Performance of SDEdit, ScoreALD, and DPS for deconvolution and inpainting

Braden Becker, Lord Crawford
Hasso Plattner Institute of Design, Stanford University

Motivation

Diffusion models have proven to be robust and useful for a variety of image generation, denoising, and correction tasks. Diffusion models can be trained to predict either the noise at any given time step *t* or the score: the gradient of the log-probability of the noise. Applying diffusion models to restoring or altering images is a major use case and requires specialized techniques to prevent the final image from straying from the base input while still providing good image correction or completion.

In this paper we present several experiments on solving inverse problems using a pre-trained diffusion model and implementing various methods from literature. We conducted tests on both deconvolution and inpainting with random noise and a boxed area removed. For each trial we recorded the Peak Signal to Noise Ratio (PSNR) and the Learned Perceptual Image Patch Similarity (LPIPS)

Related Work

This paper presents several experiments on solving inverse problems using a pre-trained diffusion model. The methods implemented are taken from across existing literature and are as follows: SDEdit (Meng et al., 2022), ScoreALD (Jalal et al., 2021) and DPS (Chung et al., 2023). Starter code for implementation and guidance was provided by the EE367 Computational Imaging teaching team at Stanford University.

References

Chung, H., Kim, J., Mccann, M. T., Klasky, M. L., and Ye, J. C. (2023). Diffusion posterior sampling for general noisy inverse problems. In ICLR.

Ho, J., Jain, A., and Abbeel, P. (2020). Denoising diffusion probabilistic models. In NeurIPS.

Jalal, A., Arvinte, M., Daras, G., Price, E., Dimakis, A. G., and Tamir, J. (2021). Robust compressed sensing mri with deep generative priors.

Meng, C., He, Y., Song, Y., Song, J., Wu, J., Zhu, J.-Y., and Ermon, S. (2022). Sdedit: Guided image synthesis and editing with stochastic differential equations.

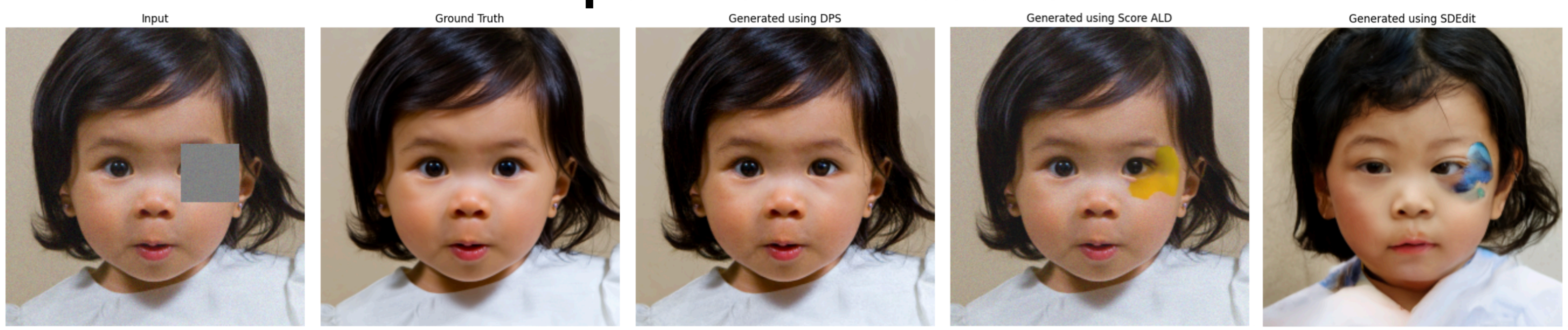
Technique

This research tested the efficacy of three methods on deconvolution and inpainting: SDEdit, ScoreALD, and DPS.

- **SDEdit** is the simplest of the methods and only requires a trained diffusion model. We first add noise to the image to remove fine details and textures while retaining low-frequency information. We then feed this intermediate image into the diffusion model and sequential denoise the image. The crucial parameter is t_0, the time at which we stop adding noise and start denoising.
- **ScoreALD** builds upon SDEdit by steering the denoising process at each step. We calculate both the likelihood gradient and score function and then scale them depending on the time remaining. The limitation of this method is that the likelihood gradient must be approximated and therefore the accuracy is limited.
- **DPS** uses Tweedie's formula to estimate the clean image from the current image and the score function. DPS has excellent performance but requires stepping through the neural network at each time step, massively increasing the computational cost compared to the other methods.

Each method was implemented in Python using the torch library and all testing was performed in Google's Colab using T4 GPUs for additional compute.

Experimental Results



	SDEdit		ScoreALD		DPS	
	PSNR	LPIPS	PNSR	LPIPS	PNSR	LPIPS
Deconv Box	20.296	0.184	25.325	0.128	28.662	0.057
Deconv Random	20.551	0.212	25.065	0.14	28.864	0.066
Inpaint Box	19.837	0.223	27.262	0.088	34.268	0.022
Inpaint Random	14.583	0.333	13.49	0.964	15.984	0.224

The test achieving the highest PSNR of 34.268 and lowest LPIPS of 0.022 used DPS and was tasked with inpainting within a box. The least performant combination was SDEdit attempting to inpaint random noise, resulting in a PSNR of 14.583 and an LPIPS of 0.333. DPS outperforms ScoreALD and SDEdit across all tests and configurations.

There were several parameters across methods that allowed for additional tuning of the results. ScoreALD uses an annealing method to scale the magnitude of the steps in proportion to the proximity of the manifold. The ideal parameters of the annealing method are specific to deconvolution or inpainting respectively. In contrast DPS uses a scaling parameter to meter the relationship between the ground truth and the output of the diffusion model. If this factor is too high the output image will retain the global structure but have high frequency noise, too low and the output will hallucinate major features. This scale factor was adjusted automatically depending on the type of test being run.

Overall we find that DPS is a more robust choice for solving these inverse problems. The computational cost is significant, but on modern hardware does not represent an insurmountable challenge.