Solving Inverse Imaging Problems with Generative Diffusion Models

Kory Yang

Abstract—This project explores the application of score-based diffusion models as generative priors for solving inverse problems in computational imaging. Utilizing a pretrained diffusion model trained on the FFHQ dataset , we implement the variance-preserving (VP) forward noising process and the iterative DDPM sampling algorithm to unconditionally generate high-fidelity facial images from pure Gaussian noise. We extend this framework to solve inverse imaging problems – specifically image inpainting and deconvolution –by implementing three distinct posterior sampling methodologies: SDEdit, ScoreALD, and Diffusion Posterior Sampling (DPS). While SDEdit utilizes a stochastic differential equation approach starting from a partially noised state , ScoreALD and DPS incorporate measurement-driven likelihood gradients to guide the reverse diffusion process. We provide a comparative analysis of these methods using Peak Signal-to-Noise Ratio (PSNR) and Learned Perceptual Image Patch Similarity (LPIPS) metrics, highlighting the trade-offs between reconstruction accuracy and generative stability.

Index Terms—Computational Imaging, Diffusion Models

1 Introduction

In the field of computational imaging, many challenges are categorized as inverse problems. An inverse problem involves recovering an unknown signal or image from a set of noisy or degraded measurements. While the "forward" process—such as an image becoming blurry due to a camera shake or losing data through a mask—is mathematically straightforward, reversing that process is "ill-posed", meaning that there are often infinitely many clean images that could have resulted in the same blurry or masked observation, making the search for the "correct" reconstruction a significant technical hurdle.

The ability to solve these problems have many applications in modern consumer technology. Features seen in modern photo-editing software like Apple Intelligence's Clean Up feature, Google's Magic Eraser, or Adobe's Generative Fill rely on solving inpainting and denoising problems to seamlessly remove unwanted objects or restore low-quality photos. Beyond aesthetics, these techniques also have more applications in medical imaging, satellite surveillance, and archival restoration, and more.

Traditional methods for solving inverse problems often relied on handcrafted mathematical priors. However, this project explores the shift toward learned generative priors—specifically Denoising Diffusion Probabilistic Models (DDPMs). By utilizing a pretrained model as a prior, we can steer the generation process to find solutions that are not only statistically likely but also strictly consistent with the observed physical measurements.

This report details the implementation of several key methodologies:

- **Generative Modeling**: Implementing the variance-preserving forward process and the iterative sampling required for unconditional image generation.
- **Stochastic Editing (SDEdit)**: A method that leverages partial noising to guide the generation process toward a desired observation.
- **Posterior Sampling**: Implementing Score-based Annealed Langevin Dynamics (ScoreALD) and Diffusion Posterior Sampling (DPS), which utilize likelihood gradients to maintain data fidelity during the denoising trajectory.

By comparing these approaches, we demonstrate how diffusion models can act as universal priors, capable of reconstructing high-quality images from degraded measurements without being explicitly trained for each specific task.

- *M. Shell is with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332. E-mail: see http://www.michaelshell.org/contact.html*
- *J. Doe is with Anonymous University.*

2 Related Work

The foundation of this project lies in the recent shift toward using generative models as priors for inverse problems. The following works represent the core methodologies and theoretical frameworks we implemented:

2.1 Denoising Diffusion Probabilistic Models (DDPM)

Ho et al. (2020) introduced the variance-preserving (VP) formulation of diffusion models, where a forward Markov chain gradually adds Gaussian noise to data [1]. They demonstrated that learning to reverse this process—parameterized as a noise-predicting neural network—allows for high-quality image generation from pure noise.

2.2 SDEdit (Score-Distillation Editing)

Meng et al. (2022) proposed a method for guided image synthesis that bypasses the need for full reverse diffusion [2]. By adding a specific amount of noise to a corrupted or guide image and then denoising it, SDEdit leverages the diffusion prior to "project" the observation back onto the natural image manifold.

2.3 Score-based Annealed Langevin Dynamics (Score-ALD)

Jalal et al. (2021) developed a more rigorous posterior sampling approach. Their method incorporates measurement information through likelihood gradients during the reverse diffusion process, using annealed dynamics to ensure the sample remains consistent with the observation [3].

2.4 Diffusion Posterior Sampling (DPS)

Chung et al. (2023) introduced a stabilized approach for solving both linear and non-linear inverse problems. DPS utilizes normalized likelihood gradients to guide the generative process, providing a robust framework for tasks like deconvolution and inpainting without requiring task-specific training [4].

3 METHODS

Our implementation utilizes a variance-preserving (VP) score-based diffusion model as a prior to solve various imaging tasks. The methodology is divided into the fundamental diffusion framework and the specific posterior sampling techniques used for inverse problems.

3.1 Diffusion Framework and Forward Process

The core of our model follows the VP formulation introduced by Ho et al. [1] The forward process transitions a clean image x_0 to a noisy state x_t over T steps using the noise schedule β_t. First, we aim to prove that the variance-preserving forward noise model defined recursively as

$$x_t = \sqrt{1-\beta_t}x_{t-1} + \sqrt{\beta_t}z_{t-1}, \quad t = 1, 2, \ldots, T$$

can be expressed in a closed-form that depends only on the initial state x_0 and the time step t:

$$x_t = \sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}z$$

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

Starting with the definition for the recursive variance-preserving forward noise model and substituting α_t, we get:

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

Expanding x_{t-1} using the recursive definition:

$$x_t = \sqrt{\alpha_t}(\sqrt{\alpha_{t-1}}x_{t-2} + \sqrt{1-\alpha_{t-1}}z_{t-2}) + \sqrt{1-\alpha_t}z_{t-1}$$

$$x_t = \sqrt{\alpha_t\alpha_{t-1}}x_{t-2} + \sqrt{\alpha_t(1-\alpha_{t-1})}z_{t-2} + \sqrt{1-\alpha_t}z_{t-1}$$

Because z_{t-1} and z_{t-2} are independent standard Gaussian random variables $z \sim \mathcal{N}(0, I)$, their linear combination is also a Gaussian distribution. For two independent Gaussians $X \sim \mathcal{N}(0, \sigma_1^2 I)$ and $Y \sim \mathcal{N}(0, \sigma_2^2 I)$, the sum $X + Y \sim \mathcal{N}(0, (\sigma_1^2 + \sigma_2^2)I)$.

$$\begin{aligned}\text{Var}[\sqrt{\alpha_t(1-\alpha_{t-1})}z_{t-2} + \sqrt{1-\alpha_t}z_{t-1}] &= \alpha_t(1-\alpha_{t-1}) + (1-\alpha_t) \\ &= 1 - \alpha_t\alpha_{t-1}\end{aligned}$$

Because the z_{t-1} and z_{t-2} noise terms are both independent standard Gaussian random variables, we can rewrite the expression as a single Gaussian distribution $z' \sim \mathcal{N}(0, I)$ with an equivalent variance:

$$x_t = \sqrt{\alpha_t\alpha_{t-1}}x_{t-2} + \sqrt{1-\alpha_t\alpha_{t-1}}z'$$

Recursively substituting the terms until $t = 1$ gives us:

$$x_t = \sqrt{\prod_{i=1}^{t}\alpha_i}x_0 + \sqrt{1-\prod_{i=1}^{t}\alpha_i}z$$

$$x_t = \sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}z$$

By deriving this closed-form expression, we demonstrate that any noisy state x_t can be sampled directly from the clean image x_0 in a single step.

3.2 Iterative Reverse Sampling

To generate images, we implement the reverse diffusion process. While the standard DDPM formulation often uses a noise predictor ϵ_θ, our implementation utilizes a score-predicting network $s_\theta(x_t, t)$ which predicts the gradient of the log-density $\nabla_{x_t} \log p_t(x_t)$. To prove the equivalence of these formulations, see the proofs below:

3.2.1 Proof of Equivalence Equivalence of DDPM Reverse Process Forms

We aim to prove that the following form of the DDPM reverse diffusion process:

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

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

Is equivalent to:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}(x_t + (1-\alpha_t)s_\theta(x_t, t))$$

First we substitute in $\hat{x}_0$:

$$x_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}x_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\left(\frac{1}{\sqrt{\bar{\alpha}_t}}(x_t + (1-\bar{\alpha}_t)s_\theta(x_t, t))\right)$$

Given that $\bar{\alpha}_t = \alpha_t\bar{\alpha}_{t-1}$, this means $\sqrt{\bar{\alpha}_{t-1}}/\sqrt{\bar{\alpha}_t} = 1/\sqrt{\alpha_t}$:

$$x_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}x_t + \frac{(1-\alpha_t)}{\sqrt{\alpha_t}(1-\bar{\alpha}_t)}(x_t + (1-\bar{\alpha}_t)s_\theta(x_t, t))$$

Distributing and grouping similar terms:

$$x_{t-1} = \left(\frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t} + \frac{(1-\alpha_t)}{\sqrt{\alpha_t}(1-\bar{\alpha}_t)}\right)x_t + \frac{(1-\alpha_t)}{\sqrt{\alpha_t}}s_\theta(x_t, t)$$

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\frac{\alpha_t(1-\bar{\alpha}_{t-1}) + (1-\alpha_t)}{1-\bar{\alpha}_t}\right)x_t + \frac{(1-\alpha_t)}{\sqrt{\alpha_t}}s_\theta(x_t, t)$$

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\frac{\alpha_t - \alpha_t\bar{\alpha}_{t-1} + 1 - \alpha_t}{1-\bar{\alpha}_t}\right)x_t + \frac{(1-\alpha_t)}{\sqrt{\alpha_t}}s_\theta(x_t, t)$$

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\frac{\alpha_t - \bar{\alpha}_t + 1 - \alpha_t}{1-\bar{\alpha}_t}\right)x_t + \frac{(1-\alpha_t)}{\sqrt{\alpha_t}}s_\theta(x_t, t)$$

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\frac{1-\bar{\alpha}_t}{1-\bar{\alpha}_t}\right)x_t + \frac{(1-\alpha_t)}{\sqrt{\alpha_t}}s_\theta(x_t, t)$$

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}x_t + \frac{(1-\alpha_t)}{\sqrt{\alpha_t}}s_\theta(x_t, t)$$

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

3.2.2 Proof of Equivalence Between Score-Based and Noise-Based Reverse Diffusion Step

We aim to prove that the one-step reverse diffusion step

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

is equivalent to the formulation with the noise-predicting network ϵ_θ:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta(x_t, t) \right)$$

We start with the formulation of the variance-preserving forward diffusion process:

$$x_t = \sqrt{\bar{\alpha}_t} \cdot x_0 + \sqrt{1 - \bar{\alpha}_t} \cdot \epsilon$$

Where $\epsilon \sim \mathcal{N}(0, I)$ is the noise. Rearranging the equation to solve for the clean image x_0:

$$x_0 = \frac{x_t - \sqrt{1 - \bar{\alpha}_t} \epsilon}{\sqrt{\bar{\alpha}_t}}$$

Then if we have $\epsilon_\theta(x_t, t)$ which predicts the noise instead of ϵ, then x_0 becomes $\mathbb{E}[x_0|x_t]$, the best estimate of the clean image given the noisy image. This gives us:

$$\mathbb{E}[x_0|x_t] = \frac{x_t - \sqrt{1 - \bar{\alpha}_t} \epsilon_\theta(x_t, t)}{\sqrt{\bar{\alpha}_t}}$$

Using Tweedie's formula:

$$\mathbb{E}[x_0|x_t] = \frac{1}{\sqrt{\bar{\alpha}_t}} (x_t + (1 - \bar{\alpha}_t) \nabla_{x_t} \log p_t(x_t))$$

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

Substituting our expression for $\mathbb{E}[x_0|x_t]$ into this formula gives us:

$$\nabla_{x_t} \log p_t(x_t) = \frac{\sqrt{\bar{\alpha}_t} \left(\frac{x_t - \sqrt{1 - \bar{\alpha}_t} \epsilon_\theta(x_t, t)}{\sqrt{\bar{\alpha}_t}} \right) - x_t}{1 - \bar{\alpha}_t}$$

$$\nabla_{x_t} \log p_t(x_t) = \frac{\left(x_t - \sqrt{1 - \bar{\alpha}_t} \epsilon_\theta(x_t, t) \right) - x_t}{1 - \bar{\alpha}_t}$$

$$\nabla_{x_t} \log p_t(x_t) = \frac{-\epsilon_\theta(x_t, t)}{\sqrt{1 - \bar{\alpha}_t}}$$

Since the score function $s_\theta(x_t, t) = \nabla_{x_t} \log p_t(x_t)$, we can substitute $s_\theta(x_t, t)$ in our original one-step reverse diffusion equation:

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

Substituting the derived identity for $s_\theta(x_t, t)$:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t + (1 - \alpha_t) \left(\frac{-\epsilon_\theta(x_t, t)}{\sqrt{1 - \bar{\alpha}_t}} \right) \right)$$

Which simplifies into:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta(x_t, t) \right)$$

Since we've established the mathematical equivalence between noise-prediction and score-based models, our implementation utilizes the score function $s_\theta(x_t, t)$ for the reverse diffusion process in later tasks.

3.3 SDEdit (Stochastic Differential Equation Editing)

This approach provides a computationally efficient way to solve inverse problems by leveraging the "projective" nature of diffusion models. We initialize the process by adding a specific amount of noise to the degraded observation y up to a fixed intermediate timestep t_0 (typically $t = 500$). We then implement a standard reverse diffusion loop from t_0 to $t = 0$, allowing the model to "hallucinate" the missing or blurry information while remaining anchored to the original structure of the observation.

3.4 Score-based Annealed Langevin Dynamics (ScoreALD)

Unlike the heuristic approach of SDEdit, ScoreALD is a more rigorous posterior sampling method. We implemented a modified sampling loop that incorporates the measurement information at every timestep by combining the prior score $\nabla_{x_t} \log p_t(x_t)$ with a likelihood gradient $\nabla_{x_t} \log p(y|x_t)$. By using annealed dynamics, we ensure that the generated samples gradually converge toward the set of solutions that are both physically consistent with y and statistically likely under the FFHQ prior.

3.5 Diffusion Posterior Sampling (DPS)

To address the potential instabilities in standard likelihood guidance, we implemented the DPS algorithm. Similar to ScoreALD, DPS utilizes gradients of the measurement error. However, our implementation includes a normalized gradient update step, which stabilizes the reverse process by ensuring the guidance signal does not overwhelm the diffusion prior. This allows the model to handle both inpainting and deconvolution tasks with high fidelity, maintaining stable sampling even in the presence of significant degradation.

4 Evaluation and Comparison of Methods

To evaluate the performance of the implemented diffusion-based methods for image restoration, we utilize two distinct quantitative metrics: Peak Signal-to-Noise Ratio (PSNR) and Learned Perceptual Image Patch Similarity (LPIPS).

- **Peak Signal-to-Noise Ratio (PSNR)**: Measures pixel-level fidelity by calculating the ratio between the maximum possible signal power and the mean squared error (MSE). Higher values indicate a more accurate reconstruction.
- **Learned Perceptual Image Patch Similarity (LPIPS)**: Assesses perceptual quality by comparing deep features of the images. Lower scores indicate that the result is "visually closer" to the ground truth, which is critical for diffusion models that may prioritize realistic textures over exact pixel matching.

The pairing of PSNR and LPIPS allows us to measure the inherent trade-off in diffusion-based restoration: maintaining strict data consistency (PSNR) while leveraging the prior to produce high-fidelity, natural-looking images (LPIPS). These metrics are used to compare the heuristic approach of SDEdit against the more rigorous posterior sampling of ScoreALD and DPS.

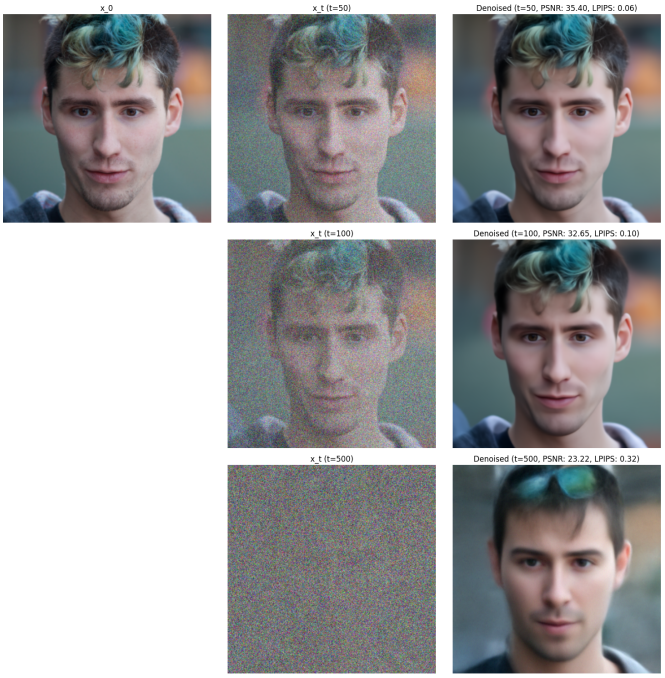

Fig. 1. Image denoising using a pretrained diffusion model for three different noise levels.

Fig. 2. Examples of images generated from pure gaussian noise using a pretrained diffusion model.

TABLE 1
PSNR and LPIPS values for the denoised images in Figure 1

Noise Level	PSNR	LPIPS
$t = 50$	35.40	0.06
$t = 100$	32.65	0.10
$t = 500$	23.22	0.32

5 Results

5.1 Image Denoising using a Pretrained Diffusion Model

We evaluate the effectiveness of the score-based diffusion model as a Gaussian denoiser by implementing a single-step reverse process. The implementation involves establishing a forward noising process to corrupt a clean test image x_0 with additive Gaussian noise. We then implement the reverse process using a single-step model prediction to reconstruct the clean image from its noisy counterpart.

As shown in Figure 1, the single-step denoising results are evaluated across three noise levels: $t = 50$, $t = 100$, and $t = 500$. The corresponding Peak Signal-to-Noise Ratio (PSNR) and Learned Perceptual Image Patch Similarity (LPIPS) for each reconstruction are summarized in Table 1. The denoising performance is strongest at lower noise levels, with higher PSNR and lower LPIPS values. However, as the noise increases (e.g., at $t = 500$), the diffusion model's ability to recover the original signal diminishes. The resulting reconstructions appear blurrier and exhibit a noticeable loss of details compared to the ground truth.

5.2 Image Generation using a Pretrained Diffusion Model

Now, we explore the generative capabilities of the pretrained diffusion prior by implementing unconditional generation. Unlike the single-step denoising performed previously, this process involves initializing the image as pure Gaussian noise and iteratively applying the DDPM denoising procedure. This iterative approach allows the model to turn random noise into a coherent face samples in FFHQ data manifold. As shown in Figure 2, the model is capable of generating a wide variety of facial images from nothing but noise.

5.3 Solving Inverse Problems with SDEdit

Now we leverage the DDPM denoising capabilities of our pretrained model to implement the SDEdit (Score-Distillation Editing) method for solving two linear inverse problems: image inpainting and deconvolution. SDEdit leverages the diffusion prior by starting the reverse process from a partially noised version of the degraded measurement. By adding Gaussian noise to the masked or blurry input up to a specific timestep t and then performing the iterative DDPM denoising procedure from there, the model "projects" the corrupted observation back onto the manifold of face images while maintaining the structural constraints of the measurement.

We provide the ground truth, the masked or blurred input, and the results of running SDEdit at three different noise levels ($t = 200$, $t = 500$, $t = 800$) in Figure 3 and Figure 4. The PSNR and LPIPS for all six reconstructions are detailed in Table 2.

Our results show the critical trade-off inherent in the choice of the noise level t: At lower noise levels ($t = 200$), reconstructions exhibit high data fidelity (higher PSNR and

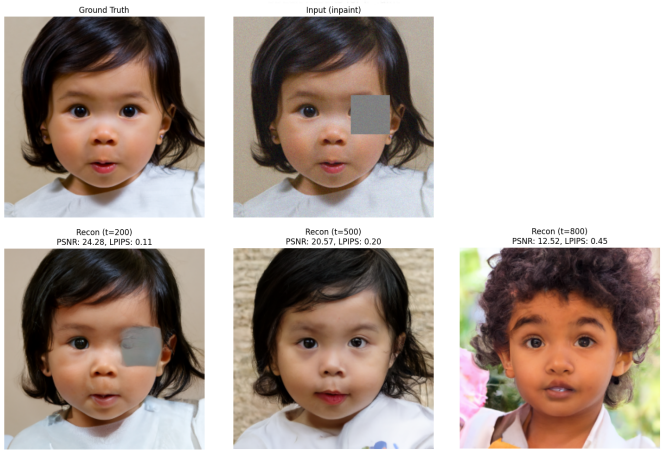

Fig. 3. Results of SDEdit on an inpainting problem for three different noise levels.

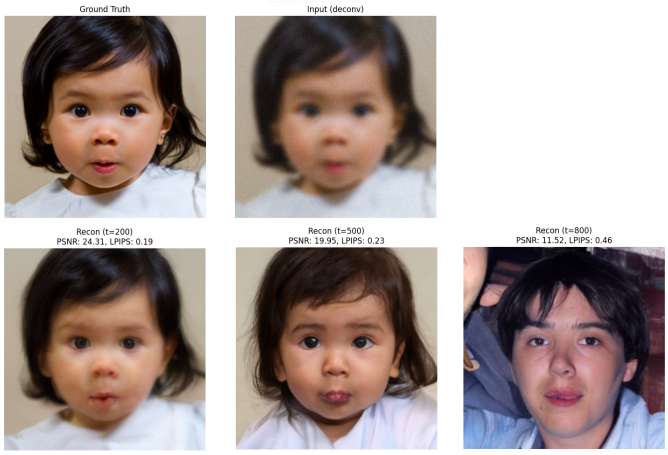

Fig. 4. Results of SDEdit on a deconvolution problem for three different noise levels.

lower LPIPS) because the starting point is close to the original measurement. However, because there was not much noise to remove, the resulting output may not successfully remove the artifacts in the original noisy measurement. This is most noticeable in the SDEdit output for inpainting at noise level $t = 200$.

At higher noise levels ($t = 800$), the model will successfully remove the mask for inpainting and deblur for deconvolution, but with too much noise, the original structural constraints from the measurement are lost, so the resulting image deviates greatly from the original measurement. In both the examples for inpainting and deconvolution, the PSNR value is lower and the LPIPS value is high, indicating that the output image bears little resemblance to the original image.

Picking a middle ground for the noise level ($t = 500$) successfully balances removing the noise from the measurement while remaining a certain level of resemblance to the provided measurements. Note that even in this case, the face in the output differs slightly from the original for both inpainting and deconvolution. In order to produce reconstructions that better match the original, we explore two more methods in the following sections.

TABLE 2
PSNR and LPIPS values for the SDEdit outputs in Figures 3 and 4.

Noise Level	Inpainting		Deconvolution	
	PSNR	LPIPS	PSNR	LPIPS
$t = 200$	24.28	0.11	24.31	0.19
$t = 500$	20.57	0.20	19.95	0.23
$t = 800$	12.52	0.45	11.52	0.46

TABLE 3
PSNR and LPIPS values for the ScoreALD outputs in Figure 5.

	Inpainting (Box Mask)	Inpainting (Random Mask)	Deconvolution
PSNR	23.24	20.43	22.65
LPIPS	0.13	0.23	0.13

5.4 Solving Inverse Problems with ScoreALD

In this section, we evaluate the ScoreALD (Score-based Annealed Langevin Dynamics) algorithm as a posterior sampling approach for image restoration. Unlike the partial-noising approach of SDEdit, ScoreALD incorporates measurement information directly into the diffusion sampling trajectory through likelihood gradients and annealed dynamics. This allows the model to compute a different sample from the feasible set of solutions each time it is executed, providing a more rigorous exploration of the posterior distribution.

As illustrated in Figure 5, we evaluate ScoreALD across three distinct scenarios: image inpainting with a box mask, image inpainting with a random mask, and image deconvolution. The PSNR and LPIPS metrics for each of these tasks are summarized in Table 3.

The face results produced by ScoreALD bear a much stronger resemblance to the original ground truth image than the results from SDEdit. The PSNR and LPIPS values are also better across all problems, indicating that ScoreALD outperforms SDEdit even when we pick optimal noise levels for SDEdit.

By integrating the noisy measurements at each step, we can consistently nudge the estimated image back toward the original observation. This prevents the issue encountered in SDEdit, where the initial noise can "wash out" the underlying measurement, causing the model to lose track of the original image structure during the reconstruction.

5.5 Solving Inverse Problems with DPS

In this section, we evaluate the DPS (Diffusion Posterior Sampling) algorithm as a posterior sampling approach for solving inverse problems. Similar to ScoreALD, DPS leverages the posterior sampling interpretation of the diffusion model to incorporate measurement information during the reverse diffusion process. However, DPS distinguishes itself by utilizing normalized likelihood gradients, which ensures more stable sampling across the trajectory.

As illustrated in Figure 6, we evaluated DPS in three distinct scenarios: image inpainting with a box mask, image inpainting with a random mask, and image deconvolution.

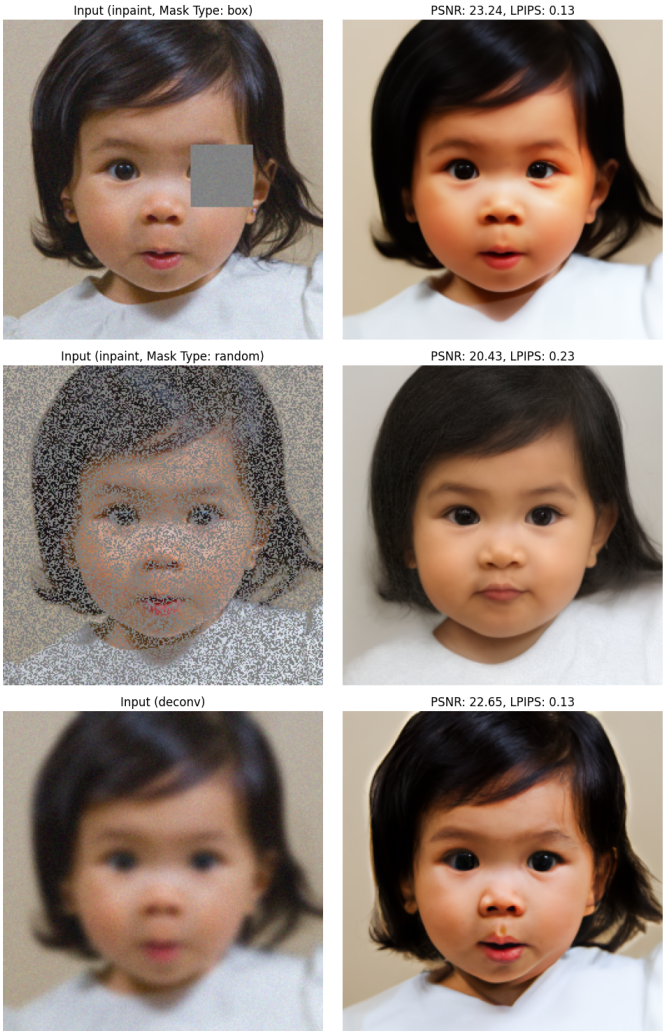

Fig. 5. Results of ScoreALD on inpainting and deconvolution problems.

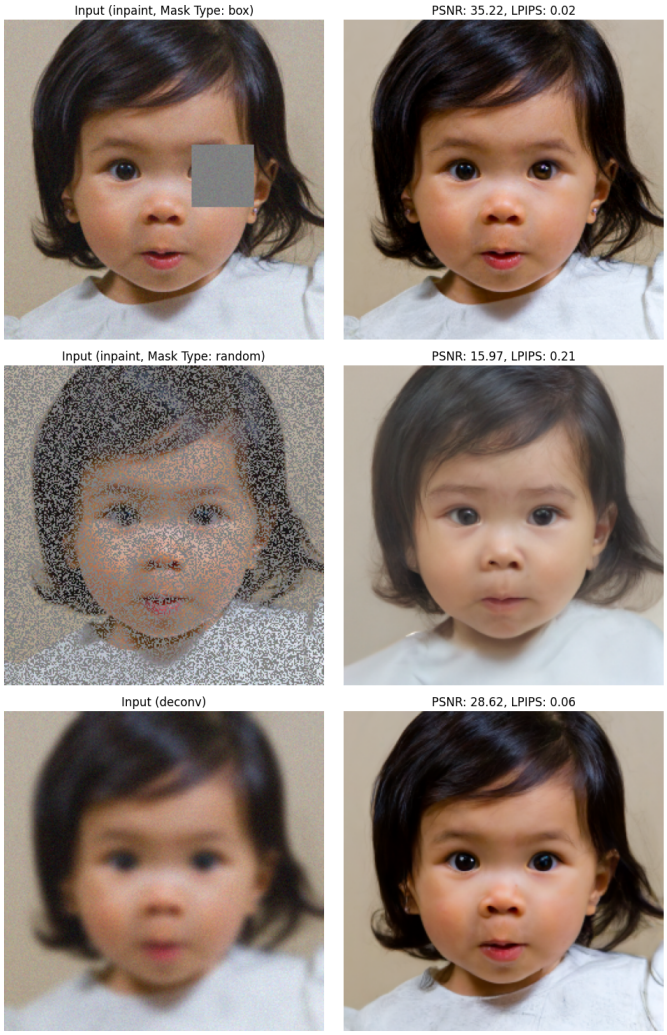

Fig. 6. Results of DPS on inpainting and deconvolution problems.

The PSNR and LPIPS metrics for each of these tasks are summarized in Table 4. The reconstruction quality was found to be sensitive to the scale parameter ζ (tested within the 0.1–1.0 range). We used the following values of ζ: $\zeta = 1.0$ for inpainting with a box mask, $\zeta = 0.1$ for inpainting with a random mask, and $\zeta = 0.3$ for deconvolution.

For inpainting with a box mask and deconvolution, DPS outperforms ScoreALD (and all other methods explored in this paper), achieving the highest PSNR and lowest LPIPS values. Qualitatively, the generated outputs exhibit the highest fidelity to the ground truth. For the random mask inpainting task, while PSNR and LPIPS suggest DPS underperforms relative to ScoreALD, the facial structure in Figure 6 appears more anatomically consistent with the ground truth than the ScoreALD output in Figure 5. This gap is likely due to "washed out" colors in the DPS output — a byproduct of the noisy input measurements — which penalizes pixel-wise metrics like PSNR, despite the superior structural output of DPS.

TABLE 4
PSNR and LPIPS values for the DPS outputs in Figure 6.

	Inpainting (Box Mask)	Inpainting (Random Mask)	Deconvolution
PSNR	35.22	15.97	28.62
LPIPS	0.02	0.21	0.06

6 Conclusion

6.1 Discussion and Comparison of Methods

Across the various approaches to inverse imaging problems, we observed a clear progression in performance. While SDEdit is computationally simple and effective for basic editing, its reliance on the noise initialization often led to a loss of resemblance to the ground truth. In contrast, ScoreALD and DPS showed that by integrating the measurement at every step of the reverse diffusion process, the output improves drastically. Specifically, the normalized gradient updates in DPS provided the most stable reconstructions, effectively balancing the diffusion prior with the physical constraints of inpainting and deconvolution.

6.2 Limitations and Future Work

Despite the quality of the results, several limitations remain. One limitation is computational efficiency - performing 1,000 iterative steps for a single image reconstruction is time-intensive, making real-time application difficult. Future work could involve implementing faster sampling schedules. Additionally, since the pretrained model was trained exclusively on the FFHQ dataset, its capabilities are currently limited to human faces and do not generalize well to other image domains without further fine-tuning.

6.3 Conclusion

This project explored the application of score-based diffusion models as priors for inverse imaging problems. By transitioning from unconditional generation to complex posterior sampling, we demonstrated that a pretrained diffusion model can be effectively leveraged to solve inverse imaging problems, including denoising, inpainting, and deconvolution. The posterior sampling methods such as ScoreALD and DPS yielded the strongest overall performance, indicating that maintaining a relationship between the generative prior and the observed measurements will allow us to achieve the best results.

REFERENCES

[1] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33. Curran Associates, Inc., 2020, pp. 6840–6851. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf

[2] 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," in *International Conference on Learning Representations (ICLR)*, 2022.

[3] A. Jalal, M. Arvinte, G. Daras, E. Price, A. G. Dimakis, and J. I. Tamir, "Robust compressed sensing mri with deep generative priors," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

[4] H. Chung, J. Kim, M. T. Mccann, M. L. Klasky, and J. C. Ye, "Diffusion posterior sampling for general noisy inverse problems," in *International Conference on Learning Representations (ICLR)*, 2023.