

ECS795P Deep Learning and Computer Vision, 2024

Coursework 2: Diffusion Models

1. What is the difference between an auto-encoder, a generative adversarial network (GAN), and a diffusion model? Consider: (1) model structure; (2) optimized objective functions, and (3) how different components of each models are trained. **(9 % of CW2)**
2. The UNet is one of the most important components of a diffusion system because it facilitates the actual diffusion process. Consider an unconditional diffusion model that has a UNet model with the following parameters:
 - Input size: $512 * 512$
 - Number of channels ($n_channel$): 128 (This is the number of channels in the initial feature map that we transform the image into)
 - Channel multipliers (ch_mults): 1, 2, 4 (This is the list of channel numbers at each resolution. The number of channels is $ch_mults[i] * n_channels$ at layer i)
 - Up/Down sampling factors: 2
 - Number of up/down blocks: 3

(1) Please write down the dimensions of the intermediate feature maps after each of the following UNet blocks in turn: the 3 downsample blocks, the 1 middle block, the 3 upsample blocks.

(2) Within UNet, there are attention modules and time-step embeddings. Briefly describe how they integrate with UNet.

Hint:

Input dimension of UNet: $[512, 512, 128]$

Output of the first Downsample Block : $[256, 256, 128]$

You can use the UNet image in the slides to help you solve the problem.

(14% of CW2)

3. If an image undergoes a forward process of noise addition and then a sampling process of denoising, (1) would the resultant image remain identical to the original? (2) Why or why not? (3) How would the results differ with more noising and denoising steps? **(5 % of CW2)**
4. Samplers are one of the key components in Diffusion models. Answer the following question about the sampler: (1) What sampler was used for training in this coursework? (2) What is the main difference between DDPM and DDIM (from the aspect of stochastic) and what are the benefits of DDIM over DDPM? **(7 % of CW2)**