Super Resolution

- Converting low resolution image to high resolution

Previous work on super resolution

- Interpolation techniques for upsampling - bilinear, nearest neighbor, bicubic
- SR using CNNs - good alternative to traditional filtering techniques
- SRResNets - deeper CNNs for better performance
- SRGANs - generative models for more realistic super resolution

Previous work on super resolution

NN interpolation
PSNR/SSIM: 24.02/0.74

SRResNet
PSNR/SSIM: 25.85/0.82

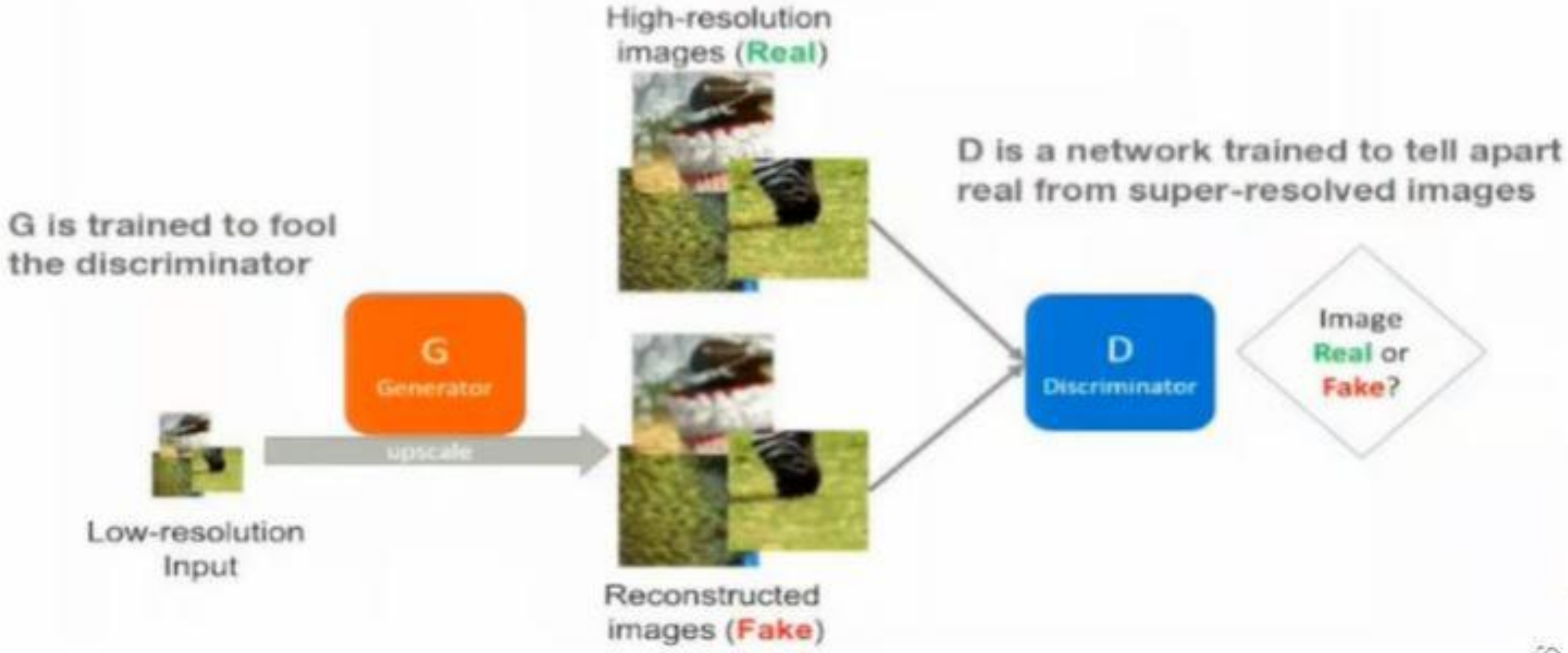
SRGAN
PSNR/SSIM: 22.71/0.70

Previous work on super resolution

NN interpolation	SRResNet	SRGAN
PSNR/SSIM: 24.02/0.74	PSNR/SSIM: 25.85/0.82	PSNR/SSIM: 22.71/0.70

SRGANs

- Generator : Low resolution to super resolution image
- Discriminator: Tell super resolution images from original high resolution images



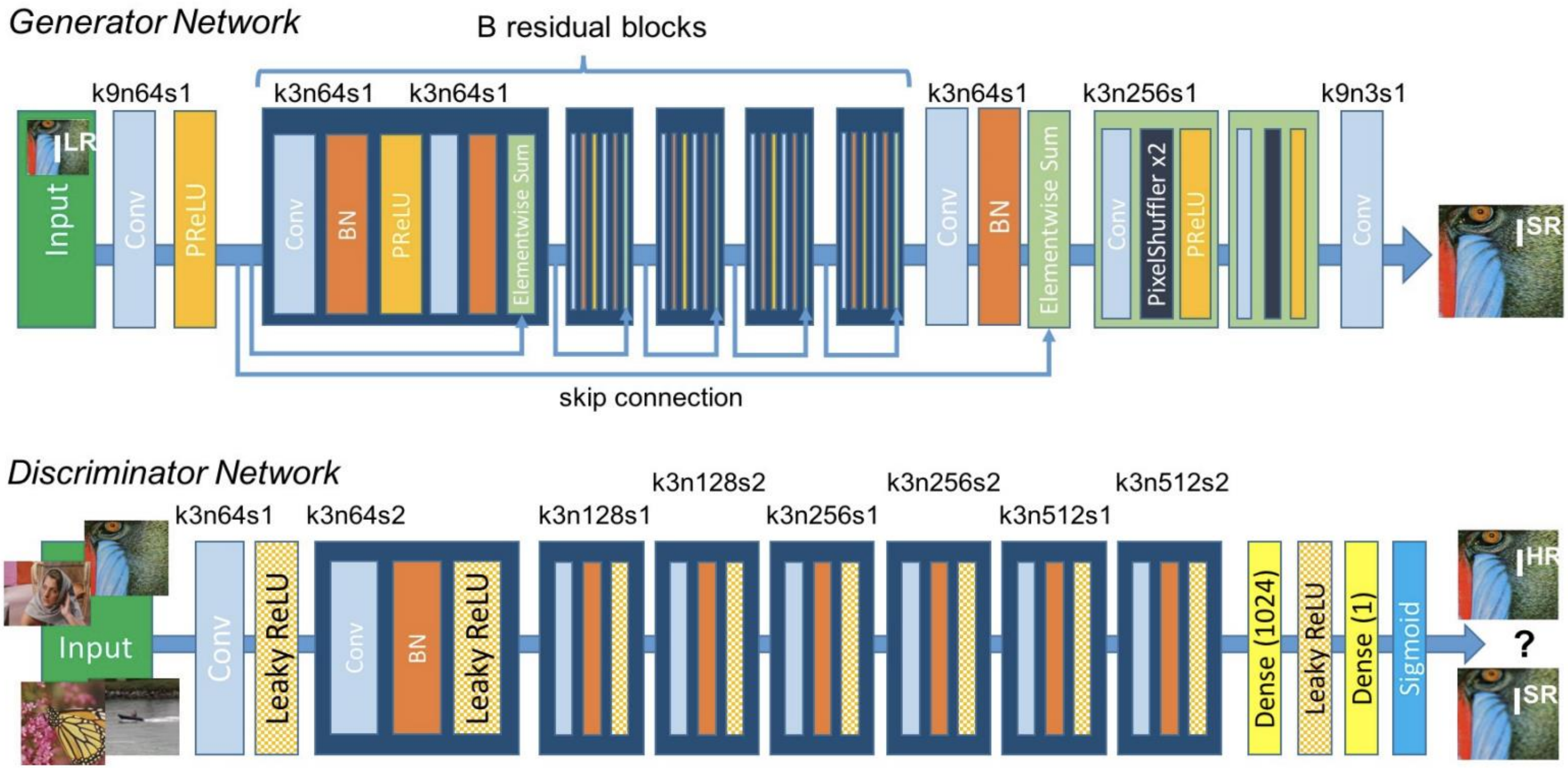


Figure 4: Architecture of Generator and Discriminator Network with corresponding kernel size (k), number of feature maps (n) and stride (s) indicated for each convolutional layer.

Loss function

- Traditional MSE loss - pixel to pixel comparison

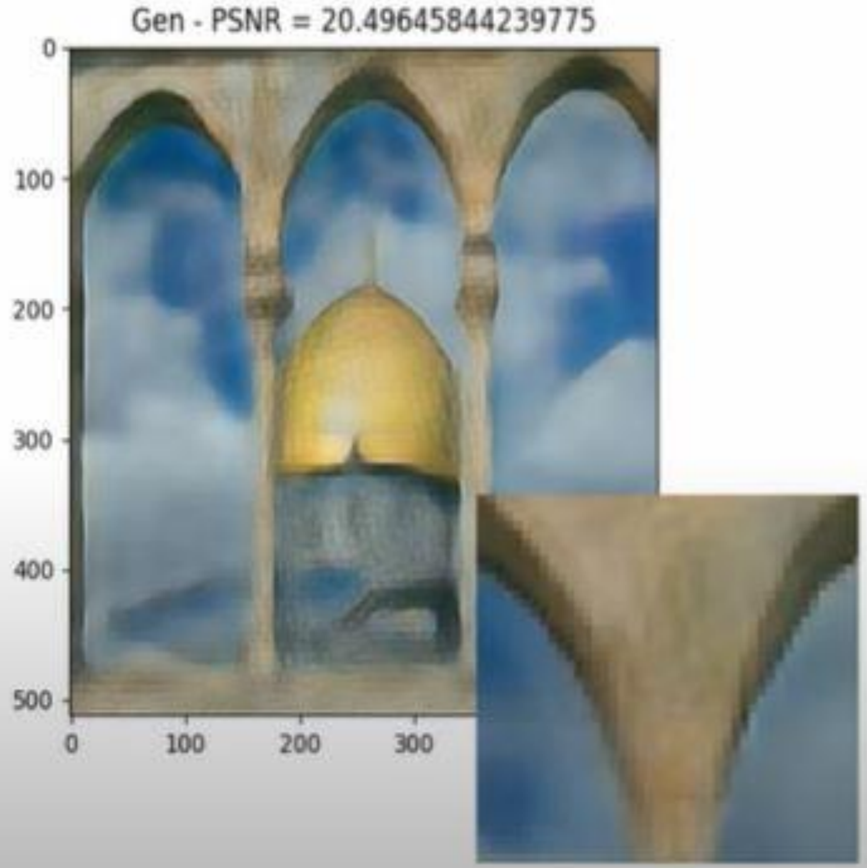
- Content VGG based loss - feature space MSE loss

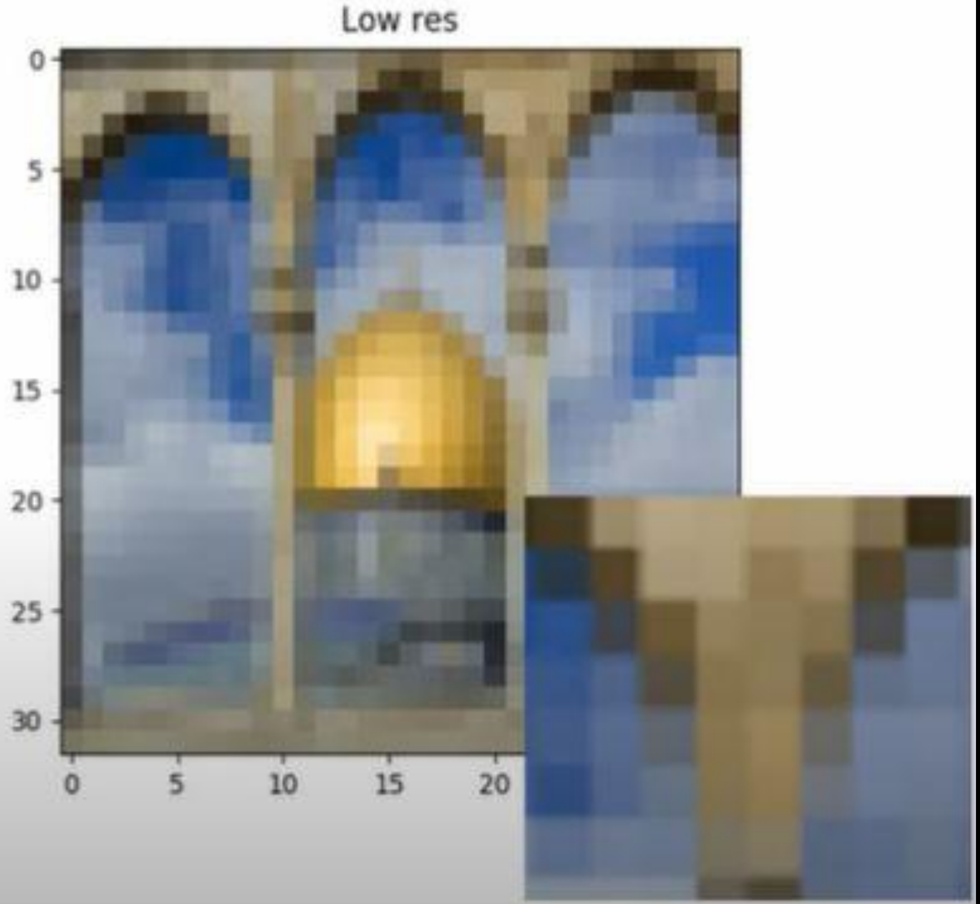
- Perpetual loss: Weighted sum of content loss and adversarial loss

Dataset

- DIV2K dataset: Diverse 2K high resolution images
- 800 training images
- 1024x1024 resolution
- Low resolution images obtained by downsampling the high resolution images

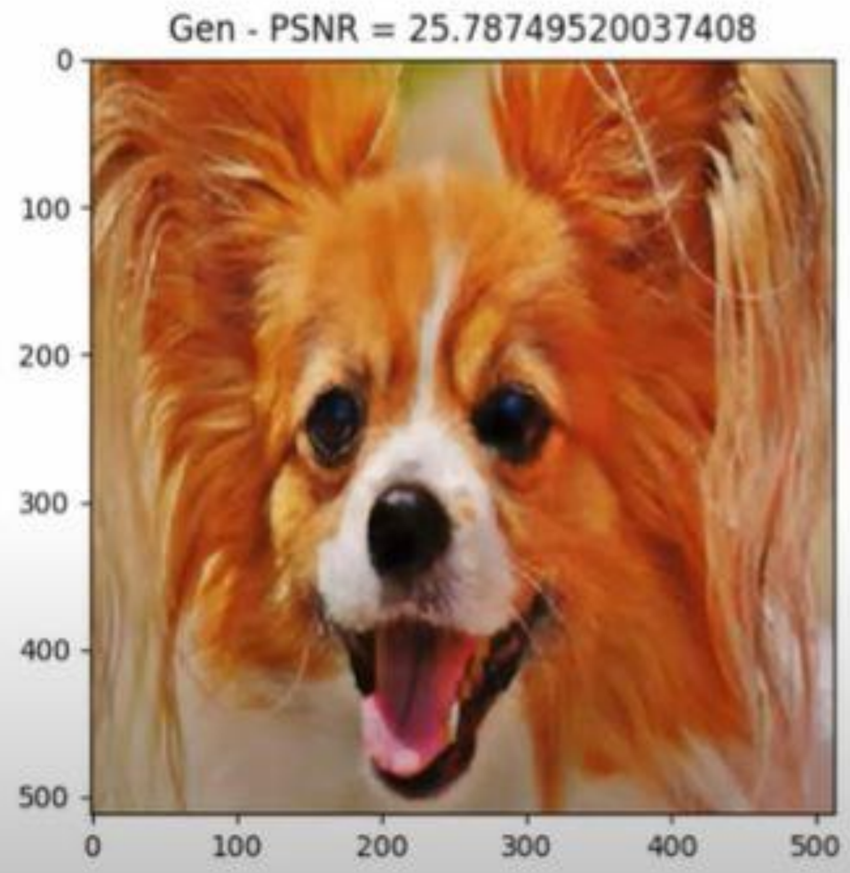
Breakdown analysis results - 16x

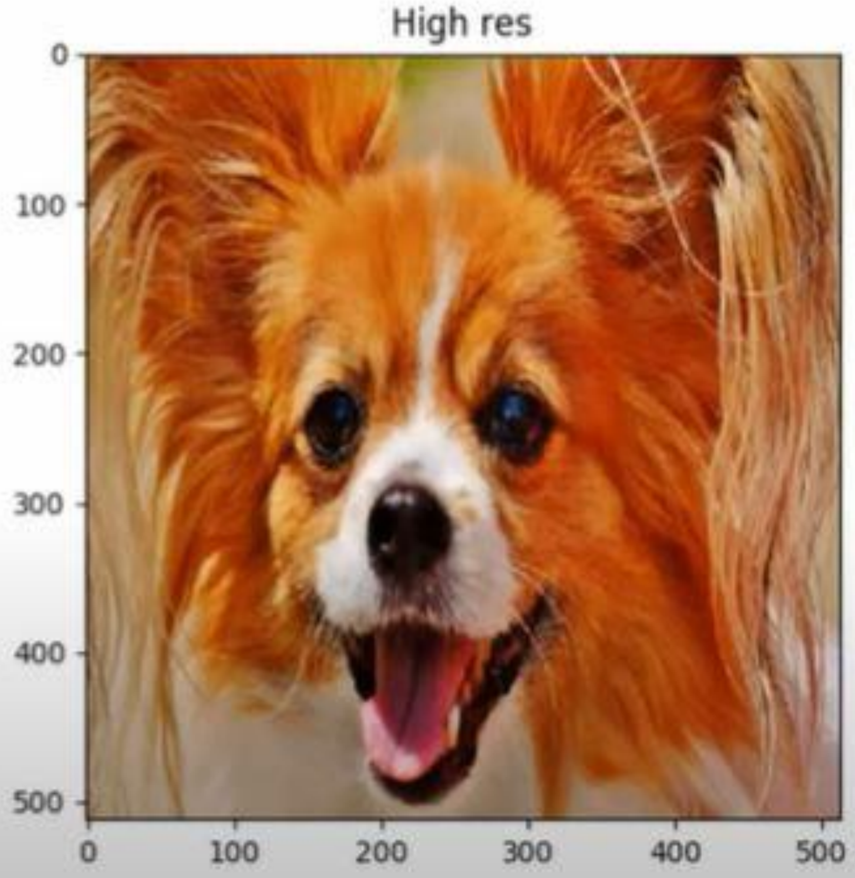




Breakdown analysis results - 16x

Breakdown analysis results - 4x





PSNR

PSNR is a commonly used metric in image processing and computer vision to measure the quality of an image. It is calculated by comparing the pixel values of the original image to the pixel values of a distorted or compressed version of that image. The distortion or compression can be caused by various factors such as noise, blur, compression artifacts, or image resizing.

The PSNR value is calculated as the ratio of the maximum possible pixel value (i.e., the maximum value that can be represented by the image bit-depth) to the mean squared error (MSE) between the original and distorted images. The MSE is the average of the squared differences between the pixel values of the original and distorted images. The PSNR value is expressed in decibels (dB) and can range from 0 (worst quality) to infinity (perfect quality). Higher PSNR values indicate better image quality.

Here is an example of how PSNR can be used to evaluate the quality of an image:

Suppose we have a 512x512 grayscale image with 8-bit pixel depth (i.e., the maximum pixel value is 255). We apply some noise to this image, resulting in a distorted version of the original image. We can then calculate the PSNR value between the original and distorted images using the following formula: PSNR = 10 * log10 (255^2 / MSE)

Where MSE is the mean squared error between the original and distorted images.
Let's say that the MSE value is 100. We can then calculate the PSNR value as:

PSNR = 10 * log10 (255^2 / 100) = 48.16 Db

This PSNR value indicates that the quality of the distorted image is relatively good, as it is close to the maximum value of 60 dB.

However, PSNR alone may not be sufficient to evaluate the quality of an image, as it does not take into account perceptual factors such as color, texture, and contrast. Therefore, other metrics such as SSIM Structural Similarity **Index** and LPIPS (**Learned Perceptual Image Patch Similarity**) may be used in conjunction with PSNR to provide a more comprehensive evaluation of the image quality.

- SSIM (Structural Similarity Index) is a metric used to measure the similarity between two images in terms of their structural content and perceptual quality. It is often used in image processing and computer vision to evaluate the performance of image restoration, enhancement, or compression algorithms.

- SSIM is based on the premise that the human visual system is highly sensitive to changes in structural information and less sensitive to changes in pixel intensity or color. Therefore, SSIM measures the similarity between two images based on their structural information, luminance (brightness), and contrast, rather than their pixel values directly.

- The SSIM index is calculated as the product of three components: luminance (l), contrast (c), and structure (s). Each component is a measure of the similarity between the corresponding features of the two images being compared. The overall SSIM index is calculated as the weighted average of these three components:
- SSIM = (l^a) * (c^b) * (s^c)

- where a, b, and c are constants that control the relative importance of each component, and l, c, and s are the luminance, contrast, and structure components, respectively.

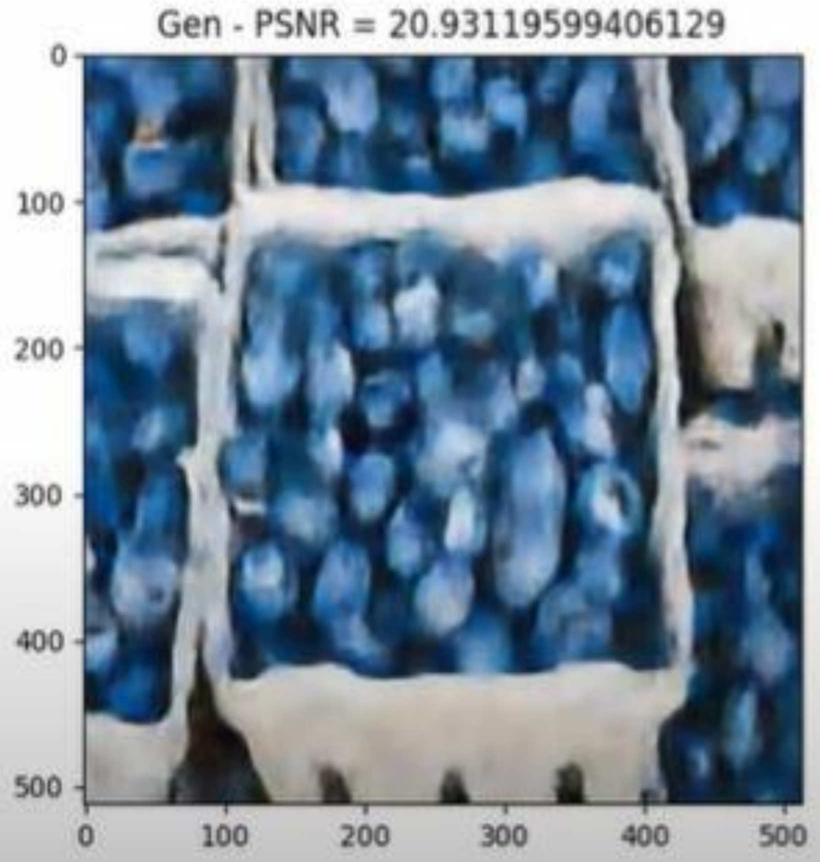
- The luminance component measures the similarity of the mean brightness of the two images. The contrast component measures the similarity of the standard deviation of the pixel values in the two images. The structure component measures the similarity of the correlation between the corresponding features of the two images.

- The SSIM index ranges from -1 to 1, with higher values indicating greater similarity between the two images. A value of 1 indicates perfect similarity, while a value of -1 indicates complete dissimilarity.

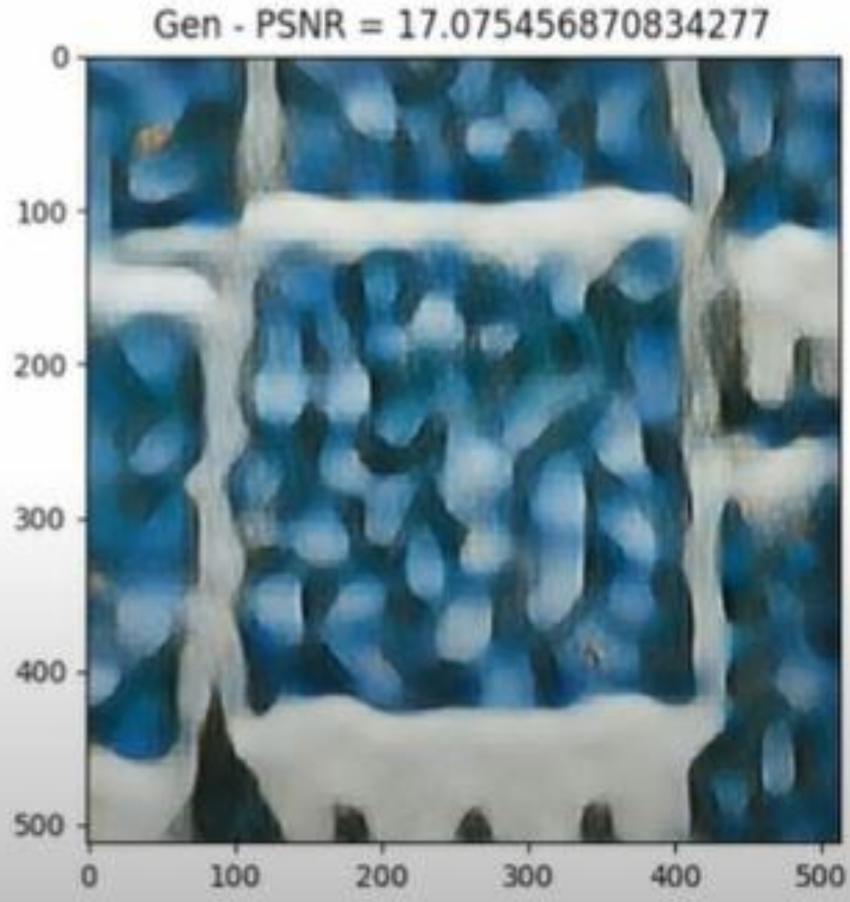
- In summary, SSIM is a useful metric for evaluating the similarity between two images in terms of their structural content and perceptual quality. It provides a quantitative measure of the similarity between the two images that takes into account the sensitivity of the human visual system to structural information.

Comparison of SRResNet with SRGAN - 16x

SRResNets



SRGANs



Comparison of SRResNet with SRGAN - 8x