Comparing Deep Learning Approaches for Image Denoising

BY

DWIJESH JIGARBHAI SHAH (DS1991)

HAARIKA AMBATI (HA539)

KUSH JAYANK PANDYA (KP1127)

Image Denoising

- Image denoising removes corruption from images caused by various factors.
- Images are often contaminated by noise during capture, compression, and transmission due to environmental and other influences.
- Noise can cause information loss and distortions that affect image analysis and tracking.
- Denoising is challenging because noise is tied to the image's high-frequency details.
- The goal is to balance noise suppression without losing important information.
- Image denoising is essential for effective image processing systems.

Types of Corruption

- Understanding these noise types is crucial for effective image denoising
- In this project, we will focus on different types of noise like: -
  a) Salt Noise
  b) Pepper Noise
  c) Salt and Pepper Noise

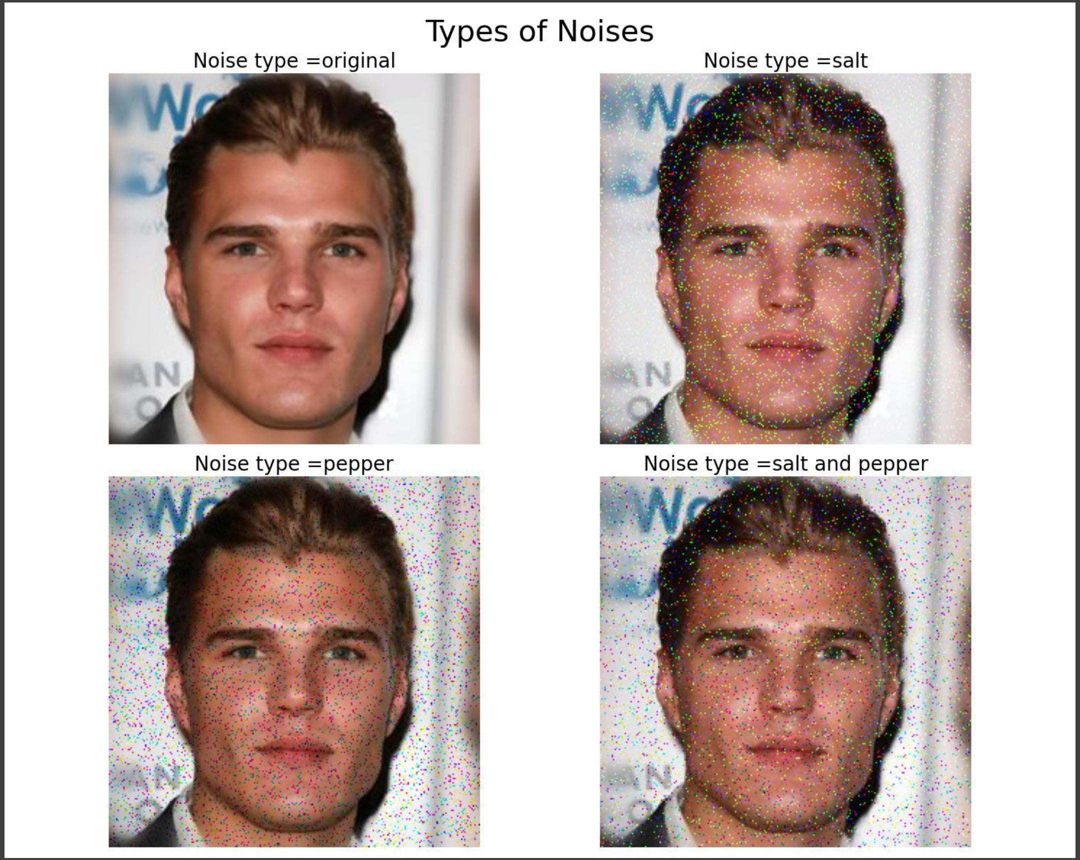


Image Similarity Metrics - Conventional

MAE/L1: Summation of the absolute difference between the predicted and actual values divided by the total number of data points. A lower MAE indicates better denoising performance.

SSIM: Used to evaluate the quality of the denoised image by comparing it with the original image. A higher SSIM indicates better denoising performance.

PSNR: It is an engineering term for the ratio of a signal's maximum possible power to the power of corrupting noise that compromises the representation of the signal's fidelity. A higher PSNR indicates better denoising performance.

What is our approach of comparing ?

- We want to compare the histogram distribution of the original image and the restored Image by seeing if the Neural Network model can preserve it or not.

Image Histogram Vector

For our Novel comparison Metric, we will be using this Histogram Vector, it's a concatenation of 3 channels of histogram values of an Image.

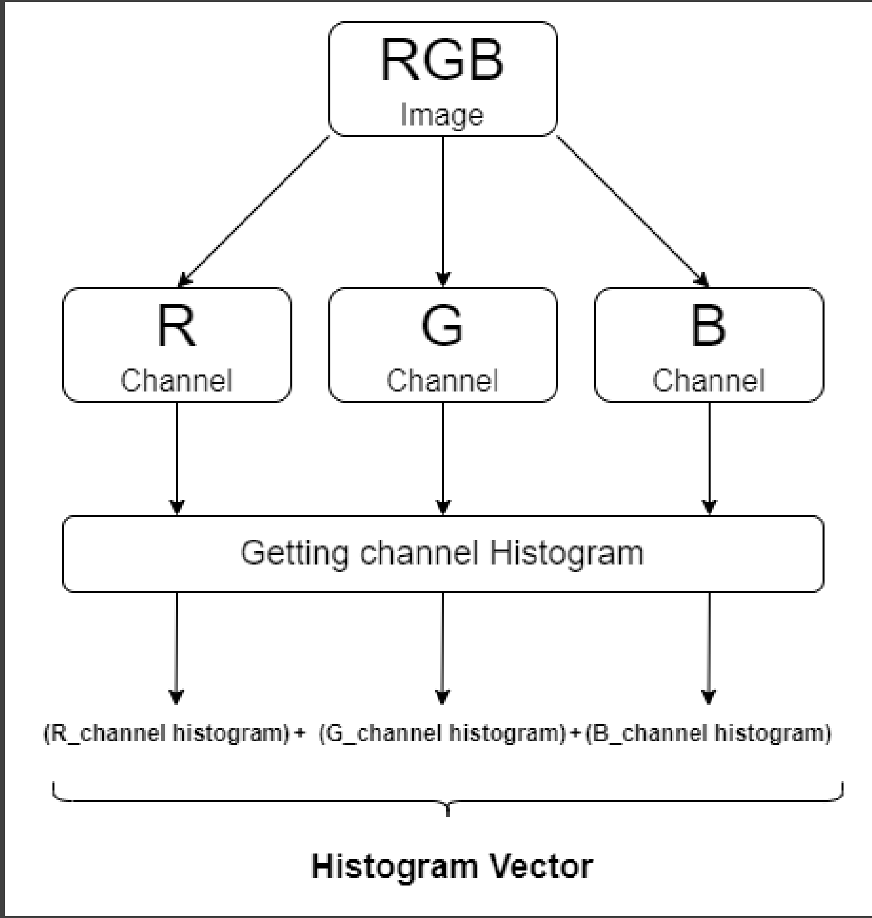


Image Similarity Metrics - Novel

Cosine of Histogram: The cosine similarity between two vectors is defined as the cosine of the angle between them, represented by the dot product of the two vectors divided by the product of their magnitudes.

RMSE of Histogram: We use the conventional L2 norm of two vectors(Image Histogram Vectors).

$$RMSE = \sqrt{\sum_{i=0}^{(3X255)-1} (H^{O}_{f=1} - H^{R}_{f=1})^2}$$

$$Cosine = \frac{\vec{H^{O}}.\vec{H^{R}}}{||H^{O}||.||H^{R}||}$$

Datasets used for validation

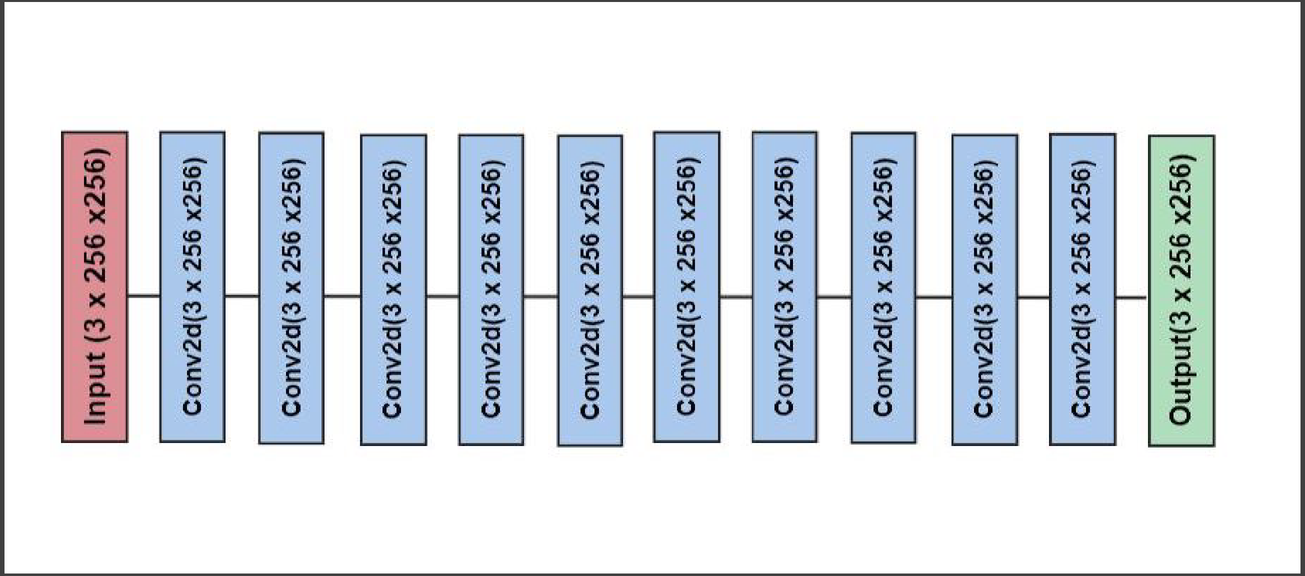
CelebA-HQ resized - This dataset is derived from the CelebA-HQ dataset which contains resized 256 x 256 image resolutions and has applications in the fields of face recognition, facial expression analysis, etc.

LSUN bedroom- This dataset contains millions of images from different indoor scenes of bedrooms, which is mainly used for scene understanding tasks. We are going to use a subset of it for comparison.

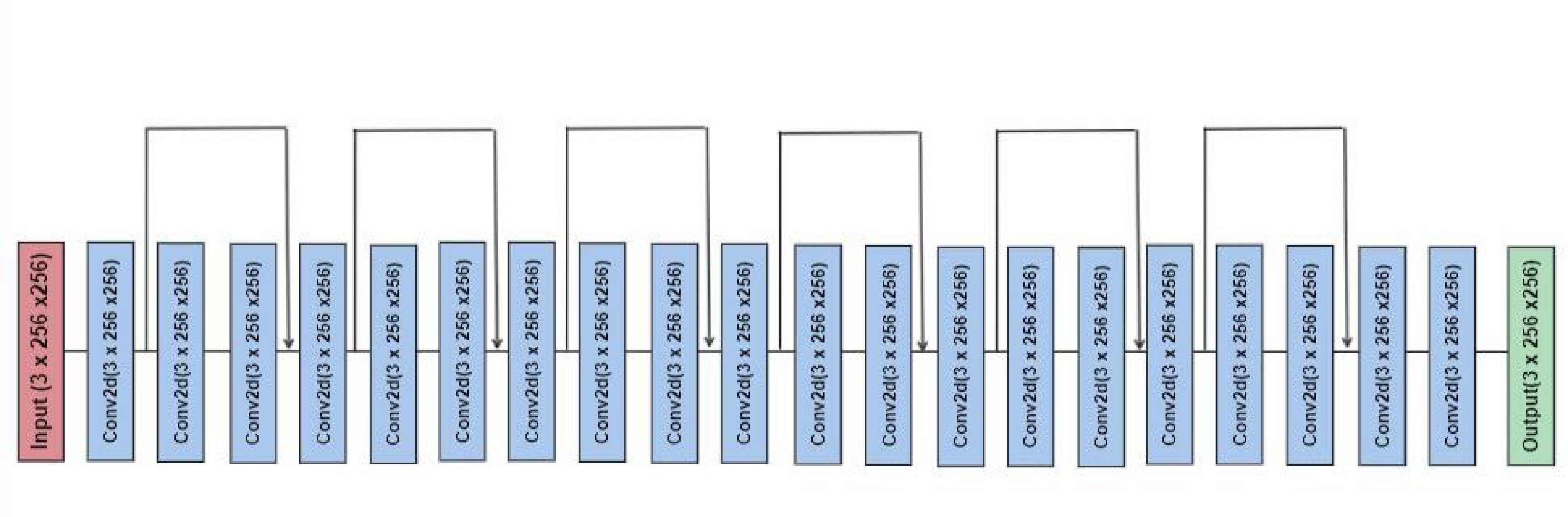
ImageNet mini- extensive image set publicly available 2009 database of labeled images designed for visual object recognition research. We are going to use a subset of it for comparison.

Models used

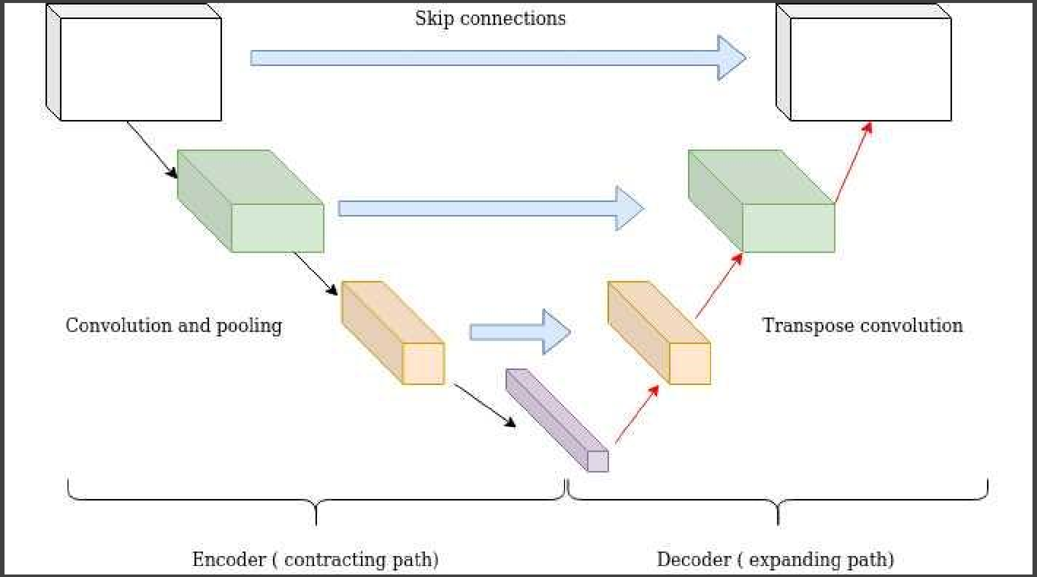
- Basic Convolution Neural Network
- Residual Network
- Convolution Skip Encoder
- U-Net
- Conditional GANs


Input (3 x 256 x256)
Conv2d(3 x 256 x256)
Conv2d(3 x 256 x256)
Conv2d(3 x 256 x256)
Conv2d(3 x 256 x256)
Conv2d(3 x 256 x256)
Conv2d(3 x 256 x256)
Conv2d(3 x 256 x256)
Conv2d(3 x 256 x256)
Conv2d(3 x 256 x256)
Conv2d(3 x 256 x256)
Output(3 x 256 x256)

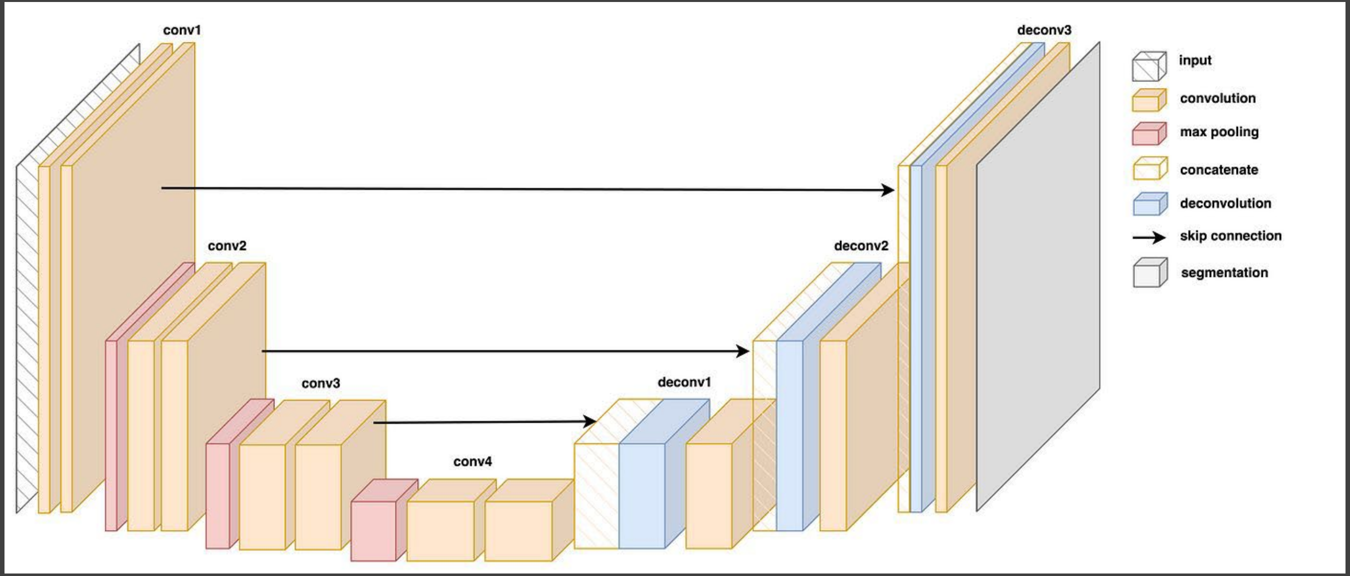
Basic CNN



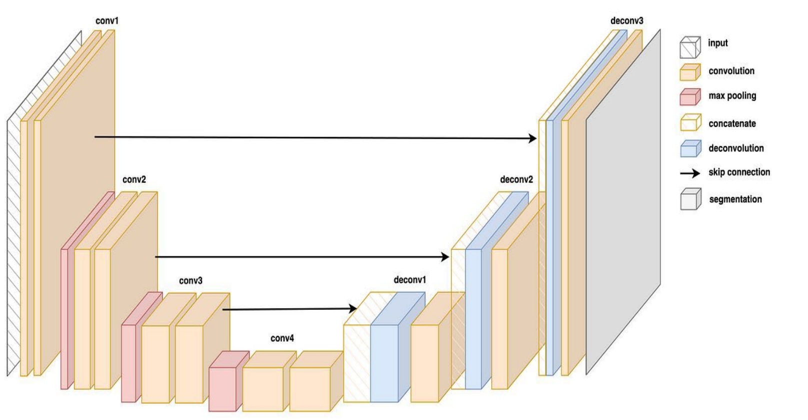
Residual Network



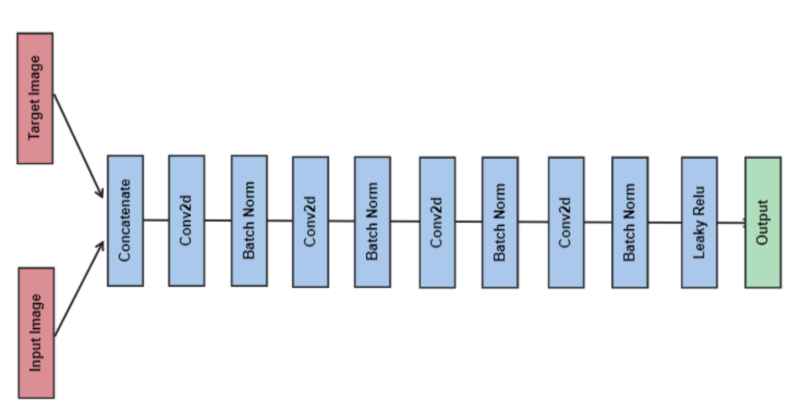
Symmetric Skip AutoEncoder


conv1
conv2
conv3
conv4
deconv1
deconv2
deconv3
input
convolution
max pooling
concatenate
deconvolution
skip connection
segmentation

U-Net



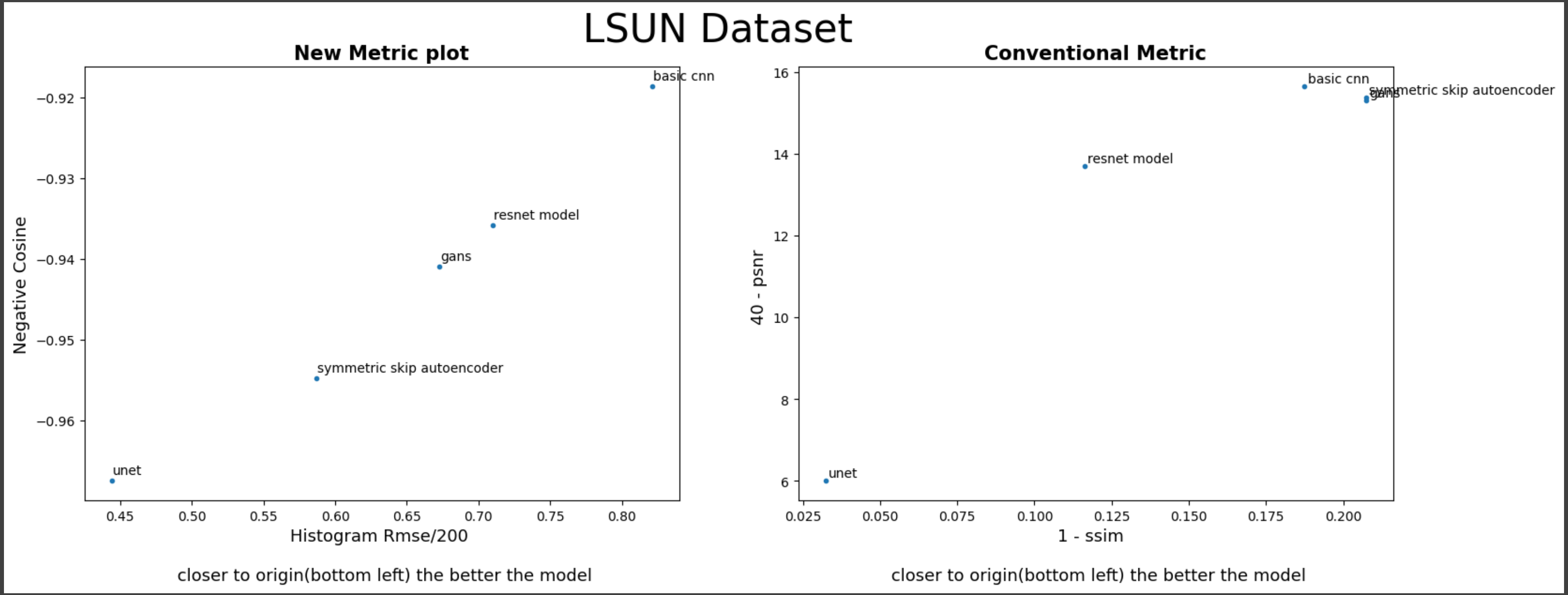
Generator (U-net)



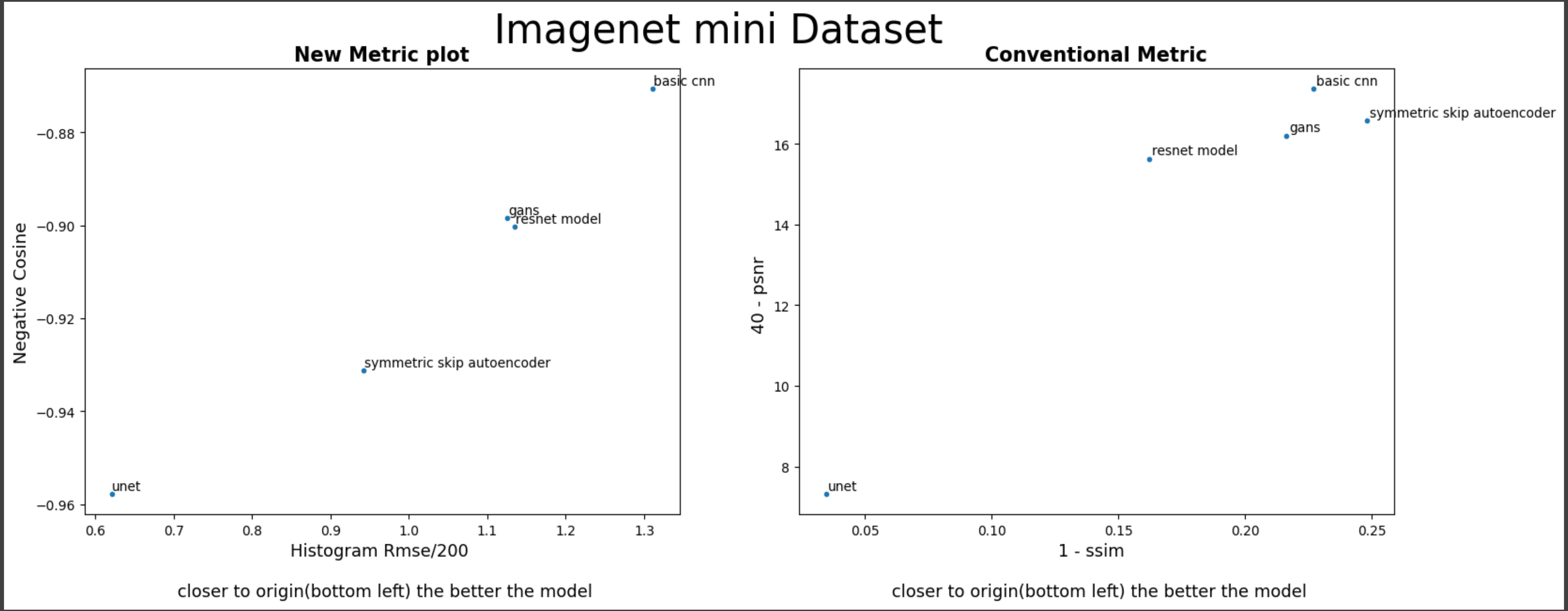
Discriminator

Conditional Gans

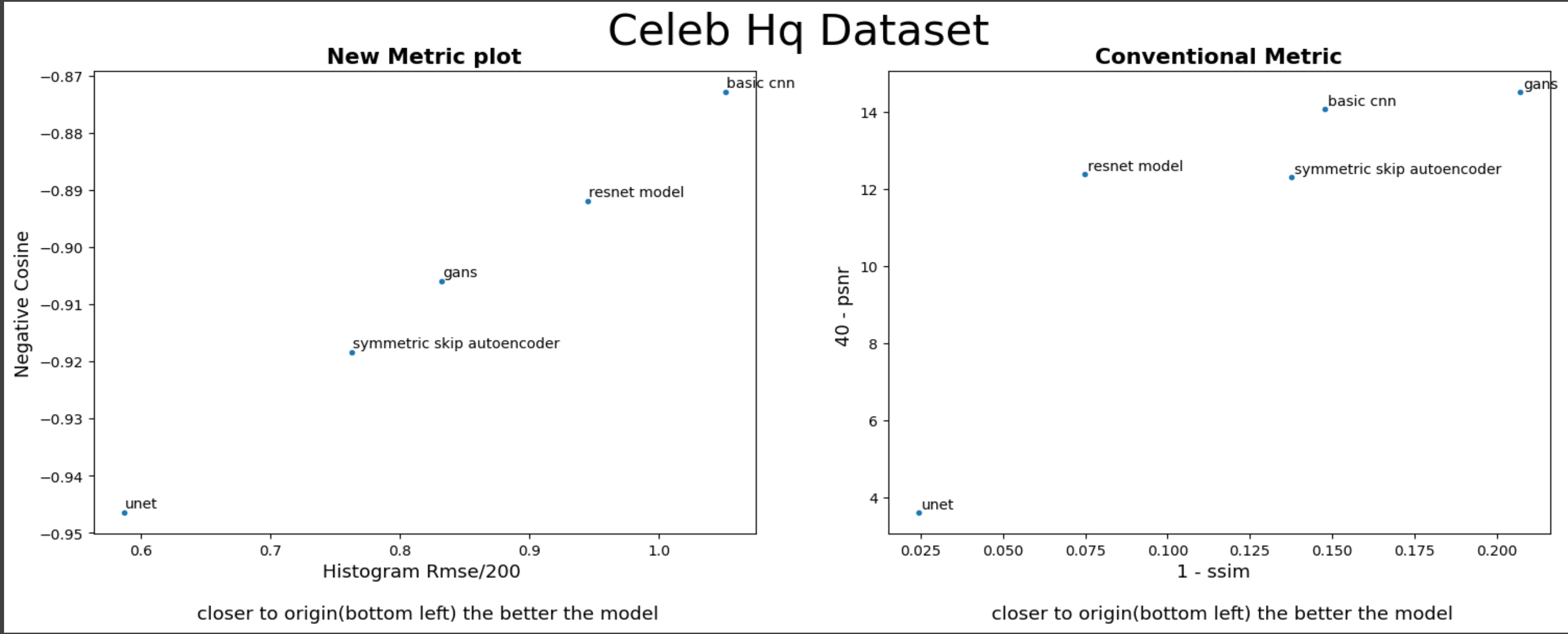
Results



Results - LSUN



Results – ImageNet mini



Results – Celeb Hq

Observation and Conclusion

- All the Autoencoder approach has seen better new metric results compared to conventional metric results. This may be because the skip connection between the encoder and decoder helps them preserve the structure, fine details, and histogram distribution.

- Normally Trained U-Net has always performed better than other models in conventional as well as the new metric.

- There is not much difference in the results of Basic CNN and Resnet.

- Although the generator of Gans is a U-net architecture, we haven't seen the same level of performance from it as we saw from a normal U-net. The reason for this may be that the GANS take more time and computation to optimize, and also, how weights and biases converge to minima depends on your training approach.

Thank You!

“No Structure, even artificial one, enjoys the process of entropy. It is the ultimate fate of everyone, and everything to resist it.”

Philip K Dick