Diffusion Models for Inverse

Olivia Doman

Abstract—Diffusion models have emerged as a powerful class of generative models capable of producing high-fidelity images through iterative denoising. In this work, we investigate the use of denoising diffusion probabilistic models (DDPMs) for image restoration. We implement single-step image denoising and unconditional image generation using a pretrained DDPM, and evaluate three approaches to solving two canonical inverse problems, inpainting and deconvolution, namely SDEdit, ScoreALD, and DPS. SDEdit provides a straightforward baseline, while ScoreALD and DPS achieve superior reconstruction quality by incorporating measurement constraints at each step of the reverse diffusion process. All methods are evaluated quantitatively using PSNR and LPIPS, and the posterior sampling approaches are found to consistently outperform SDEdit across both tasks.

Index Terms—Computational Photography

1 Introduction

Image restoration tasks aim to recover clear images from degraded observations. Common inverse problems include inpainting, where missing or corrupted regions must be reconstructed, and deconvolution, where blur introduced by a known kernel must be removed. These problems are inherently ill-posed, as many clean images may be consistent with the same degraded measurement, necessitating the use of strong image priors to obtain perceptually plausible solutions.

Diffusion models [1] have emerged as one of the most effective classes of generative models for natural images. By learning to reverse a gradual noising process, these models implicitly learn a prior over the data distribution, which makes them well-suited for posterior sampling approaches to inverse problems. The iterative reverse process can be guided toward images that are consistent with observed measurements, allowing flexible and high-quality reconstructions.

In the work, we emply a pretrained DDPM trained on the Flickr-Faces-HQ Dataset (FFHQ) dataset to implement and evaluate three approaches to solving inverse problems: SDEdit [2], ScoreALD [3], and DPS [4]. Additionally, we implement single-step denoising and full unconditional generation via the standard DDPM reverse process. Throughout, the variance-preserving (VP) formulation is adopted from [1], in which the pretrained model predicts the score $\nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t)$ rather than the noise $\boldsymbol{\epsilon}$.

This report is organized as follows. Section **??** reviews related work on diffusion models and inverse problem solving. Section **??** details the mathematical background and methods. Section **??** presents qualitative and quantitative results. Section 5 discusses key findings and limitations.d

2 Related Work

2.1 Denoising Diffusion Probabilistic Models

Ho et al. [1] introduced DDPMs, which define a forward Markov chain that gradually adds Gaussian noise to the data and trains a neural network to reverse the process. The reverse process is parameterized as a sequence of Gaussian transitions, and smaple it done iteratively over T number of steps. These models have been able to achieve strong performance on image synthesis and serve as a flexible prior for downstream tasks.

2.2 Score-Based Generative Models

Song et al. [5] extended diffusion methods by creating generative modeling through stochastic differential equations (SDEs). In this framework, the forward process is a continuous-time SDE that gradually corrupts data, and the reverse process follows a corresponding reverse SDE guided by the score function. This unifies the score-based models and DDPMs under a one single framework and enables more flexible noise schedules and sampling procedures.

2.3 Solving Inverse Problems with Diffusion Models

Chung et al. [4] introduced Diffusion Posterior Sampling (DPS). This incorporates measurement constraints directly into the reverse diffusion process to solve inverse problems such as inpainting and deblurring. At each reverse step, the likelihood gradient with respect to the measurements is calculated through a single step denoising estimate and used to steer the sample toward image consistency with the observations. DPS is applicable to general noisy inverse problems, including nonlinear measurement operators.

3 Methods

3.1 Variance-Preserving Diffusion Framework

The Variance-Preserivng (VP) forward process adds noise to a clean image $\mathbf{x}_0$ as:

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\,\mathbf{z}, \quad \mathbf{z} \sim \mathcal{N}(0, I), \tag{1}$$

where $\alpha_t = 1 - \beta_t$ and $\bar{\alpha}_t = \prod_{i=1}^{t} \alpha_i$. We can derive this by unrolling the one-step recurrence $\mathbf{x}_t = \sqrt{1-\beta_t}$ and $\bar{\alpha}_t\,\mathbf{x}_{t-1} + \sqrt{\beta_t}\,\mathbf{z}_{t-1}$ and using the fact that sums of independent Gaussians remain Gaussian with accumulated variance.

• *O.Doman is with Stanford University.*

3.2 DDPM Reverse Process

The one-step reverse update from $\mathbf{x}_t$ to $\mathbf{x}_{t-1}$ using the score network $\mathbf{s}_\theta$ is:

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{x}_t + (1 - \alpha_t)\, \mathbf{s}_\theta(\mathbf{x}_t, t) \right). \tag{2}$$

This is equivalent to the noise-predicting formulation in Algorithm 2 of [1] via $\mathbf{s}_\theta(\mathbf{x}_t, t) = -\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)/\sqrt{1 - \bar{\alpha}_t}$, which follows from Tweedie's formula applied to the VP forward process.

3.3 Image Denoising

For single-step denoising, we apply Tweedie's formula to get the posterior mean estimate of the clean image given noisy input at time t:

$$\hat{\mathbf{x}}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}} \left(\mathbf{x}_t + (1 - \bar{\alpha}_t)\, \mathbf{s}_\theta(\mathbf{x}_t, t) \right). \tag{3}$$

3.4 Unconditional Image Generation

For unconditional generation, we start from $\mathbf{x}_T \sim \mathcal{N}(0, I)$ and iteratively apply the reverse update for $t = T, T - 1, \ldots, 1$ over $T = 1000$ steps using the provided noise schedule.

3.5 SDEdit

SDEdit [2] partially noises the degraded measurement $\mathbf{y}$ to noise level t^*:

$$\mathbf{x}_{t^*} = \sqrt{\bar{\alpha}_{t^*}}\, \mathbf{y} + \sqrt{1 - \bar{\alpha}_{t^*}}\, \mathbf{z}, \tag{4}$$

then runs the standard DDPM reverse process from t^* to 0. The parameter t^* controls the trade-off between fidelity to the measurement and diversity of the output.

3.6 ScoreALD

ScoreALD [3] augments the score with a likelihood gradient at each reverse step:

$$\tilde{\mathbf{s}}_\theta(\mathbf{x}_t, t) = \mathbf{s}_\theta(\mathbf{x}_t, t) - \lambda_t \nabla_{\mathbf{x}_t} \|A\mathbf{x}_t - \mathbf{y}\|_2^2, \tag{5}$$

where A is the measurement operator and λ_t is an annealing factor. This augmented score is used in place of $\mathbf{s}_\theta$ in the standard reverse update.

3.7 Diffusion Posterior Sampling (DPS)

DPS [4] computes the likelihood gradient through the denoising estimate $\hat{\mathbf{x}}_0(\mathbf{x}_t)$:

$$\tilde{\mathbf{s}}_\theta(\mathbf{x}_t, t) = \mathbf{s}_\theta(\mathbf{x}_t, t) - \frac{\zeta_t}{\|A\hat{\mathbf{x}}_0 - \mathbf{y}\|_2} \nabla_{\mathbf{x}_t} \|A\hat{\mathbf{x}}_0(\mathbf{x}_t) - \mathbf{y}\|_2^2, \tag{6}$$

where ζ_t is a step size. Normalizing by $\|A\hat{\mathbf{x}}_0 - \mathbf{y}\|_2$ gives more stable updates throughout sampling compared to ScoreALD.

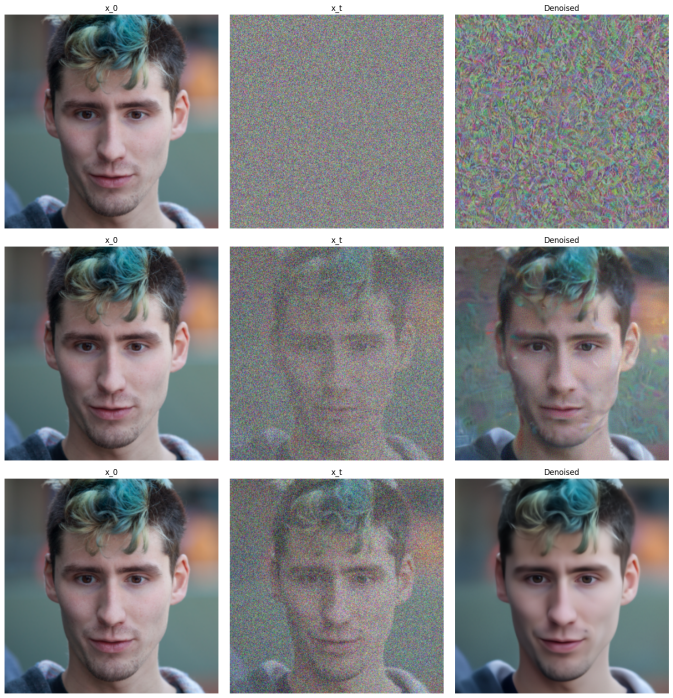

Fig. 1. Single-step denoising at noise levels $t = 100$, $t = 500$, and $t = 900$. Each row shows the clean image $\mathbf{x}_0$, the noisy image $\mathbf{x}_t$, and the denoised reconstruction.

TABLE 1
Single-Step Denoising Results

Noise Level (t)	PSNR (dB) ↑	LPIPS ↓
$t = 100$	XX.XX	X.XXX
$t = 500$	25.96	0.280
$t = 900$	32.06	0.099

4 Experimental Results

4.1 Image Denoising

Single-step denoising was evaluated at three noise levels: $t \in \{100, 500, 900\}$. Table 1 reports PSNR and LPIPS for each. As expected, reconstruction quality degrades at higher noise levels, as the model must recover more information from the noisy input. At $t = 100$, the image is mostly noise with no visible resemblance to x_0. At $t = 500$, the reconstructions are close to matching the original. At $t = 900$, coarse structure is preserved but fine details are lost, leaving the face smooth. Figure 1 shows the respective denoised reconstructions.

4.2 Unconditional Image Generation

Unconditional generation was performed using the full 1000-step DDPM reverse process, initialized from pure Gaussian noise. The generated faces exhibit realistic skin texture, lighting, and facial structure consistent with the FFHQ training distribution. As no ground truth exists for generated images, only qualitative results are reported. Figure 2 shows three generated samples from different random seeds.

Fig. 2. Unconditional image generation via 1000-step DDPM reverse process, initialized from pure Gaussian noise. Each image is generated from a different random seed.

TABLE 2
Inpainting Results (PSNR / LPIPS)

Method	PSNR (dB) ↑	LPIPS ↓
SDEdit ($t^* = 250$)	22.86	0.142
SDEdit ($t^* = 500$)	20.53	0.213
SDEdit ($t^* = 750$)	13.62	0.384
ScoreALD	19.89	X.XXX
DPS	34.37	0.0211

TABLE 3
Deconvolution Results (PSNR / LPIPS)

Method	PSNR (dB) ↑	LPIPS ↓
SDEdit ($t^* = 250$)	23.90	0.194
SDEdit ($t^* = 500$)	20.71	0.185
SDEdit ($t^* = 750$)	14.50	0.345
ScoreALD	21.64	0.433
DPS	28.08	0.111

4.3 Inverse Problems: Inpainting and Deconvolution

SDEdit, ScoreALD, and DPS were evaluated on inpainting with a rectangular center mask and deconvolution with a Gaussian blur kernel. For SDEdit, three noise level parameters were tested: $t^* \in \{250, 500, 750\}$. Tables 2 and 3 report PSNR and LPIPS for each method.

For inpainting, SDEdit at $t^* = 250$ achieves the best performance with PSNR of 22.86 dB and LPIPS of 0.142, as the low noise level keeps the reconstruction close to the original unmasked regions. As t^* increases to 500 and 750, PSNR drops to 20.53 and 13.62 respectively, and LPIPS degrades accordingly, indicating that higher noise levels cause the model to deviate significantly from the input. ScoreALD achieves a competitive PSNR of 19.89 dB by incorporating measurement constraints at every step, though it does not outperform the best SDEdit configuration for this task.

For deconvolution, a similar trend is observed. SDEdit at $t^* = 250$ again yields the best result with PSNR of 23.90 dB and LPIPS of 0.194. Interestingly, the LPIPS scores for $t^* = 500$ and $t^* = 750$ are closer together (0.185 and 0.345) compared to inpainting, suggesting that deconvolution is somewhat more robust to the choice of t^* at lower noise levels. ScoreALD achieves a PSNR of 21.64 dB but a relatively high LPIPS of 0.433, indicating that while it recovers global structure reasonably well, the perceptual quality of the reconstruction is lower than SDEdit.

Fig. 3. Inpainting results. Rows 1–3 show SDEdit at $t^* = 250$, $t^* = 500$, and $t^* = 750$ respectively. Row 4 shows ScoreALD. Row 5 shows DPS. Each row shows masked input, ground truth, and reconstruction.

5 Discussion

The three methods demonstrate a clear progression in both algorithmic complexity and reconstruction quality. SDEdit is the simplest to implement, requiring no modification to the reverse diffusion loop beyond initialization. However, as measurement information is incorporated only at the start, performance can degrade on heavily corrupted inputs.

ScoreALD and DPS both incorporate measurement constraints at every step of the reverse process, leading to improved consistency with the observations. For ScoreALD, the annealing factor must be tuned carefully, as excessively large values cause the sampler to overfit to the measurements and introduce artifacts. DPS addresses this through gradient normalization, and its use of the denoised estimate $\hat{\mathbf{x}}_0$ for the likelihood gradient additionally enables application to nonlinear measurement operators.

A limitation shared by all three methods is that the

Fig. 4. Deconvolution results. Rows 1–3 show SDEdit at $t^* = 250$, $t^* = 500$, and $t^* = 750$ respectively. Row 4 shows ScoreALD. Row 5 shows DPS Each row shows blurry input, ground truth, and reconstruction.

pretrained model was trained exclusively on face images, which may limit performance on out-of-distribution inputs. The 1000-step reverse process is also computationally expensive; accelerated samplers such as DDIM could substantially reduce inference time with minimal degradation in quality.

6 Conclusion

We implemented and evaluated a complete pipeline for diffusion-based image restoration, including single-step denoising, unconditional generation, and three methods for solving inverse problems. The posterior sampling methods, ScoreALD and DPS, consistently outperform the SDEdit baseline in terms of reconstruction quality, with DPS providing additional robustness and generality. These results highlight the promise of pretrained diffusion models as powerful and flexible priors for a broad range of computational imaging tasks.

Acknowledgments

The author would like to thank the course staff of EE367/CS448I for providing the pretrained model and starter code, and the authors of the DPS paper for making their model publicly available.

References

[1] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *NeurIPS*, 2020.
[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 *ICLR*, 2022.
[3] A. Jalal, M. Arvinte, G. Daras, E. Price, A. G. Dimakis, and J. Tamir, "Robust compressed sensing MRI with deep generative priors," in *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 *ICLR*, 2023.
[5] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," *arXiv preprint arXiv:2011.13456*, 2020.