Diffusion Model: Exploring Methods for Inverse Problems

UC San Diego JACOBS SCHOOL OF ENGINEERING

Runpeng Jian, Yinbo Chen, Xiaolong Wang
University of California, San Diego

UC LEADS

Introduction

Diffusion models are Text-to-Image generative models, which can generate diverse types of output according to given prompts. Many famous applications, such as DALL-E 3, Imagen and Stable Diffusion, have been using diffusion models as their foundation models.

Figure 1. Prompt: "A man skiing on a snowy mountain", generated by DALLE 3.

Figure 2. Prompt: "A kid is playing skateboard on a rainy day", generated by Stable Diffusion

This research follows two proposed paradigms discussed at CVPR 2023: the replacement-based method and the reconstruction-based method. By exploring existing methods, we aim to discover new possibilities for the application of diffusion models in the field of inverse problems.

Background

Inverse Problems

Figure 3. Illustrations of inverse tasks, including super resolution [1], denoising, inpainting [1], and deblurring [1].

Forward Process

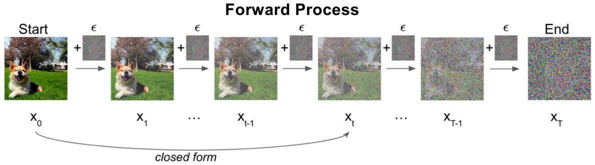

Figure 4. An illustration of the forward process of diffusion models.

Reverse Process

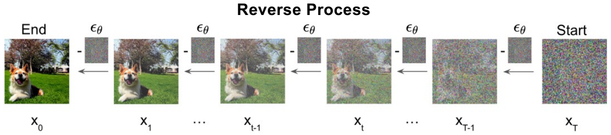

Figure 5. An illustration of the reverse process of diffusion models.

Train a Noise Predictor

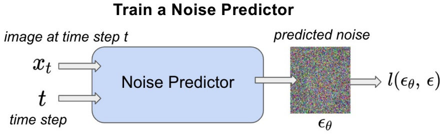

Figure 6. Training a noise predictor by comparing the predicted noise and the real noise added to x_t

Methodology

Reconstruction-Based Paradigm

Score-Based Generative Model

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

Conditional Diffusion Model:

$$\nabla_{x_t} \log p(x_t) \xrightarrow{\text{add a condition } y} \nabla_{x_t} \log p(x_t \mid y) = \nabla_{x_t} \log p(y \mid x_t) + \nabla_{x_t} \log p(x_t)$$

Sample Process:

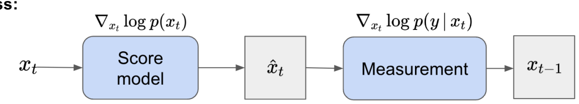

Figure 7. A measurement of $\hat{x}_t$ is applied to guide the model toward the target data distribution.

Replacement-Based Paradigm

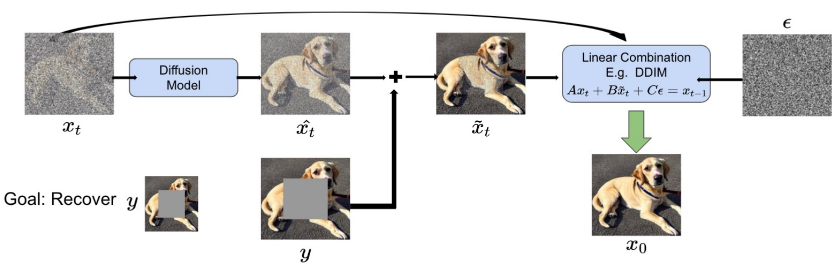

Figure 8. An overview of the replacement-based paradigm .

Preliminary Results and Discussion

Preliminary Results

Sampler	PSNR(dB)	SSIM	KID	LPIPS	top-1
DPS	21.27	0.67	15.28	0.26	58.2
ΠGDM	20.3	0.82	4.5	0.12	67.8
DDRM	20.72	0.83	2.5	0.14	68.6
RED-diff	23.29	0.87	0.86	0.1	72.0

Table 1. Performance of different samplers for ImageNet inpainting using pretrained unconditional guided diffusion model [2].

Discussion

- PSNR measures error in reconstructed images;
- SSIM evaluates similarity in brightness, contrast, and structure;
- KID assesses distance between features of real and generated images;
- LPIPS checks perceptual similarity;
- Top-1 accuracy measures how often a model correctly identifies an image's class on the first try.
- Among the four samplers—DPS, ΠGDM, DDRM, and RED-diff—the results presented in the table clearly indicate that RED-diff excels in the evaluated metrics, including PSNR, SSIM, KID, LPIPS, and top-1 accuracy.

Conclusion

1. Reconstruction-Based Methods leverage the properties of score-based models, in which the score reflects the gradients of data distributions, and of conditional diffusion models, wherein the measurement component can guide the model towards the target data distribution during the sampling process.
2. Replacement-Based Methods utilize the known part of the input as an additional term in each iteration, compared to the traditional reverse process. The output of each linear combination is used as the input for the next iteration.

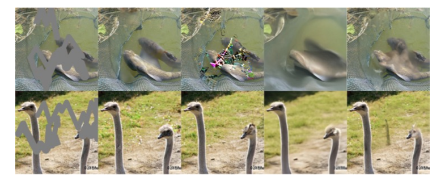

Figure 9. Comparison of RED-diff, ΠGDM, DPS, DDRM for Inpainting representative ImageNet examples [2].

Future Work

There are well-performing methods beyond the two paradigms discussed in this poster. We plan to undertake a comparative analysis of models presented in existing literature. Following this, we intend to develop an incremental approach based on the methodologies of these models.

References

1. Chung, Hyungjin, et al. "Diffusion posterior sampling for general noisy inverse problems." *arXiv preprint arXiv:2209.14687* (2022).
2. Mardani, Morteza, et al. "A Variational Perspective on Solving Inverse Problems with Diffusion Models." *arXiv preprint arXiv:2305.04391* (2023).

Acknowledgements

I would like to give special thanks to Professor Wang and my PhD mentor Yinbo Chen for their support and guidance during the research project. I would like to thank Betty Ramirez, Edgar Beas and the rest of the UC LEADS for their help in making the program go smoothly. I would like to thank UCSD STARS program for their effort during the summer research.