1-1 大纲

概述

① 扩散模型是什么

② 扩散模型的大致流程

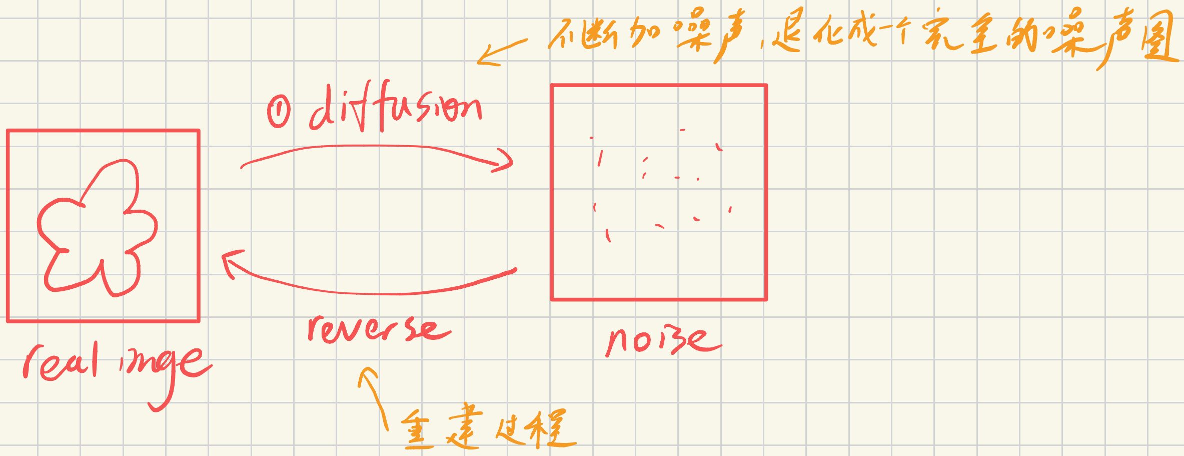
③ 扩散模型的优势

里程碑工作(公认较为简单易懂)

DDPM: Denosing Diffusion Probabilistic Models (NIPS'20)

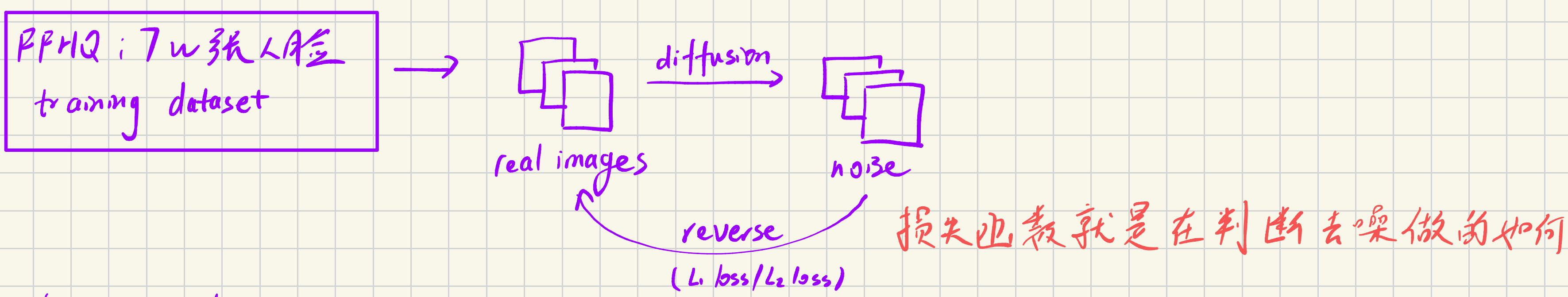
是什么 扩散模型 → 本质上是一个生成模型

流程

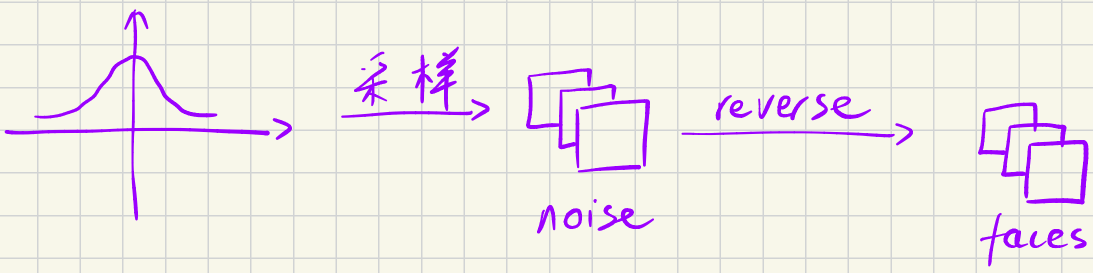


训练和测试

① training phase

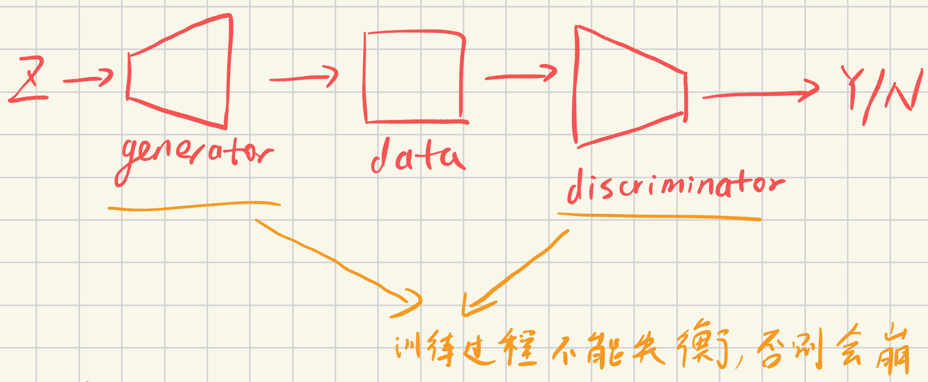


② inference phase



扩散模型比GAN强在哪?

训练 ①生成器
②判别器

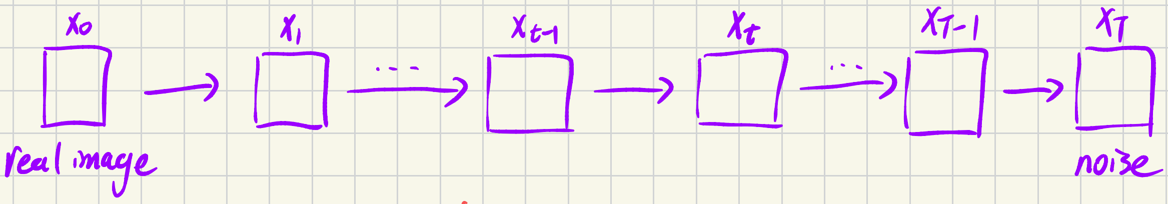


GAN是generator和discriminator对抗的过程，损失函数无法反映训练的阶段

但是diffusion没有这个问题，因为它是训练去噪的过程，损失函数反映了去噪的好坏

1-2 扩散阶段

$\begin{cases} \text{Diffusion 扩散阶段} \\ \text{Reverse} \end{cases}$



$x_t = f(x_{t-1})$?

$x_t = \sqrt{1-\beta_t}\, x_{t-1} + \sqrt{\beta_t}\, z_t \ , \quad z_t \sim N(0,1)$

加权加噪声

权重由β_t指定，随着加噪的进行β_t要逐渐变大

? 因为干净的图上有一点噪声就会很明显

随着噪声越多，需要更多的噪声来体现加噪

要求 β_t要足够大，能够使得加噪后变成一个彻底的噪声图，不能看出有特征轮廓

$\beta: 10^{-4} \to 2\times10^{-2}$ 线性变化 $T=2000$

? 太消耗时间，能不能一步得到扩散图

$x_t = f(x_{t-1}) \quad x_t = g(x_0)$

$1-\beta_t = \alpha_t \quad \begin{cases} x_t = \sqrt{\alpha_t}\, x_{t-1} + \sqrt{1-\alpha_t}\, z_t \\ x_t = \sqrt{\alpha_{t-1}}\, x_{t-2} + \sqrt{1-\alpha_{t-1}}\, z_{t-1} \end{cases}$

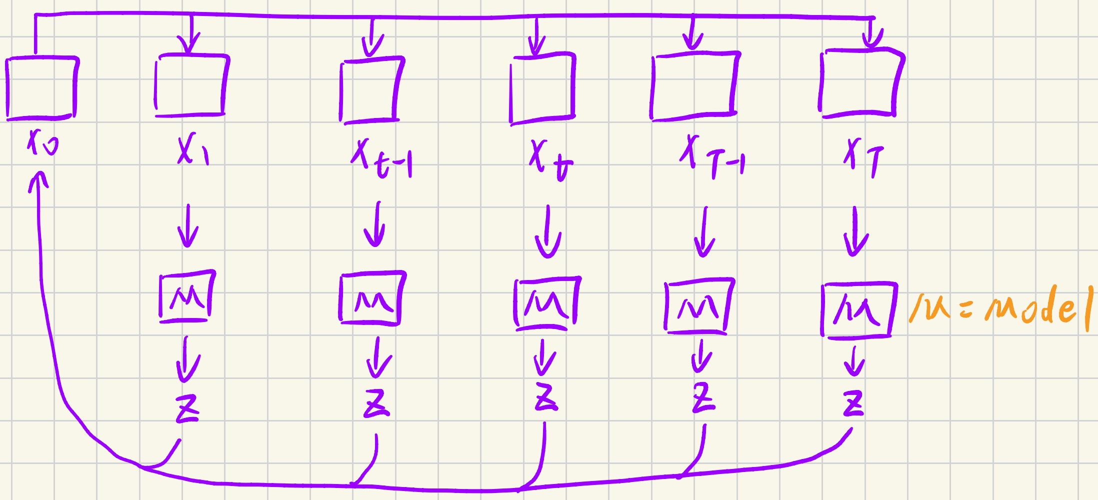
$\Rightarrow x_t = \sqrt{\alpha_t}\left(\sqrt{\alpha_{t-1}}\, x_{t-2} + \sqrt{1-\alpha_{t-1}}\, z_{t-1}\right) + \sqrt{1-\alpha_t}\, z_t$

$= \sqrt{\alpha_t\alpha_{t-1}}\, x_{t-2} + \underbrace{\underbrace{\sqrt{\alpha_t - \alpha_t\alpha_{t-1}}\, z_{t-1}}_{\sim N(0,\alpha_t-\alpha_t\alpha_{t-1})} + \underbrace{\sqrt{1-\alpha_t}\, z_t}_{\sim N(0,1-\alpha_t)}}_{\sim N(0,1-\alpha_t\alpha_{t-1})} \qquad z_t \sim N(0,1) \quad z_{t-1} \sim N(0,1)$

$= \sqrt{\alpha_t\alpha_{t-1}}\, x_{t-2} + \sqrt{1-\alpha_t\alpha_{t-1}}\, z \ , \quad z \sim N(0,1)$

$x_t = \sqrt{\alpha_t \cdots \alpha_1}\, x_0 + \sqrt{1-\alpha_t \cdots \alpha_1}\, z, \quad z \sim N(0,1)$

$\boxed{= \sqrt{\bar\alpha_t}\, x_0 + \sqrt{1-\bar\alpha_t}\, z} \qquad \bar\alpha_t = \prod_{i=1}^{T} \alpha_i$



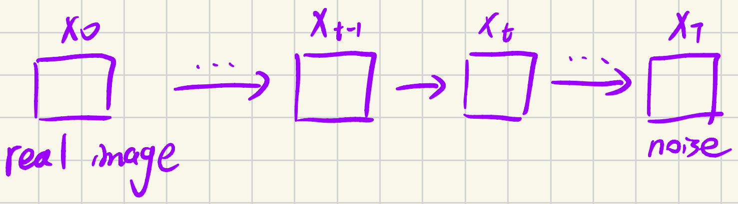
$x_T = \sqrt{\bar\alpha_T}\, x_0 + \sqrt{1-\bar\alpha_T}\, z \qquad x_T \approx z$

$\sqrt{\bar\alpha_T} \approx 0 \quad \bar\alpha_T \approx 0$

重建的过程是一步一步推出来的

1-3 重建阶段
1-4

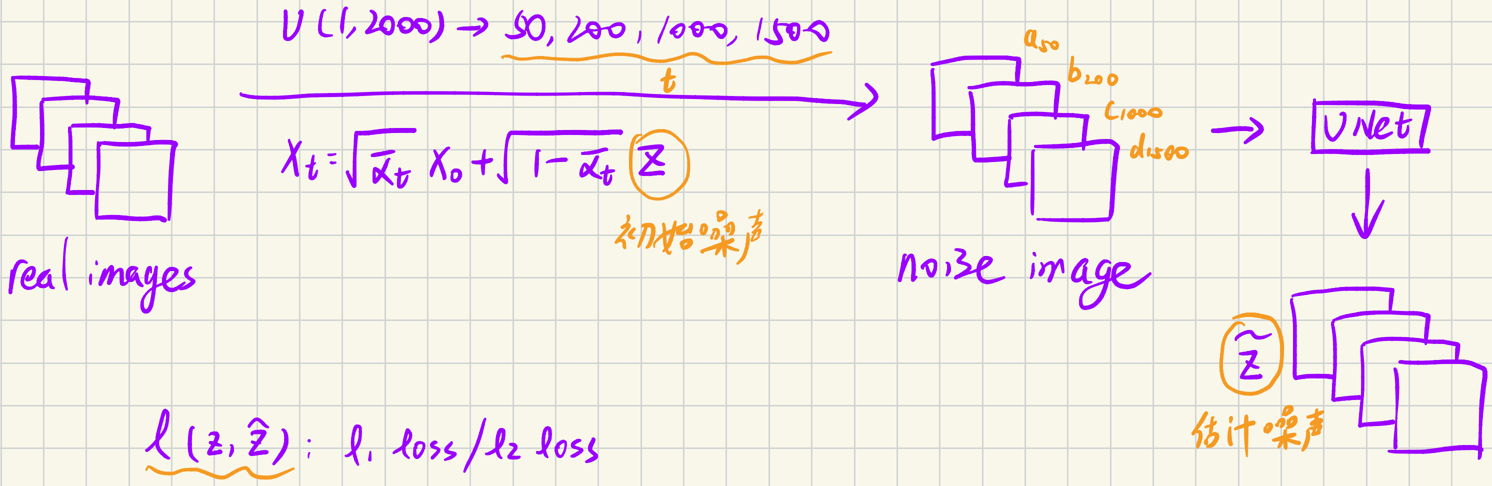
- Diffusion
- Reverse



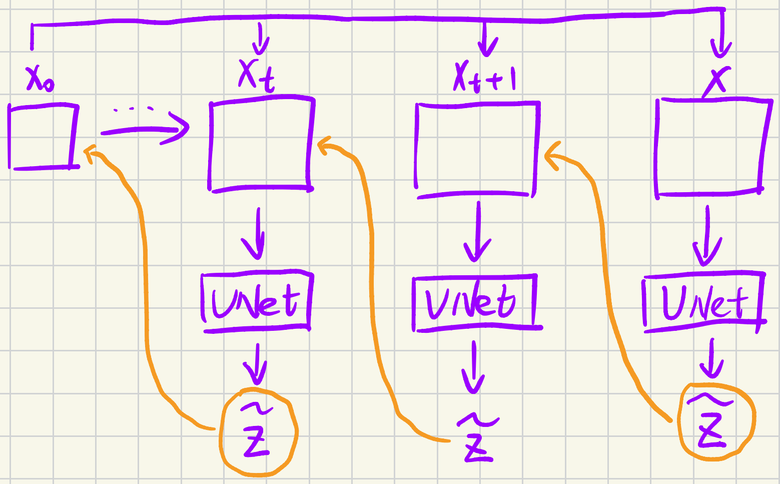
$$x_t = \sqrt{\alpha_t}\, x_{t-1} + \sqrt{1-\alpha_t}\, z_t$$

$$x_t = \sqrt{\bar{\alpha}_t}\, x_0 + \sqrt{1-\bar{\alpha}_t}\, z$$

batch size = 4 T=2000



$\ell(z, \tilde{z})$: l_1 loss / l_2 loss



$$x_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(x_t - \sqrt{1-\bar{\alpha}_t}\, \tilde{z}\right)$$

这样直接导回去的结果是模糊的，最好一步一步推上去

$\tilde{z} = \mathrm{UNet}(x_t)$ $\boxed{x_{t-1} = f(x_t, \tilde{z}) \quad ?}$ ☆

$\mathrm{UNet}(x_t, t)$

条件去噪 β强度不同

$$q(x_{t-1} \mid x_t) = \frac{q(x_t, x_{t-1})}{q(x_t)} = \boxed{\frac{q(x_t \mid x_{t-1})\, q(x_{t-1})}{q(x_t)}}$$

$x_{t-1} \to x_t$

$x_0 \to x_t$

$$x_t = \sqrt{\alpha_t}\, x_{t-1} + \sqrt{1-\alpha_t}\, z_t \sim N\left(\sqrt{\alpha_t}\, x_{t-1}, (1-\alpha_t) I\right)$$

$$x_t = \sqrt{\bar{\alpha}_t}\, x_0 + \sqrt{1-\bar{\alpha}_t}\, z \sim N\left(\sqrt{\bar{\alpha}_t}\, x_0, (1-\bar{\alpha}_t) I\right)$$

$$q(x_t \mid x_{t-1}) \sim N\left(\sqrt{\alpha_t}\, x_{t-1}, (1-\alpha_t) I\right)$$

$$q(x_t) \sim N\left(\sqrt{\bar{\alpha}_t}\, x_0, (1-\bar{\alpha}_t) I\right)$$

$$q(x_{t-1}) \sim N\left(\sqrt{\bar{\alpha}_{t-1}}\, x_0, (1-\bar{\alpha}_{t-1}) I\right) \qquad N(\mu, \sigma^2) \propto \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right)$$

$$\frac{q(x_t \mid x_{t-1})\, q(x_{t-1})}{q(x_t)} = q(x_{t-1} \mid x_t)$$

$$\propto \exp\left\{-\frac{1}{2}\left(\frac{(x_t - \sqrt{\alpha_t}\, x_{t-1})^2}{1-\alpha_t} + \frac{(x_{t-1} - \sqrt{\bar{\alpha}_{t-1}}\, x_0)^2}{1-\bar{\alpha}_{t-1}} - \frac{(x_t - \sqrt{\bar{\alpha}_t}\, x_0)^2}{1-\bar{\alpha}_t}\right)\right\} \qquad \alpha_t + \beta_t = 1$$

$$\propto \exp\left\{-\frac{1}{2}\left(\underbrace{\left(\frac{\alpha_t}{\beta_t} + \frac{1}{1-\bar{\alpha}_{t-1}}\right)}_{A} x_{t-1}^2 - 2\underbrace{\left(\frac{\sqrt{\alpha_t}\, x_t}{\beta_t} + \frac{\sqrt{\bar{\alpha}_{t-1}}\, x_0}{1-\bar{\alpha}_{t-1}}\right)}_{B} x_{t-1} + \underbrace{?}_{C}\right)\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left(A x_{t-1}^2 + B x_{t-1} + C\right)\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\cdot A\left(x_{t-1} + \frac{B}{2A}\right)^2 + ?\right\}$$

$$\mu = -\frac{B}{2A} \quad \sigma^2 = \frac{1}{A} \quad A = \frac{\alpha_t}{\beta_t} + \frac{1}{1-\bar\alpha_{t-1}} \quad B = -2\left(\frac{\sqrt{\alpha_t}\, x_t}{\beta_t} + \frac{\sqrt{\bar\alpha_{t-1}}\, x_0}{1-\bar\alpha_{t-1}}\right)$$

$$\sigma^2 = \frac{1}{A} = \frac{1}{\frac{\alpha_t}{\beta_t} + \frac{1}{1-\bar\alpha_{t-1}}} = \frac{1}{\frac{\alpha_t - \alpha_t\cdot\bar\alpha_{t-1} + \beta_t}{\beta_t(1-\bar\alpha_{t-1})}} \qquad \frac{\alpha_t + \beta_t = 1}{\alpha_t\cdot\bar\alpha_{t-1} = \bar\alpha_t} \qquad \boxed{\frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}\cdot\beta_t}$$

$$\mu = -\frac{B}{2A} = \left(\frac{\sqrt{\alpha_t}\, x_t}{\beta_t} + \frac{\sqrt{\bar\alpha_{t-1}}\, x_0}{1-\bar\alpha_{t-1}}\right)\cdot\frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}\cdot\beta_t$$

$$\boxed{= \sqrt{\alpha_t}\cdot\frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}\, x_t + \frac{\sqrt{\bar\alpha_{t-1}}\,\beta_t}{1-\bar\alpha_t}\, x_0}$$

?

$$x_t = \sqrt{\bar\alpha_t}\, x_0 + \sqrt{1-\bar\alpha_t}\, z$$

$$\boxed{x_0 = \frac{1}{\sqrt{\bar\alpha_t}}\left(x_t - \sqrt{1-\bar\alpha_t}\,\tilde z\right)}$$

$$= \sqrt{\alpha_t}\,\frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}\, x_t + \frac{\sqrt{\bar\alpha_{t-1}}\cdot\beta_t}{1-\bar\alpha_t}\cdot\frac{1}{\sqrt{\bar\alpha_t}}\left(x_t - \sqrt{1-\bar\alpha_t}\,\tilde z\right)$$

$$= \frac{x_t}{\sqrt{\alpha_t}}\left(\frac{\alpha_t - \bar\alpha_t + \beta_t}{1-\bar\alpha_t}\right) - \frac{\tilde z}{\sqrt{\alpha_t}}\left(\frac{\beta_t}{\sqrt{1-\bar\alpha_t}}\right)$$

$$\boxed{= \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-\bar\alpha_t}}\,\tilde z\right)}$$

$$q(x_{t-1}\mid x_t) \sim \mathcal{N}\left(\frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-\bar\alpha_t}}\,\tilde z\right),\ \frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}\cdot\beta_t\right)$$

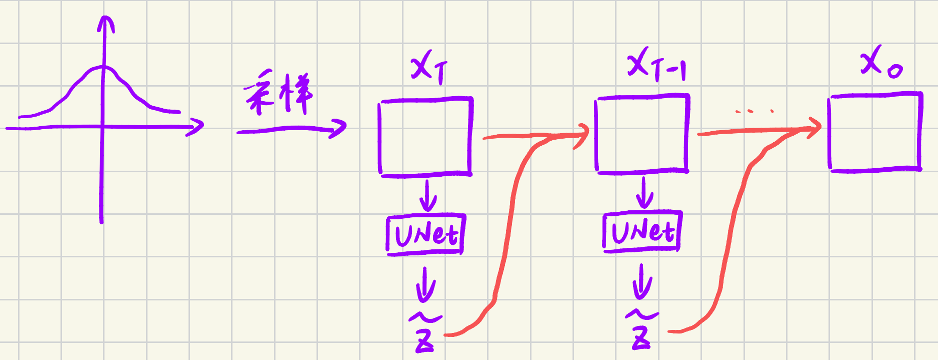
$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-\bar\alpha_t}}\,\tilde z\right) + \sqrt{\frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}\cdot\beta_t}\; z$$

写法参考: $x_t = \sqrt{\bar\alpha_t}\, x_0 + \sqrt{1-\bar\alpha_t}\, z \sim N(\sqrt{\bar\alpha_t}\, x_0, (1-\bar\alpha_t) I)$

$\tilde z = \text{UNet}(x_t, t)$

$z \sim N(0, I)$

从标准正态分布里随机采样的



? 为什么还要加新的噪声　　为了模仿布朗运动，有确定的部分，也有不确定的部分

1-5 整理总结

① Diffusion

$$x_t \sim N(\sqrt{\alpha_t}\, x_{t-1}, (1-\alpha_t) I)$$

$$x_t \sim N(\sqrt{\bar\alpha_t}\, x_0, (1-\bar\alpha_t) I)$$

$$\boxed{x_t = \sqrt{\alpha_t}\, x_{t-1} + \sqrt{1-\alpha_t}\, z_t}$$

α_t：预设好的 不可学习

z_t：正态分布中 采样的噪声

$$\boxed{x_t = \sqrt{\bar\alpha_t}\, x_0 + \sqrt{1-\bar\alpha_t}\, z}$$

一步一步退化效率低

$$\boxed{\begin{aligned} &x \sim N(\mu, \sigma^2 I) \\ &z \sim N(0, I) \\ &x = \mu + \sigma z \sim N(\mu, \sigma^2 I) \end{aligned}}$$

重参数化

② Reverse

$$x_{t-1} = f(x_t, \underbrace{UNet(x_t, t)}_{\hat z})$$

$$\hat z = UNet(x_t, t)$$

$$x_{t-1} \sim N(\mu, \sigma^2)$$

$$\sigma^2 = \boxed{\frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}\beta_t}$$

$$x_0 = \frac{1}{\sqrt{\bar\alpha_t}}(x_t - \sqrt{1-\bar\alpha_t}\,\hat z)$$

UNet估计出来的x_0

$$\mu = \sqrt{\alpha_t}\,\frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}\boxed{x_t} + \frac{\sqrt{\bar\alpha_{t-1}}\,\beta_t}{1-\bar\alpha_t}\boxed{x_0}$$

$$= \frac{1}{\sqrt{\alpha_t}}\left(\boxed{x_t} - \frac{\beta_t}{\sqrt{1-\bar\alpha_t}}\boxed{\hat z}\right)$$

确实是在一步步将噪声去掉

x_t的权重会越来越小

x_0的权重会越来越大

x_0也会越来越准

x_T x_{0T} $x_{0(T-1)}$ ⋯ x_{01}

权重下降

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-\bar\alpha_t}}\hat z\right) + \underbrace{\sqrt{\frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}\beta_t}}_{\text{方差上界}}\, z$$

Algorithm 1 Training

1: **repeat** 图像分布

2: 一些图像 $\mathbf{x}_0 \sim \boxed{q(\mathbf{x}_0)}$

3: $t \sim \text{Uniform}(\{1, \ldots, T\})$ 生成 t，为 $x_0 \to x_t$

4: $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 采样一个 z

5: Take gradient descent step on

$$\nabla_\theta \left\| \epsilon - \underbrace{\epsilon_\theta}_{\text{UNet}}(\underbrace{\sqrt{\bar\alpha_t}\mathbf{x}_0 + \sqrt{1-\bar\alpha_t}\epsilon}_{x_t}, t) \right\|^2$$

6: **until** converged

损失函数 训练UNet对z的估计能力

Algorithm 2 Sampling

1: $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$

2: **for** $t = T, \ldots, 1$ **do**

3: $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ $\boxed{\text{if } t > 1, \text{ else } \mathbf{z} = \mathbf{0}}$ 最终求干净的x_0时不加噪声

4: $\underbrace{\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar\alpha_t}}\underbrace{\epsilon_\theta(\mathbf{x}_t, t)}_{\text{估计噪声}}\right)}_{\mu_t} + \sigma_t \mathbf{z}$ 不一定和公式一样

5: **end for**

6: **return** $\mathbf{x}_0$

$$x_T \to x_{T-1} \to x_{T-2} \cdots \to x_0$$

μ 相当于是确定了路径的方向

σ 是扰动，扰动不是太大，也没问题