Discussion: Comparing Original and Reconstructed Images

Clarity:

The clarity of reconstructed images is a crucial indicator of how well the VAE has learned the underlying features of the input images.

Observation:

In the provided visualizations, the reconstructed images appear blurred and lack the sharpness seen in the original images.

The digits are not distinctly recognizable in many cases, which indicates that the VAE struggles to capture finer details.

Analysis:

This could be due to the small latent space dimensions (2, 5, or 10), which limit the amount of information that can be encoded.

A larger latent dimension might be required to retain more detail during the encoding process.

Distortion:

Distortion refers to the differences between the original and reconstructed images, such as missing details, blurred edges, or altered shapes.

Observation:

The reconstructions show significant blurring and a loss of details, making the digits in the reconstructed images hard to distinguish.

The overall shape of the numbers is somewhat preserved but lacks the sharp edges and clarity of the originals.

Analysis:

Some level of distortion is expected due to the compression-decompression process of the VAE.

However, the extent of blurring suggests that the model might benefit from additional training or a better reconstruction loss function like MeanSquaredError instead of binary\_crossentropy.

Variability:

Variability checks if the VAE can consistently reconstruct images with different styles, backgrounds, and orientations.

Observation:

The original images vary significantly in digit style, orientation, and background noise.

The reconstructions show a consistent blurring pattern, with less distinction between different digits and backgrounds.

Analysis:

The VAE struggles with images that have complex backgrounds or rotated digits, suggesting that the latent representation does not capture these variations well.

Training for more epochs or using a higher latent dimension might help the model better capture the diversity in the dataset.

Generalization:

Generalization is the model's ability to produce coherent reconstructions of images it has not seen before.

Observation:

The reconstructions suggest that the model has learned some general features of the SVHN dataset, such as the overall shape and structure of numbers.

However, the significant blurring indicates that it has not fully captured the finer details needed for clearer reconstructions.

Analysis:

The model’s ability to generalize to unseen data could improve with a more extensive training regimen, an increased number of epochs, or a more complex architecture.

Improving the balance between the KL divergence loss and reconstruction loss could also enhance the quality of the learned representations