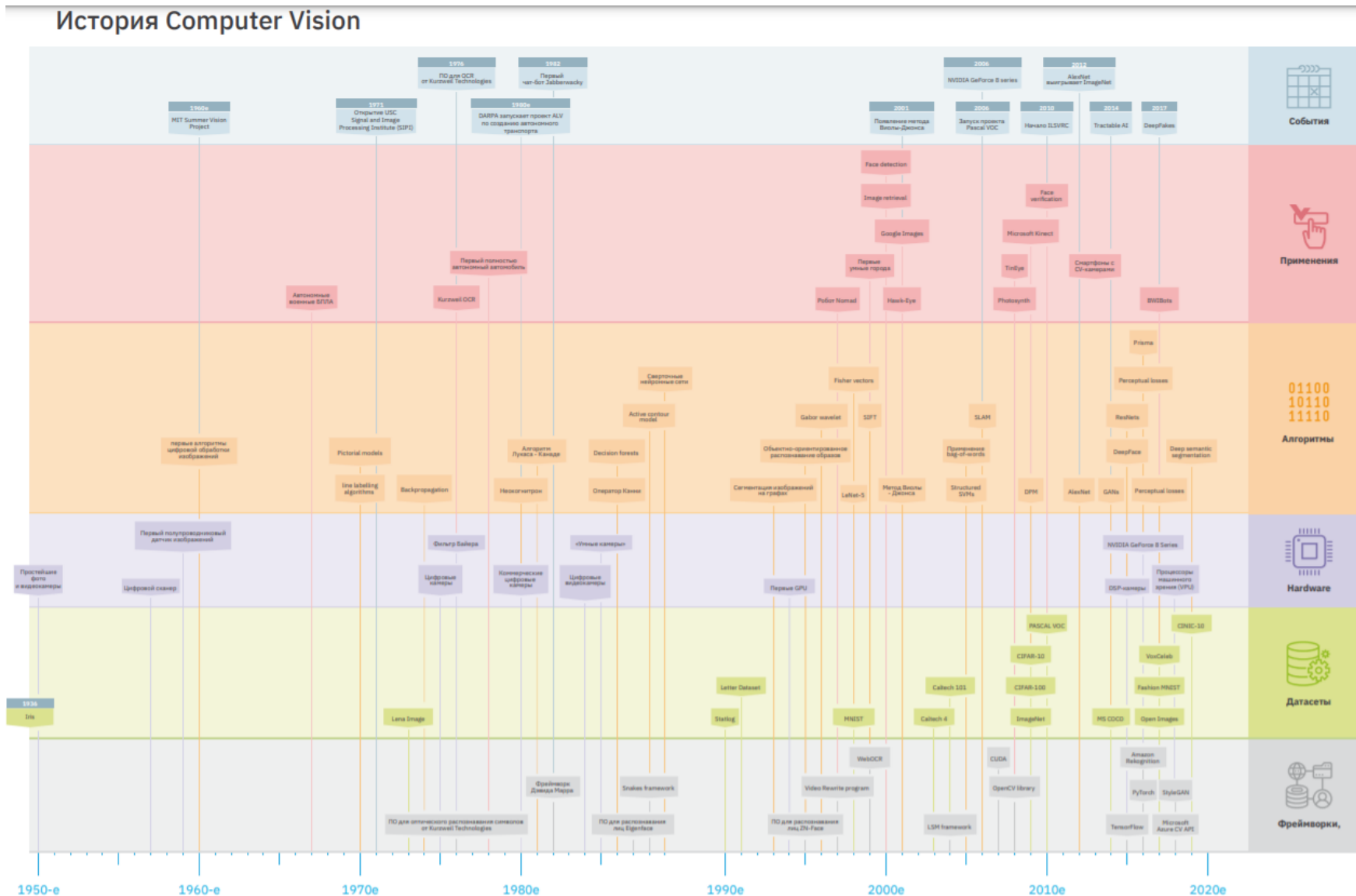# AN OVERVIEW OF DEEPFAKE
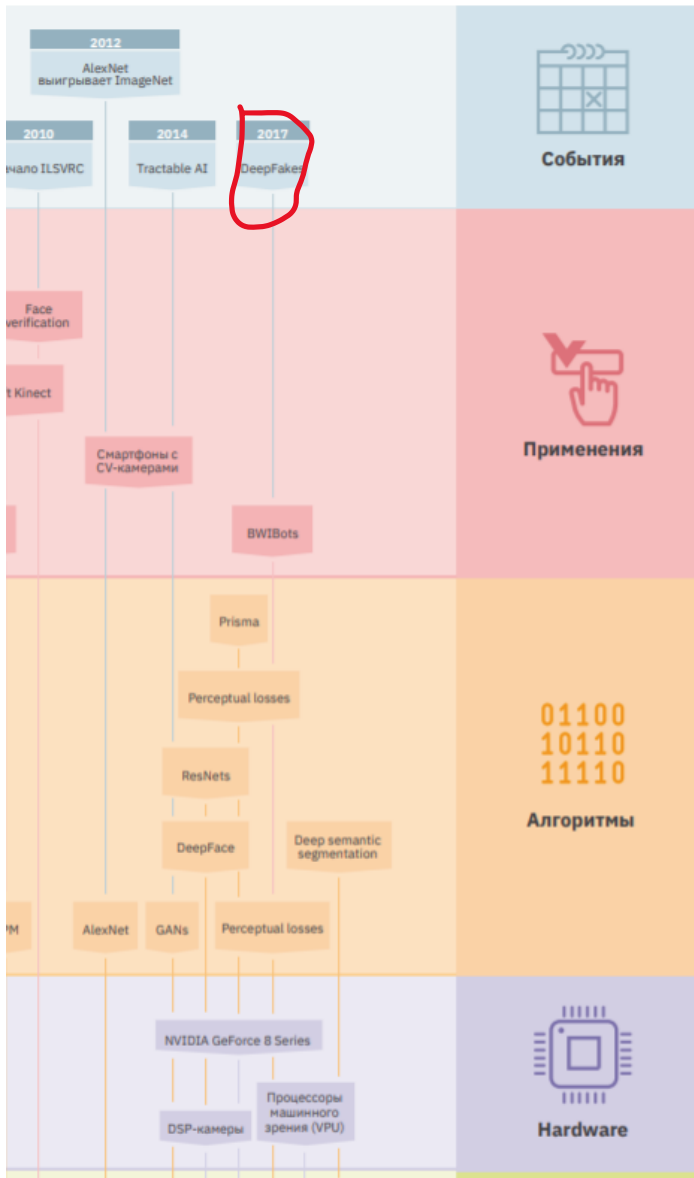
YONGYU CHEN

FOR THE FINAL LAB OF BIOMETRIC TECHNOLOGY

The
history of
the
developm
ent of
Computer
Vision
[Russia AI
report –
computer
vision]

Deepfakes was invented in 2017. [AI report]
(While wiki says that it was created in 90s,
but was applied in industry in recent year)

Deepfakes are synthetic media[2] in which a person in an existing image or video is replaced with someone else's likeness. [Wikipedia]

The main machine learning methods used to create deepfakes are based on deep learning and involve training generative neural network architectures, such as autoencoders or generative adversarial networks (GANs). [wiki]

TECHNIQUES

- Deepfakes rely on a type of neural network called an auto-encoder. These consist of an encoder, which reduces an image to a lower dimensional latent space, and a decoder, which reconstructs the image from the latent representation. Deepfakes utilize this architecture by having a universal encoder which encodes a person into the latent space. The latent representation contains key features about their facial features and body posture. This can then be decoded with a model trained specifically for the target. This means the target's detailed information will be superimposed on the underlying facial and body features of the original video, represented in the latent space. [wiki]

- A popular upgrade to this architecture attaches a generative adversarial network to the decoder. A GAN trains a generator, in this case the decoder, and a discriminator in an adversarial relationship. The generator creates new images from the latent representation of the source material, while the discriminator attempts to determine whether or not the image is generated. This causes the generator to create images that mimic reality extremely well as any defects would be caught by the discriminator. Both algorithms improve constantly in a zero sum game. This makes deepfakes difficult to combat as they are constantly evolving; any time a defect is determined, it can be corrected. [wiki]

IT CAN BE REALIZED BY A GAN (GENERATIVE ADVERSARIAL NETWORK)
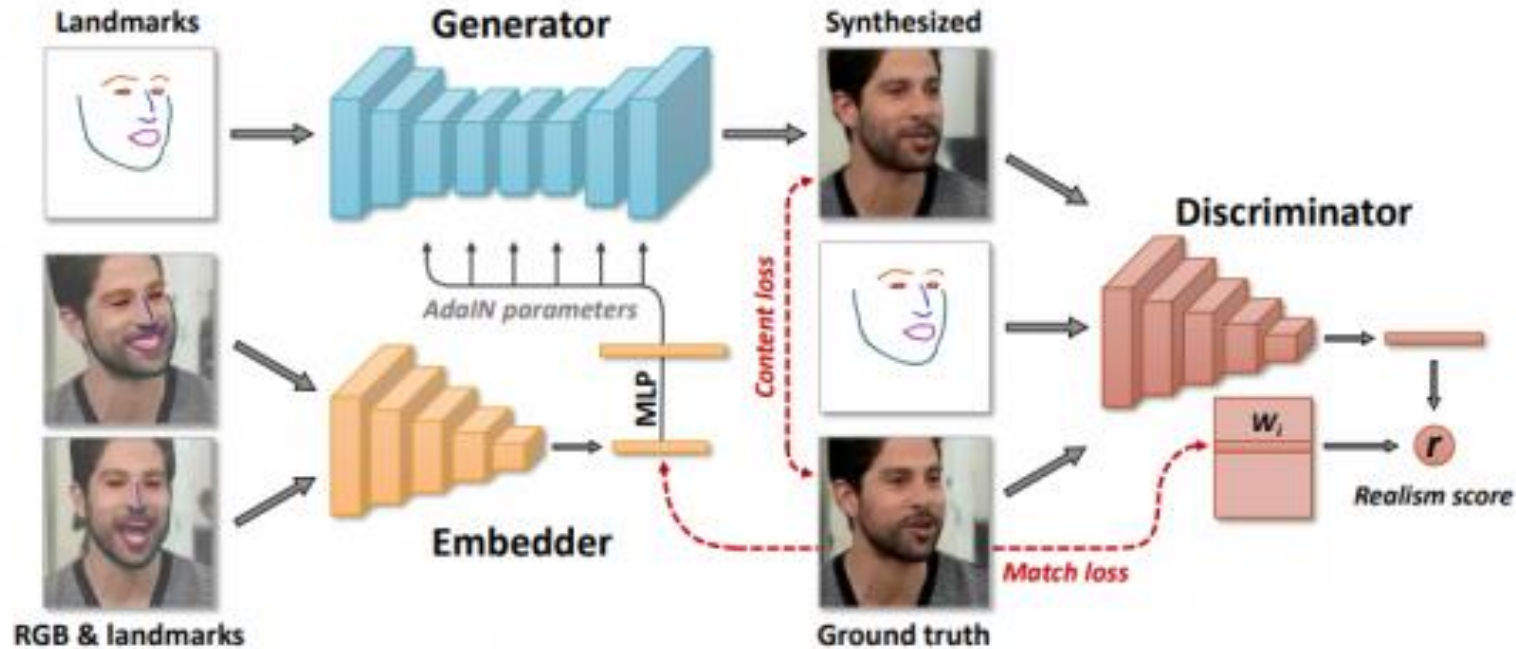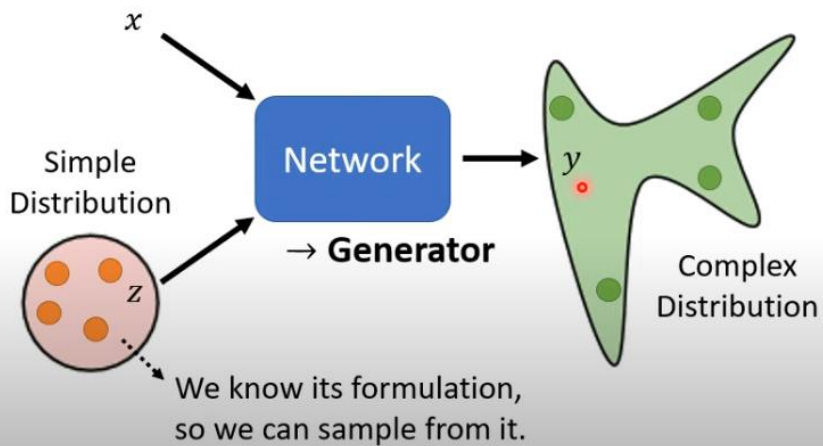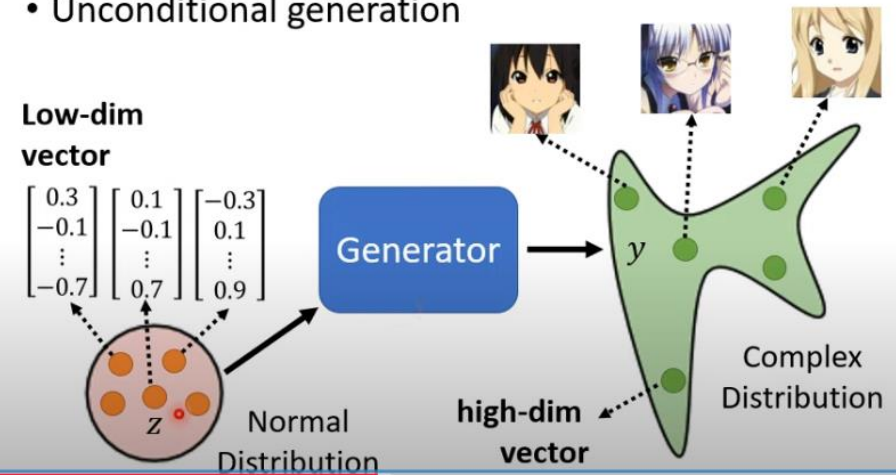

Photo ref: https://arxiv.org/pdf/1905.08233.pdf

A BRIEF DESCRIPTION OF GAN GENERATOR

https://www.youtube.com/watch?v=4OWp0wDu6Xw

DISCRIMINATOR

https://www.youtube.com/watch?v=jNY1WBb8I4U
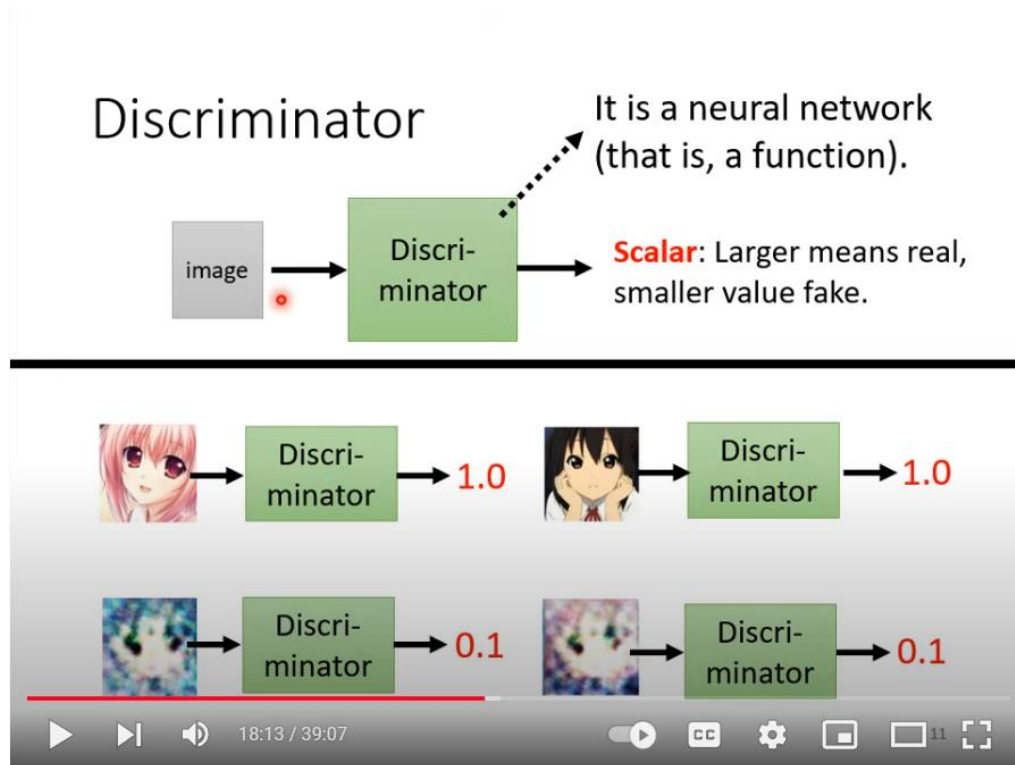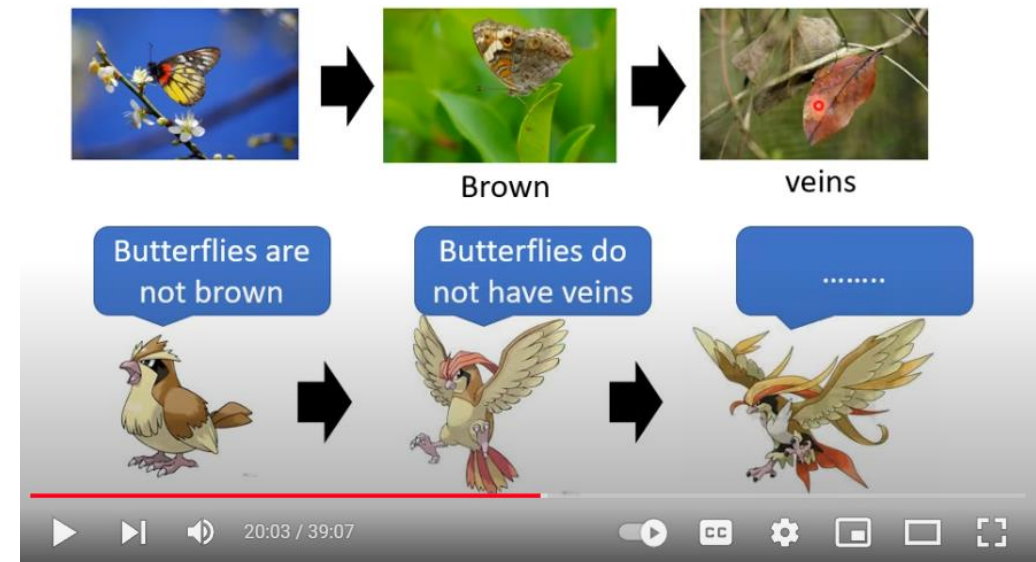
GAN – MATHEMATICAL DESCRIPTION

https://www.youtube.com/watch?v=MP0BnVH2yOo
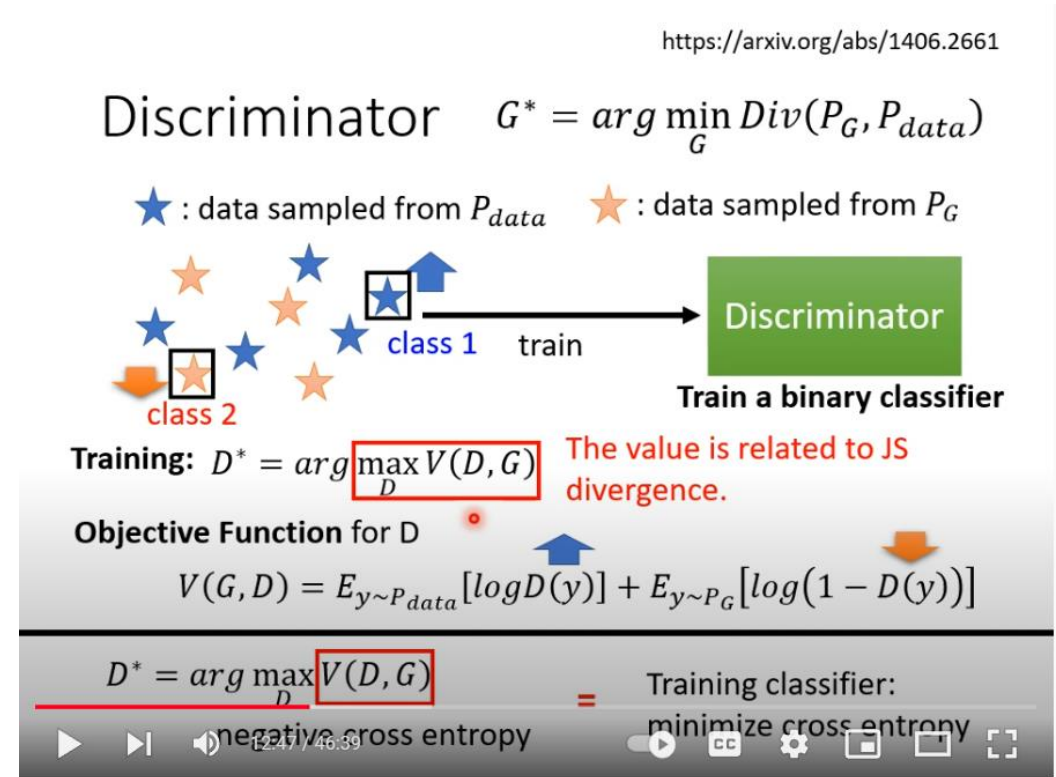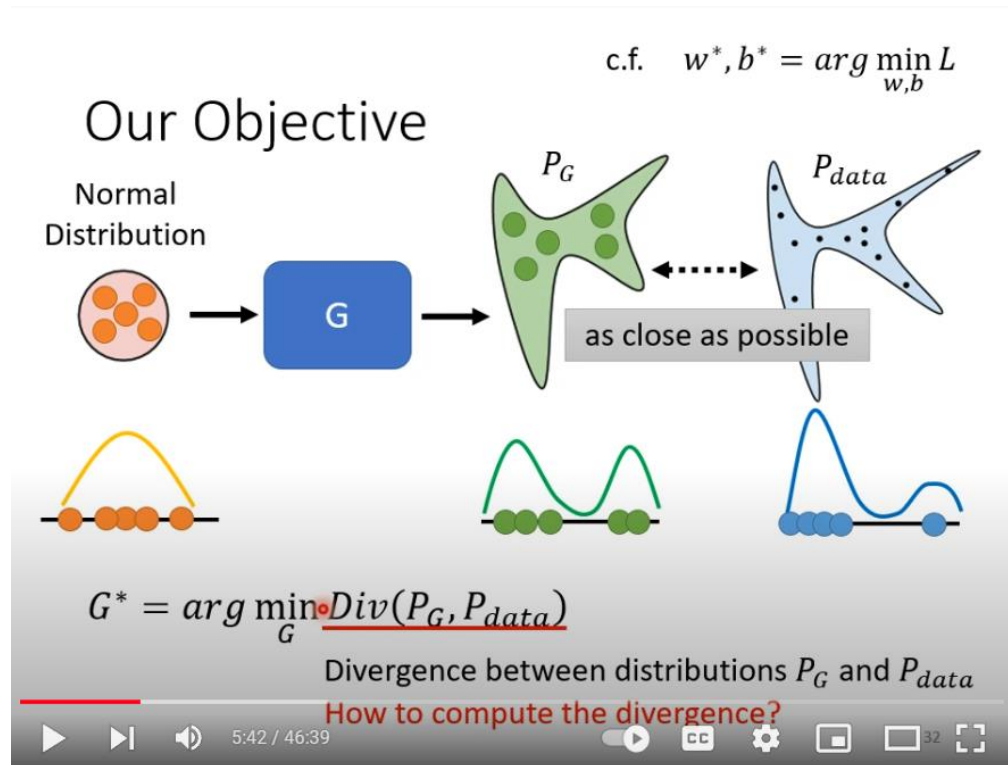
POLITICAL FIGURES

Belgium's Socialistische Partij Anders (sp.a) posted this video on Facebook back in May 2018 showing Trump taunting Belgium for remaining in the Paris climate agreement.
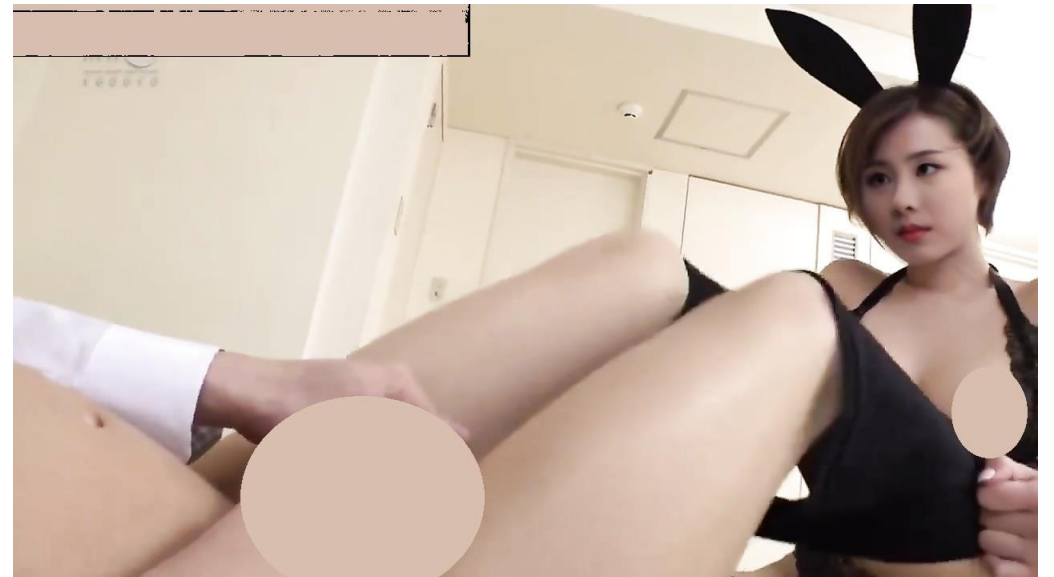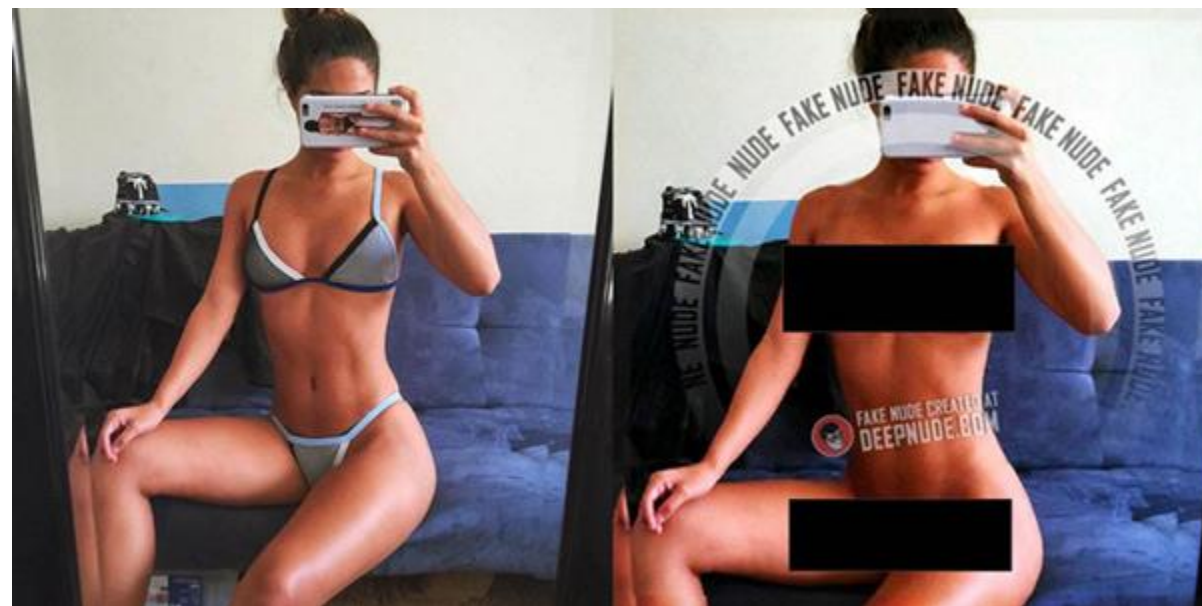
https://www.youtube.com/watch?v=gLoI9hAX9dw

CLASSIC AND EROTIC MOVIES

The face was replaced by the face of
a famous Chinese actor

REMOVING CLOTHES FROM IMAGES

LEGAL RESPONSE?

- **Fraud?**

- **Credibility and authenticity?**

https://en.wikipedia.org/wiki/File:Dictators_-_Kim_Jong-Un_by_RepresentUs.webm

THANK YOU FOR YOUR ATTENTION