GAN et al.

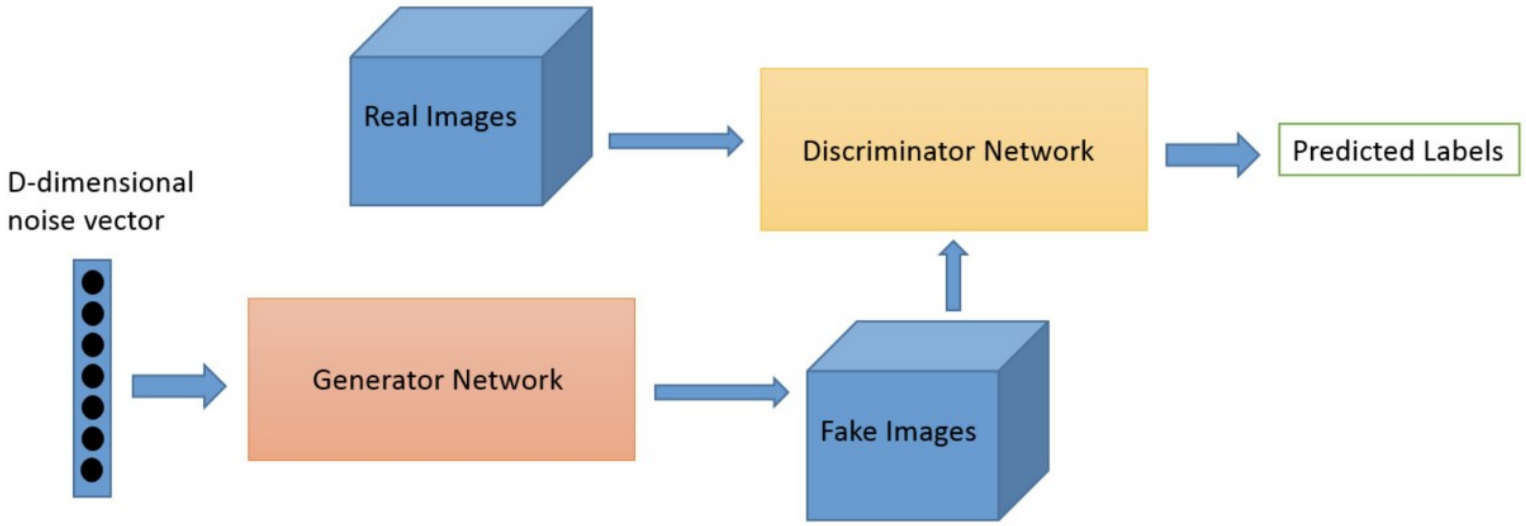
Generative adversarial network (GAN)

- Invented by Ian Goodfellow and his colleagues in 2014
- Include a generator network and a discriminator network

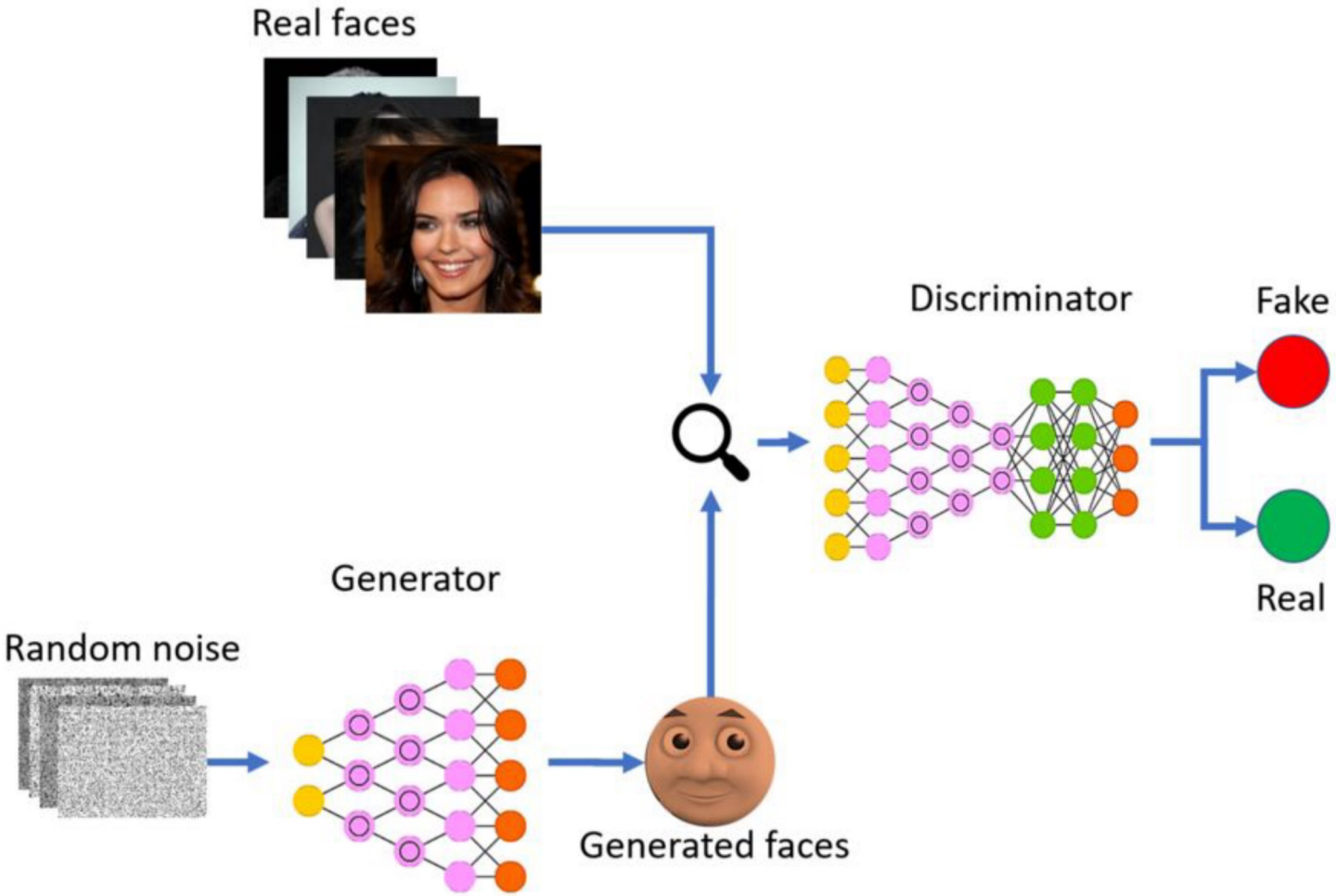
GAN example

- https://thiscatdoesnotexist.com/
- https://thispersondoesnotexist.com/
- https://artbreeder.com/
  - https://towardsdatascience.com/artbreeder-draw-me-an-electric-sheep-841babe80b67
  - https://www.youtube.com/watch?v=kSLJriaOumA
  -

Generative adversarial network (GAN)



- Invented by Ian Goodfellow and his colleagues in 2014
- Include a generator network and a discriminator network

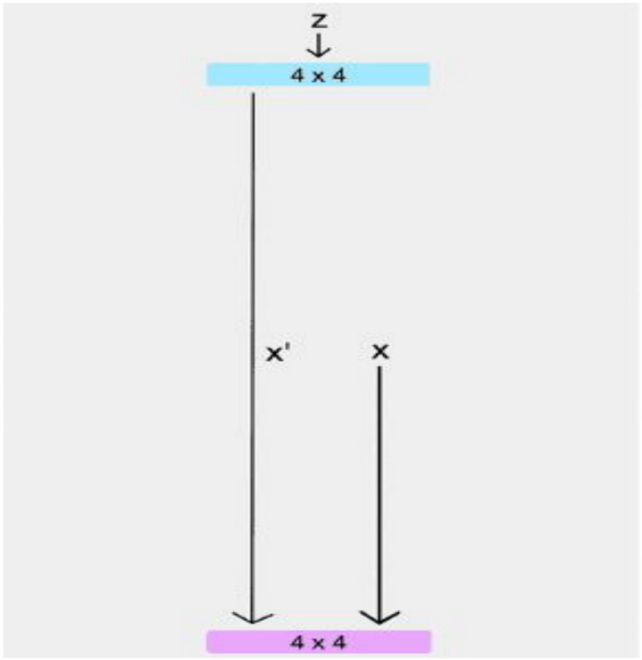

Real faces
Discriminator
Fake
Real
Generator
Random noise
Generated faces

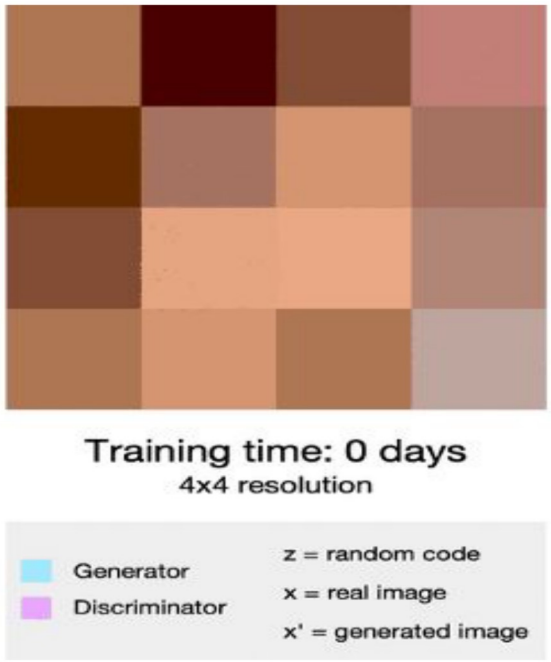
GAN generated images

ProGAN

- Researchers had trouble generating high-quality large images (e.g. 1024×1024) until 2018, when NVIDIA first tackles the challenge with ProGAN.
- The key innovation of ProGAN is the progressive training — it starts by training the generator and the discriminator with a very low-resolution image (e.g. 4×4) and adds a higher resolution layer every time.
- This technique first creates the foundation of the image by learning the base features which appear even in a low-resolution image, and learns more and more details over time as the resolution increases.

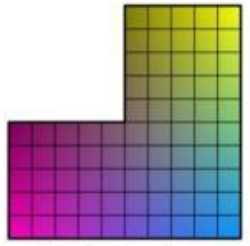
ProGAN


z
4 x 4
x'
x
4 x 4


Training time: 0 days
4x4 resolution
Generator
Discriminator
z = random code
x = real image
x' = generated image

ProGAN

- ProGAN generates high-quality images but, as in most models, its ability to control specific features of the generated image is very limited.
- This is because of feature entanglement
  - the features are entangled and therefore attempting to tweak the input, even a bit, usually affects multiple features at the same time
  - With entangled representations, the data distribution may not necessarily follow the normal distribution where we want to sample the input vectors z from. For example, the data distribution would have a missing corner like this which represents the region where the ratio of the eyes and the face becomes unrealistic.



(a) Distribution of features in training set

StyleGAN

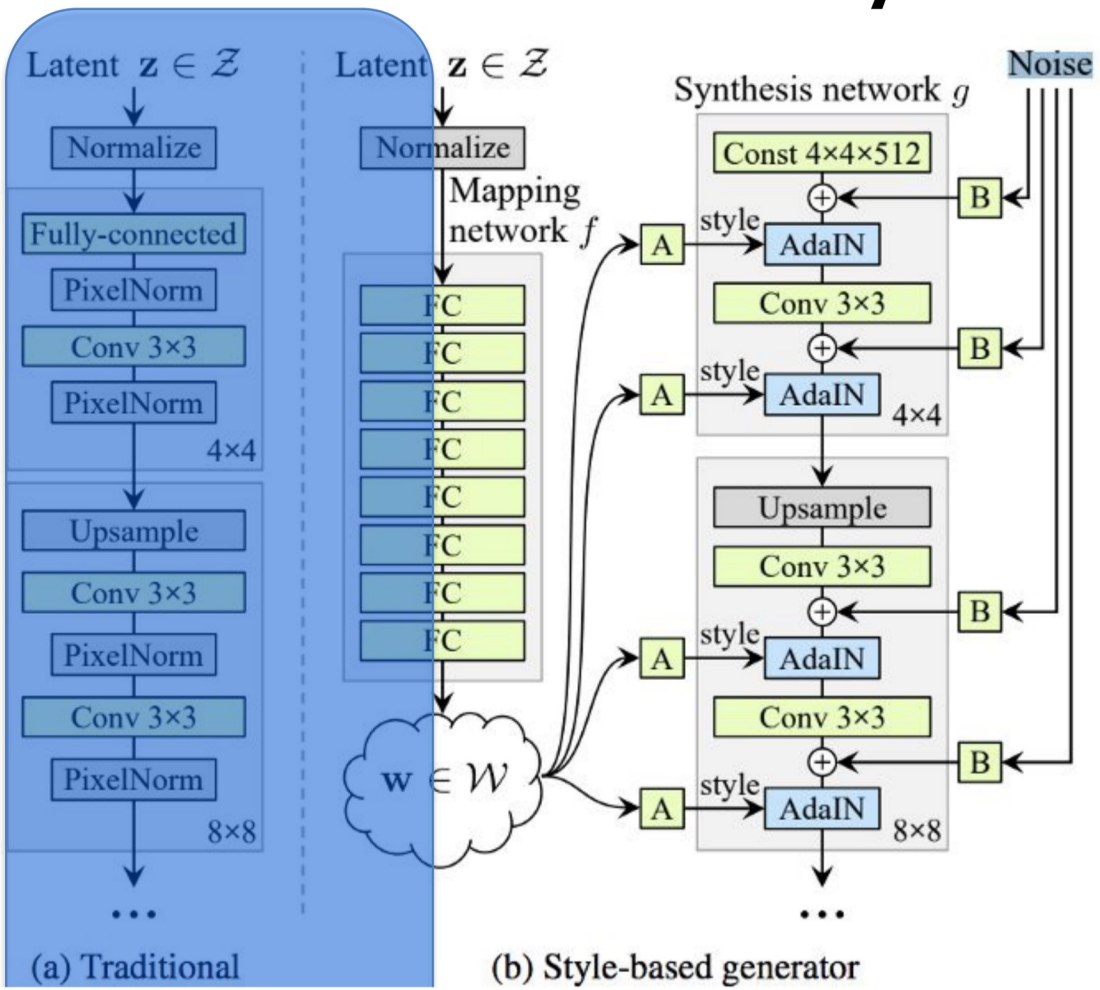
- Based on ProGAN
  - networks are trained on lower resolution initially (4x4), then bigger layers are gradually added after it's stabilized
- Add a mapping network that encodes the input vectors into an intermediate latent space, w

(a) Distribution of features in training set

(b) Mapping from $\mathcal{Z}$ to features

(c) Mapping from $\mathcal{W}$ to features

StyleGAN



A scaled noise is added to each channel and changes a bit the visual expression of the features of the resolution level it operates on.

StyleGAN

- The StyleGAN generator uses the intermediate vector in each level of the synthesis network, which might cause the network to learn that levels are correlated.
- It is capable of combine multiple images in a coherent way (as shown in the video below). The model generates two images A and B and then combines them by taking low-level features from A and the rest of the features from B.
- https://youtu.be/kSLJriaOumA
-

styleGAN Examples

- Portraits
  - https://twitter.com/cyrildiagne/status/1094685986691399681
- Magic the gathering card
  - https://iforcedabot.com/i-have-no-mana-and-i-must-tap/
- A colab example:
  - https://colab.research.google.com/drive/1e7EDCUvLSAyntKPgyMjK3J4RbfwGjJGz?usp=sharing

Other GAN models

- Image-to-Image Translation using CycleGAN Model
  - https://towardsdatascience.com/image-to-image-translation-using-cyclegan-model-d58cfff04755
  - https://junyanz.github.io/CycleGAN/
- Decrappification, DeOldification, and Super Resolution – the DeOldify model
  - https://www.fast.ai/2019/05/03/decrappify/
- Comparison of different GAN models for making anime faces
  - https://www.gwern.net/Faces

Other Text to Image Techniques

An introduction video:

https://www.youtube.com/watch?v=aqW-8WtYJvw

PARTI

Google’s new next to image model:

https://www.youtube.com/watch?v=qS-iYnp00uc

https://www.lesswrong.com/posts/tMr3HJwitJCbQ5HTc/google-s-new-text-to-image-model-parti-a-demonstration-of

Other Resources

- Disco Diffusion
  - https://colab.research.google.com/github/alembics/disco-diffusion/blob/main/Disco_Diffusion.ipynb
- https://pollinations.ai/c/Anything