Diffusion Models as Generative Priors for Inverse Problems

Niharika Gunturu

Abstract—Diffusion models [1] are a class of generative models that synthesize data by progressively denoising a noisy input through a Markov process consisting of a forward and reverse process. In this work, we explore their application to inverse problems such as inpainting and deconvolution. Specifically, we apply three diffusion-based methods to these inverse tasks: Score Distillation Editing (SDEdit) [2], Score Annealed Langevin Dynamics (ScoreALD) [3], and Diffusion Posterior Sampling (DPS) [4]. We use a pre-trained diffusion model trained on the FFHQ dataset [4] with a variance-preserving formulation. Our findings indicate that DPS achieves the highest PSNR and LPIPS values, making it the most effective for solving inverse problems with high perceptual similarity. While SDEdit performs suboptimally for inverse problems at higher noise levels, it excels as an image editing framework, allowing controlled modifications based on a prior.

Index Terms—Diffusion Models, Inverse Problems, Inpainting, Deconvolution, Diffusion Posterior Sampling, Langevin Dynamics, ScoreALD, SDEdit

1 Introduction

Diffusion models [1] are a class of generative models that learn to synthesize data by gradually denoising a noisy input. In this work, we explore their capabilities in **image generation**, **denoising**, **inpainting**, and **deconvolution**.

Specifically, we use diffusion model-based methods for the following tasks:

- Unconditional Image Generation from Noise
- Single Step Image Denoising

We also apply diffusion based methods to inverse problems - specifically Inpainting and Deconvolution via three methods:

- **Score Distillation Editing** (SDEdit) [2]
- **Score Annealed Langevin Dynamics** (ScoreALD) [3]
- **Diffusion Posterior Sampling** (DPS) [4]

For these tasks, we use a pre-trained diffusion model [4] that is trained on the Flickr-Faces-HQ Dataset (FFHQ) dataset, and use the variance preserving formulation of diffusion models [1].

2 Related Work

Diffusion models [1] generate images from noise by leveraging a Markov process consisting of a forward and reverse process. The forward process gradually adds Gaussian noise to an image through multiple timesteps, effectively transforming it into pure noise. The reverse process, parameterized by a deep neural network, learns to iteratively denoise and reconstruct a coherent image by estimating the noise at each step.

- *Niharika Gunturu is with the Department of Applied Physics, Stanford University, Stanford, CA, 94305.*
E-mail: ngunturu@stanford.edu

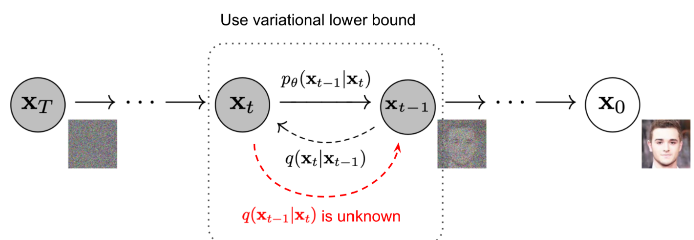

Fig. 1. Figure from Ho et al. 2020 [1], additionally annotated by Lilian Weng.

The canonical task of diffusion models is denoising, which enables two key applications: unconditional image generation by progressively refining pure noise into a coherent image and single-shot denoising, where a corrupted image is restored in a single reverse step.

The forward process can be written as follows:

$$x_t = \sqrt{1-\beta_t}x_{t-1} + \sqrt{\beta_t}z_{t-1} \tag{1}$$

We can reparametrize this equation with $\alpha_t = 1-\beta_t$ and $\bar{\alpha}_t = \Pi_{i=1}^{t}\alpha_i$:

$$\begin{aligned} x_t &= \sqrt{\alpha_t}x_{t-1} + \sqrt{1-\alpha_t}\epsilon_{t-1} \quad \text{where } \epsilon_{t-1}, \epsilon_{t-2}, \dots \sim \mathcal{N}(0, \mathbb{I}) \\ &= \sqrt{\alpha_t\alpha_{t-1}}x_{t-2} + \sqrt{1-\alpha_t\alpha_{t-1}}\bar{\epsilon}_{t-2} \\ &= \cdots \\ &= \sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon \end{aligned}$$

Forward Process: Generating Noisy Images for a Given Image:

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t\mathbb{I})) \tag{2}$$

Single Step Reverse Process: Generating Images from Noise in a Single Step:

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

Now, we substitute the following for $\hat{x}_0$:

$$\hat{x}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t + (1-\bar{\alpha}_t)s_\theta)$$

Which, on simplification yields the following:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(x_t + (1-\alpha_t)s_\theta\right) \tag{3}$$

Score-based model's equivalence to error-based model: The variance preserve form of Tweedie's formula gives us:

$$x_t = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon$$
$$\hat{x}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}\left[x_t + (1-\bar{\alpha}_t)\nabla_{x_t} \log p(x_t)\right]$$

On simplification, we have:

$$\sqrt{\bar{\alpha}_t}\mathbb{E}[x_0|x_t] - x_t = (1-\bar{\alpha}_t)\nabla_x \log p_t(x_t) \tag{4}$$

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left[x_t + (1-\alpha_t)s_\theta(x_t, t)\right] \tag{5}$$

$$= \frac{1}{\sqrt{\alpha_t}}\left[x_t + (1-\alpha_t)\frac{(\sqrt{\bar{\alpha}_t}\mathbb{E}[x_0|x_t] - x_t)}{1-\bar{\alpha}_t}\right] \tag{6}$$

$$= \cdots \tag{7}$$

$$= \frac{1}{\sqrt{\alpha_t}}\left[x_t - \frac{\sqrt{1-\alpha_t}}{\sqrt{1-\bar{\alpha}_t}}\epsilon_\theta\right] \tag{8}$$

2.1 Image Generation from Noise

Here, we include results of unconditional image generation from noise, using the pretrained diffusion model [4]:

Fig. 2. Examples of faces generated by the pretrained model starting from noise.

2.2 Single Shot Denoising

Single-shot denoising skips the full, iterative reverse process by directly estimating noise in a single step, using learned priors.

Single Shot Denoising Algorithm:

$$s \leftarrow \frac{-\epsilon}{\sqrt{1-\bar{\alpha}_t}}; \qquad \hat{x}_0 \leftarrow \sqrt{\frac{1}{\bar{\alpha}_t}}\left(x_t + (1-\bar{\alpha}_t)s\right) \tag{9}$$

Here, we include results of single shot denoising for different levels of noise. We see that as the noise increases, we are not able to denoise the image with a high level of perceptual similarity to the original image.

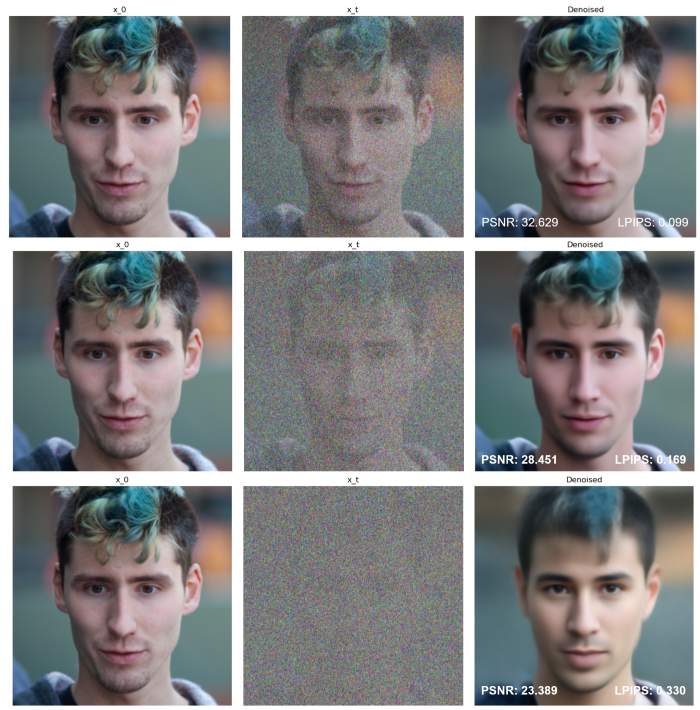

Fig. 3. Ground Truth (left), Noisy Input (middle), Model Result (right) for noise levels t= 100, 250 and 500 (top to bottom)

3 Diffusion Models for Inverse Problems

An inverse problem is generally framed as:

$$y = Ax^* + w \tag{10}$$

where y are a given set of measurements that one has access to, A defines the transform on the original data x^* - which is to be recovered. The term w refers noise that is voerlayed on the data, which for the purposes of our work we assume to be $\mathcal{N}(0, 1)$.

Specifically, we focus on two subcategories of inverse problems: inpainting and deconvolution and apply three diffusion-based approaches to solve them.

3.1 Inverse Problems via Score Distillation Editing (SDEdit) [2]

Here we first take the input x_0 and propagate it using the forward process to a certain timestep t_{start}, after which we iteratively apply the SDEdit Algorithm, to calculate $\hat{x}_0(x_t)\ \forall\ t = T-1, \cdots, 0$.

SDEdit Algorithm:

$$s \leftarrow \frac{-\epsilon}{\sqrt{1-\bar{\alpha}_t}} \tag{11}$$

$$\hat{x}_0 \leftarrow \sqrt{\frac{1}{\bar{\alpha}_t}}\,(x_t + (1-\bar{\alpha}_t)s) \tag{12}$$

$$\tag{13}$$

$$x_{t-1} \leftarrow q(\hat{x}_{t-1}|x_t, x_0) + \sigma z \qquad \text{where } w \sim \mathcal{N}(0,1) \tag{14}$$

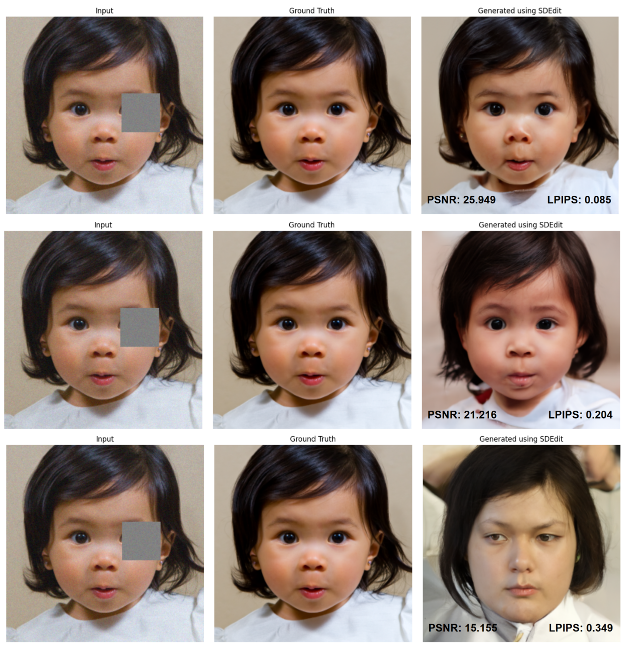

Fig. 4. Results for SDEdit applied to inpainting for t=250 (top), 500 (middle) and 750 (bottom)

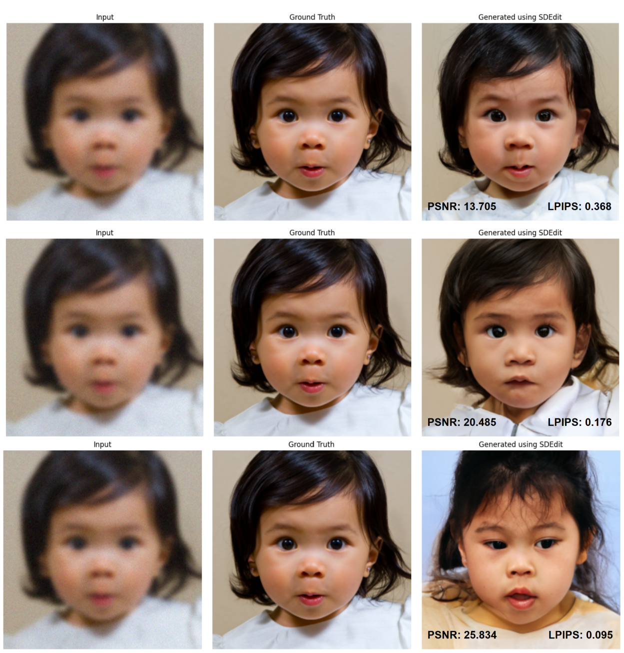

Fig. 5. Results for SDEdit applied to deconvolution for t=250 (top), 500 (middle) and 750 (bottom)

3.2 Inverse Problems via Score Annealed Langevin Dynamics (ScoreALD) [3]:

The ScoreALD algorithm modifies the reverse process of the diffusion model to also sample from the posterior distribution $\mu(x|y)$.

Sampling the posterior $\mu(x|y)$ using Langevin Dynamics:

$$x_{t+1} \leftarrow x_t + \eta_t \nabla_{x_t} log\ \mu(x_t|y) + \sqrt{2\eta_t}\zeta_t \tag{15}$$

$$x_{t+1} \leftarrow x_t + \eta_t \left(f(x_t;\beta_t) + \frac{A^H(y - Ax_t)}{\gamma_t^2 + \sigma^2} \right) + \sqrt{2\eta_t}\zeta_t \tag{16}$$

where $\zeta_t \sim \mathcal{N}(0,1)$.

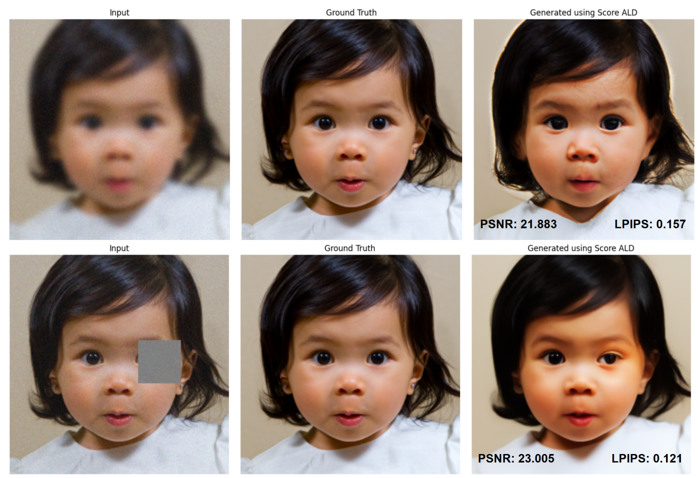

Fig. 6. ScoreALD results for deconvolution (top) and inpainting (bottom); $\sigma = 0.05$

3.3 Inverse Problems via Diffusion Posterior Sampling (DPS) [4]:

The DPS algorithm, like ScoreALD algorithm also modifies the reverse process of the diffusion model to sample from the posterior distribution $\mu(x|y)$. However, the DPS algorithm approximates this distribution with a term that ultimately works out to reflect the "loss"term for the inverse problem, as a function of $\hat{x}_0$.

Diffusion Posterior Sampling Algorithm:

$$x_{i-1} \leftarrow x'_{i-1} - \eta_i \nabla_{x_i} ||y - \mathcal{A}(\hat{x}_0)||^2; \tag{17}$$

$$\text{where } \ \zeta_i = \frac{c}{||y - \mathcal{A}(\hat{x}_0(x_i))||^2} \tag{18}$$

where for $\forall i = N-1, \cdots, 0$ we have $x_N \sim \mathcal{N}_c(0, \mathbb{I})$ and:

$$\hat{s} \leftarrow s_\theta(x_i, i); \qquad \hat{x}_0 \leftarrow \frac{1}{\sqrt{\alpha}}(x_i + (1-\bar{\alpha}_i\hat{s}) \tag{19}$$

$$x'_{i-1} \leftarrow \frac{\sqrt{\alpha_i}(1-\bar{\alpha}_{i-1})}{(1-\bar{\alpha}_i)}x_i + \frac{\sqrt{\bar{\alpha}_{i-1}}\beta_i}{1-\bar{\alpha}_i}\hat{x}_0 + \sigma_i z \tag{20}$$

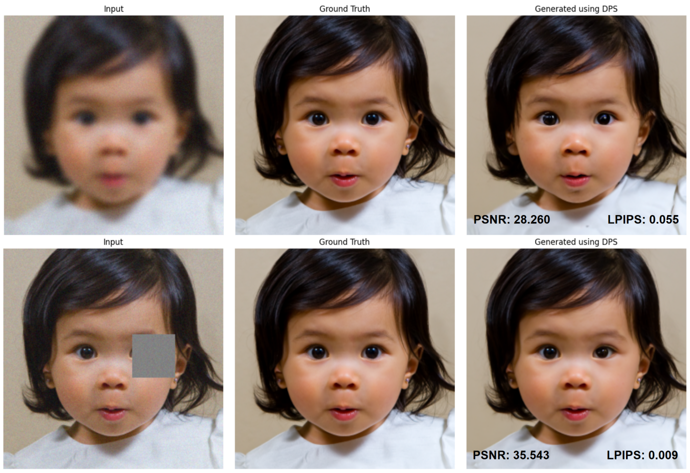

Fig. 7. DPS results for deconvolution (top) and inpainting (bottom)

4 Conclusion

Through this work, we explored the application of diffusion models to inverse problems, inpainting and deconvolution to be specific. We apply three algorithms in specific: SDEdit [2], ScoreALD [3] and DPS [4].

We notice that the DPS approach results in the highest PSNR and LPIPS values, and thus highest perceptual similarity, and is thus best suited to solve inverse problems.

On the other hand, the SDEdit approach has its own merits. Although the results it yields for the inverse problems are subpar at higher noise levels, it can be used as an image editing framework to generate new images based on a prior.

Acknowledgments

The author would like to thank the teaching staff of the Win 2025 edition of EE367 for their teaching, support and guidance.

References

[1] P. A. Jonathan Ho, Ajay Jain, "Denoising diffusion probabilistic models." *In NeurIPS*, 2020. [Online]. Available: https://doi.org/10.48550/arXiv.2006.11239

[2] Y. S. J. S. J. W. J.-Y. Z. S. E. Chenlin Meng, Yutong He, "Guided image synthesis and editing with stochastic differential equations." 2022. [Online]. Available: https://doi.org/10.48550/arXiv.2108.01073

[3] G. D. E. P. A. G. D.-J. I. T. Ajil Jalal, Marius Arvinte, "Robust compressed sensing mri with deep generative priors." 2021. [Online]. Available: https://doi.org/10.48550/arXiv.2108.01368

[4] M. T. M. M. L. K. J. C. Y. Hyungjin Chung, Jeongsol Kim, "Diffusion posterior sampling for general noisy inverse problems." *In ICLR*, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2209.14687