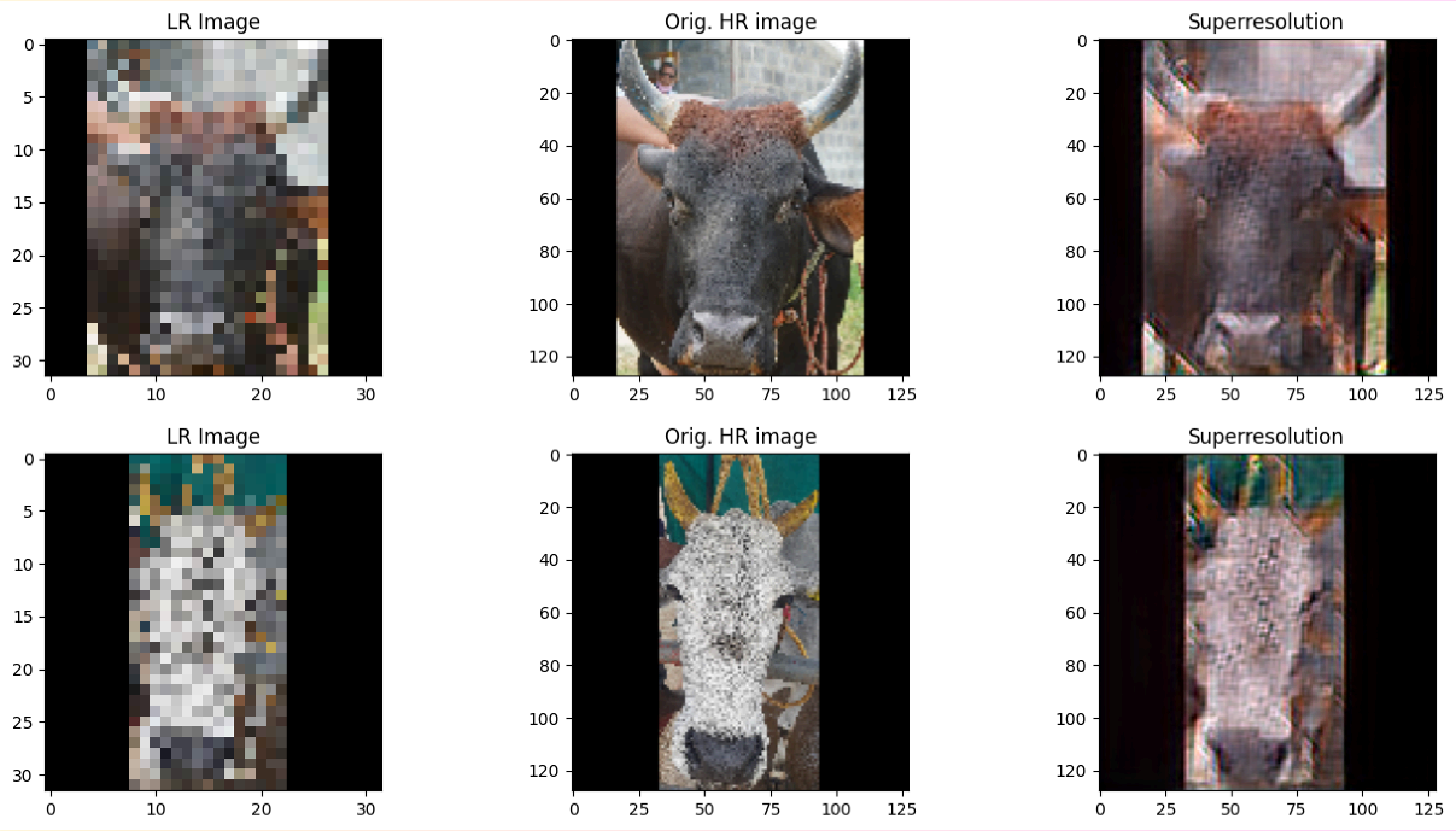
- Input of the generator is the image that is downscaled by 4 times to its high resoluted original image.
- The Generator Network consisting of 16 identical residual blocks is used for the feature extraction of the low resoluted image.
- Then upsampling the image dimensions to generate an image which is four times of the image that is given as input.

- As a result the generator provides an output of the input image which has a size equal to that of the original high resolution image.
- To the discriminator network, the generated image is given as input.
- The features are compared with the original image features that are extracted and is compared for its validity.
- The discriminator gives the validity of the image i.e., the probability of it being the real image.

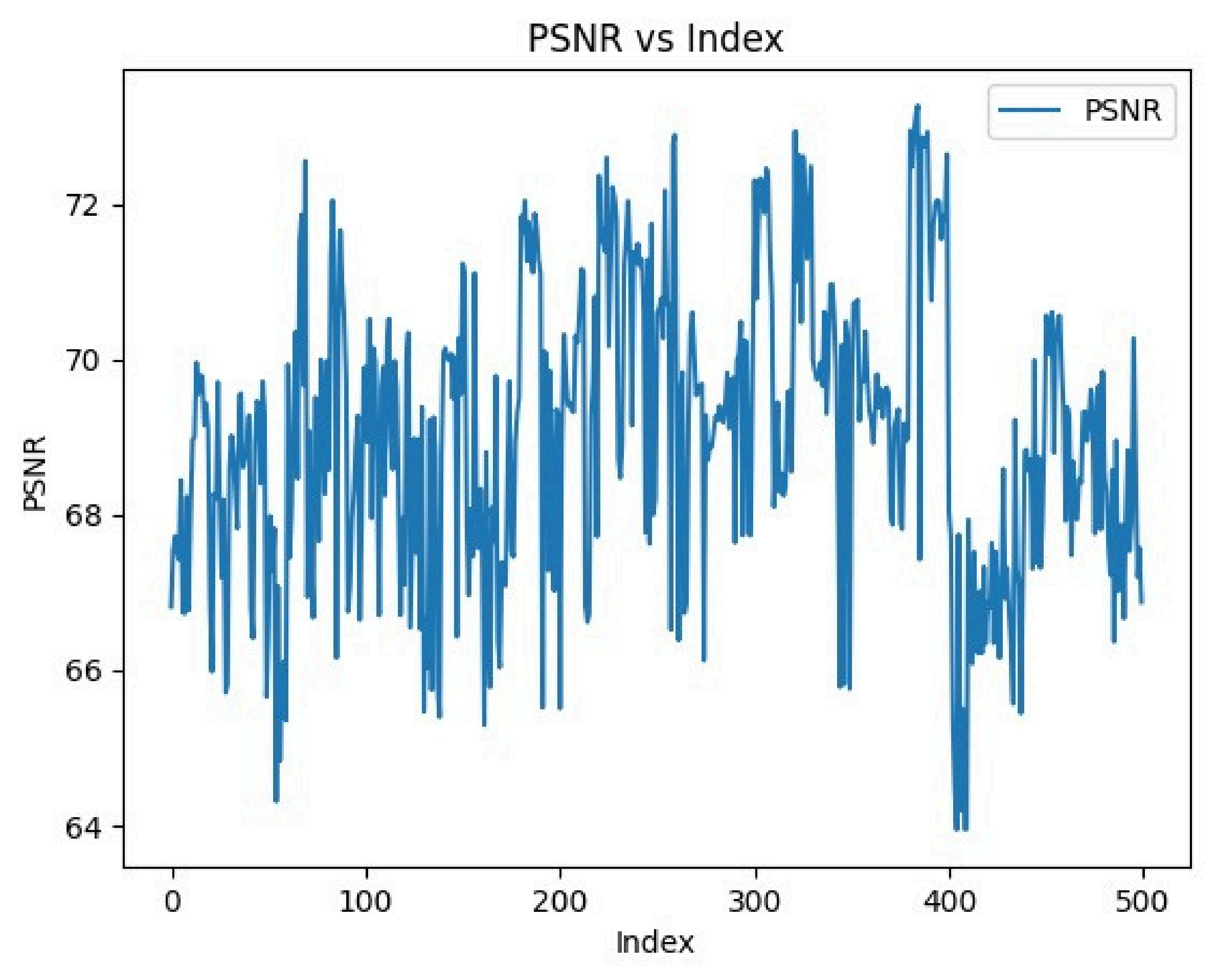


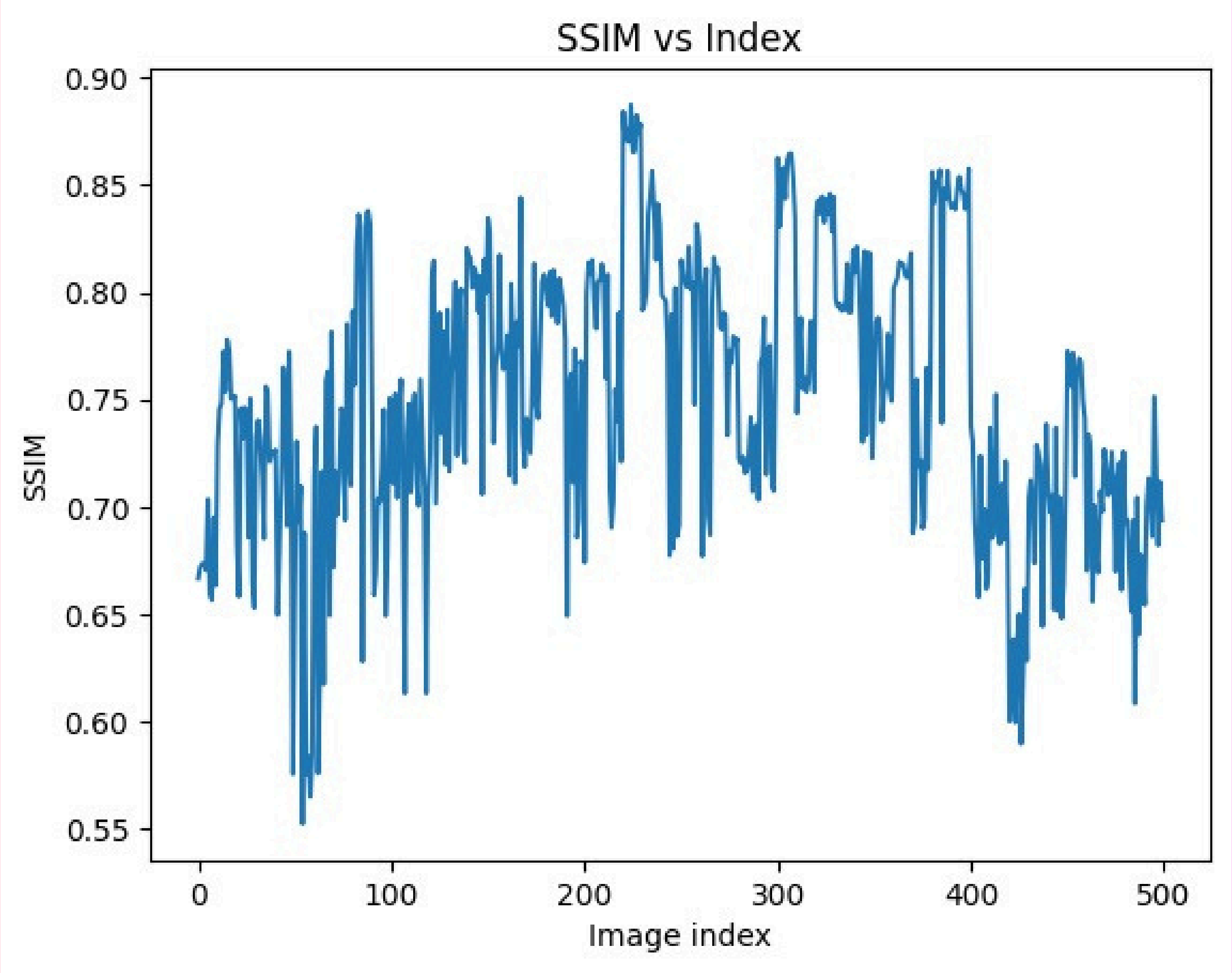
Both the discriminator and generator are trained alternatively in each epoch using the losses generated, and it is continued until the generator is able to successfully trick the discriminator.
This process is repeated over a large number of epochs to get realistic super-resolution images.





Results from the SRGAN model implemented





Average PSNR=69.02

Average SSIM=0.7499

- PSNR-Peak Signal to Noise Ratio which is calculated by comparing original image to genrated image. Higher the PSNR value depicts that image is less distorted and more closed
- SSIM-Structural Similarity Index, which ranges between -1 to 1 measures structural similarity between images using 3 factors namely - luminance, contrast and structure. Value closer to 1 implies for a more perfect image.

Conclusion

We have implemented SRGAN using keras and related python libraries. The finally generated images for the test data have decent PSNR and SSIM average on being run for 125 epochs. Increasing the number of epochs will improve the values of PSNR and SSIM and the generated images will also be more realistic.

The main features of this model are its loss function - perceptual loss function which uses pre-trained models to compare features and validate the generated images. Also, the model has many practical applications in fields such as medical imaging, remote sensing, and video processing, where high-resolution images are critical for accurate analysis and interpretation.

THANK YOU