A Comparison of Learning-Based Techniques for Image-Denoising and Reconstruction

Christian Kubicka

Abstract—This work explores the practicalities of using pretrained diffusion models for image denoising, image generation, and image reconstruction with the explicit goal of finding the best approach for accomplishing the aforementioned tasks. For the purposes of general image denoising and image generation this paper evaluates DDPM [1]. For image reconstruction, this paper explores three techniques: SDEdit [2], ScoreALD [3], and DPS [4]. Overall, this paper finds that, for image reconstruction, using DPS yields the largest PSNR value and lowest LPIPS value indicating this technique provides the best image quality among the techniques tested.

Index Terms—Diffusion Models, Noise Scheduling

1 Introduction

IMAGING is one of the primary methods used to gather information from the world around us. Whether captured from space telescopes, low-earth orbit satellites or simply from a smartphone, images provide vital insights in a number of different fields. However, there are a variety of phenomena (thermal noise, photon noise, etc.) which can cause the data within an image to become corrupted, lessening its informative value. It is for this possibility that image denoising and reconstruction techniques were developed. To this end, this paper explores four total techniques: DDPM [1] for image denoising, and SDEdit, ScoreALD, and DPS for image reconstruction.

2 Related Work

The working principle behind DDPM [1] based diffusion models is the iterative removal of noise. In the image denoising case, a noisy image is passed to the model and an additional parameter is used to tell the model how much noise should be removed. The model then predicts the noise that was added to a hypothetical clean version of the image and subtracts that noise from the existing image. This process produces a ”single-shot denoised image”. However, depending on the level of noise within the image, the reconstruction may be subpar. In order to rectify this, new noise is added to the image and the process is repeated. It is important to note that when the newly created noise is added to the image it standard deviation should be less than the noise which was just removed from the image. By repeating this process multiple times, the resulting image is often of a higher quality than if the process was repeated only a single time.

All of the techniques mentioned in this paper operate on this basic principle. The variation arises from what is referred to as the noise schedule. The noise schedule dictates how much noise is removed by the model at each time step as well as how much noise is reintroduced to the image at each time step. Each of the techniques discussed essentially just introduces small perturbations to this process.

• *C. Kubicka is with the Department of Electrical Engineering, Stanford University, Stanford, CA, 94305.*
E-mail: kubickce@stanford.edu

3 Methods & Experimental Results

3.1 Unconditional Sampling

3.1.1 DDPM-Based Denoising

The first technique which was evaluated was single-shot denoising through the use of DDPM. Results for three different noise levels can be seen in Fig. 1. The specific noise levels tested are t= 50, t=100, and t=500. For a value of t= 50, this technique yields a PSNR of 35.35 with an LPIPS value of 0.056. For a value of t= 100, this technique yields a PSNR value of 32.62 with an LPIPS value of 0.092. For a value of t= 500, this technique yields a PSNR value of 23.40 with an LPIPS value of 0.311.

3.1.2 Unconditional Image Generation

Unconditional image generation is essentially the same process as DDPM-Based Denoising, however, instead of providing an image with noise added, pure noise is supplied instead. The DDPM process then moves this noisy image towards the closes in-distribution image. The distribution is defined based on what data is used to train the diffusion model. In this work, the pretrained diffusion model which was used was trained on the Flickr-Faces-HQ [5] dataset, thus the randomly sampled image is of a human face. A series of randomly generated examples can be seen in Fig. 2.

3.2 Posterior Sampling

3.2.1 SDEdit

The purpose of SDEdit is to reconstruct original images from blurry images or images which contain missing portions. SDEdit accomplishes this task by adding sufficient noise to the image, then simply diffusing normally. This, in turn, seeds the diffusion process with a prior (the corrupted image) and allows the diffusion process to map that corrupted image to the nearest uncorrupted image which is still within distribution. Through the course of this experiment, three

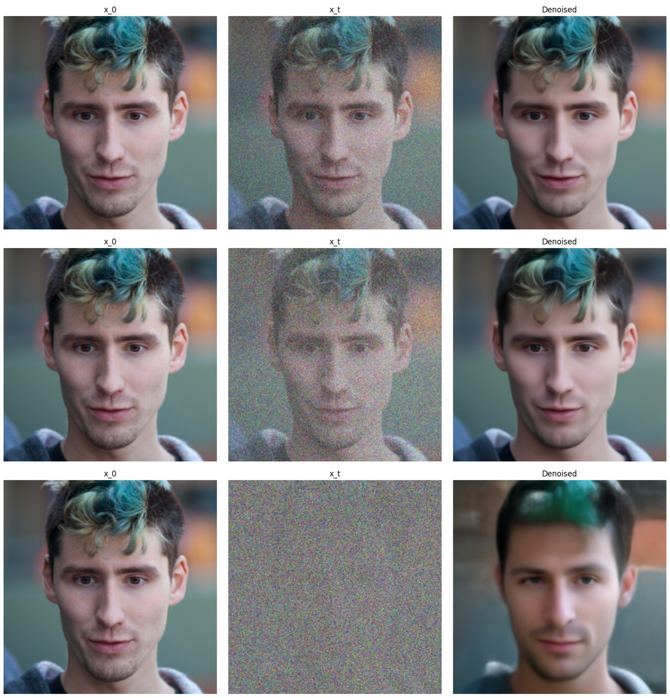

Fig. 1. Results of single-shot denoising at various noise levels.

Fig. 2. Results of unconditional face generation.

different starting times teps were tested (t=500, t=250, and t=750) on two different corruption patterns (blurring, and patch removal). The generated images can be seen in Fig. 3. Tbl. 1 describes what the PSNR and LPIPS values are for certain configurations.

TABLE 1
Table showcasing the PSNR and LPIPS values for various SDEdit configurations.

	t=250 Inpaint	t=250 Deconv	t=500 Inpaint
PSNR	23.34	23.70	19.92
LPIPS	0.14	0.16	0.19

	t=500 Deconv	t=750 Inpaint	t=750 Deconv
PSNR	19.93	14.27	13.23
LPIPS	0.23	0.35	0.41

3.2.2 ScoreALD

ScoreALD incorporates the posterior in a differnt manner to SDEdit. Instead of denoising on the measurement directly, ScoreALD performs the standard denoising process, however at the end of a step the partially noisy image is then stepped in the direction of the gradient of the loss between the one-shot denoised image and the corrupted measurement. The size of this step is essentially controlled by a hyperparameter γ_t which is time step dependent. During testing it was found that sweeping γ from 10 to 15 over the course of the diffusion yielded the best results. For inpainting, the PSNR was 25.00, and the LPIPS value was 0.11. For deconvolution, the PSNR was 26.59, and the LPIPS value was 0.09. A set of sample images can be seen in Fig. 4.

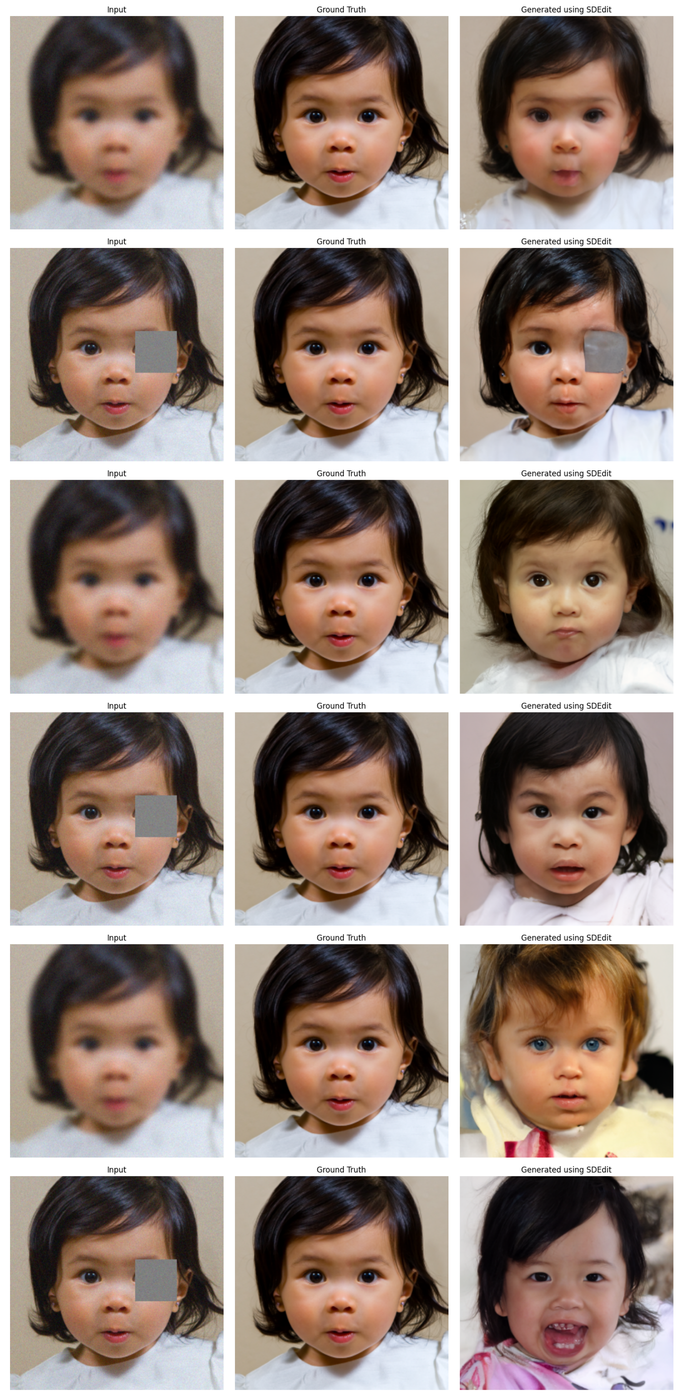

Fig. 3. Results of SDEdit for various configurations. From top to bottom, the images are: (t=250, deconv), (t=250,inpainting), (t=500, deconv), (t=500,inpainting), (t=750, deconv), (t=750,inpainting)

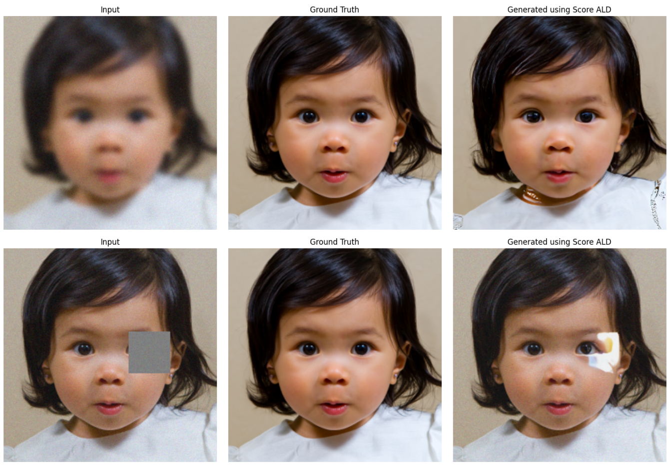

Fig. 4. Results of ScoreALD for various corruption patterns. The top image features a deconvolution based reconstruction while the bottom image features an inpainting based reconstruction.

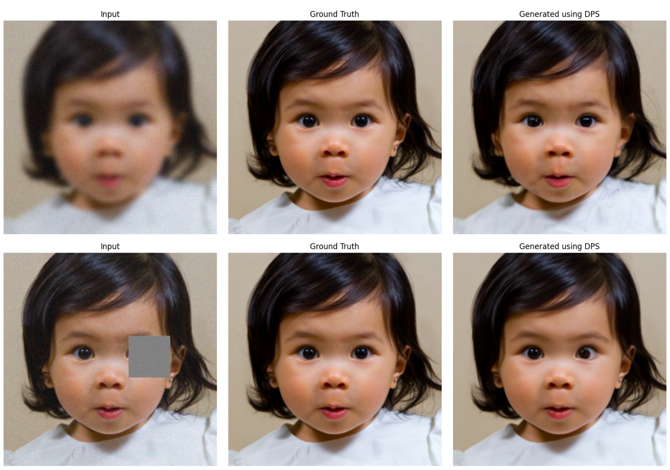

Fig. 5. Results of DPS for various corruption patterns. The top image features a deconvolution based reconstruction while the bottom image features an inpainting based reconstruction.

3.2.3 DPS

DPS expands on ScoreALD by simply altering the step size taken in the direction of the gradient. DPS uses the normalized gradient multiplied by a hyperparameter while ScoreALD is only multiplied by a time-dependent hyperparameter. Given the similarity of the techniques an extensive discussion will be omitted. During testing it was found that a scale factor of 0.3 yielded the best results. For deconvolution the PSNR was 28.38 with an LPIPS value of 0.05. For inpainting the PSNR was 33.84 with an LPIPS value of 0.02. A set of sample images can be seen in Fig. 5.

4 Discussion

4.1 Unconditional Sampling

4.1.1 DDPM-Based Denoising

While single shot denoising does appear to work well for small amounts of noise the approach tends to struggle with moderate to large amounts of noise. Additionally, even in the images with relatively little noise, the denoised image tends to have very few high-frequency details.

4.1.2 Unconditional Image Generation

The results of unconditional image generation leave much to be desired. There are many significant visual artifacts which degrade the quality of the image.

4.2 Posterior Sampling

4.2.1 SDEdit

The SDEdit algorithm is highly sensitive to the initial noise added to the corrupted measurement. For example, as seen in the figure, when noise representative of time step 500 is added prior to diffusion, the quality of the image is somewhat acceptable despite some minor disfigurement. However, when a different amount of noise is added (noise representative of time step 250 or time step 750) the quality of the image degrades significantly.

4.2.2 ScoreALD

The ScoreALD algorithm improves the quality of the recovered image drastically compared to SDEdit. This is likely due to the mechanism by which the prior is incorporated into the diffusion process. Rather than diffusing directly on the prior as in SDEdit (a process which will by definition not incorporate all information from the prior) the prior is used to guide the diffusion of a pure-noise image indirectly. This approach allows the algorithm to preserve much more data from the prior thereby improving image quality.

4.2.3 DPS

The main difference between DPS and ScoreALD is the choice of the gradient scale. DPS uses a normalized gradient multiplied by a hyperparameter while ScoreALD uses an unnormalized gradient multiplied by a scale factor. However, this minor algorithmic change does contribute to a modest increase in image quality.

5 Conclusion

Overall, DPS is the best performer out of the three algorithms tested. Furthermore, a qualatative analysis of the images matches this conclusion.

References

[1] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *CoRR*, vol. abs/2006.11239, 2020. [Online]. Available: https://arxiv.org/abs/2006.11239

[2] C. Meng, Y. Song, J. Song, J. Wu, J. Zhu, and S. Ermon, "Sdedit: Image synthesis and editing with stochastic differential equations," *CoRR*, vol. abs/2108.01073, 2021. [Online]. Available: https://arxiv.org/abs/2108.01073

[3] A. Jalal, M. Arvinte, G. Daras, E. Price, A. G. Dimakis, and J. I. Tamir, "Robust compressed sensing MRI with deep generative priors," *CoRR*, vol. abs/2108.01368, 2021. [Online]. Available: https://arxiv.org/abs/2108.01368

[4] H. Chung, J. Kim, M. T. Mccann, M. L. Klasky, and J. C. Ye, "Diffusion posterior sampling for general noisy inverse problems," 2024. [Online]. Available: https://arxiv.org/abs/2209.14687

[5] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," *CoRR*, vol. abs/1812.04948, 2018. [Online]. Available: http://arxiv.org/abs/1812.04948

Christian Kubicka Christian Kubicka is an M.S. student in Electrical Engineering at Stanford University. His research interests lie in using hardware-software co-design for robotics applications. His current research focuses on system-level design of micro robots with high levels of autonomy. He previously received a B.S. from Clarkson University in Computer Engineering and Software Engineering.