**LAB 1-DDPM**

**314552036 周子翔**

**Task1**

1. **Complete and Explain of all #TODO code** 
   1. **Todo1**

A screenshot of a computer program

AI-generated content may be incorrect.

* It stacks several TimeLinear layers, which are linear layers modulated by a time embedding (so the network can use timestep information).
* Each hidden layer’s size is specified by dim\_hids.
* The final output layer maps the last hidden features to the output dimension (usually the same as the input, for noise prediction).

A screen shot of a computer code

AI-generated content may be incorrect.

* The input [x](vscode-file://vscode-app/private/var/folders/4h/tgp5rzwn16g42y719p2gqx4m0000gn/T/AppTranslocation/284880D8-3402-4D36-9D71-2D5213D1D17B/d/Visual%20Studio%20Code.app/Contents/Resources/app/out/vs/code/electron-browser/workbench/workbench.html) (noisy data) is passed through each layer, with the current timestep [t](vscode-file://vscode-app/private/var/folders/4h/tgp5rzwn16g42y719p2gqx4m0000gn/T/AppTranslocation/284880D8-3402-4D36-9D71-2D5213D1D17B/d/Visual%20Studio%20Code.app/Contents/Resources/app/out/vs/code/electron-browser/workbench/workbench.html) provided to each TimeLinear layer.
* After each layer, a SiLU activation is applied for non-linearity.
* The final output is produced by the last linear layer, which predicts the noise for the given input and timestep.
  1. **Todo2**

A screen shot of a computer code

AI-generated content may be incorrect.

* For a given clean image x0 and timestep [t](vscode-file://vscode-app/private/var/folders/4h/tgp5rzwn16g42y719p2gqx4m0000gn/T/AppTranslocation/284880D8-3402-4D36-9D71-2D5213D1D17B/d/Visual%20Studio%20Code.app/Contents/Resources/app/out/vs/code/electron-browser/workbench/workbench.html), it adds noise according to the DDPM formula.
* alphas\_prod\_t is the cumulative product of alphas up to timestep [t](vscode-file://vscode-app/private/var/folders/4h/tgp5rzwn16g42y719p2gqx4m0000gn/T/AppTranslocation/284880D8-3402-4D36-9D71-2D5213D1D17B/d/Visual%20Studio%20Code.app/Contents/Resources/app/out/vs/code/electron-browser/workbench/workbench.html).
* The output xt is a noisy version of x0 at timestep [t](vscode-file://vscode-app/private/var/folders/4h/tgp5rzwn16g42y719p2gqx4m0000gn/T/AppTranslocation/284880D8-3402-4D36-9D71-2D5213D1D17B/d/Visual%20Studio%20Code.app/Contents/Resources/app/out/vs/code/electron-browser/workbench/workbench.html).
  1. **Todo 3**

A screen shot of a computer program

AI-generated content may be incorrect.

* Predicts the noise in xt using the network.
* Calculates the mean and variance for the posterior distribution at timestep [t](vscode-file://vscode-app/private/var/folders/4h/tgp5rzwn16g42y719p2gqx4m0000gn/T/AppTranslocation/284880D8-3402-4D36-9D71-2D5213D1D17B/d/Visual%20Studio%20Code.app/Contents/Resources/app/out/vs/code/electron-browser/workbench/workbench.html).
* Samples the previous timestep x\_{t-1} using the mean and variance, adding noise unless at the final step.
  1. **Todo 4**

A screen shot of a computer code

AI-generated content may be incorrect.

* Starts from pure noise.
* Iteratively denoises through all timesteps using p\_sample.
* Returns the final denoised sample.
  1. **Todo 5**

A screen shot of a computer program

AI-generated content may be incorrect.

* Randomly selects a timestep for each sample in the batch.
* Adds noise to the clean data using q\_sample.
* Predicts the noise using the network.
* Computes the mean squared error between predicted and true noise.

1. **Fig of your implementation of q\_sample (10pts)**

A group of black and white objects

AI-generated content may be incorrect.

1. **Fig of loss curve and evaluation result (10 pts)**

A diagram of a graph

AI-generated content may be incorrect.A diagram of blue dots

AI-generated content may be incorrect.A blue dots in a circle

AI-generated content may be incorrect.A graph with blue lines

AI-generated content may be incorrect.A diagram of a spiral

AI-generated content may be incorrect.

The DDPM implementation successfully learns the 2D swiss roll distribution, as shown by the sample plot and a reasonable Chamfer Distance.

**Task 2**

1. **Complete and Explain of all #TODO code.**
   1. **Todo 1**

A screenshot of a computer

AI-generated content may be incorrect.

* Adds noise to the clean image according to the DDPM formula for a given timestep.
  1. **Todo 4**

A computer screen shot of a program

AI-generated content may be incorrect.

* It creates a smooth schedule for the cumulative product of alphas using a cosine function.
* Betas are then derived from the ratio of consecutive alphā values.
* This schedule can improve sample quality compared to linear/quadratic schedules.
  1. **Todo 5**

A computer screen with text

AI-generated content may be incorrect.

A computer screen shot of a program code

AI-generated content may be incorrect.

* Extracts scheduler parameters for the current timestep.
* Calculates the mean and variance for the reverse process.
* Adds noise unless at the final step.
* Returns the denoised sample.

A computer code with many colorful text

AI-generated content may be incorrect.

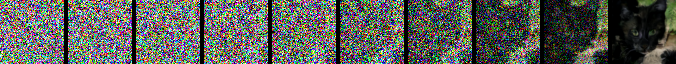
* Uses the predicted clean image to compute the mean for the reverse process.
* Adds noise as in DDPM.

A screen shot of a computer code

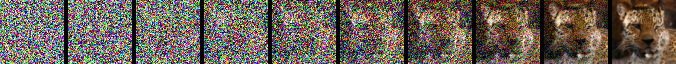
AI-generated content may be incorrect.

* Implements the reverse step when the network directly predicts the posterior mean.

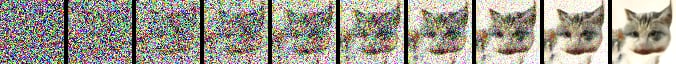
1. **Show the trajectory fig and compare results across the three beta schedulers.**
   1. **Linear**

****

* 1. **Quadratic**

****

* 1. **Cosine**

****

1. **Show the results of different predictors and discuss.**
   1. **Noise**
   2. **x₀**

**A black and white image of a person's face

AI-generated content may be incorrect.A close up of a baby leopard

AI-generated content may be incorrect.A blurry image of a dog

AI-generated content may be incorrect.A close up of a tiger

AI-generated content may be incorrect.A cat with a hand on its head

AI-generated content may be incorrect.A close up of a tiger's face

AI-generated content may be incorrect.A close up of a cat

AI-generated content may be incorrect.A blurry image of a leopard

AI-generated content may be incorrect.**

* 1. **posterior mean**

**A square with two dots

AI-generated content may be incorrect.A square with a white square in the middle

AI-generated content may be incorrect.A blurry image of a pink square

AI-generated content may be incorrect.A square with a white square with a red and blue border

AI-generated content may be incorrect.A white square with colorful dots

AI-generated content may be incorrect.A white square with three dots

AI-generated content may be incorrect.A white square with a red border

AI-generated content may be incorrect.A square with a few dots

AI-generated content may be incorrect.**

1. **Screenshot of the Best FID of your training result, explain the training setting.**

**The best FID score is 28.7 with linear schedulers and x0 predictors**

A black numbers on a white background

AI-generated content may be incorrect.

* **Beta Schedule:** Linear (--mode linear)
* **Predictor:** x₀ (--predictor x0)
* **Batch Size:** 16 (--batch\_size 16)
* **Image Resolution:** 64x64 (--image\_resolution 64)
* **Number of Training Steps:** 50,000 (--train\_num\_steps 50000)
* **Diffusion Timesteps:** 1,000 (--num\_diffusion\_train\_timesteps 1000)
* **Beta Range:** β₁ = 1e-4, β\_T = 0.02 (--beta\_1 1e-4 --beta\_T 0.02)
* **Seed:** 63 (--seed 63)
* **Optimizer:** Adam, learning rate 2e-4
* **Scheduler:** LambdaLR, warmup steps 200
* **Dataset:** AFHQ, max 3,000 images per category
* **Checkpoint/Results Directory:** Saved to Google Drive (/content/drive/MyDrive/results/predictor\_x0/beta\_linear/...)
* **Sampling Frequency:** Every 2,000 steps, model samples and saves checkpoint