01 背景介绍

- 论文标题截图：
- 论文链接：https://proceedings.neurips.cc/paper/2021/hash/49ad23d1ec9fa4bd8d77d02681df5cfa-Abstract.html
- 代码链接：https://github.com/openai/guided-diffusion
- 录用信息：NIPS'21

Diffusion Models Beat GANs on Image Synthesis

Prafulla Dhariwal*
OpenAI
prafulla@openai.com

Alex Nichol*
OpenAI
alex@openai.com

下午3:05 11月27日周日 94%

VictorYuki bilibili 墨迹设置 画笔 橡皮擦 清屏 退出Pencil模式

02 论文摘要

- 论文摘要截图：

- **提出问题：**
- 提高扩散模型生成性能

- **提出解决方案：** 通过消融实验找到的
- 对于无条件图像生成，找到一个更好架构
- 对于有条件图像生成，用分类器引导

原始的扩散模型就是一个无条件的图像生成器

- **优势&实验结果：**
- 更好的架构可以带来更好的生成效果
- 分类器引导的方法可以进一步提高生成效果

性能

Abstract

We show that diffusion models can achieve image sample quality superior to the current state-of-the-art generative models. We achieve this on unconditional image synthesis by finding a better architecture through a series of ablations. For conditional image synthesis, we further improve sample quality with classifier guidance: a simple, compute-efficient method for trading off diversity for fidelity using gradients from a classifier. We achieve an FID of 2.97 on ImageNet 128×128, 4.59 on ImageNet 256×256, and 7.72 on ImageNet 512×512, and we match BigGAN-deep even with as few as 25 forward passes per sample, all while maintaining better coverage of the distribution. Finally, we find that classifier guidance combines well with upsampling diffusion models, further improving FID to 3.94 on ImageNet 256×256 and 3.85 on ImageNet 512×512. We release our code at https://github.com/openai/guided-diffusion.

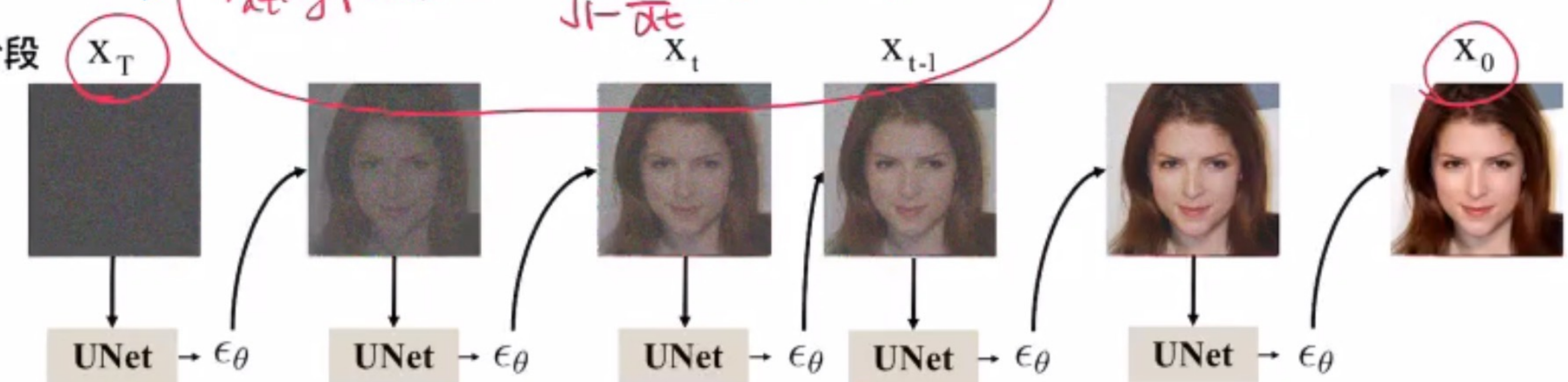
03 相关工作

$$\epsilon_\theta(x_t,t) = -\sqrt{1-\bar{\alpha}_t}\,\nabla_{x_t}\log p(x_t)$$

- 相关工作：DDPM采样过程/DDIM采样过程
- 用score来描述?

$$\nabla_{x_t}\log p(x_t) = -\frac{1}{\sqrt{1-\bar{\alpha}_t}}\epsilon_\theta(x_t,t)$$

- 采样阶段



DDPM:

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

DDIM:

$$x_{t-1} = \sqrt{\bar{\alpha}_t}\,x_0 + \sqrt{1-\bar{\alpha}_t-\sigma_t^2}\,\epsilon_\theta(x_t,t) + \sigma_t\epsilon$$

$$x_0 \approx \frac{x_t - \sqrt{1-\bar{\alpha}_t}\,\epsilon_\theta(x_t,t)}{\sqrt{\bar{\alpha}_t}}$$

推理时如何引导，用梯度下降引导吗

03 相关工作

- 相关工作：条件扩散模型

y: label

加条件

$$\nabla_{x_t}\log p(x_t) \Longrightarrow \underline{\nabla_{x_t}\log p(x_t|y)}$$

$$p(x_t|y) = \frac{p(y|x_t)\,p(x_t)}{p(y)}$$

与 x_t 无关

$$\nabla_{x_t}\log p(x_t|y) = \boxed{\nabla_{x_t}\log p(y|x_t)} + \underline{\nabla_{x_t}\log p(x_t)}$$

对无条件梯度做修正的梯度

$\approx s_\theta(x_t, t)$

关键在于如何求修正的梯度？

04 提出方法

- 用分类器引导生成

$P_\phi(y|x_t)$

$$
\begin{tikzcd}
x_t \arrow[r] & \boxed{\phi} \arrow[r] & out \arrow[r, leftrightarrow] & y \\
t \arrow[ru] & & &
\end{tikzcd}
$$

04 提出方法

- 用分类器引导生成
- https://github.com/openai/guided-diffusion/blob/main/scripts/classifier_sample.py#L54

34 28 56

```
def cond_fn(x, t, y=None):
    assert y is not None
    with th.enable_grad():
        x_in = x.detach().requires_grad_(True)
        logits = classifier(x_in, t)
        log_probs = F.log_softmax(logits, dim=-1)
        selected = log_probs[range(len(logits)), y.view(-1)]
        return th.autograd.grad(selected.sum(), x_in)[0] * args.classifier_scale
```

xt
预测结果
概率
找到对应的GT的概率
类似步长？学习率？
求梯度
所有GT概率相加
期望越大越好
xt

$\phi(x_t)=y$

$\nabla_{x_t}\log p(y|x_t)$

$\nabla_{x_t}\log p(x_t)$

05 实验结果

由于同时对质量和多样性均有影响，那么FID一定存在一个平衡的中间值

FID IS

FID越小越好

IS越大越好

单张图片质量 fidelity

diversity 图像集合

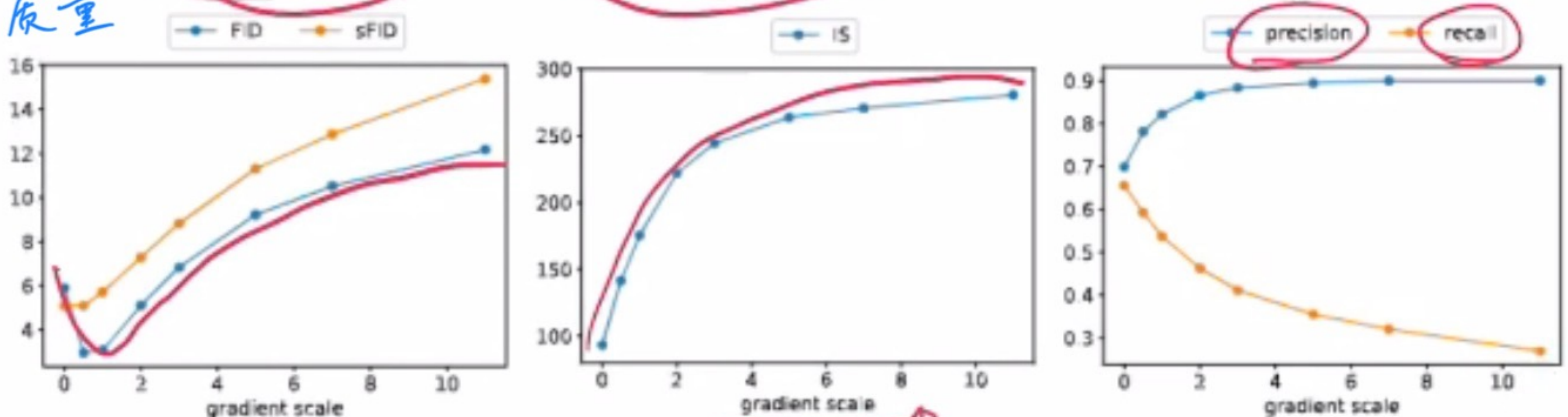


Figure 4: Change in sample quality as we vary scale of the classifier gradients for a class-conditional ImageNet 128×128 model.

05

实验结果

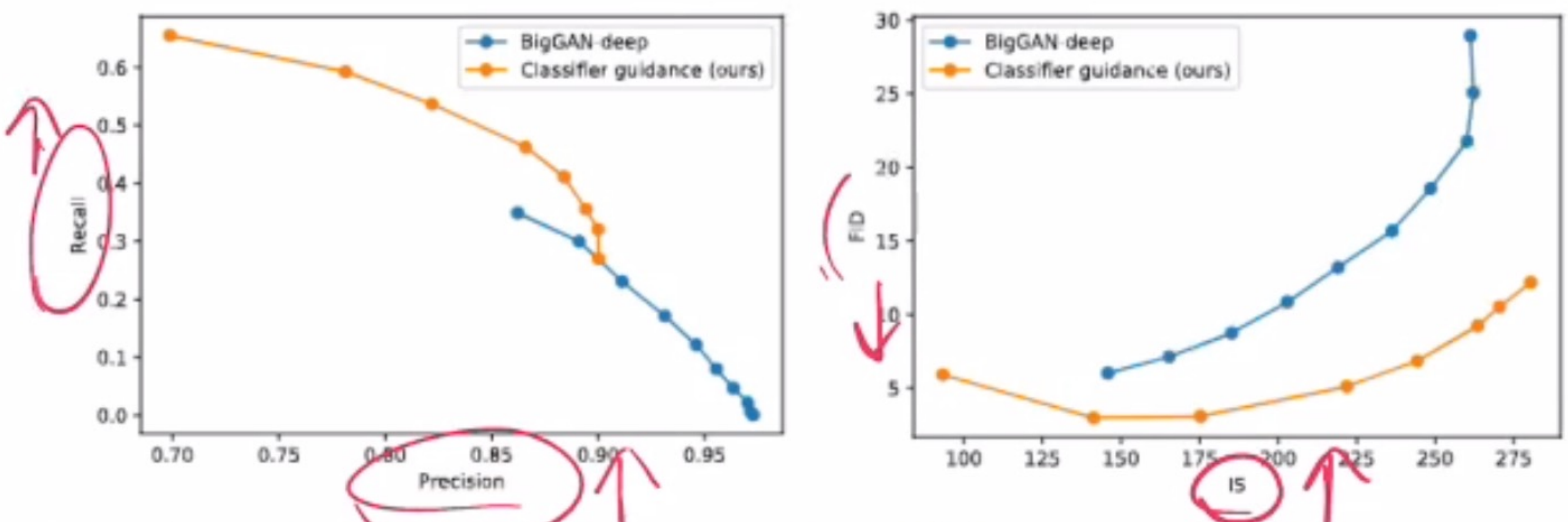


Figure 5: Trade-offs when varying truncation for BigGAN-deep and gradient scale for classifier guidance. Models are evaluated on ImageNet 128×128. The BigGAN-deep results were produced using the TFHub model [12] at truncation levels [0.1, 0.2, 0.3, ..., 1.0].

06 总结与收获

- 总结：
- 梯度引导是条件扩散模型的重要方法之一

$$\nabla_{x_t}\log p(y|x_t)$$