Generative models and cinema

Raphaël Deknop

July 2020

Introduction

This document is aimed at the general public and intends to provide an intuitive -yet precise- understanding on what Generative Adversarial Networks are capable of. In the last decade, machine learning and artificial intelligence have taken more and more place in our lives. From autonomous vehicles to face recognition or even beating world-class players in online games, it is no wonder that this technology has an enormous potential and can be dangerously powerful.

1. Discriminative and generative models

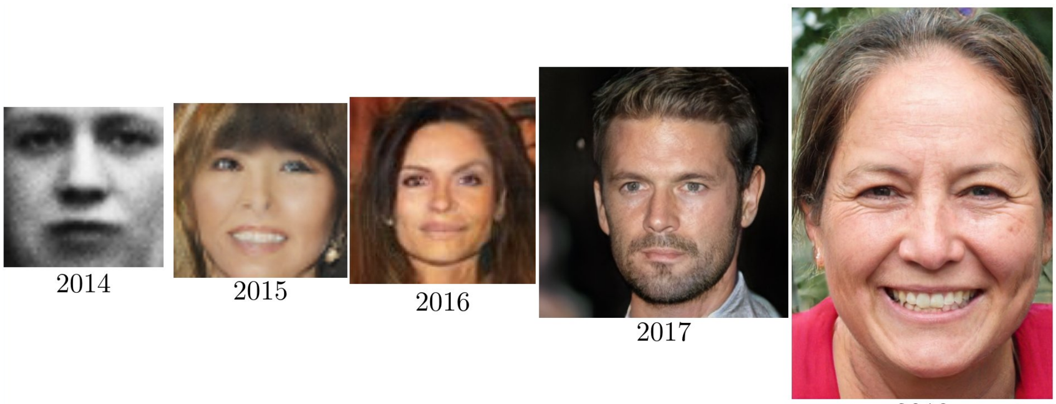
In machine learning, two main approaches can be followed. On the one hand, the **discriminative** method can tackle classifying problems such as assigning the correct label to an image, computing a highly probable output from a (previously unseen) complex input and much more. On the other hand, this model cannot *generate* similar data. An intuitive illustration of discriminative modelling in human behaviour is being able to distinguish chinese characters while remaining unable to correctly draw one of them.

This is where the **generative** one differs. As its name suggests, a model using such an approach can generate data resembling what it has been fed. Here, as long as the input data is carefully selected, no label is required. Let's say that you want your model to generate a Shakespeare-like poem, it would be counterproductive to train it with some -unlabelled- Edgar Allan Poe or Oscar Wilde writings.

2. Generative Adversarial Networks

Although those two approaches were just presented separately, nothing is forbidding a network to follow both of them. This is where Generative Adversarial Networks come in. In a GAN, the generative network produces candidates while the discriminative network evaluates them. More precisely, it outputs the probability of it being generated. To put it simply, the goal of the generator is to fool the discriminator into thinking that the data provided is real.

This type of network, presented in 2014 by Ian Goodfellow and his collaborators [Goo+14], has been an enormous step in machine learning. To illustrate that statement, look at this series of generated human faces:



2014 2015 2016 2017 2018

Figure 1: Progress of GANs on human face generation. Source: Ian Goodfellow's Twitter

The evolution here is astonishing. Let's remember that before 2014, GANs simply didn't exist. There are a bunch of websites showing GAN-generated images of people, artworks, cats or even horses produced by StyleGAN2 [Kar+19] for the curious out there. It is effortless -and an interesting exercise- to come up with potential applications and realise that there are plenty of them.

3. Going further and wider applications

There exists plenty of neural network types, generative adversarial networks are just one of them. The reason I presented them here is because I find them to be lying on the right spot between intuitive understanding and current state of the art in AI research. I haven't been technical here, as it is not the intention of this article.

Interestingly enough, GANs have their limitation. For example, it is impossible to tweak parameters on a generated human face picture in order to change, say, hair colour, face expression or nose type. That is made possible with Adversarial Latent Autoencoders (ALAEs) [PAD20] which make use of latent spaces.

4. Discussion and key takeaways

References

[Goo+14] Ian J. Goodfellow et al. *Generative Adversarial Networks*. 2014. arXiv: `1406.2661 [stat.ML]`.

[Kar+19] Tero Karras et al. *Analyzing and Improving the Image Quality of StyleGAN*. 2019. arXiv: `1912.04958 [cs.CV]`.

[PAD20] Stanislav Pidhorskyi, Donald Adjeroh, and Gianfranco Doretto. *Adversarial Latent Autoencoders*. [preprint]. 2020. arXiv: `2004.04467 [cs.LG]`.