Project Description: Study of generative models through an open style transfer competition

Rémy Sabathier
CentraleSupélec
Paris-Saclay University

Omar Souaidi
CentraleSupélec
ENS Paris-Saclay

Abstract

Generative Adversarial Networks (called GANs) is a type of deep learning models introduced by Ian Goodfellow and other researchers in 2014. In this type of models, two neural networks are competing against each other's in order to learn how to generate samples with the same statistics as the original distribution. On its early release, GAN were identified as highly unstable during the training (because of the competition between the two networks) and hard to evaluate. These two issues are still active domain in Artificial Intelligence: we propose to study the key aspects of these two problems by participating to an open competition hosted on Kaggle: 'I'm something of a painter myself' .

1. Motivation and problem definition

1.1 Competition

We propose to participate to a competition hosted on Kaggle: I'm something of a painter myself. The competition is a style transfer challenge based on Monnet paintings. As such, given a dataset of real-life images, we are asked to design a GAN that can give the style of Monnet paintings to an image of our choice.

After implementing a model, we can send our proposition to measure the performance of our network compared to other participants (as we are writing this description, 143 teams are enrolled). A constraint on the training time of the model is imposed (less than 3 hours training time on TPUs or 5 hours on GPUs).

1.2 Evaluation

In generative adversarial learning, evaluating a model is not as simple as in classification for instance where the progression of a model can be easily quantified witch metrics. Evaluating GANs performance is still considered as active research subject. This project is also motivated on understanding the current most used metrics for GAN evaluation (Fréchet Inception Distance [2], Inception Score [3]). The Kaggle competition propose to evaluate candidate on the MiFID: *memorization informed Fréchet Inception Distance*, which is a modified version of the Fréchet Inception Distance which takes training set memorization into account. For each generated training sample, the memorization distance is the minimum cosine distance between this element and the elements from the training set.

1.3 Dataset & Computing Units

Figure 1: Example from the Monnet painting training set. (256x256x3) in TF records format.

To train our GAN, the competition proposes a collection of 300 Monet paintings. We then have 7363 images that can be used to test the style transfer. One interesting auxiliary aspect of the competition is the use of the TF records format for the dataset and TPU processors for training.

2. Methodology

The first objective of this project is to propose an implementation of a *Cycle-Consistent Adversarial Network* (also called Cycle-GAN) introduced by Jun-Yan Zhu et al. in 2017 [1] which learn a mapping to translate an image from a distribution domain A to another distribution domain B without having a paired dataset. This network was used in many domains such as style transfer, photo enhancement, object transfiguration. Cycle GANs are composed of two generative and two discriminative networks (one for transformation from domain A to B and another for transformation from B to A).

This model will allow us to try techniques to improve quality and stability of GANs. On this purpose we will especially try different optimizer to look at their impact on output quality and training stability. We propose to try a new optimizer 'AdaBelief' [4] proposed at NeurIPS 2020 conference. The authors found that the use of this optimizer for GAN training gave more stability and improve the quality of the generator (the test was performed on Cifar10 dataset).

The second objective of the project is to study the most used score/metrics in the literature for GAN evaluation and understand their limit/possible improvements. We will propose precise analysis about this metrics on the style transfer task of the competition.

During this project we may not only focus on the Monnet competition samples and try our implementation on other dataset if this can enlighten some key points on our two main focus.

References

[1] Jun-Yan Zhu et al. Unpaired Image-to-Image using Cycle-Consistent Adversarial Networks, 2017

[2] Martin Heussel et al. GANs trained by a two time-scale rule convergence to a local Nash equilibrium 2017

[3] Tim Salimans, Ian GoodFellow et al. Improved techniques for training GANs 2016

[4] Juntang Zhuang et al. Adabelief Optimizer: Adaptating Stepsizes by the belief in observed gradients 2020

Competition Link

https://www.kaggle.com/c/gan-getting-started/overview