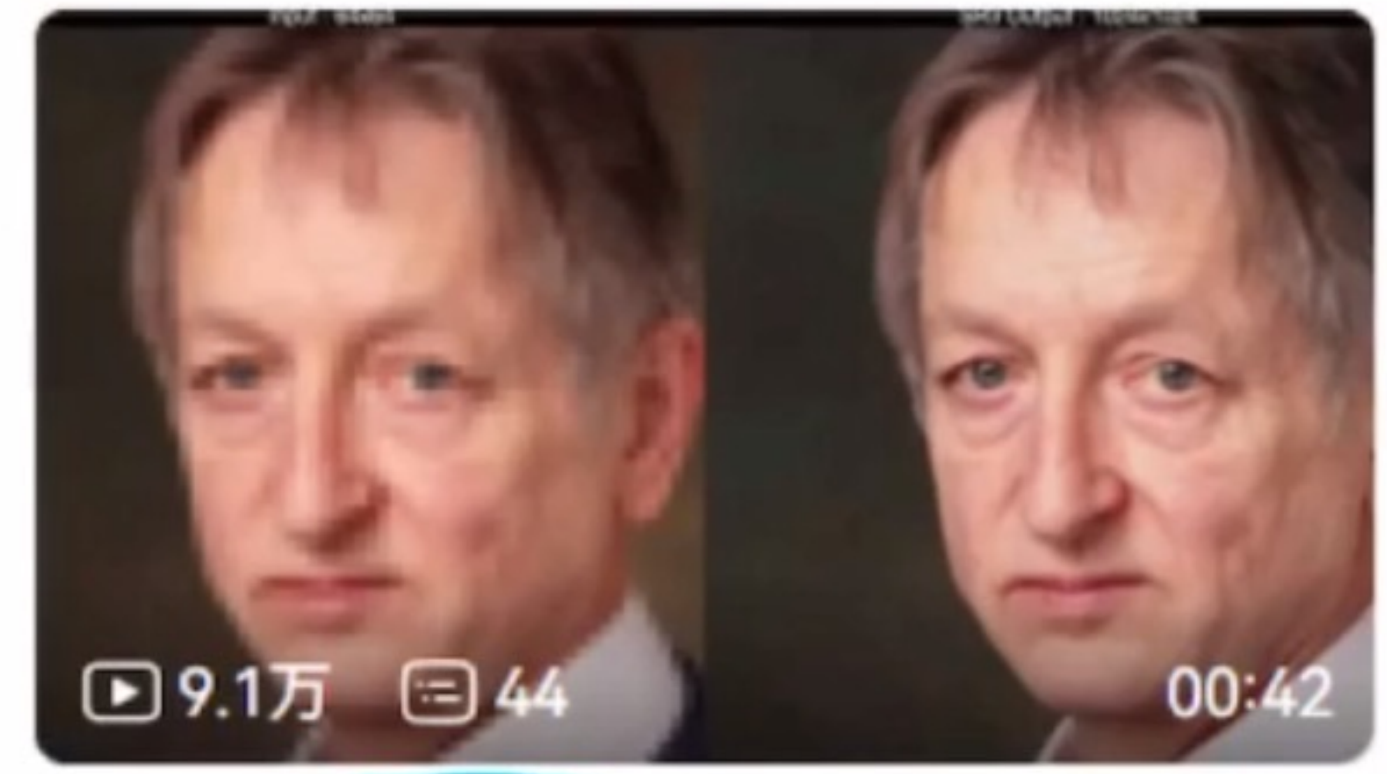
# 01 背景介绍

- 论文链接：https://arxiv.org/abs/2104.07636
- 代码链接：无官方代码
- 录用信息：TPAMI'22




谷歌提出：基于条件扩散模型的图像超分辨率！放大 16 倍，效果绝绝...

AI算法与图像处理 · 2021-10-21

- 论文标题截图：

**Image Super-Resolution via Iterative Refinement**

Chitwan Saharia[†], Jonathan Ho, William Chan, Tim Salimans, David J. Fleet, Mohammad Norouzi
{sahariac,jonathanho,williamchan,salimans,davidfleet,mnorouzi}@google.com
Google Research, Brain Team

# 02 论文摘要

- **提出问题：**
- 利用DDPM来做图像超分辨率

- **提出解决方案：**
- 将超分任务描述为一个有条件生成

- **优势&实验结果：**
- 实验表明该方案可以做图像超分，而且在某些指标上有相比基于GAN的方法的较好的性能

把超分辨率描述成一个生成问题

- 论文摘要截图：

名字里3R


## Abstract

*We present SR3, an approach to image Super-Resolution via Repeated Refinement. SR3 adapts denoising diffusion probabilistic models [17, 48] to conditional image generation and performs super-resolution through a stochastic iterative denoising process. Output generation starts with pure Gaussian noise and iteratively refines the noisy output using a U-Net model trained on denoising at various noise levels. SR3 exhibits strong performance on super-resolution tasks at different magnification factors, on faces and natural images. We conduct human evaluation on a standard 8× face super-resolution task on CelebA-HQ, comparing with SOTA GAN methods. SR3 achieves a fool rate close to 50%, suggesting photo-realistic outputs, while GANs do not exceed a fool rate of 34%. We further show the effectiveness of SR3 in cascaded image generation, where generative models are chained with super-resolution models, yielding a competitive FID score of 11.3 on ImageNet.*


基于DDPM

迭代

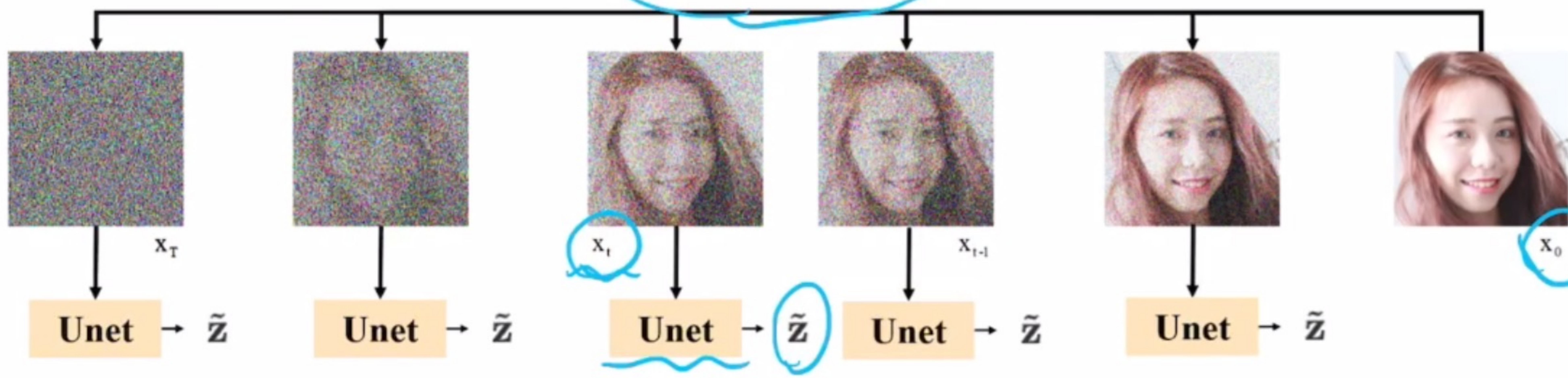
训练去噪的Net

性能好

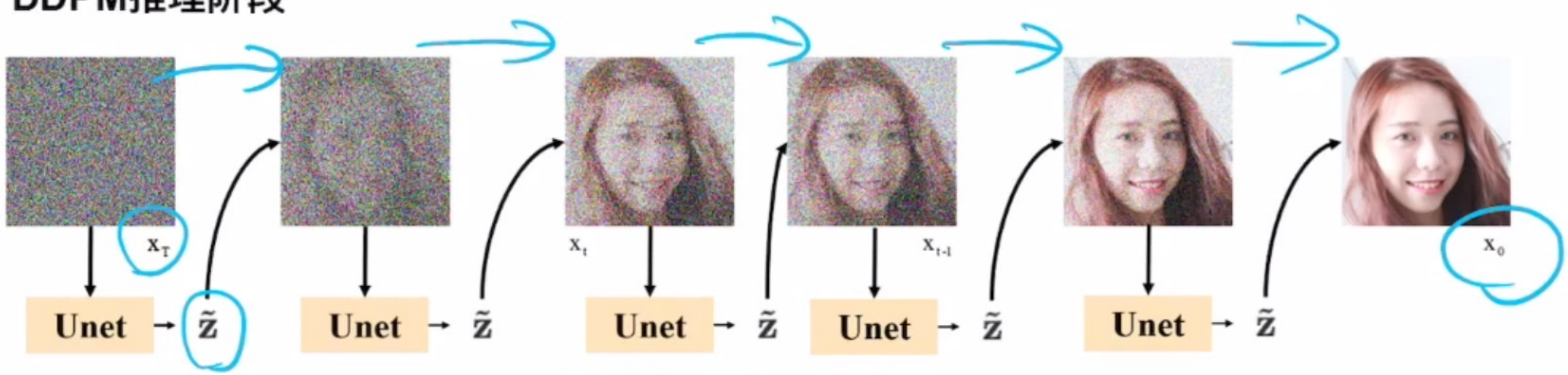
# 03 相关工作

## ➤ DDPM训练阶段

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\mathbf{z}$$



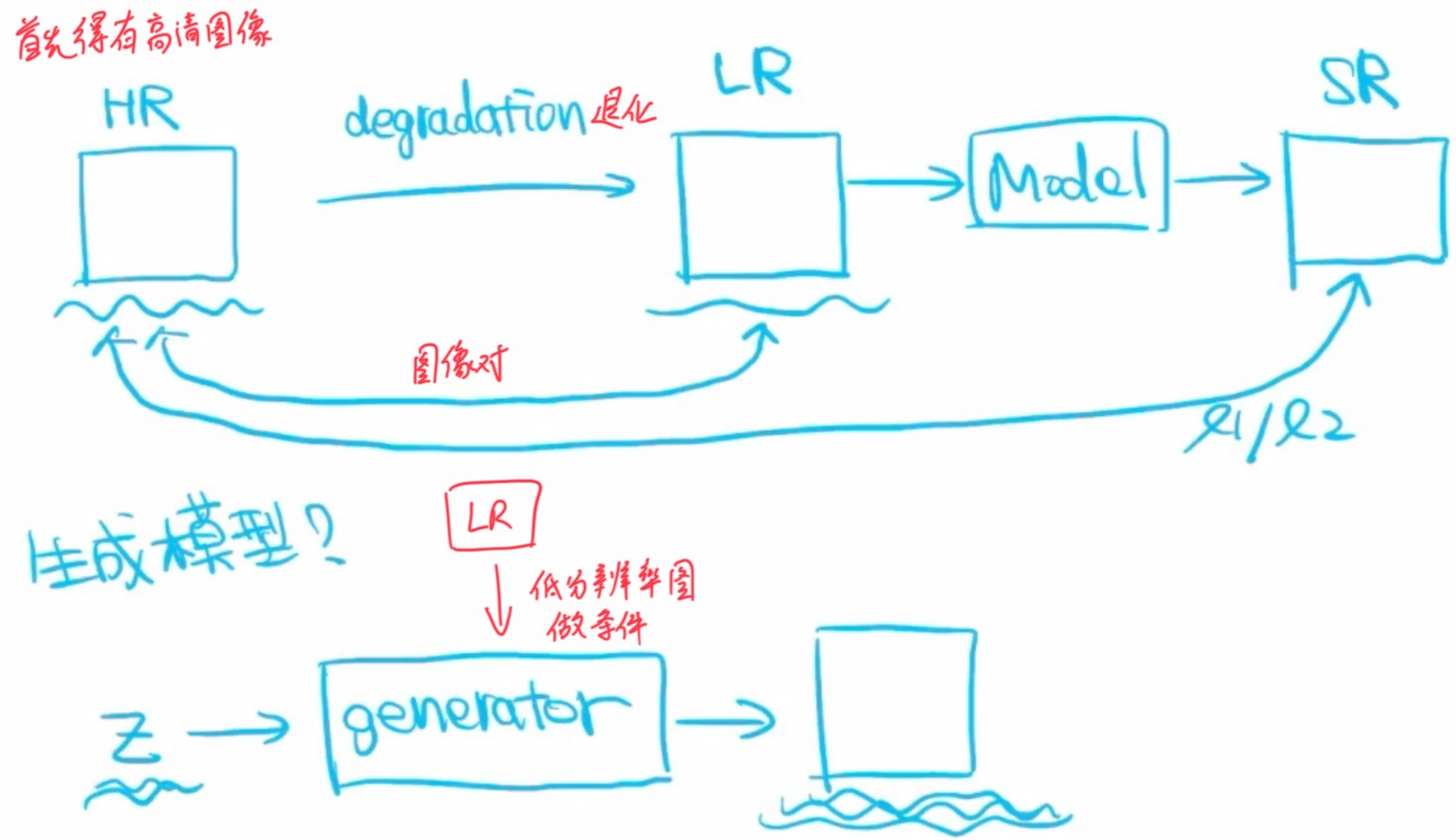

## ➤ DDPM推理阶段




$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\tilde{\mathbf{z}}_\theta(\mathbf{x}_t, t)\right) + \sqrt{\frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}\beta_t}\,\mathbf{z}$$

## 04 提出方法

- 怎么用扩散模型做超分辨率?



# 04 提出方法

- 怎么用扩散模型做超分辨率?


$X_t$ + LR

输入
RGB⊕RGB
6通道

输出
RGB
3通道

# 04 提出方法

- 怎么用扩散模型做超分辨率?

随机生成 → 小图控制的生成结果


LR → HR

# 04 提出方法

- 改动1:
  - 将LR作为condition，与噪声图concat之后送给UNet重建，即Unet现在是输入6通道，输出3通道
- 改动2:
  - 不再直接取bar_alpha_t，而是取均匀分布[bar_alpha_{t-1}, bar_alpha_{t}]
- 改动3:
  - 不再输入t给UNet，而是直接输入noise level，也就是改动2中均匀采样的值

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

$\tilde{z} = \mathrm{UNet}(x_t, \bar{\alpha}_t)$

与其给第t步，不如给出噪声强度

推理中，可以不按照给定的 $\alpha_0 \cdots \alpha_T$ 去做

训练 $\beta = 1-\alpha$ $10^{-4} \to 2\times10^{-2}$ $T=2000$

推理 $\beta' = 1-\alpha'$ $10^{-4} \to 0.1$ $T=100$

步数变少，每一步去噪强度增加

最终 $\bar{\alpha}_T \approx 0$

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

把离散去噪变成连续去噪

$t \to \bar{\alpha}_t$

$t \to U[\bar{\alpha}_{t-1}, \bar{\alpha}_t]$

# 05 实验结果

- 主观效果上来说，虽然扩散模型没有使用GAN，但是重建图的细节质量是很不错的

# 05 实验结果

- 定量结果：

此的方法不太SOTA

| Metric | PULSE [28] | FSRGAN [7] | Regression | SR3 |
|---|---|---|---|---|
| PSNR ↑ | 16.88 | 23.01 | **23.96** | 23.04 |
| SSIM ↑ | 0.44 | 0.62 | **0.69** | 0.65 |
| Consistency ↓ | 161.1 | 33.8 | 2.71 | **2.68** |

**Table 1:** PSNR & SSIM on 16×16 → 128×128 face super-resolution. Consistency measures MSE ($\times 10^{-5}$) between the low-resolution inputs and the down-sampled super-resolution outputs.

| Model | FID ↓ | IS ↑ | PSNR ↑ | SSIM ↑ |
|---|---|---|---|---|
| Reference | 1.9 | 240.8 | - | - |
| Regression | 15.2 | 121.1 | **27.9** | **0.801** |
| SR3 | **5.2** | **180.1** | 26.4 | 0.762 |

SRGAN

**Table 2:** Performance comparison between SR3 and Regression baseline on natural image super-resolution using standard metrics computed on the ImageNet validation set.

| Method | Top-1 Error | Top-5 Error |
|---|---|---|
| Baseline | 0.252 | 0.080 |
| DRCN [22] | 0.477 | 0.242 |
| FSRCNN [13] | 0.437 | 0.196 |
| PsyCo [35] | 0.454 | 0.224 |
| ENet-E [44] | 0.449 | 0.214 |
| RCAN [64] | 0.393 | 0.167 |
| Regression | 0.383 | 0.173 |
| SR3 | **0.317** | **0.120** |

**Table 3:** Comparison of classification accuracy scores for 4× natural image super-resolution on the first 1K images from the ImageNet Validation set.

| Model | FID-50k |
|---|---|
| **Training with Augmentation** | |
| SR3 | 13.1 |
| SR3 (w/ Gaussian Blur) | 11.3 |
| **Objective $L_p$ Norm** | |
| SR3 ($L_2$) | 11.8 |
| SR3 ($L_1$) | 11.3 |

↓ ✓ ↓ 0.5

**Table 5:** Ablation study on SR3 model for class-conditional 256×256 ImageNet.

## 06 总结

- **总结：**
- 将图像超分辨率问题描述成一个条件生成问题
- 方法简单，实验部分一般

- **扩展：**
- （NIPSW’22）Palette: Image-to-Image Diffusion Models
- https://github.com/Janspiry/Image-Super-Resolution-via-Iterative-Refinement
- 论文附录里面有模型结构参数的介绍