Deep learning CV development trends

Anonymous Author(s)
Affiliation
Address
email

Abstract

Object generation and detection are the most important problems in computer vision tasks. Since the Convolutional Neural Network (CNN)-based method proposed by AlexNet became mainstream in the field of computer vision, there has been a lot of research on neural networks, and algorithms with various structures have been created. In order to achieve fast and accurate detection effects, there is a big challenge to break away from the traditional CNN framework. Due to Transformer's theoretical support and technological advances in the field of natural language processing, it has gained the attention of researchers, and it has been proven that Transformer's methods can be used in computer vision tasks, and in some cases, it has been proven to surpass traditional CNN methods. In addition, there is a challenge to improve control and its quality in generative models. Various methods have been attempted to provide additional input data and improve the quality of generative models, and interesting results have been obtained in diffusion models. In this paper, we discuss the main features, advantages and disadvantages of the algorithm for vision transformer and stable diffusion.

1 Introduction

1.0.1 Vit

1.0.2 Diffusion

2 Background

2.0.1 Object Detection

2.0.2 Generative Adversarial Network (GAN)

2.0.3 Variational AutoEncoder (VAE)

2.0.4 Diffusion model

3 Vision transformer

4 Stable diffusion

References

[1]

Submitted to NeurIPS 2021 AI for Science Workshop.