Image Processing using Diffusion Models

Won Sup Song
Center for Global & Online Education, Stanford University

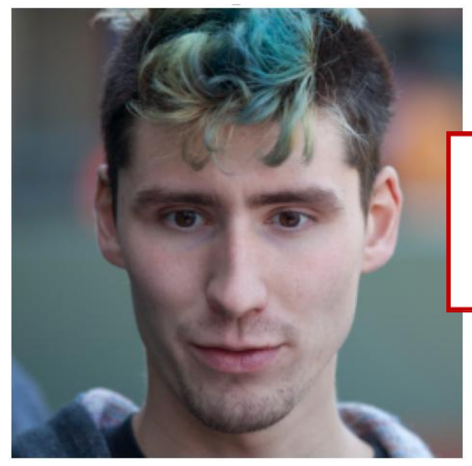
Motivation

This project explores the fundamentals of diffusion models and their capabilities in two key areas: **(1) Image Generation** and **(2) Image Restoration**, specifically inpainting and deconvolution.

Background & Related Work

***Key Measurement Metircs:**
- PSNR (Peak Signal-to-Noise Ratio):
  → Measures noise level/reconstruction quality (higher the better)
- LPIPS (Learned Perceptual Image Patch Similarity):
  → Measures perceptual similarity (lower the better)

Forward Noising Process (Variance Preserving(VP)):



"Noising"

$$x_t = \sqrt{\overline{\alpha_t}}\, x_0 + \sqrt{(1-\overline{\alpha_t})}\, z, \qquad for\ t = 1, 2, \dots, T$$

$$where\ \alpha_t = 1 - \beta_t, \overline{\alpha_t} = \prod_{i=0}^{t} \alpha_i, and\ z \sim \mathcal{N}(0,1)$$

$^*\beta_t$ *is the noise schedule*

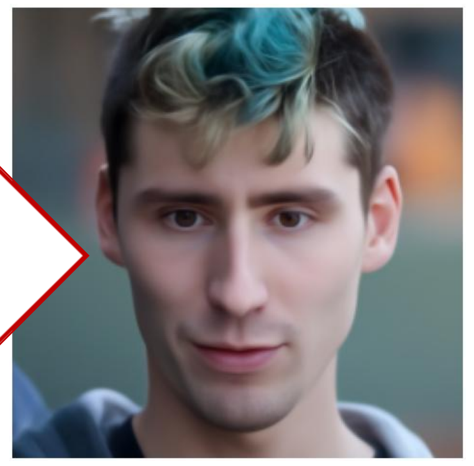
*T = 1000 steps

Reverse Denoising Process (DDPM):

"Denoising"

$$\hat{x}_0 = \frac{1}{\sqrt{\overline{\alpha_t}}}(x_t + (1-\overline{\alpha_t}) s_\theta(x_t, t))\ for\ t = T, T-1, \dots, 0$$

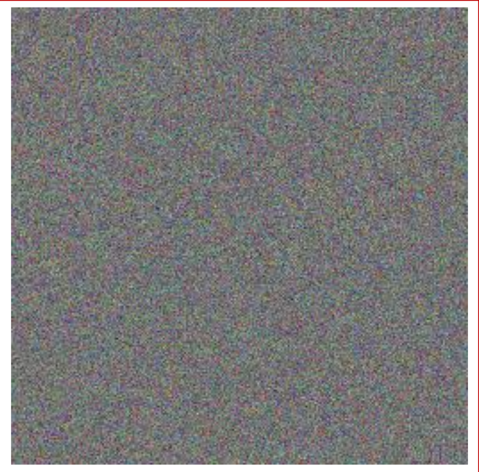
$$x_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t} x_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t} \hat{x}_0$$



PSNR: 31.70
LPIPS: 0.07

Image Restoration Methods:
- SDEdit method [4]: "Naïve" Approach (take measurement after timestep t)
- Score-ALD [3]: additional denoising step: $x_{t-1} = x_{t-1} - \frac{1}{\sigma^2 + {\gamma_t}^2} \nabla_{x_t} \|\mathcal{A}(x_t) - y\|^2$
- DPS [1]:additional denoising step: $x_{t-1} = x_{t-1} - \frac{\zeta_t}{2\sigma^2} \nabla_{x_t} \|\mathcal{A}(\hat{x}_0) - y\|^2$

Image Generation using DDPM Method [2] -> No "Guide" Image

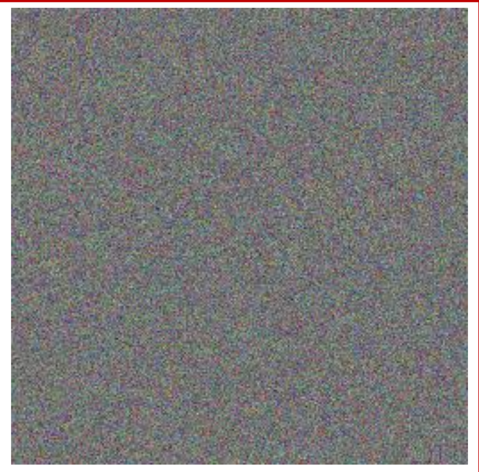


$x_T \sim \mathcal{N}(0,1)$

"Denoising"

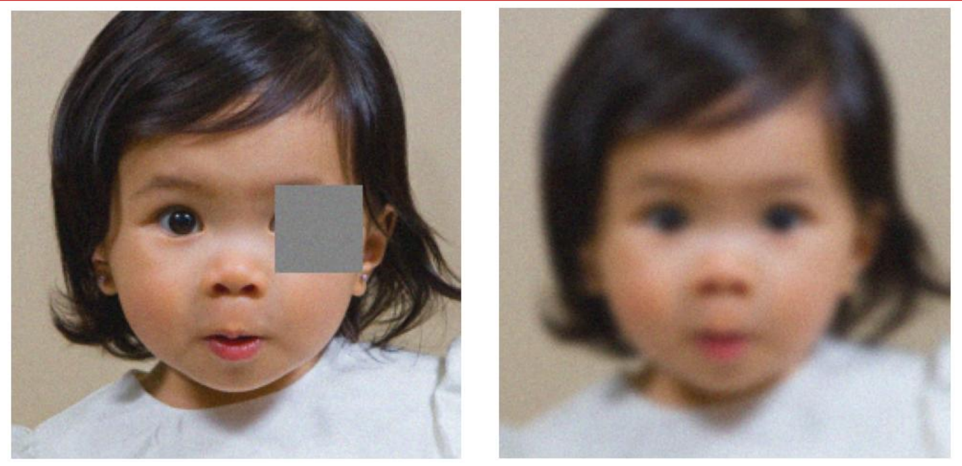
Every random noise will create different set of "denoised" images

Image Impainting & Deconvolution [1, 3, 4] -> With "Guide" Image



$x_T \sim \mathcal{N}(0,1)$

+



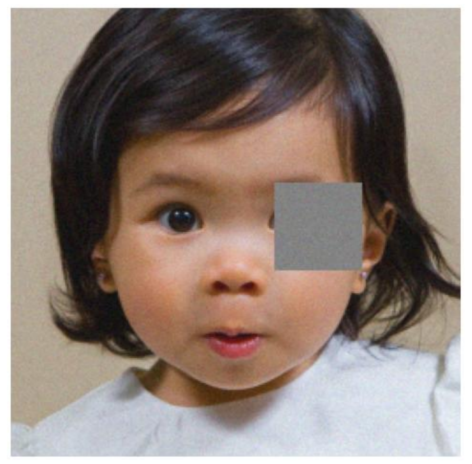
"Guide" Images -> Measurement

"Denoising"

Ground Truth

Measurement | **SDEdit** | **Score-ALD** | **DPS**

Impainting



Start time: 300
PSNR: 22.76
LPIPS: 0.14

Start time: 500
PSNR: 20.72
LPIPS: 0.18

Start time: 800
PSNR: 14.05
LPIPS: 0.316

PSNR: 22.75
LPIPS: 0.12

PSNR: 34.93
LPIPS: 0.02

Deconvolution

Start time: 300
PSNR: 24.45
LPIPS: 0.21

Start time: 500
PSNR: 22.34
LPIPS: 0.20

Start time: 800
PSNR: 13.32
LPIPS: 0.37

PSNR: 22.49
LPIPS: 0.14

PSNR: 28.36
LPIPS: 0.07

References

[1] Chung, Kim, Mccann, Klasky, and Ye (2023). Diffusion posterior sampling for general noisy inverse problems. In ICLR.
[2] Ho, Jain, and Abbeel (2020). Denoising diffusion probabilistic models. In NeurIPS.
[3] Jalal, Arvinte, Daras, Price, Dimakis, and Tamir (2021). Robust compressed sensing mri with deep generative priors.
[4] Meng, He, Y. Song, J. Song, Wu, Zhu, and Ermon (2022). Sdedit: Guided image synthesis and editing with stochastic differential equations.