Reconstruct an image from the Gaussian noise by the Diffusion model

Kuang Wenwei, Wei Tianqi, Chen Siwei, Li Dongheng

01 MOTIVATION

What is Denoising Diffusion Probabilistic Model (DDPM)?

香 港 大 學 THE UNIVERSITY OF HONG KONG

- The diffusion probability model is a **Markov chain-based probability model** that splits the mapping relationship between the noise and the target waveform into T steps, forming a Markov chain.

- We train for the diffusion process of this chain (from the target audio to the noise) and then decode it by the **inverse process** (from the noise to the target audio)

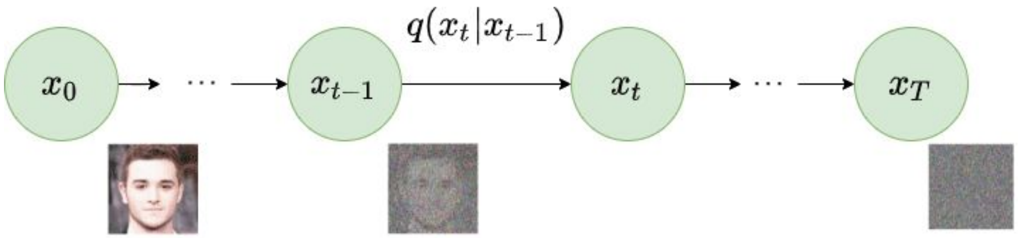
Why not use GAN?

- Using GAN to train networks is not easy to **converge**
- The **diversity** of GAN model is relatively poor
- The Diffusion model explains in a simpler way how the model should be learned and produced

02

Model Training and Sampling

Forward Process



1. Mathematical Formula for future training process

$$\alpha_t = 1 - \beta_t$$

$$x_t = \sqrt{\alpha_t} x_{t-1} + \sqrt{1 - \alpha_t} z_1$$

$$x_{t-1} = \sqrt{\alpha_{t-1}} x_{t-2} + \sqrt{1 - \alpha_{t-1}} z_2$$

$$\rightarrow x_t = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} z_t$$

2. Save noise as labels for training process

During the forward process, we can acquire noise Zt at each step that follows a Gaussian Distribution. We save the Zt’s value at each step for the purpose of loss calculation during the training process.

Training and Sampling (Reverse Process)

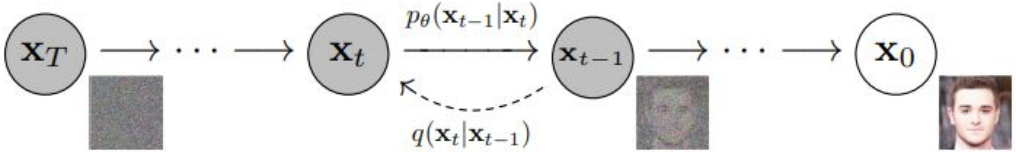


Figure 2: The directed graphical model considered in this work.

Use Xt to directly acquire X0? the performance is really bad.
-> we use iteration to get Xt-1 using Xt, Xt-2 using Xt-1… until we get X0.

$$\mathrm{q}(x_{t-1}|x_t)$$

$$Forward\ Process : \mathrm{q}(x_t|x_{t-1})$$

$$We\ use\ Bayes\ Theorem : \mathrm{q}(x_{t-1}|x_t, x_0) = \mathrm{q}(x_t|x_{t-1}, x_0)\frac{\mathrm{q}(x_{t-1}|x_0)}{\mathrm{q}(x_t|x_0)}$$

$$what\ we\ get : \mu_t = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-\overline{\alpha_t}}}z_t\right),\ \sigma^2\ is\ fixed$$

Algorithm 1 Training

1: **repeat**
2: $\quad \mathbf{x}_0 \sim q(\mathbf{x}_0)$
3: $\quad t \sim \mathrm{Uniform}(\{1, \ldots, T\})$
4: $\quad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
5: $\quad$ Take gradient descent step on
$$\nabla_\theta \left\| \boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, t) \right\|^2$$
6: **until** converged

Algorithm 2 Sampling

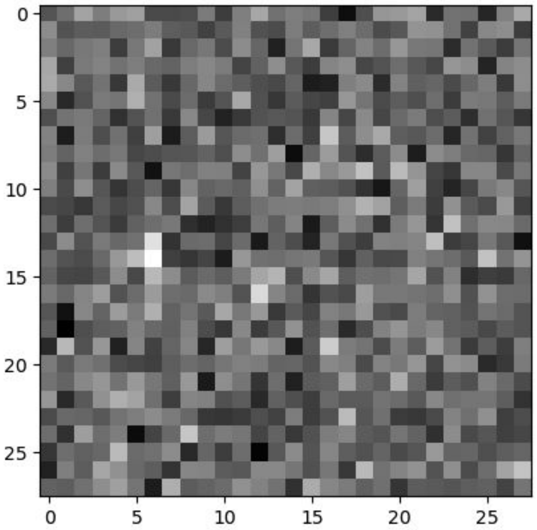
1: $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
2: **for** $t = T, \ldots, 1$ **do**
3: $\quad \mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$, else $\mathbf{z} = \mathbf{0}$
4: $\quad \mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\right) + \sigma_t \mathbf{z}$
5: **end for**
6: **return** $\mathbf{x}_0$

03 Experiments

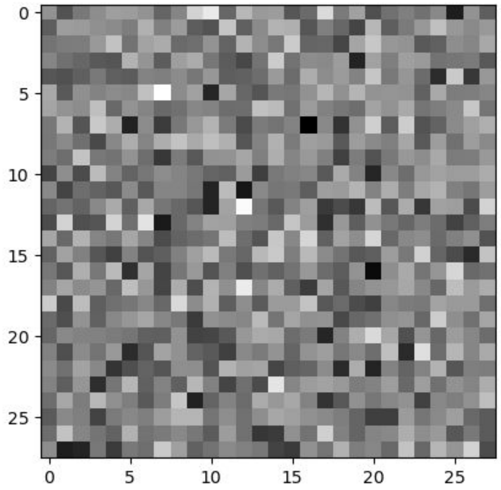
Reconstruct an image from a noise matrix

```
img = torch.randn((1,1,28,28),device=device)
```

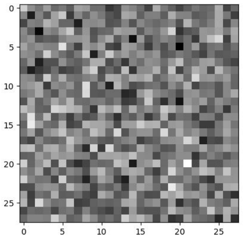
$$\text{out}_i \sim \mathcal{N}(0, 1)$$

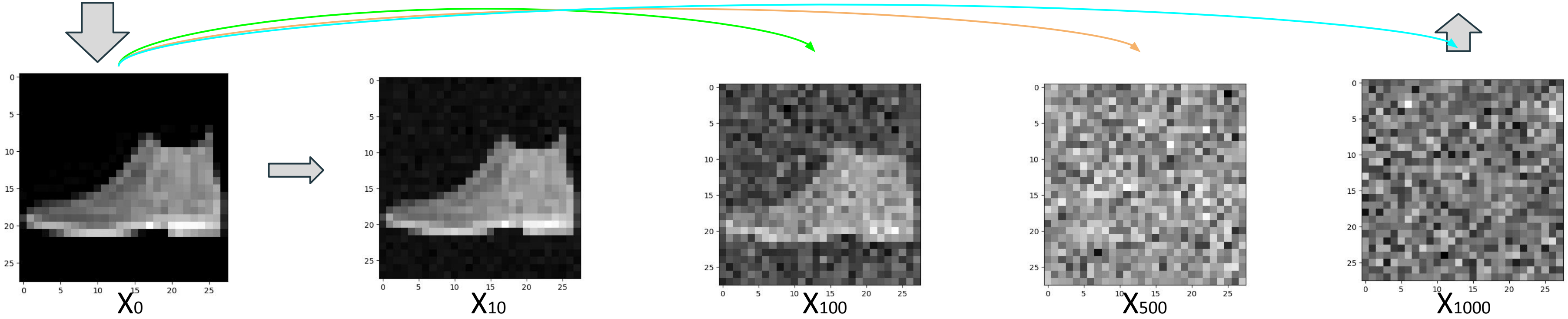


Sampling



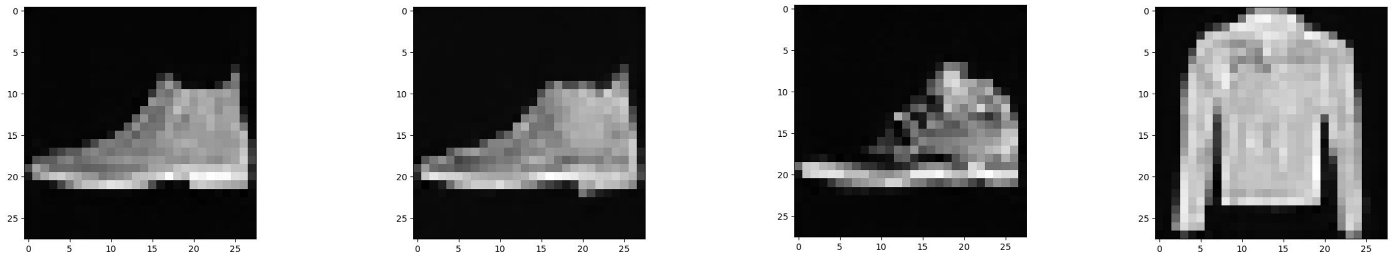
Reconstruct an image from an image





$D_{KL}(p(x_0)||P(X\sim N(0,1))=1.8643$ $D_{KL}(p(x_{10})||P(X\sim N(0,1))=1.542$ $D_{KL}(p(x_{100})||P(X\sim N(0,1))=0.501$ $D_{KL}(p(x_{500})||P(X\sim N(0,1))=0.0993$ $D_{KL}(p(x_{1000})||P(X\sim N(0,1))=0.00101$

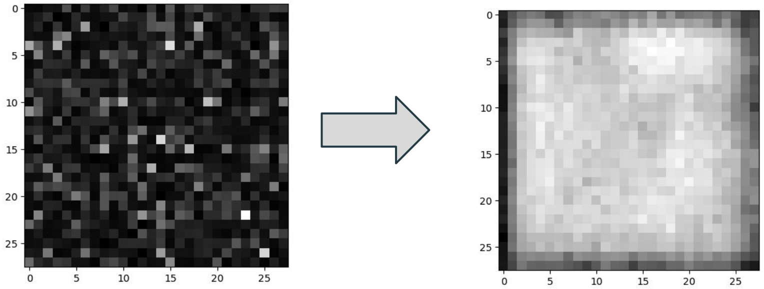
Reconstruct:



Reconstruct a noise matrix which != N(0,1)

img = -torch.log(torch.rand((1,1,28,28))) / 0.5 **(Exponential distribution)**

$D_{KL}(p(img)||P(X\sim N(0,1))=3.7697$



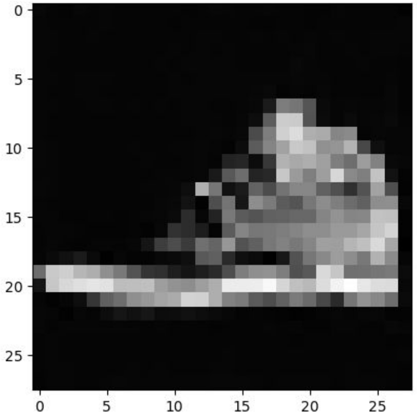
https://www.kaggle.com/code/wenweikuang/ddpm-example

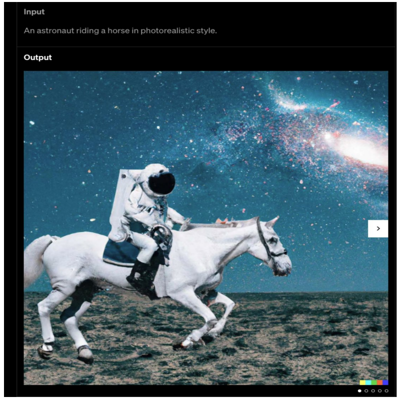
04

Limitation and Futures

Limitations

1. The sample is now done on the image pixel at high latitude.
2. The results are uncontrollable.





Thank you!

Boost speed

With the increased computational speed of diffusion models for image denoising and recovery, models can be processed and optimised in real time.

Research into how to speed things up is the way forward.