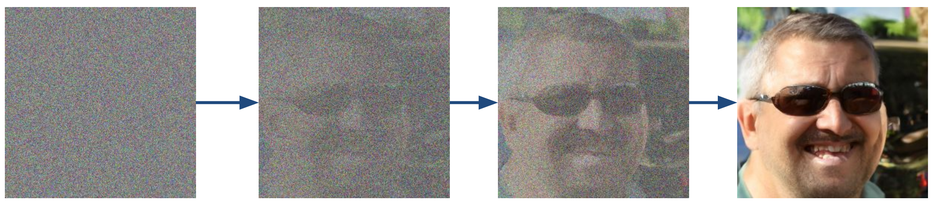
Diffusion Models for Image Priors

Nick Ping
Stanford University

Motivation

- Diffusion models can be used to generate images through iterative denoising
- Image deconvolution and inpainting are inverse problems which may use diffusion models as image priors



Related Work

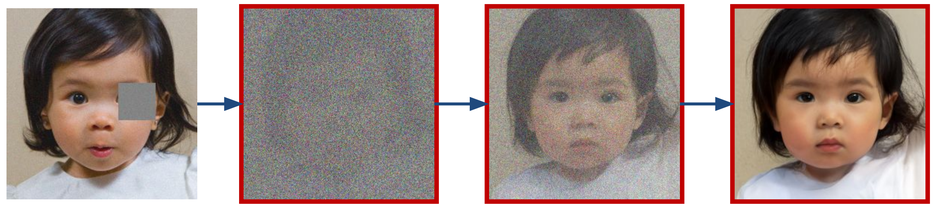
- Conditional Image Generation
- Additional Image Editing tasks:
  - Upsampling, Customization, etc.
- Video Generation
- Cross-Modal Generation
- Imitation Learning

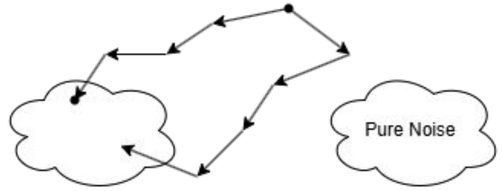
References

[1] Meng et al., Sdedit: Guided image synthesis and editing with stochastic differential equations, *CoRR*, 2022
[2] Jalal et al., Robust compressed sensing mri with deep generative priors, *CoRR*, 2021
[3] Chung et al., Diffusion posterior sampling for general noisy inverse problems, *ICLR*, 2023

Methods

- Comparing SDEdit [1], Score ALD [2], and DPS [3]





Score ALD

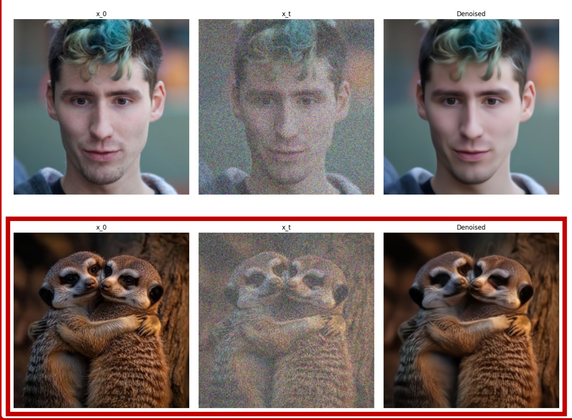
$\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
for $t = T, \dots, 1$ **do**
$\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$, else $\mathbf{z} = \mathbf{0}$
$\hat{\mathbf{x}}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t + (1 - \bar{\alpha}_t)\mathbf{s}_\theta(x_t, t))$
$\mathbf{x}_{t-1} = \frac{\sqrt{\alpha_t}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1 - \alpha_t)}{1 - \bar{\alpha}_t}\hat{\mathbf{x}}_0 + \sqrt{1 - \alpha_t}\mathbf{z}$
$x_{t-1} = x_{t-1} - \frac{1}{2(\sigma^2 + \gamma_t^2)}\nabla_{\mathbf{x}_t}\left\|\mathcal{A}(\mathbf{x}_t) - \mathbf{y}\right\|^2$
end for
return $\mathbf{x}_0$

DPS

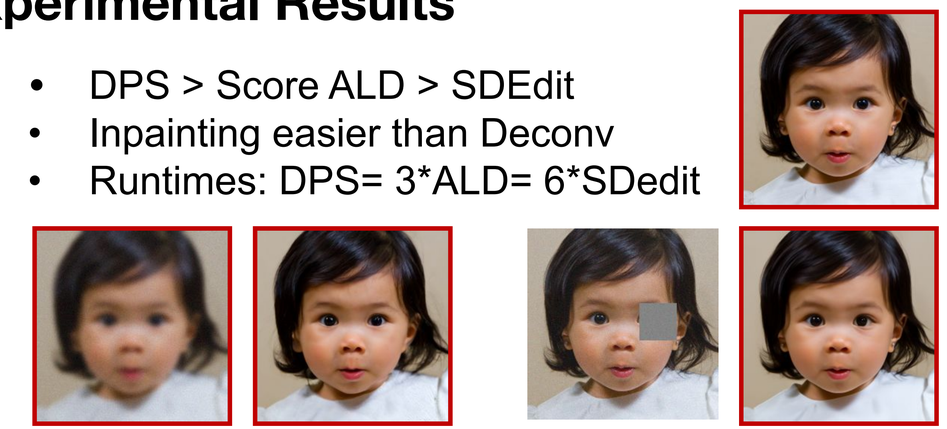
$\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
for $t = T, \dots, 1$ **do**
$\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$, else $\mathbf{z} = \mathbf{0}$
$\hat{\mathbf{x}}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t + (1 - \bar{\alpha}_t)\mathbf{s}_\theta(x_t, t))$
$\mathbf{x}'_{t-1} = \frac{\sqrt{\alpha_t}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1 - \alpha_t)}{1 - \bar{\alpha}_t}\hat{\mathbf{x}}_0 + \sqrt{1 - \alpha_t}\mathbf{z}$
$x_{t-1} = x'_{t-1} - \zeta_t \nabla_{\mathbf{x}_t}\left\|\mathcal{A}\left(\hat{\mathbf{x}}_0\right) - \mathbf{y}\right\|^2$
end for
return $\mathbf{x}_0$

Experimental Results

- In-Distr > Out-of-Distr
- 32.5/0.10 vs 27.6/0.16



- DPS > Score ALD > SDEdit
- Inpainting easier than Deconv
- Runtimes: DPS= 3*ALD= 6*SDedit



Deconvolution Method	PSNR	LPIPS
SDEdit	20.276	0.209
Score ALD	22.186	0.161
DPS	28.665	0.063

Inpainting Method	PSNR	LPIPS
SDEdit	21.035	0.183
Score ALD	28.014	0.059
DPS	34.732	0.027