Autoencoders: MNIST&CIFAR10 image reconstruction

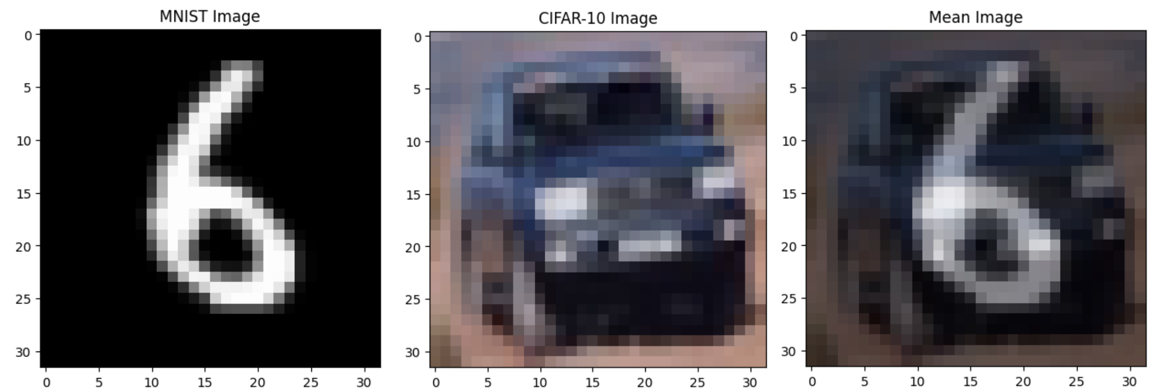
1. Introduction

This report describes the process of training an autoencoder to reconstruct images from the MNIST and CIFAR-10 datasets. The steps include data processing, model architecture definition, training, image reconstruction, and evaluation using *PSNR* and *SSIM* metrics.

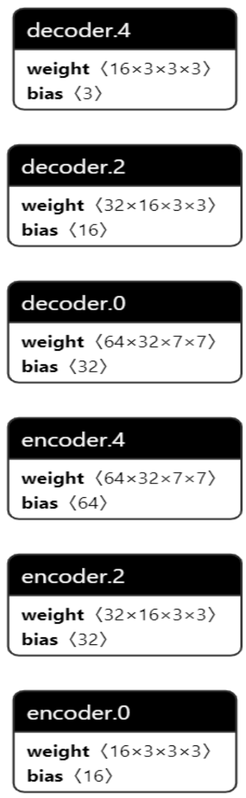
2. Data Processing

Datasets: The MNIST and CIFAR-10 datasets are used.
Steps:

- Randomly select one image from each dataset. Resize the MNIST image to 32x32 pixels and extend it to 3 channels to match CIFAR-10 images.
- Compute the mean of these two images element-wise.

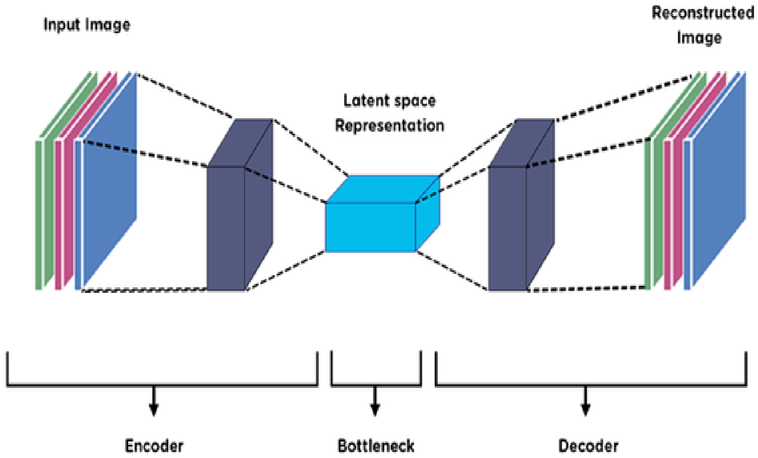


3. Model Architecture



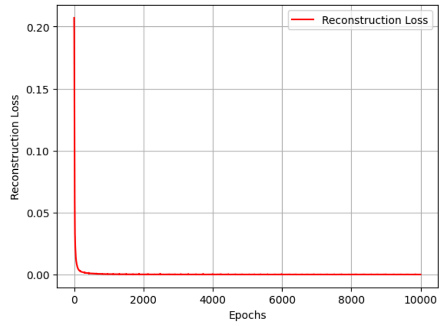
Autoencoder Architecture:

- **Encoder:** Compresses the input image into a latent representation using convolutional layers.
- **Decoder:** Reconstructs the input image from the latent representation using transposed convolutional layers.



4. Reconstruction

The autoencoder is trained using the mean image computed earlier, with MSE as loss module and Adam as optimizer with learning rate equal to 0.001.
The loss curve for 10000 epochs of the network is reported as Follows, which shows the well convergence of the autoencoder.

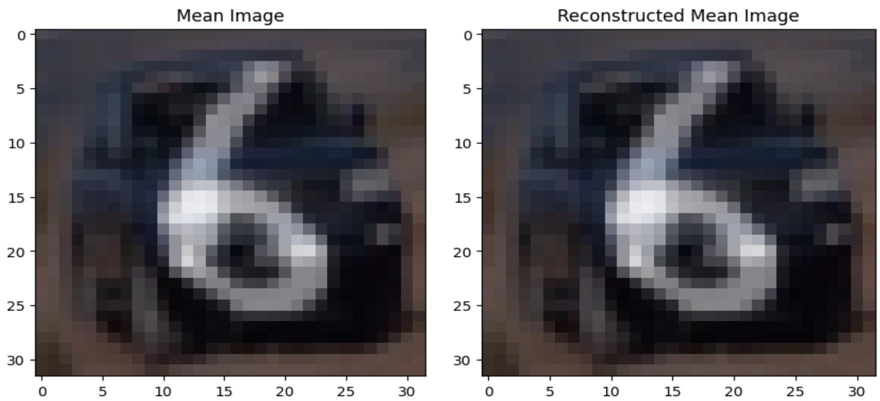


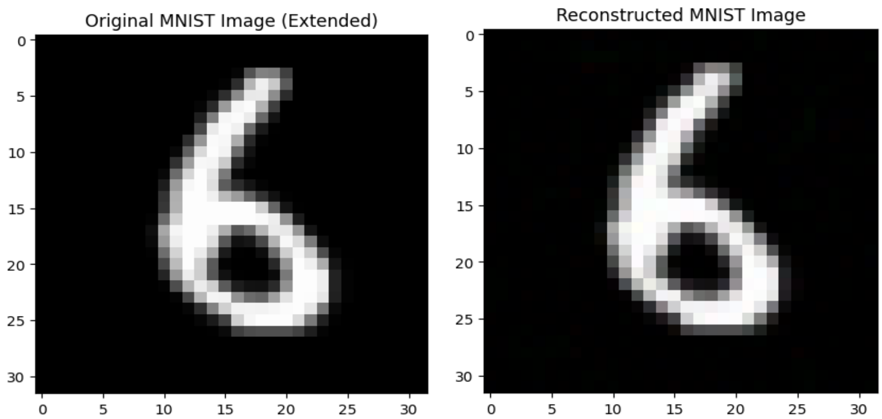
5. Quality Comparison

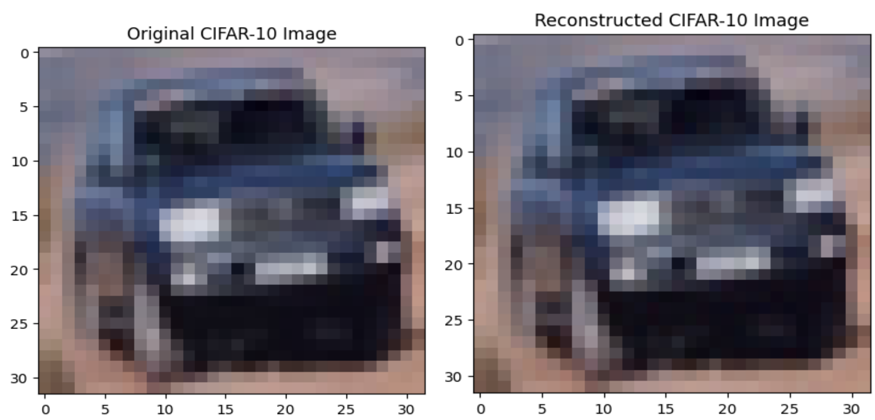
Then we plot reconstructed images for mean input images.
Also for reconstructing the MNIST and CIFAR10 image as well, we multiplied the reconstructed mean image by 2 and subtract it by one of the base images(MNIST or CIFAR10) as follow:

- MNIST_image = (mean_recunstructed_image * 2) - CIFAR10_image
- CIFAR10_image = (mean_recunstructed_image * 2) - MNIST_image

The results are as below:







6. Evaluation Metrics

We calculated the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) for the reconstructed images compared to the originals. Peak Signal-to-Noise Ratio (PSNR) is a metric used to evaluate the quality of a reconstructed image by comparing it to the original image. It is a measure of the ratio between the maximum possible power of a signal and the power of corrupted or distorted noise that affects the fidelity of its representation.Typically, a high PSNR value (usually above 30 dB) indicates good quality reconstruction, while a lower value suggests that there is more noise and distortion in the image. Structural Similarity Index (SSIM) is another method used to measure the quality of a reconstructed image by comparing it to the original image. SSIM takes into account not only the intensity of the pixels in the images but also their structure, texture, and luminance. SSIM calculates the similarity between the original and reconstructed images by measuring three components: luminance comparison, contrast comparison, and structural comparison. These components are combined to give an overall SSIM index, where a value of 1 indicates a perfect match between the two images.

Dataset	SSIM	PSNR
MNIST	0.9954	44.5539
CIFAR-10	0.9988	44.5539

The reported result in the above table shows our model could reconstruct images with high quality and similarity as the original image.