DDPM: Denoising Diffusion Probabilistic Model

LDM: Latent Diffusion Model

U-Net

前向扩散过程 (加噪)

反向扩散过程 (去噪).

CLIP

大纲

讲解 文本 → bitmap → SVG 技术路线

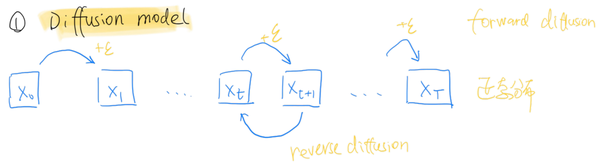
1、文生图 stable diffusion

2、bitmap → SVG

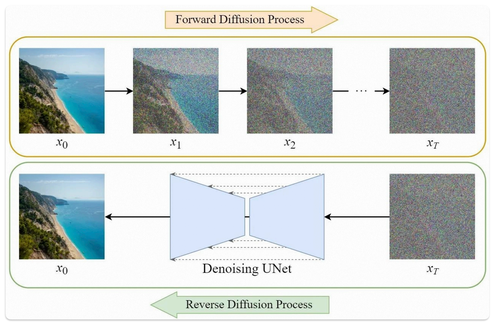
3、后续优化路径

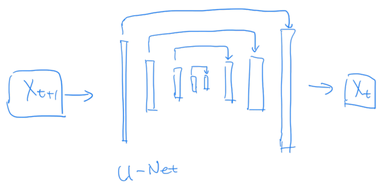
1、文生图 (SD)

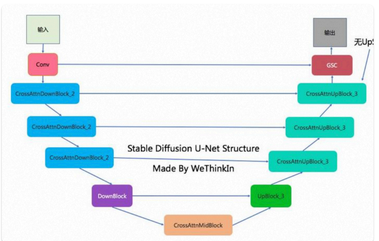
① Diffusion model



用同一模型迭代许多次 (denosing diffusion)

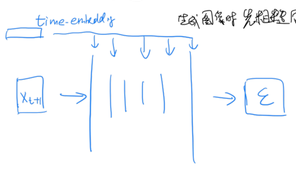






DDPM
- ① 预测ε ($x_t \xrightarrow{+\varepsilon} x_{t+1}$)
- ② time-embedding (类似transformer中位置编码)

生成图片的先粗略后精细



$$p(x_t \mid x_{t+1}) = \| \varepsilon - f_\theta(x_{t+1}, t) \|$$

classifier guided diffusion

$$x_t \xrightarrow{f} x_{t-1} \xrightarrow{f} x_{t-2} \cdots x_0$$

classifer → g (类似梯)

$$p(x_{t-1} \mid x_t) = \| \varepsilon - f_\theta(x_t, t, y) \|$$

classifier clip img text

② CLIP 模型

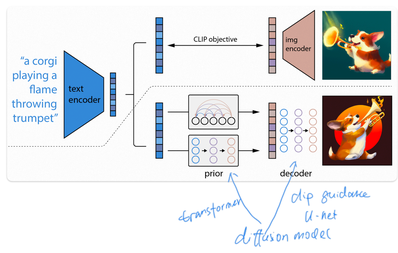
③ SD 模型

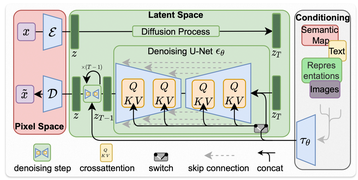
训练数据 (图像 (文本)) : (x, y)

clip → z_t clip → z_t

prior 模型 : $p(z_i \mid y)$

decoder 模型 : $p(x \mid z_i, y)$





2、bitmap → SVG

转换步骤:

① 颜色量化
- 减少颜色数量, 图像分割为单色图层
- 渐变

② 轮廓检测与路径生成
- 边缘检测: Canny 等
- 贝塞尔曲线拟合 路径简化

③ 矢量图形合成
- 坐标映射
- 颜色填充

3、后续优化路径

1、图像风格 prompt
- 抽象几何分
- 抽象类 图案域

```
prompt_prefix = "Simple, classic image of"
prompt_suffix = "with flat color blocks, beautiful, minimal details, solid colors only"
negative_prompt = "lines, framing, hatching, background, textures, patterns, details, outlines"
```

2、迭代
- ocr分
- 美学分
- vqa分 自动问题, 纠错问题

3. Lora

4、数据清洗