AI 3603 Generator Model
Homework 5

Adhere to the Code of Academic Integrity. You may discuss background issues and general strategies with others and seek help from course staff, but the implementations that you submit must be your own. In particular, you may discuss general ideas with others but you may not work out the detailed solutions with others. It is never OK for you to see or hear another student's code and it is never OK to copy code from published/Internet sources. Moss (Measure Of Software Similarity) will be used for determining the similarity of programs to detect plagiarism in the class (`https://theory.stanford.edu/~aiken/moss/`) If you encounter some difficulties or feel that you cannot complete the assignment on you own, discuss with your classmates in Discussion forum on Canvas, or seek help from the course staff.

You are required to complete this homework **_individually_**. Please submit your assignment following the instructions summarized in Section 7.

1 Task 1: The random generation with diffusion

In the attachment, we provide a set of diffusion programs, including all parts of diffusion components, forward and reverse process, trainer and tester. In addition, we provide a cat face dataset. Use this dataset to train diffusion as a cat face image generator. After training, when the model is fed random noise, the model outputs cat face images. The FID is calculated between the randomly generated 1000 cat face images and the training dataset, which is used as the evaluation metric. The steps of the experiment are detailed as follows:

- Train the diffusion model with cat face dataset: Set the parameters of model and trainer in **main.py** and use the **Class trainer** to train the model and save checkpoints and results.

- Randomly generate 1000 cat face photos using the trained model: Use the **trainer.inference** to randomly generate cat faces.

- Calculate the FID metric. The process of FID is described as follows:

$$FID(x,g) = ||\mu_x - \mu_g||_2^2 + Tr(\Sigma_x + \Sigma_g - 2(\Sigma_x \Sigma_g)^{0.5}),$$

where x denotes real images, g denotes generated images, μ_x denotes mean values of real images, μ_g denotes mean values of generated images, Σ_x and Σ_g denote covariance matrix of real images and generated images. To calculate FID, we use the project (https://github.com/mseitzer/pytorch-fid). First, pip install pytorch-fid. Then use the command "pytorch-fid path/to/generation path/to/dataset". Check out the pytorch-fid project for more details.

- Try to adjust different parameters to make the randomly generated images of higher quality. Please analyze the impact of each parameter in the report. Please describe your idea that how these parameters are tuned.

python environment you need:
   pip install torch
   pip install torchvision
   pip install einops
   pip install transformers
   pip install ema_pytorch
   pip install accelerate
   pip install PIL
   pip install tqdm

Figure 1: random generation by diffusion

2 Task 2: Fusion generation with diffusion

Take two source images from the training dataset, map the two images into the latent space, sum the two mappings in the latent space, and then input to the trained diffusion to see whether the generated results have the characteristics of the source images, 50 pairs were selected for fusion generation, and the FID between source images and fusion images was taken as the evaluation metrics.

- Map selected source images to the latent space and sum them: Use the **trainer.inference2** for fusion generation.

- Generate the images using the summed latent maps: Use the **trainer.inference2** for fusion generation.

- Calculate the FID metric: Use the command "pytorch-fid path/to/fusion path/to/source"

Try to select different source images, adjust the fusion proportion and perparameters to make the fusion generation effect better.

Figure 2: fusion generation by diffusion

3 Task 3: Read and Analyze the Code

In this section, a complete DDPM implementation project is given for the generation task. You are required to read the code and understand the DDPM training process. You are required to write comments in the **model.py** to point out the function or your understanding of each part of the code. Please fill in every """"comments: """ with your understanding of the code.

4   Task 4: Use the styleGAN

Here is another cat body dataset, again using diffusion for training and random generation testing. Also here is a styleGAN project, please try random generation with this method again. Compare the results of the two, and consider why there is such a difference. Please analyze the differences between them by combining the principles of the paper.

- Train and test the diffusion model with cat body dataset and save the result.

- Train and test the styleGAN (https://github.com/lucidrains/stylegan2-pytorch) with cat body dataset and save the result.

- Calculate the FID metric: Use the command "pytorch-fid path/to/generation path/to/dataset" and discuss them.

Figure 3: random generation by styleGAN using cat body dataset

5   Online Resources

1. Cat face dataset(https://jbox.sjtu.edu.cn/l/A1kRcw): download Dataset.

2. Cat body dataset(https://jbox.sjtu.edu.cn/l/b1Fq10): download Dataset.

3. DDPM (https://github.com/lucidrains/denoising-diffusion-pytorch): the source project of we provide.

4. FID (https://github.com/mseitzer/pytorch-fid): use to calculate FID.

5. styleGAN (https://github.com/lucidrains/stylegan2-pytorch): a GAN project from NVIDIA.

6. HPC Computation: You can conduct experiments on SJTU HPC with given accounts.

Other generator model project are also acceptable for this homework.

6   Discussion and Question

You are encouraged to discuss your ideas, ask and answer questions about this homework. If you encounter any difficulty with the assignment, try to post your problem **on Canvas** for help. The classmates and the course staff will try to reply.

7   Submission Instructions

1. Zip all your files to a file named as **HW5_ID_name.zip**:

   The randomly generated results are stored in the **submission** folder.

   The fusion results are stored in the **source** and **fusion** folders respectively.

   The Report is saved in pdf format.

   Code with comments should also be included.

2. Upload the file to the homework 5 page on the Canvas.

3. **Deadline**: 2025.01.02