CS323: Deep Learning for Visual Computing
Reading Assignment 5

David Felipe Alvear Goyes

April 30, 2023

1 GANs

Generative Adversarial Networks consist of two neural networks, a generative model that generates samples from random noise or a prior distribution, and a discriminative model that determines if the sample is from the generative model distribution or the data distribution. These two neural networks are trained together in a two-player min-max game to generate new, realistic samples from an underlying data distribution.

The goal of the generator is to create new data samples based on the real data distribution, this model learns to produce iteratively realistic samples to fool the discriminator. The objective of the discriminator is to distinguish between real samples and samples generated by the generator while training the discriminator to learn to become better at identifying fake samples. The training of the models goes up to the generator is capable to create realistic samples.

GANs are trained to receive a prior distribution or a simple random variable of a specific probability distribution (Uniform distribution, Gaussian distribution), the goal of the neural network is to transform the random variable learning a complex function through a neural network to obtain a new random variable that represents the targeted sample distribution.

1.1 Explain the different loss terms in the GANs loss function

The training of adversarial networks can be considered as a min-max two players game, where the discriminator is trained to detect fake data, then the objective of the training is to minimize the final classification error, this is the basic metric of both networks. The generator is trained to fool the discriminator by taking input random variables and generating a distribution that is closer to the true distribution. The objective of the network is to maximize the final classification error.

The loss function in GANs consists of two main parts: the generator and discriminator loss. The loss for the discriminator is given below and it aims to maximize the probability of classifying real and generated samples. x are the real images, z random noise, $p_data(x)$ is the real data distribution, $p_z(z)$ is the noise distribution, $G(z)$ represent the generated samples, and $D(x), D(G(z))$ represent the discrimination predictions. The first term evaluates the performance of real samples and the second term measure the discriminator's performance of generated samples.

$$L_D = -\frac{1}{2}\mathbb{E}_{x\sim p_{data}(x)}[\log D(x)] - \frac{1}{2}\mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))]$$

The generator loss tries to maximize the probability of the discriminator classifying the generated images as real samples. The generator loss is given below.

$$L_G = -\mathbb{E}_{z\sim p_z(z)}[\log D(G(z))]$$

1.2 Explain the concept of mode collapse in GANs

The mode collapse happens when the generator model starts to produce a limited variety of outputs, focusing on a small number of modes. The generator discovers a few modes of data distribution to fool the discriminator and leading to less diverse and less realistic outputs. The reason for this problem is that the generator finds to generate sample data that fools the discriminator, and it stops the exploration of the rest of the data distribution. Mode collapse can lead to poor generalization, reduced diversity in new samples, and the inability of the model to learn true data distribution.

1.3 Explain the concept of latent space in GANs and latent space interpolation

The latent space in Generative adversarial networks is an N-dimensional vector that follows a probability distribution that serves as an input to the generator model. It represents a compressed or encoded representation of the data distribution. The objective of the generator is to learn how to map the latent vector to samples in the data space. This vector captures essential information about the data distribution that then will be used to create diverse samples.

Latent interpolation is used to create an interpolation between points in the sample space to create diverse and smooth transitions between samples. The latent space corresponds to generated images or samples, where using the latent interpolation we can create intermediate points between two generated samples. The latent space provides information on the underlying structure of the data distribution.

1.4 Is training a generative adversarial network stable?

Given the adversarial nature of the optimization problem training GANs, the competition of the discriminator and the generator lead to different problems in that the GANs training become unstable. We mentioned that one of the problems is the Mode Collapse, where the generator starts to produce unrealistic modes that fool the discriminator, resulting in a less diverse and realistic output. The other problem is the vanishing gradients where the discriminator becomes too good at distinguishing generated and real samples, leading to a small gradient update for the generator. The other problem in training GANs is the oscillation which occurs when the models are competing with each other to converge to the equilibrium, this lead to a lack of stability in training, and the quality of the generated samples fluctuate with time.