Inverse Imaging with Diffusion Model Prior

Sikandar Y. Mashayak

Abstract—The motivation of this project is to learn and explore diffusion model based techniques to solve inverse imaging problems, such as deconvolution and in-painting. In this project, we implement and compare three techniques for modeling conditional terms of priors, namely score-distillation editing, ScoreALD, and diffusion posterior sampling. For a diffusion model, we use a pre-trained diffusion model trained with FFHQ dataset and DDPM sampling approach.

Index Terms—Stanford EE367 Computational Imaging

1 Introduction

Inverse imaging problems in most cases are ill-posed problems. For a given reference image $\mathbf{x} \in \mathbb{R}^N$ and a measurement matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$, the vectorized measurements $\mathbf{b} \in \mathbb{R}^M$ are formed as

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \eta, \tag{1}$$

where η is an additive noise term. In general imaging applications the number of measurements is smaller than the size of the target (unknown) image, i.e., $M < N$. Therefore Eq. 1 is under-determined and there are infinitely many solutions for $\mathbf{x}$ that could satisfy Eq. 1. Hence we need to add a constraint or a prior on the unknown $\mathbf{x}$ to determine the feasible solutions among the infinitely many possible solutions.

The inverse imaging problem with a prior is

$$\underset{\mathbf{x}}{\text{minimize}} \quad \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \lambda\Psi(\mathbf{x}), \tag{2}$$

where $\Psi(\mathbf{x})$ is the prior and λ is its relative weight compared to the data fidelity term. There are various options to choose prior. In this project, we solve in-painting and deconvolution inverse imaging problems with different diffusion model based priors, namely score-distillation editing (SDEdit) [1], ScoreALD [2], and diffusion posterior sampling [3], and compare their accuracies.

2 Related Work

In conventional approaches, based on the knowledge of the target image characteristics, some popular priors are as follows. $\Psi(\mathbf{x}) = \|\nabla\mathbf{x}\|_2$ for blurry target images to promote smoothness, $\Psi(\mathbf{x}) = \|\nabla\mathbf{x}\|_1$ for sparse images to promote sparsity, $\Psi(\mathbf{x}) = \text{TV}(\mathbf{x})$ [4] for natural images with sparse gradients. TV norm is one of the most popular regularization prior for natural images. Furthermore, for a general prior condition an arbitrary denoiser, such as DnCNN can also be used as a prior.

In this project, we use the most recent and powerful diffusion model based approach as a prior for inverse problem as described in the following Method section.

• *Stanford Online Student.*
E-mail: mashayak@stanford.edu

3 Method

Based on the Bayesian perspective of inverse problem, the target distribution of $\mathbf{x}$ could be conditioned on the measurement of $\mathbf{b}$ as

$$p_t(\mathbf{x}|\mathbf{b}) \propto \mathbf{p_t}(\mathbf{b}|\mathbf{x})\mathbf{p_t}(\mathbf{x}). \tag{3}$$

The optimal solution for Eq. 3 could be determined as a maximum-a-posterior (MAP) solution

$$\mathbf{x}_{\text{MAP}} = \arg\min_x -\log\left(p(\mathbf{b}|\mathbf{x},\sigma)\right) - \log\left(p(\mathbf{x})\right). \tag{4}$$

Since, the diffusion model based sampling approach is equivalent to solving inverse stochastic difference equation. We can modify inverse SDE by replacing unconditional probability of $p(\mathbf{x})$ with conditional probability of $p(\mathbf{x}|\mathbf{b})$ which results in the modified inverse SDE equation

$$\begin{aligned} d\mathbf{x} =& \mathbf{f}(\mathbf{x},\mathbf{t}) \\ &- g^2(t)\left(\nabla_x \log p_t(\mathbf{x}) + \nabla_{\mathbf{x}} \log \mathbf{p_t}(\mathbf{b}|\mathbf{x})\right) dt \\ &+ g(t) d\tilde{\mathbf{w}}. \end{aligned} \tag{5}$$

However, the score function of conditional probability $\nabla_x \log p_t(\mathbf{b}|\mathbf{x})$ is only valid at target $\mathbf{x}_0$, i.e.,

$$\nabla_x \log p_t(\mathbf{b}|\mathbf{x_0}) \neq \nabla_{\mathbf{x}} \log \mathbf{p_t}(\mathbf{b}|\mathbf{x_t}). \tag{6}$$

It is non-trivial to estimate $\nabla_x \log p_t(\mathbf{b}|\mathbf{x})$ accurately. Hence, we test two approximate approaches to estimate this conditional score terms.

First is ScoreALD [2],

$$\begin{aligned} \nabla_x \log p_t(\mathbf{b}|\mathbf{x_0}) &\approx \nabla_x \log p_t(\mathbf{b}|\mathbf{x_t}) \\ &\approx -\frac{1}{\sigma^2 + \gamma_t^2}\left(\mathbf{A^T}(\mathbf{b} - \mathbf{A}\mathbf{x})\right), \end{aligned} \tag{7}$$

which simply approximates conditional score term with the score term evaluated at current time step in inverse diffusion step, and controls the approximation error through annealing term γ_t^2. ScoreALD mainly applies to linear inverse problems.

Second is DPS [3],

$$\nabla_x \log p_t(\mathbf{b}|\mathbf{x_0}) \approx \nabla_{\mathbf{x}} \log \mathbf{p_t}\left(\mathbf{b}|\mathbf{x_0} = \mathbb{E}\left[\mathbf{x_0}|\mathbf{x_t}\right]\right), \tag{8}$$

which approximates the conditional score term with the score term evaluated at estimated mean at a given inverse diffusion step. DPS is shown to work with both linear and non-linear inverse problems.

In this project, to solve inverse SDE we use denoising diffusion probabilistic model [5] with slight modifications as described in the following subsection 3.1.

3.1 Denoising Diffusion Probabilistic Models

3.1.1 Variance Preserving Formulation

In this project, we only consider the variance-preserving (VP) formulation of diffusion models []. In VP formulation, the forward noise model from step $t-1$ to step t is given as

$$\mathbf{x}_t = \sqrt{1-\beta_t}\mathbf{x}_{t-1} + \sqrt{\beta_t}\mathbf{z}_{t-1}, \quad t = 1, 2, ..., T, \tag{9}$$

where $\mathbf{z}_{t-1} \sim \mathcal{N}(0, I)$ are i.i.d. Gaussian random variables in each step and β_t is the noise schedule. Substituting $\alpha_t = 1 - \beta_t$, in Eq. 9 we get

$$\mathbf{x}_t = \sqrt{\alpha_t}\mathbf{x}_{t-1} + \sqrt{1-\alpha_t}\mathbf{z}_{t-1}, \tag{10}$$

which could be expanded to

$$\mathbf{x}_t = \sqrt{\alpha_t}\sqrt{\alpha_{t-1}}\mathbf{x}_{t-2} + \sqrt{1-\alpha_t}\mathbf{z}_{t-1} + \sqrt{\alpha_t}\sqrt{1-\alpha_{t-1}}\mathbf{z}_{t-2},$$

$$\mathbf{x}_t = \sqrt{\alpha_t\alpha_{t-1}}\mathbf{x}_{t-2} + \sqrt{1-\alpha_t\alpha_{t-1}}\mathbf{z}, \tag{11}$$

where we use the relation that for weighted sum of two standard Gaussian variables $Z = w_1 Z_1 + w_2 Z_2$, the resultant Gaussian distribution is $Z \sim \mathcal{N}(0, w_1^2 + w_2^2)$. By expanding Eq. 11 to $\mathbf{x}_0$ terms we get

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\mathbf{z}, \tag{12}$$

where $\bar{\alpha}_t = \prod_{i=1}^{t} \alpha_i$, and $\mathbf{z} \sim \mathcal{N}(0, I)$.

3.1.2 Reverse Diffusion Step

For the case of VP formulation, DDPM sampling step, i,.e., reverse diffusion step, is defined as

$$\begin{aligned}
\hat{\mathbf{x}}_0 &= \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t + (1-\bar{\alpha}_t)\mathbf{s}_\theta(\mathbf{x}_t, t)\right) \\
\mathbf{x}_{t-1} &= \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_0.
\end{aligned} \tag{13}$$

We can further simplify Eq. 13 by substituting the formulation for $\hat{\mathbf{x}}_0$ as

$$\begin{aligned}
\mathbf{x}_{t-1} =& \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{\sqrt{\bar{\alpha}_t}(1-\bar{\alpha}_t)}\mathbf{x}_t \\
&+ \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)(1-\bar{\alpha}_t)}{\sqrt{\bar{\alpha}_t}(1-\bar{\alpha}_t)}\mathbf{s}_\theta(\mathbf{x}_t, t), \\
=& \frac{\sqrt{\bar{\alpha}_t}\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1}) + \sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{\sqrt{\bar{\alpha}_t}(1-\bar{\alpha}_t)}\mathbf{x}_t \\
&+ \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)(1-\bar{\alpha}_t)}{\sqrt{\bar{\alpha}_t}(1-\bar{\alpha}_t)}\mathbf{s}_\theta(\mathbf{x}_t, t), \\
=& \frac{\sqrt{\alpha_t\bar{\alpha}_{t-1}}\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1}) + \sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{\sqrt{\alpha_t\bar{\alpha}_{t-1}}(1-\bar{\alpha}_t)}\mathbf{x}_t \\
&+ \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)(1-\bar{\alpha}_t)}{\sqrt{\alpha_t\bar{\alpha}_{t-1}}(1-\bar{\alpha}_t)}\mathbf{s}_\theta(\mathbf{x}_t, t).
\end{aligned} \tag{14}$$

Eq. 14 simplifies to

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t + (1-\alpha_t)\mathbf{s}_\theta(\mathbf{x}_t, t)\right). \tag{15}$$

TABLE 1
Single-step denoising results.

	t=100		t=300		t=500	
	PSNR	LPIPS	PSNR	LPIPS	PSNR	LPIPS
Sample 1	31.20	0.09	26.12	0.20	21.96	0.35
Sample 2	32.59	0.09	27.26	0.21	23.47	0.32

Furthermore, as per Tweedie's formulation for VP forward process,

$$\begin{aligned}
\mathbb{E}\left[\mathbf{x}_{t-1}|\mathbf{x}_t\right] &= \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x_t} + (1-\alpha_t)\mathbf{s}_{\theta,t}\right), \\
\mathbf{s}_{\theta,t} &= \frac{\sqrt{\alpha_t}\mathbb{E}\left[\mathbf{x}_{t-1}|\mathbf{x}_t\right] - \mathbf{x_t}}{(1-\alpha_t)}, \\
\mathbf{s}_{\theta,t} &= -\frac{\epsilon_{\theta,t}}{\sqrt{1-\alpha_t}}.
\end{aligned} \tag{16}$$

Substituting Eq. 16 in Eq. 15 we get

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \sqrt{1-\alpha_t}\epsilon_\theta(\mathbf{x}_t, t)\right). \tag{17}$$

4 Experimental Results

First, we implement forward noise addition and inverse denoising methods based on DDPM sampling. Here we test our implementation by doing single step denoising and unconditional image generation. In Fig. 1, we can observe that the as we increase the number of forward noise steps, the single step denoising estimate deviate more from the ground truth input. This is reflected in both qualitative comparison of images in Fig. 1, and PSRN and perception metrics as shown in Table 1.

Furthermore, as shown in Fig. 2 when we generate different samples from pure noise by 1000 denoising time steps, we notice that the diffusion model predicts samples that look like human faces.

Fig. 1. Single-step denoising at different noise levels.

Finally, we estimate and compare inverse problem solution for image in-painting and deconvolution with 3 different techniques, SDEdit, ScoreALD, and DPS.

Fig. 2. Unconditional image generation.

TABLE 2
SDEdit results.

	t=400		t=500		t=600	
	PSNR	LPIPS	PSNR	LPIPS	PSNR	LPIPS
Box mask	20.57	0.23	18.48	0.25	16.33	0.34
Random mask	17.43	0.28	16.04	0.33	14.17	0.48
Blur	20.18	0.23	18.78	0.29	17.10	0.36

TABLE 3
Comparison of ScoreALD and DPS results.

	ScoreALD		DPS	
	PSNR	LPIPS	PSNR	LPIPS
Box mask	22.70	0.18	31.77	0.05
Random mask	20.35	0.31	22.00	0.16
Blur	22.00	0.19	24.62	0.11

For SDEdit, as shown in Fig. 3 and Table 2, we test prediction with different noise time steps and observe that for small time steps the prediction is close to the input measurement and for larger time steps, the prediction looks more realistic but further from the input measurement. Here we show outputs with 3 different noise steps of 400, 500, and 600.

In Fig. 4 and Table 3, we compare the predication quality of ScoreALD and DPS. As we can see from this example, DPS appears to solve both in-painting and devolution inverse problems better than ScoreALD in terms of qualitative and quantitative measures. We notice that prediction from DPS are much more consistent with the ground truth compared to ScoreALD predictions.

5 Conclusion

In this project, we learn basics of diffusion models and apply it to solve inverse imaging problems. We implement and experiment with inverse stochastic denoising step using denoising diffusion probabilistic model (DDPM). We implement conditional priors to solve inverse imaging problems using ScoreALD and DPS approximations. Based on our experiments to solve image in-painting and deconvolution problems, we find that DPS method outperforms ScoreALD.

References

[1] C. Meng, Y. He, Y. Song, J. Song, J. Wu, J.-Y. Zhu, and S. Ermon, "SDEdit: Guided Image Synthesis and Editing with Stochastic Differential Equations," 2022.

[2] A. Jalal, M. Arvinte, G. Daras, E. Price, A. G. Dimakis, and J. I. Tamir, "Robust Compressed Sensing MRI with Deep Generative Priors," 2021.

[3] H. Chung, J. Kim, M. T. Mccann, M. L. Klasky, and J. C. Ye, "Diffusion Posterior Sampling for General Noisy Inverse Problems," May 2024.

[4] E. F. L. Rudin, S. Osher, "Nonlinear total variation based noise removal algorithms - ScienceDirect," 1992.

[5] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," 2020.

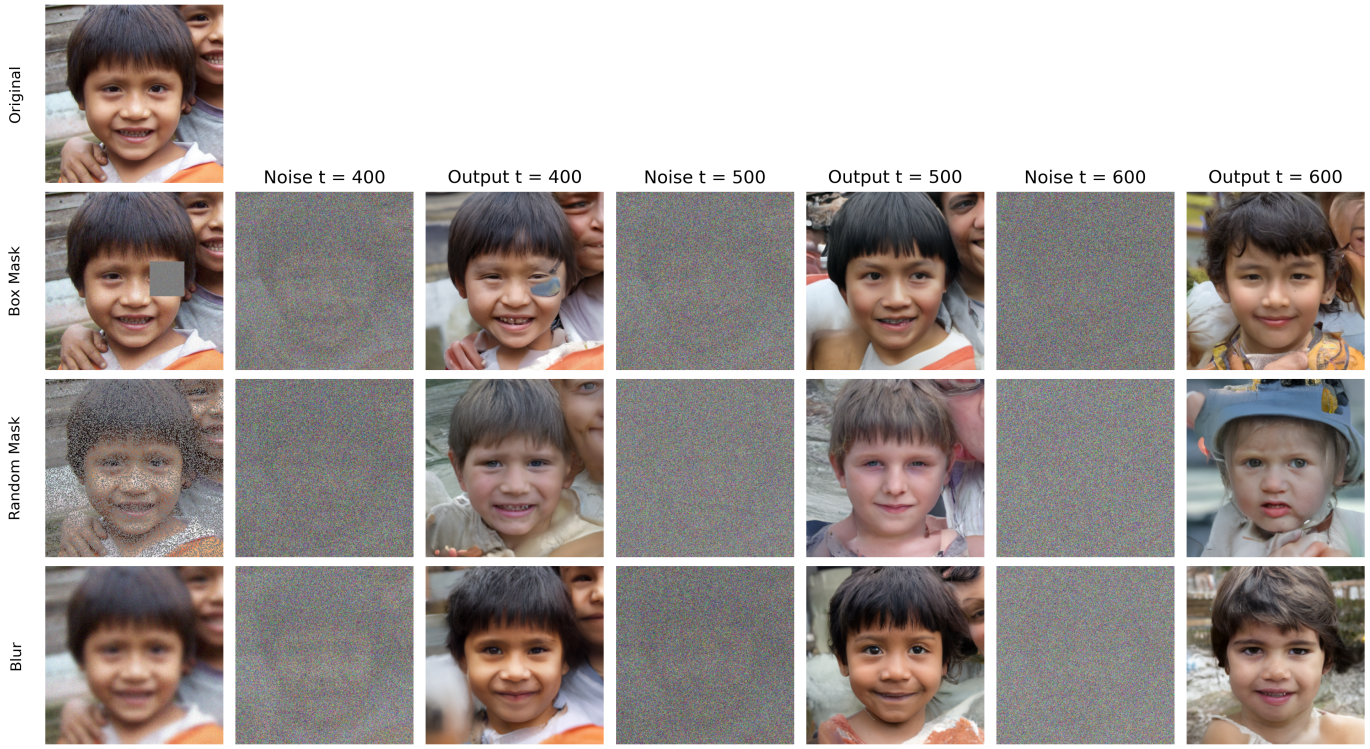

Fig. 3. SDEdit denoising at different noise levels.

Fig. 4. ScoreALD and DPS inverse imaging results.