

CMPE 597 Sp. Tp. Deep Learning
Spring 2025 Assignment 2
Due May 18th by midnight

Data

In this assignment, the Quick Draw dataset will again be used. You may still download the dataset at the same link <https://drive.google.com/file/d/1oG0HnQEb7VVwpdYsn8rShfAQ8Vq4hA9E/view?usp=sharing>. You can use deep learning libraries to implement the architectures in this assignment.

Tasks

I. Autoencoders

In this task, you need to implement and train the following autoencoders. You may use mean squared error (MSE) loss.

1. Consider a gated recurrent neural network as the encoder and decoder. Please treat the grayscale images as multivariate time series. You may try Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU). You need to determine the number of layers and the number of neurons in the units.
2. Consider convolutional encoder and decoder. You must determine the number of layers, kernel sizes, number of feature maps, and the activation functions.
3. Please report the change in the MSE loss during training.
4. Plot the embeddings obtained from the encoder in 2D using t-SNE ¹. Comment on the discriminability of the embeddings learned by the autoencoders above.

II. Variational Autoencoders

In this task, you must implement and train the following variational autoencoders (VAEs). You may use the architectures you designed in the first task.

1. Consider a VAE with the gated RNN encoder and a convolutional decoder. Train the VAE and plot the changes in its objective during training. Show the change in both KL divergence and reconstruction terms on the same plot.
2. Now, consider a VAE with a convolutional encoder and decoder. Compare its convergence with the one with the RNN encoder.

¹<https://lvdmaaten.github.io/tsne/>

3. Generate new samples using the decoders in question 1 and 2. Compare their visual quality and diversity. You may use metrics such as Inception Score (IS) and Frechet Inception Distance (FID) to evaluate the visual quality.
4. Take the best-performing VAE architecture and implement a conditional VAE. Generate and visualize five new samples from each rabbit, yoga, and snowman category. Using the classifier you trained in Assignment 1, predict the labels of the generated samples. Comment on the confidence of the classifier.

Submission

- A PDF report addressing the above points, including your name, student number, references, and drive link to the code, should be submitted.
- Please include the link to your code on the first page of your report.
- Your experiments should be reproducible.
- A readme should be provided.

IMPORTANT NOTES:

- There is no report template. However, your reports should address all the bullet points above.
- You should not forget to cite your references. If you followed a GitHub repository, do not forget to cite it.
- Please note that you should be ready to answer questions regarding your code.
- Grading of this assignment will not be based on performance. However, you should be able to show that your networks can learn.