Exercise - Variational Autoencoder

Computer Vision Course, ETH Zurich

Overview

In this assignment, we will implement a Variational AutoEncoder (VAE), a class of generative models that enables us to sample new output not seen in the training data. Your task is to implement the base VAE model and its variant, Conditional Variational AutoEncoder (CVAE), which allows the sampling to be conditioned on additional signals, such as a class label.

NOTE: Do not modify the complete code provided. Simply add your codes in parts marked with TODO. There is no need to write a report for this exercise. Remember to keep all outputs in every cell of your notebook; do not clean them in your submission!

1 Setup

As you will need a GPU for training in this assignment, we recommend using Google Colab for GPU access. Alternatively, you can use your local GPU if they are available.

If you use Google Colab, upload the accompanying Python notebook file to the website. The libraries required by this assignment are natively available on Colab so you do not need to set up a new environment. If you use your local GPU, you will need the latest version of Pytorch and Matplotlib.
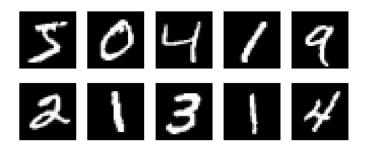
The notebook has self-contained descriptions for scoring and explanations. The following sections only serve as a general guideline for this assignment.

2 Variational Autoencoder (VAE) (50 points)

In this task, we will train a model to generate images of hand-written digits from the MNIST dataset. The goal is to obtain a model that we can use to sample a valid image of the digits.

Your tasks include:

- Implement VAE Encoder and Decoder (20 pts.)

- Implement the reparameterization trick (5 pts.)

- Implement VAE losses (5 pts.)

- Implement the training loop (5 pts.)

- Reach acceptable reconstruction quality (5 pts.)

- Implement the sampling code and reach acceptable generation quality (10 pts.)

3 Conditional Variational Autoencoder (CVAE) (50 points)

With the previous model, we cannot control which digit we want to generate e.g. generate different images of digit 5. In this section, we will implement the CVAE model that allows us to select which digit we want to generate.

Your tasks include:

- Implement CVAE Encoder and Decoder (20 pts.)

- Implement the reparameterization trick (5 pts.)

- Implement the training loop (5 pts.)

- Reach acceptable reconstruction quality (5 pts.)

- Implement the conditional sampling code and reach acceptable generation quality (15 pts.)