Paper review of "Self-Attention Generative Adversarial Networks"

Merlin Rousseau

24th March 2023

1 Article resume

The article "Self-Attention Generative Adversarial Networks" [1] proposes a new approach for improving the performance of generative adversarial networks (GANs). GANs are a popular class of deep learning models used for generating images that mimics real images. A GAN consists of two neural networks : a generator which learns to produce data that is similar to real data and a discriminator which aims to distinguish between real and generated data. The authors propose a novel architecture that includes a self-attention mechanism.

The basic idea behind self-attention mechanism is to allow each element in a sequence to attend to all other elements in the same sequence, and then use these attentions to compute a weighted representation of the entire sequence that models relations between elements. In this paper, the self-attention mechanism allows the generator and discriminator networks to focus on specific regions of the image, based on their relative importance to the generation process. This is done by computing a set of attention weights for each pixel in the feature maps of the generator and discriminator networks. These attention weights are then used to compute a weighted sum of the feature maps which is passed through a convolutional layer to produce the final feature maps that include spatial relationships between pixels. The self-attention mechanism allows the generator to capture long-range dependencies between the input image regions, leading to more realistic output images. It also reduces the sensitivity of the GAN to the spatial location of objects in the input image.

In order to evaluate the performance of the proposed self-attention GAN (SAGAN), the authors conducted experiments on ImageNet dataset. They compared the performance of SAGAN to that of Conditional GAN (cGAN) which was the state-of-the-art GAN. They found that SAGAN outperforms cGAN on both image quality and diversity. In addition, SAGAN is more stable during training and is less subject to mode collapse.

2 Interest in this publication

During my studies in master 2 at Centrale Lille, I carried out a research project which consisted in reconstructing the part of a face hidden by a surgical mask to perform face recognition. The face completion was obviously done by using a GAN and I remember having trouble ensuring the stability of the GAN during training.

Moreover, when I was on an internship at the Institut Mines Telecom Nord Europe, I developed a model for image-text matching. This architecture used transformer layers, which implement the self-attention mechanism, to better model relationships between regions of an image and words in a sentence and perform more accurate alignments between those image regions and words.

I chose this paper because it tackles subjects that I encountered during my schooling and introduces the use of the self attention mechanism on GAN architectures to improve their performances. This approach has surpassed the performance of the state of the art which makes it all the more interesting.

References

[1] Han Zhang et al. "Self-Attention Generative Adversarial Networks". In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri and Ruslan Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, 2019, pp. 7354–7363. URL: https://proceedings.mlr.press/v97/zhang19d.html.