Stanford ENGINEERING
Electrical Engineering

Diffusion Model for Image Restoration

Chih-Ying Liu [1]

[1]Electrical Engineering, Stanford

Overview

The goal of this project is to explore diffusion models for image restoration. We compare the performance and speed of a supervised super resolution method SRLDM [2] with an unsupervised method DDRM [1]. We also merged the two methods by using a supervised fine-tuned SRLDM as a denoising diffusion generative model in the DDRM pipeline.

Method

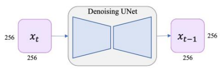
Unsupervised method: DDRM

$$y = Hx$$
$$H = U\Sigma V^T$$
$$\rightarrow \begin{cases} \bar{x} = V^T x \\ \bar{y} = \Sigma^{\dagger} U^T y \end{cases}$$

$$p_\theta^{(T)}(\bar{x_T}^{(i)}|y) = \begin{cases} \mathcal{N}(\bar{y}^{(i)}, \sigma_T^2 - \frac{\sigma_y^2}{s_i^2}) & s_i > 0 \\ \mathcal{N}(0, \sigma_T^2) & s_i = 0 \end{cases}$$

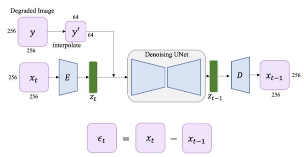
$$p_\theta^{(t)}(\bar{x_t}^{(i)}|x_{t+1}, y) = \begin{cases} \mathcal{N}(\bar{x_{\theta,t}}^{(i)} + \sqrt{1-\eta^2}\sigma_t \frac{\bar{x_{t+1}}^{(i)} - \bar{x_{\theta,t}}^{(i)}}{\sigma_{t+1}}, \eta^2\sigma_t^2) & s_i = 0 \\ \mathcal{N}(\bar{x_{\theta,t}}^{(i)} + \sqrt{1-\eta^2}\sigma_t \frac{\bar{y}^{(i)} - \bar{x_{\theta,t}}^{(i)}}{\sigma_y/s_i}, \eta^2\sigma_t^2) & \sigma_t < \frac{\sigma_y}{s_i} \\ \mathcal{N}((1-\eta_b)\bar{x_{\theta,t}}^{(i)} + \eta_b \bar{y}^{(i)}, \sigma_t^2 - \frac{\sigma_y^2}{s_i^2}\eta_b^2) & \sigma_t \leq \frac{\sigma_y}{s_i} \end{cases}$$

$\bar{x_{\theta,t}}^{(i)}$: pre-trained denoising diffusion model.



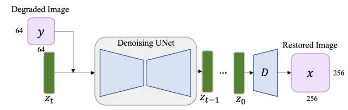
- Take SVD of degradation matrix H and perform diffusion in its spectral space
- Applicable to any linear inverse problems

SRLDM in DDRM Pipeline



- Use SRLDM as a denoising generative model in the DDRM pipeline.

Supervised method: SRLDM



- Perform denoising diffusion process in latent space.
- Supervised trained on super resolution task.

Experiment

Data

Random 2000 images from validation set of Imagenet-1K.

Baseline

The bicubic upsacled image for sr tasks and the degraded image for others.

Result

Quantitative result

- Super resolution ($64 \rightarrow 256$)

Method	PSNR ↑	SSIM ↑	FID ↓	Throughput ↑
Baseline	25.8	0.745	66.3	-
LDM	23.7	0.686	29.9	1.6
DDRM	27.4	0.799	29.6	0.1
DDRM-LDM	14.68	0.141	135	0.1

- Denoise ($\sigma = 0.1$)

Method	PSNR ↑	SSIM ↑	FID ↓	Throughput ↑
Baseline	26.3	0.634	17.47	-
LDM	23.7	0.657	29.9	1.6
DDRM	34.8	0.932	8.94	0.1
DDRM-LDM	22.1	0.588	36.5	0.1

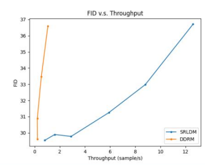
- Deblur ($\sigma = 1$)

Method	PSNR ↑	SSIM ↑	FID ↓	Throughput ↑
Baseline	25.5	0.778	27.8	-
LDM	22.8	0.664	35.6	1.6
DDRM	20.1	0.384	34.5	0.1
DDRM-LDM	18.2	0.309	45.7	0.1

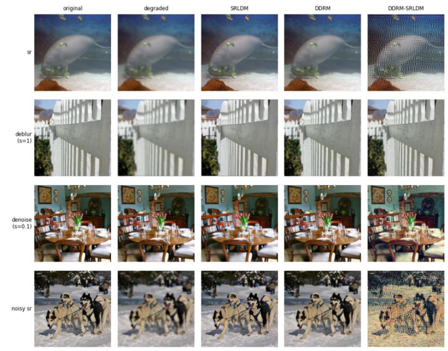
- Noisy super resolution ($64 \rightarrow 256$, $\sigma = 0.05$)

Method	PSNR ↑	SSIM ↑	FID ↓	Throughput ↑
Baseline	24.9	0.6603	97.12	-
LDM	22.7	0.553	38.3	1.6
DDRM	26.6	0.759	36.66	0.1
DDRM-LDM	13.4	0.109	168	0.1

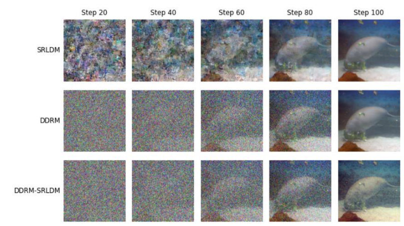
FID v.s. throughput



Generation result



Intermediate results



References

[1] Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. Denoising diffusion restoration models, 2022.

[2] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models, 2022.