

# Theory Introduction and Application Analysis of DDPM

Yunchuan Li\*

Beijing University of Posts and Telecommunications, Beijing, China

\*Corresponding author: 3312853823@bupt.edu.cn

**Abstract.** The study on DDPM demonstrates that it has potential for image generation, but its limitations must be addressed. The slow sampling speed remains a significant issue, as it limits the model's applicability in real-time settings. The authors' implementation of the DPM-solver sampler represents an important step towards addressing this problem. The results indicate that using the DPM-solver can greatly improve the sampling speed without sacrificing too much quality in the generated samples. The authors also explore some specific applications of DDPM, providing further insights into its capabilities and limitations. For instance, the study demonstrates that DDPM can be used for image inpainting and super-resolution tasks, enabling applications like image restoration and upscaling. However, the study also reveals that DDPM may struggle to generate high-quality images on some datasets or for specific tasks, indicating that there is still much room for improvement. Overall, the study highlights the strengths and weaknesses of DDPM and presents a promising direction for future research in diffusion models. The improved sampling speed achieved through the DPM-solver sampler could open up new possibilities for utilizing DDPM in real-world applications.

**Keywords:** Denoising diffusion probabilistic models (DDPMs), DPM-solver, Slow sampling, Sampling acceleration.

## 1. Introduction

Denoising Diffusion Probabilistic Models (DDPMs) are generative models based on a diffusion process that can generate high-quality images, videos, audios, and perform well in multi-modal tasks. However, DDPMs typically require hundreds of sequential steps of evaluating large networks to obtain high-quality samples, resulting in much slower sampling speeds than single-step GANs or VAEs. DPM-Solver is a fast solver for diffusion ODEs that approximates an exponentially weighted integral equal to the solutions of these ODEs [1]. Specifically, DPM-solver comes in several variants, including first-order, second-order, and third-order versions, as well as DPM-solver++, which were all proposed with convergence order guarantees.

## 2. Relevant Theories

### 2.1. DDPM System

This section provides a review of DDPMs, including their definitions and associated differential equations. Totally, DDPMs are latent variable models. The forward process or diffusion process of diffusion models, set apart from other types of latent variable models, adds Gaussian noise to the data step by step. Given a fixed Markov chain, the transition distribution is computed at each time according to the equation [2]:

$$q_t(x_t|x_0) = N\left(x_t \mid \sqrt{\bar{\alpha}_t}x_0, \sigma_t^2 I\right) \quad (1)$$

Where  $x_0, \dots, x_T \in R^D$  are D-dimensional random variables;  $\bar{\alpha}_t = \sum_{s=1}^t \alpha_s$ ;  $\sqrt{\bar{\alpha}_t}, \sigma_t > 0$ ; the signal-to-noise-ratio (SNR)  $\alpha_t / \sigma_t^2$  is strictly decreasing. In contrast, DDPMs recover  $x_{t-1}$  by learning with a deep-learning model  $m_\theta$  to extract noise from noised image  $x_t$ . This step is actually reverse step, as well as sampling step. There are various sampling algorithms aimed at accelerating sampling speed and quality. The basic sampling algorithm compute the denoised image following :

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left( x_t - \frac{1-\alpha_t}{\sqrt{1-\alpha_t}} \right) m_\theta(x_t, t) + \sigma_t z \quad (2)$$

Where  $z \sim N(0, I)$ . The model  $m_\theta$  can be trained by minimizing least square error ( $m \sim N(0,1)$ ):

$$\min_\theta \| \mathbf{m} - \mathbf{m}_\theta \left( \sqrt{\alpha_t} \mathbf{x}_0 + \sqrt{1 - \alpha_t} \mathbf{m}, t \right) \|^2 \quad (3)$$

## 2.2. DPM-solver Sampling Algorithm

Among all the fast-sampling algorithm, discretizing SDEs perform poorly in high dimensions since it is hard to converge within few steps. On the contrary, ODEs are easier the solution of diffusion ODEs are easier to acquire, with a potential for sampling acceleration. Thus, optimization work mathematically derived an extremely concise form of the solution to the diffusion ordinary differential equation model (diffusion ODEs) and designed a high-order ordinary differential equation solver with the smallest possible error based on this form, called DPM-Solver [3,4]. DPM-solver is a fast-sampling algorithm for diffusion models which can obtain high-quality samples in 10 to 25 steps, which greatly reduces the computational cost of the diffusion model. However, later research has demonstrated that high-order DPM-solvers for generate unsatisfactory samples for guided sampling, even worse than the simple first-order solver DDIM .

## 3. Experiment and Analysis

### 3.1. Experimental Setup

The author ran my DDPM on the following two datasets: CIFAR-10 and StanfordCars. To extract or predict noise in images, the author implemented and train a default U-Net with time-embedding. The author tested basic sampling algorithms and DPM-solver sampling algorithms separately. Due to computational limitations, the author set  $T = 300$  for the basic sampling algorithm; and the order parameters of DPM-solver are set to second order, and its sampling steps are set to 20. The training epochs are limited to 300.

### 3.2. Sampling Quality

Table 1 shows configurations and corresponding FID scores on CIFAR10 and StanfordCars. My FID score is computed with the random part of whole set, consisting of only 1000 samples (configuration “base” means using L1 loss, basic DDPM sampling method and epochs=300).

Despite no evidence supporting, it is speculated that due to the insufficient real samples for comparison, the FID score is generally low. Therefore, the absolute value of FID in the table below has little significance compared with other models, and is only used as a reference for comparing the sample quality of different configurations. As shown in Table 1.

**Table 1.** Configuration and FID comparing

Dataset	Configuration	T	FID
CIFAR-10	base	300	1.458
	L2 loss	300	1.926
	DPM-solver	20	3.062
StanfordCars	base	300	7.704
	DPM-solver	20	15.551

Figure 1, Figure 2 show the samples on StanfordCars of basic DDPM Sampling Algorithm and DPM-solver.



**Fig. 1** StanfordCar, basic DDPM Sampling Algorithm



**Fig. 2** StanfordCar, DPM-solver

### 3.3. Comparative Analysis of Experimental Results

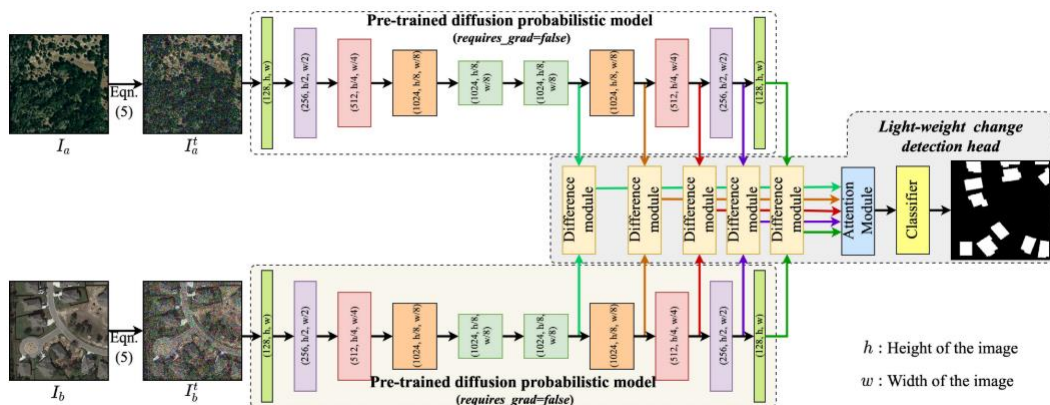
In my experiment, DPM-solver performed worse than the basic sampling algorithm on both CIFAR-10 and StanfordCars datasets. The author also tried other DPM-solver parameter configurations, such as replacing second order with third order and increasing the number of sampling steps to 50 and 100, but there was no significant improvement. This may confirm the explanation of the conclusion cited in Sec. 2.2. Although DPM-solver failed to generate images that were not inferior to the base sampling method due to some uncertain reasons, it was still able to generate some images with real characteristics in a very short time step. while in such a short time step, using the basic sampling method can only obtain noisy images. This also proves that DPM-solver has a powerful ability to accelerate sampling.

## 4. Application Research

Due to its flexibility and strength, diffusion models have recently been used to solve a variety of challenging real-world tasks, including computer vision, natural language processing, temporal data modeling, multimodal learning, robust learning, and interdisciplinary applications. Here the author explored further on several specific tasks worth more attention: remote sensing change detection, point cloud completion and generation and image restoration [5].

### 4.1. Remote sensing Change Detection

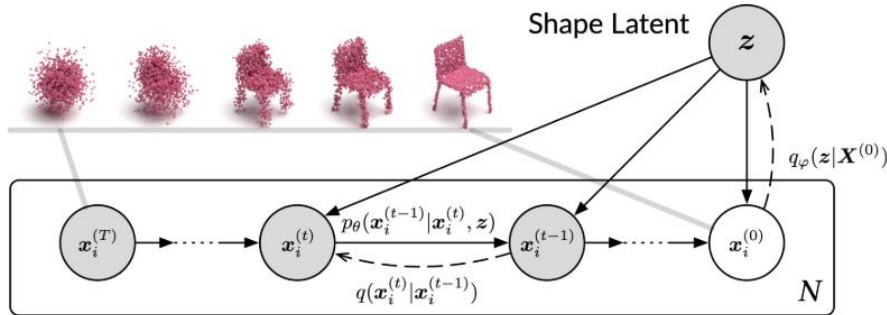
Remote sensing change detection aims to observe well-defined changes on the Earth's surface due to the enormous impact of human civilization on the Earth. Recent work has shown that by using a pre-trained probabilistic diffusion denoising model and then using to train a CD classifier trained by the multi-level feature representations for accurate change detection (shown in Figure 3), which works by means of semantic segmentation, DDPM performs significantly better than SOTA methods in F1, IoU and other metrics [6].



**Fig. 3** Process of DDPM-CD: DDPM with proposed CD approach

### 4.2. Point Cloud Completion and Generation

With the rapid development of depth sensing and 3D lidar-scanning technology, point cloud as a popular representation of 3D shape modeling has arisen more and more attention. However, incomplete point clouds are often generated due to local observation or self-occlusion [7]. Similar to the diffusion process in non-equilibrium thermodynamics, point clouds can be viewed as particles in a thermodynamic system in contact with a heat bath, diffusing from the original distribution to a noise distribution, as is shown in Fig. 4. The task of point cloud generation is learning the backward diffusion process that generate the distribution of a shape with real meaning from the noise distribution. As shown in Figure 4.

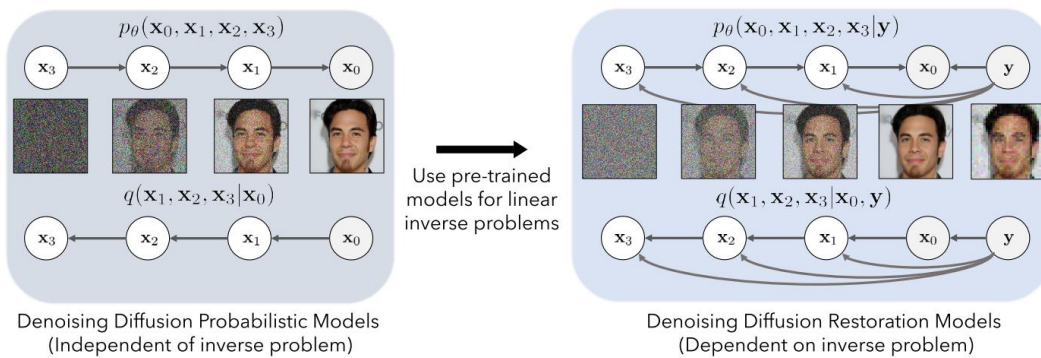


**Fig. 4** The directed graphical model of the diffusion process for point clouds

### 4.3. Image Restoration

Image corruption may often occur because of motion blur, noise, camera mis-focus and so on. Correspondingly, image restoration stands for the process of taking a damaged or noisy image and predict the clean image [8].

The task in image restoration involves linear inverse problem. Based on variational reasoning, DDRM, reformed from DDPM, utilizes a pre-trained denoising diffusion-generating model to solve the linear inverse problem and Figure 5 illustrates the DDRM method for a specific inverse problem [9]. Recent research shows that DDRM performs well on various image sets in reconstruction quality, perceptual quality, and runtime [10]. As shown in Figure 5.



**Fig. 5** DDRM method for a specific inverse problem

## 5. Conclusion

The study conducted by the author is centered around the Denoising Diffusion Probabilistic Model (DDPM) and its application in image generation. The author first provided a brief introduction to DDPM and DPM-solver and explained the mathematical implementations behind them. Next, the author re-implemented DDPM and improved its sampling speed using DPM-solver. Experimental results were presented and analyzed, showing that the use of DPM-solver significantly reduced the sampling overhead while maintaining image quality to some extent. The authors then investigated

three specific applications of DDPM, which could be of great significance in real-world scenarios. These applications include image inpainting, super-resolution, and domain translation. The results demonstrate the model's potential for these tasks, but also shed light on its limitations in generating high-quality images for specific datasets. Overall, this study provides valuable insights into the strengths and weaknesses of DDPM, as well as possible solutions to some of its limitations. The findings could have significant implications for future research on diffusion-based models and their applications in various image-related tasks.

## References

- [1] Lu C, Zhou Y, Bao F, et al. Dpm-solver: A fast ode solver for diffusion probabilistic model sampling in around 10 steps[J]. arXiv preprint arXiv:2206.00927, 2022.
- [2] Preechakul K, Chatthee N, Wizadwongsa S, et al. Diffusion autoencoders: Toward a meaningful and decodable representation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 10619-10629.
- [3] Lu C, Zhou Y, Bao F, et al. Dpm-solver++: Fast solver for guided sampling of diffusion probabilistic models[J]. arXiv preprint arXiv:2211.01095, 2022.
- [4] Xia W, Cong W, Wang G. Patch-Based Denoising Diffusion Probabilistic Model for Sparse-View CT Reconstruction[J]. arXiv preprint arXiv:2211.10388, 2022.
- [5] Campos-Taberner M, Romero-Soriano A, Gatta C, et al. Processing of extremely high-resolution Lidar and RGB data: outcome of the 2015 IEEE GRSS data fusion contest—part a: 2-D contest[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, 9(12): 5547-5559.
- [6] Goswami A, Sharma D, Mathuku H, et al. Change detection in remote sensing image data comparing algebraic and machine learning methods[J]. Electronics, 2022, 11(3): 431.
- [7] Smit M. Deducing the Location of Glass Windows in 3D Indoor Environments[J]. 2022.
- [8] Banerjee S. Study on Biologically-Inspired Recognition Using Multi-Platform and Multi-Modal Data[M]. University of Notre Dame, 2022.
- [9] Yang T, Ren P, Zhang L. Synthesizing Realistic Image Restoration Training Pairs: A Diffusion Approach[J]. arXiv preprint arXiv:2303.06994, 2023.
- [10] Wang Y, Yu J, Zhang J. Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model[J]. arXiv preprint arXiv:2212.00490, 2022.