Diffusion Models in Depth: From a Theoretical and a Practical Perspective

Gregory Sedykh

June 28th 2024

UNIVERSITÉ
DE GENÈVE

FACULTY OF SCIENCE
Department of Informatics

Abstract

Contents

1 Introduction

Diffusion models have gained widespread popularity since 2020, as models such as DALL-E, Stable Diffusion and Midjourney have proven to be capable of generating high-quality images given a text prompt. Furthermore, OpenAI's recent announcement of Sora has shown that diffusion models have now also become highly capable of generating minute long high-definition videos from a text prompt. [1]

These models date back to 2015, where the idea of a diffusion model appeared, based on diffusion processes used in thermodynamics. [5]
Denoising Diffusion Probabilistic Models (DDPMs) were a development of the original diffusion probabilistic model introduced in 2015. [3]
Subsequently, OpenAI improved upon the original DDPMs which originally did not have ideal log likelihoods [3] while also using less forward passes and therefore speeding up the sampling process. [4]
The most recent progress done by OpenAI has allowed their diffusion models to obtain better metrics and better sample quality than Generative Adversarial Networks (GANs) which were previously considered the state-of-the-art in image generation. [2]

The fairly recent apparition of diffusion models means not only that there is still a lot to be discovered about them, but also that progress is being made rapidly and that it is difficult to get a comprehensive overview of the field.
This report aims to provide this overview, going over the theoretical and practical aspects of diffusion models, and to provide a detailed analysis of the different models that have been developed so far.

2   Forward process

3 Backward process

4 Sampling

5 Conclusion

References

[1] Tim Brooks et al. "Video generation models as world simulators". In: (2024). URL: https://openai.com/research/video-generation-models-as-world-simulators.

[2] Prafulla Dhariwal and Alex Nichol. *Diffusion Models Beat GANs on Image Synthesis.* 2021. arXiv: 2105.05233 [cs.LG].

[3] Jonathan Ho, Ajay Jain, and Pieter Abbeel. *Denoising Diffusion Probabilistic Models.* 2020. arXiv: 2006.11239 [cs.LG].

[4] Alex Nichol and Prafulla Dhariwal. *Improved Denoising Diffusion Probabilistic Models.* 2021. arXiv: 2102.09672 [cs.LG].

[5] Jascha Sohl-Dickstein et al. *Deep Unsupervised Learning using Nonequilibrium Thermodynamics.* 2015. arXiv: 1503.03585 [cs.LG].