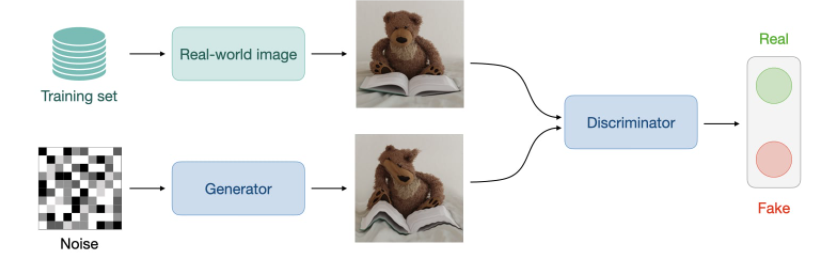
[CS230 course on CNNs](https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-convolutional-neural-networks)

Some notes about the neural networks we will be using:

* GANs = Generative adversarial Networks. Composed of a generative and a discriminative model. The generative model aims at generating the most truthful output that will be fed into the discriminative, whose role is to differentiate the generated and true image.



Those GANs are used in a lot of different applications, especially in music generation and synthesis.

GANs are a type of network that uses a computational trick.

[About CycleGANs in voice coding:](https://www.kecl.ntt.co.jp/people/kaneko.takuhiro/projects/cyclegan-vc/)

Some info about CycleGANs:

CycleGANs allow a lot of progress in Voice Conversion problems. Usually, VC methods use temporally aligned parallel data of source and target speech as training data. If those two data types are perfectly aligned, then you can obtain a **mapping function** relatively easily. However, this time alignment is really hard to perform. CycleGANs is way more powerful since it doesn’t require to perform this alignment. It can work with unparallel data.

The CycleGAN captures a spectrotemporal structure w/o any alignment procedure.

CycleGAN structure:

Goal is to learn of a mapping from source x to target y. To create such mapping, two losses Gx->y are learnt. One adversarial loss & one cycle-consistency loss.