الجامعة المصرية اليابانية للعلوم و التكنولوجيا
エジプト日本科学技術大学
EGYPT-JAPAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

ABSTRACT
"COMPUTER ENGINEERING DEPARTMENT"

2023 / 2024

On 11/06/2024

Cross-Domain Image-to-Image Translation with VQGAN and BBDM

Presented by:
Mostafa Ahmed Kotb - 120200043
Abderahman Said Ahmed - 120200075
Ahmed Hamdy Rashed - 120200225
Mariam Ayman Mostafa - 120200094
Ashrakat Saeed Elfawal - 120200091

Advisor: Dr. Ahmed Fares

Abstract

Image-to-image translation remains one of the core challenges in computer vision and image processing. Diffusion models represent one of the most promising directions for high-quality image synthesis and have quickly displayed excellent performance on a variety of relevant tasks. Yet, existing diffusion models mainly consider image-to-image translation as conditional generation processes and face the challenge of domain shift.

In this thesis, we propose a novel image-to-image translation approach based on a concept called the Brownian Bridge Diffusion Model (BBDM). Our approach conceptualizes image translation as a stochastic Brownian bridge process and overcomes the limitations of conditional generation by learning the translation between different domains directly via bidirectional diffusion.

We present diverse experiments on different benchmarks to show that our proposed BBDM model is effective. Our method consistently outperforms other approaches under both qualitative visual inspection and quantitative evaluation metrics, showing potential to be used as a robust solution for image-to-image translation tasks.

The BBDM framework is a groundbreaking method in image-to-image translation. Using ideas from the definition of stochastic processes, our framework improves not only on generated image quality but also on generalization across different domains. Second, the natural bidirectionality of the diffusion process allows our model to capture intricate relationships between the source and target domains, leading to subtle and context-relevant translations.

Finally, in addition to its performance improvement, the diffusion of BBDM provides natural interpretability in the image translation process. Due to the decompensation of the translation task into a sequence of diffusion steps, our model reveals the dynamics of image translation over domains, enabling practitioners to gain an improved understanding of the underlying data manifold.

In conclusion, the BBDM framework opens up wide opportunities for the further development of research on image-to-image translation. Future directions can focus on improving the diffusion process to support complex data distributions and generalizing the framework to cover multimodal translation. In addition, reinforcement learning and self-supervised learning methods can be integrated to improve the generalization properties of the BBDM model, leading to its application in real-world problems in various domains.

Keywords: Brownian Bridge, Image to Image Translation, Diffusion Model