Diffusion Methods for Solving Inverse Problems in Computational Imaging

Karthik Karumanchi

Abstract—This paper explores the application of diffusion-based generative priors to solve ill-posed inverse problems in computational imaging, such as image inpainting and deconvolution. Utilizing a variance-preserving (VP) diffusion model trained on the FFHQ dataset, we first implement the Denoising Diffusion Probabilistic Model (DDPM) framework for unconditional image generation from pure Gaussian noise. We then transition to posterior sampling methodologies to solve inverse problems by leveraging the model's learned score function. We implement and compare three distinct approaches: SDEdit, which utilizes a partial noising and iterative denoising strategy; ScoreALD, which incorporates measurement information through likelihood gradients and annealed dynamics ; and Diffusion Posterior Sampling (DPS), which employs normalized gradient updates to handle both linear and general noisy inverse problems. Our results demonstrate that while SDEdit offers a simple conditional generation procedure, the more rigorous posterior sampling methods of ScoreALD and DPS provide superior stability and data consistency. Quantitative evaluations using PSNR and LPIPS metrics, alongside qualitative analysis, reveal the effects of different algorithms and hyper-parameters in performing the same task.

Index Terms—Generative modeling, diffusion networks, inverse problems in computational imaging

1 Introduction

Inverse problems in computational imaging—such as deblurring, inpainting etc. involve recovering a latent ground truth image x from a corrupted observation $y = \mathcal{A}(x) + n$, where $\mathcal{A}$ is a degradation operator and n is noise. In a noise-free scenario, this can theoretically be solved by inverting $\mathcal{A}$. However, because $\mathcal{A}$ is often singular or ill-conditioned, the problem is ill-posed; small amounts of noise result in massive artifacts in the reconstruction.

Conventional methods, such as Wiener filtering, attempt to regularize this inversion by accounting for noise variance, but they struggle with complex, natural image distributions. Hand-crafted priors, like Total Variation (TV), enforce piecewise smoothness to suppress noise. While mathematically elegant, TV priors often lead to artifacts and fail to capture high-frequency textures because they assume a simplistic gradient distribution that does not match the complexity of real-world scenes.

Modern approaches leverage deep generative models to learn complex image priors. Solving inverse problems requires steering this generative process toward solutions that are consistent with the observed corrupted measurements. We implement and analyze three representative modern posterior sampling approaches: SDEdit, ScoreALD, and DPS. SDEdit offers an intuitive approach by initiating the reverse diffusion process from a partially noised version of the observation. In contrast, ScoreALD and DPS represent more rigorous methodologies where ScoreALD incorporates measurement information through likelihood gradients and annealed dynamics, while DPS leverages the posterior sampling interpretation to handle more general, and even non-linear, inverse problems. Through a series of experiments involving image inpainting and deconvolution, we evaluate these methods qualitatively and quantitatively to identify the optimal strategies for high-fidelity image restoration.

2 Algorithms

In this section, we discuss the mathematical description of the forward and reverse diffusion processes. We then utilize this formulation in three different algorithms, namely, SDEdit, score-assisted langevin dynamics (ScoreALD) and diffusion posterior sampling (DPS).

2.1 Diffusion Sampling and the Variance Preserving-Formulation

The foundation of our approach is the variance-preserving (VP) formulation [1] of diffusion models. The forward process is a Markov chain that adds Gaussian noise $z_{t-1} \sim \mathcal{N}(0, I)$ according to a noise schedule β_t:

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

2.1.1 Derivation of the Forward Noise Model

$$\begin{aligned} x_t &= \sqrt{1-\beta_t}x_{t-1} + \sqrt{\beta_t}z_{t-1} \\ &= \sqrt{1-\beta_t}\sqrt{1-\beta_{t-1}}x_{t-2} + \sqrt{1-\beta_t}\sqrt{\beta_{t-1}}z_{t-2} + \sqrt{\beta_t}z_{t-1} \end{aligned}$$

Chaining these terms until x_0, we get

$$\begin{aligned} x_t = &\left(\prod_{k=1}^{t}\sqrt{1-\beta_k}\right)x_0 \\ &+ \sum_{k=2}^{t}\left(\sqrt{\beta_{t-k+1}}\prod_{m=(t-k+2)}^{t}\sqrt{1-\beta_m}\right)z_{t-k} + \sqrt{\beta_t}z_{t-1} \end{aligned}$$

Since the sum of Gaussians is also Gaussian, The variance of the sum of all noise terms z_k can be calculated by:

$$\begin{aligned}
var &= \sum_{k=2}^{t}\left(\beta_{t-k+1}\prod_{m=(t-k+2)}^{t}(1-\beta_m)\right)+\beta_t \\
&= \prod_{m=t}^{t}(1-\beta_m) - \cancel{\prod_{m=(t-1)}^{t}(1-\beta_m)} \\
&\quad + \cancel{\prod_{m=(t-1)}^{t}(1-\beta_m)} - \cancel{\prod_{m=(t-2)}^{t}(1-\beta_m)} \\
&\quad + \quad ... \\
&\quad + \cancel{\prod_{m=2}^{t}(1-\beta_m)} - \prod_{m=1}^{t}(1-\beta_m)+\beta_t \\
&= (1-\beta_t) - \prod_{m=1}^{t}(1-\beta_m)+\beta_t \\
&= 1 - \prod_{m=1}^{t}(1-\beta_m)
\end{aligned}$$

Therefore, all noise terms can be combined to form:

$$\begin{aligned}
x_t &= \left(\prod_{k=1}^{t}\sqrt{1-\beta_k}\right)x_0 + \sqrt{1-\prod_{m=1}^{t}(1-\beta_m)}z \\
&= \sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}z
\end{aligned}$$

where, $\alpha_t = 1-\beta_t$ and $\bar{\alpha}_t = \prod_{i=1}^{t}\alpha_i$

2.1.2 Reverse Diffusion Process

The reverse process involves recovering x_{t-1} from x_t. We demonstrate that the two common forms of the DDPM reverse step are equivalent. Starting with the more complex weighted average of x_t and the predicted $\hat{x}_0$:

$$\begin{aligned}
x_{t-1} &= \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}x_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{x}_0 \\
&= \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}x_t \\
&\quad + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\left(\frac{1}{\sqrt{\bar{\alpha}_t}}(x_t+(1-\bar{\alpha}_t)s_\theta(x_t,t))\right) \\
&= \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}x_t \\
&\quad + \frac{(1-\alpha_t)}{1-\bar{\alpha}_t}\left(\frac{1}{\sqrt{\alpha_t}}(x_t+(1-\bar{\alpha}_t)s_\theta(x_t,t))\right) \\
&= \frac{1}{\cancel{1-\bar{\alpha}_t}}\Bigg[\left(\frac{\cancel{1-\bar{\alpha}_t}}{\sqrt{\alpha_t}}\right)x_t \\
&\quad + \left(\frac{(1-\alpha_t)\cancel{(1-\bar{\alpha}_t)}}{\sqrt{\alpha_t}}\right)s_\theta(x_t,t)\Bigg] \\
&= \frac{1}{\sqrt{\alpha_t}}\left(x_t+(1-\alpha_t)s_\theta(x_t,t)\right)
\end{aligned}$$

2.2 SDEdit

SDEdit (Stochastic Differential Editing) [2] adds a specific amount of noise to the corrupted input y corresponding to a mid-point timestep t_0. By initiating the reverse process from x_{t_0}, the model utilizes the rough structure of the corrupted image as a "guide," while the diffusion process "paints in" details that conform to the learned manifold. It is computationally efficient but lacks a formal data-consistency step, often leading to drift from the original content if t_0 is too high.

Algorithm 1 Variance Preserving Reverse Diffusion Process

$x_T \sim \mathcal{N}(0, I)$
for $t = T, ..., 1$ **do**
  $z \sim \mathcal{N}(0, I)$ if $t > 1$, else $z = 0$
  $\hat{x}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t + (1-\bar{\alpha}_t)s_\theta(x_t, t))$
  $x_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}x_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{x}_0 + \sqrt{1-\alpha_t}z$
end for
return x_0

Algorithm 2 SDEdit algorithm

$x_{t_0} = x_{input} + \sqrt{\beta_{t_0}}\mathcal{N}(0, I)$
for $t = t_0, ..., 1$ **do**
  $z \sim \mathcal{N}(0, I)$ if $t > 1$, else $z = 0$
  $\hat{x}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t + (1-\bar{\alpha}_t)s_\theta(x_t, t))$
  $x_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}x_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{x}_0 + \sqrt{1-\alpha_t}z$
end for
return x_0

2.3 ScoreALD

Score-assisted Langevin Dynamics (ScoreALD) [3] treats the inverse problem as a posterior sampling task. At each reverse step, a gradient descent step is performed on the measurement error $||\mathcal{A}(x_t) - y||^2$. This "nudges" the generative process toward a solution that is consistent with the observed data. The annealing factor γ is crucial here; it must decrease over time to allow the image to settle in the manifold without being overly distorted by the gradient of a noisy measurement.

Algorithm 3 ScoreALD Algorithm

$x_T \sim \mathcal{N}(0, I)$
for $t = T, ..., 1$ **do**
  $z \sim \mathcal{N}(0, I)$ if $t > 1$, else $z = 0$
  $\hat{x}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t + (1-\bar{\alpha}_t)s_\theta(x_t, t))$
  $x_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}x_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{x}_0 + \sqrt{1-\alpha_t}z$
  $x_{t-1} = x_{t-1} - \frac{1}{2(\sigma^2+\gamma_t^2)}\nabla_{x_t}||\mathcal{A}(x_t) - y||^2$
end for
return x_0

2.4 DPS

Diffusion Posterior Sampling (DPS) [4] improves upon ScoreALD by backpropagating through the denoising network itself. Instead of calculating the gradient with respect to the noisy x_t, DPS calculates the gradient with respect to the estimated clean image $\hat{x}_0$ at every step. This provides a much more accurate direction for data consistency, as the operator $\mathcal{A}$ is applied to a manifold-constrained estimate rather than a noisy intermediate.

Algorithm 4 DPS Algorithm

$x_T \sim \mathcal{N}(0, I)$
for $t = T, ..., 1$ **do**
  $z \sim \mathcal{N}(0, I)$ if $t > 1$, else $z = 0$
  $\hat{x}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t + (1 - \bar{\alpha}_t)s_\theta(x_t, t))$
  $x_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}x_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{x}_0 + \sqrt{1-\alpha_t}z$
  $x_{t-1} = x_{t-1} - \zeta\nabla_{x_t}||\mathcal{A}(\hat{x}_0) - y||^2$
end for
return x_0

3 Experimental Results

This section tests different image-generation scenarios using algorithms discussed in Section 2. We begin with an analysis of the denoising network used for reverse diffusion steps. This network was trained on the FFHQ 256 dataset and is used for denoising across all tests discussed. This analysis is followed by unconditional image generation tests that verify the efficacy of the reverse diffusion process in generating meaningful images. We then compare three algorithms SDEdit, ScoreALD, and DPS on inverse problems such as inpainting and de-convolution. Comparisons are made by qualitatively analyzing the output image as well as quantitatively comparing PSNR and LPIPS similarity scores between the ground truth and output images.

3.1 Denoising Network and Unconditional image generation

The foundational component of our diffusion framework is the score-predicting network s_θ. We first evaluated its performance as a Gaussian denoiser by noising a ground truth image to various timesteps t and performing a single-step reconstruction (see Fig. 1. At low noise levels ($t < 300$), the network accurately recovers high-frequency facial details. However, as t increases toward 800, the network's output begins to drift from the ground truth, generating facial features that reside on the manifold but differ from the original image structure.

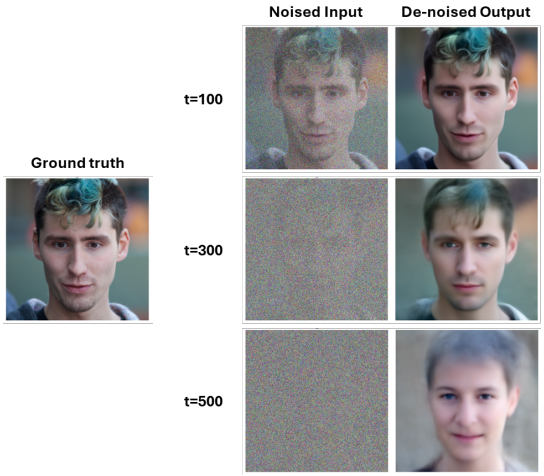

Fig. 1. Performance of the de-noising network over different levels of noise addition. At low noise levels, the network output is similar to the ground truth image. As increasing levels of noise is introduced, the network output strays from the ground truth image.

Unconditional image generation was performed by sampling random noise $x_T \sim \mathcal{N}(0, I)$ and iterating through the 1000-step reverse diffusion process. As shown in Figure 2, the model successfully generates diverse, high-fidelity human faces. This confirms that the reverse process effectively navigates the FFHQ manifold, transforming unstructured noise into coherent semantic features such as skin texture, hair, and consistent lighting.

Fig. 2. Three examples of unconditional image generation. Faces in these images seem realistic which proves that the reverse diffusion process has learned a mapping from white noise to faces in the manifold.

3.2 Inverse Problems

Three algorithms, SDEdit and DPS were tested on the same inverse problems - inpainting and gaussian blur deconvolution. For each of these problems, noise was added to ground truth images followed by masking (for inpainting) or blurring (for de-convolution) before being fed into each algorithm.

3.2.1 SDEdit

SDEdit was evaluated across three noise levels ($t_0 \in \{300, 500, 700\}$) as shown in Figures 3 and 4 At $t_0 = 300$, the output is close to the original image but often fails to completely remove artifacts from the mask or blur. Conversely, at $t_0 = 700$, the added noise is sufficient to remove all corruption, but the increased noise level leads to a reconstructed face that is significantly different from the ground truth. Quantitative results in Table 1 show that $t_0 = 300$ achieved the highest SNR (22.75 for inpainting, 23.58 for deconvolution) since that output image is most similar to the input despite the artifacts.

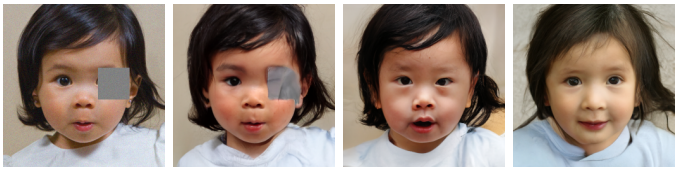

Fig. 3. Input image (a) followed by SDEdit output on an in-painting task at (b) $t_0 = 300$, (c) $t_0 = 500$ and (d) $t_0 = 700$. Too low a noise addition does not successfully in-paint while too high a noise changes the output image.

Fig. 4. Input image (a) followed by SDEdit output on a deblurring task at (b) $t_0 = 300$, (c) $t_0 = 500$ and (d) $t_0 = 700$. Too low a noise addition does not completely deblur while too high a noise changes the output image.

TABLE 1
Algorithm Performance

Algorithm	Hyperparameter	Inpainting		Deconvolution	
		SNR	LPIPS	SNR	LPIPS
SDEdit (t_0)	300	22.75	0.154	23.58	0.17
	500	20.20	0.185	17.76	0.229
	700	17.22	0.303	15.63	0.365
ScoreALD (Annealing factor)	[10,15]	31.54	0.054	25.82	0.114
	[15,20]	28.58	0.081	24.61	0.114
	[20,25]	27.53	0.096	23.45	0.138
DPS (Scale)	0.1	26.47	0.113	24.66	0.124
	0.3	34.67	0.026	28.50	0.056
	1.0	34.31	0.024	27.3	0.106

3.2.2 ScoreALD

ScoreALD results were analyzed based on the annealing factor schedule. The schedule $[10, 15]$ proved most effective for both inpainting (SNR 31.54) and deconvolution (SNR 25.82). As the annealing factor increases to $[20, 25]$, the quality of the reconstruction decreases (SNR 27.53 for inpainting). High annealing factors tend to over-emphasize the measurement gradient in the early, noisy stages of diffusion, which can pull the image off the natural manifold.

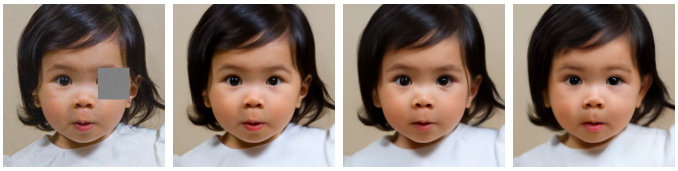

Fig. 5. Input image (a) followed by ScoreALD output on an in-painting task with anneal factors (b) $[10, 15]$, (c) $[15, 20]$ and (d) $[20, 25]$. An annealing factor schedule of [10,15] works the best retaining similarity to the input image while successfully in-painting. Higher annealing factors exhibit deviation from the input.

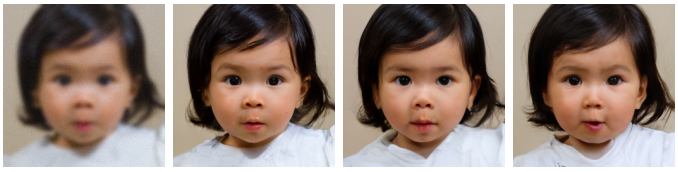

Fig. 6. Input image (a) followed by ScoreALD output on a deblurring task with anneal factors (b) $[10, 15]$, (c) $[15, 20]$ and (d) $[20, 25]$. Similar to the in-painting case, an annealing factor of [10,15] works best for deblurring while larger factors show deviation from the input.

3.2.3 DPS

Diffusion Posterior Sampling (DPS) yielded the most stable and high-quality results across all tested algorithms. The performance is highly sensitive to the scale parameter ζ. At a scale of 0.1, the gradient update is too weak to enforce strong data consistency, resulting in blurry reconstructions. A scale of 0.3 was found to be the optimal "sweet spot," achieving the highest SNR of 34.67 and lowest LPIPS of 0.026 for inpainting. Increasing the scale further to 1.0 maintained a high SNR (34.31) but slightly increased the LPIPS for deconvolution (0.106).

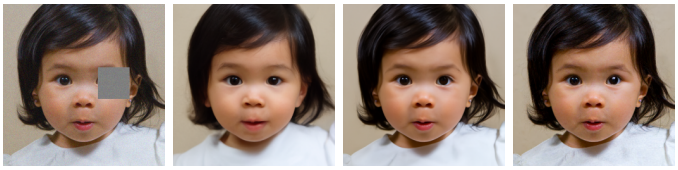

Fig. 7. Input image (a) followed by DPS output on an in-painting task with scale values (b) 0.1, (c) 0.3 and (d) 1.0. Too low a scale parameter reduces the effectiveness of in-painting. Higher scale factors (0.3 and 1.0) show good similarity to the input image.

Fig. 8. Input image (a) followed by DPS output on a deblurring task with scale values (b) 0.1, (c) 0.3 and (d) 1.0. Similar to the inpainting case, scale factors 0.3 and 1.0 successfully deblur the image while retaining features from the input image while a low scale factor does not perform as well.

4 Conclusions

In this work, we investigated the application of diffusion-based generative priors for solving ill-posed inverse problems in computational imaging, specifically inpainting and deconvolution. Our experiments on the FFHQ-256 dataset reveal a clear hierarchy of performance and trade-offs among the three algorithms studied.

SDEdit offers the highest computational efficiency by bypassing the full reverse diffusion chain. However, its reliance on a heuristic starting timestep t_0 makes it sensitive to the noise-fidelity trade-off; it lacks an explicit data-consistency mechanism, often resulting in outputs that, while high in perceptual quality, drift significantly from the ground truth measurements.

ScoreALD introduces a more rigorous approach by incorporating a gradient-based consistency step. While it produces results superior to traditional TV-regularization or Wiener filtering, the choice of the annealing factor schedule remains a critical bottleneck. Without careful tuning, the reconstruction can easily diverge from the image manifold.

Finally, Diffusion Posterior Sampling (DPS) emerged as the most robust solver. By backpropagating the measurement error through the denoising network to estimate the gradient with respect to $\hat{x}_0$, DPS maintains high data consistency without sacrificing the generative power of the prior.

Quantitatively, this resulted in the best balance of PSNR and LPIPS scores across both inpainting and deblurring tasks.

REFERENCES

[1] J. Ho, A. Jain, and P. Abbeel, *Denoising diffusion probabilistic models*. NeurIPS, 2020.

[2] C. Meng, Y. He, Y. Song, J. Song, J. Wu, J. Zhu, and S. Ermon, "Sdedit: Guided image synthesis and editing with stochastic differential equations," *ICLR*, 2022.

[3] A. Jalal, M. Arvinte, G. Daras, E. Price, A. Dimakis, and J. Tamir, "Robust compressed sensing mri with deep generative priors," *NeurIPS*, 2021.

[4] H. Chung, J. Kim, M. Mccann, K. M.L., and J. Ye, "Diffusion posterior sampling for general noisy inverse problems," *ICLR*, 2023.