Necessity of visual scepticism in the post-truth era

How Generative Adversarial Networks (GANs) changed the way we look on the world

Lenka Hámošová
Pavol Rusnák


There is
Before and After
Deep Fake videos

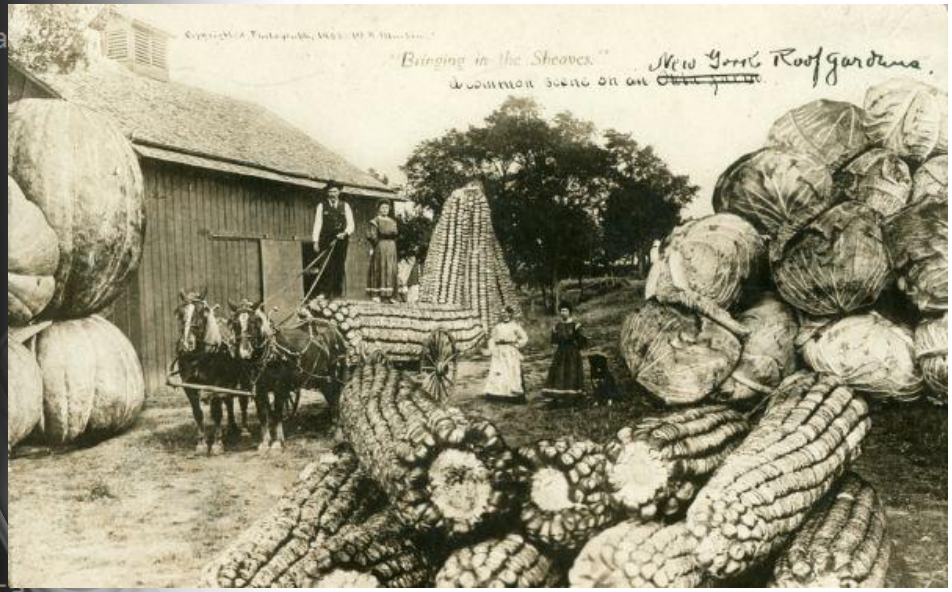
Are we currently standing on the verge of justified permanent suspicion against every piece of visual information?


1.THE BEFORE

Spirit Photography

William H. Mumler in the 1860s
- double exposure technique

Tall-Tale Postcards



Bringing in the Sheaves, 1908
Kansas City, Missouri, USA

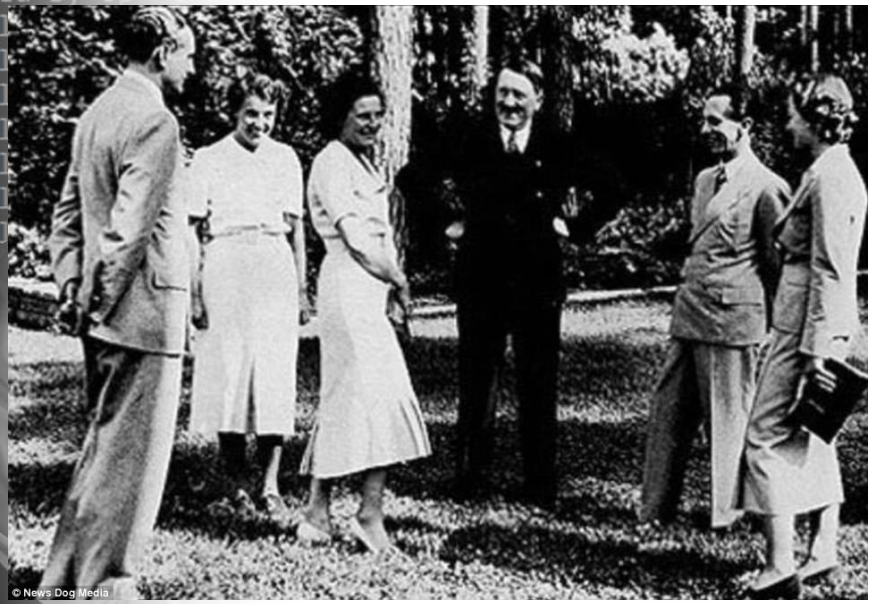
Kickers with Frogs, 1913
Waupun, Wisconsin, USA

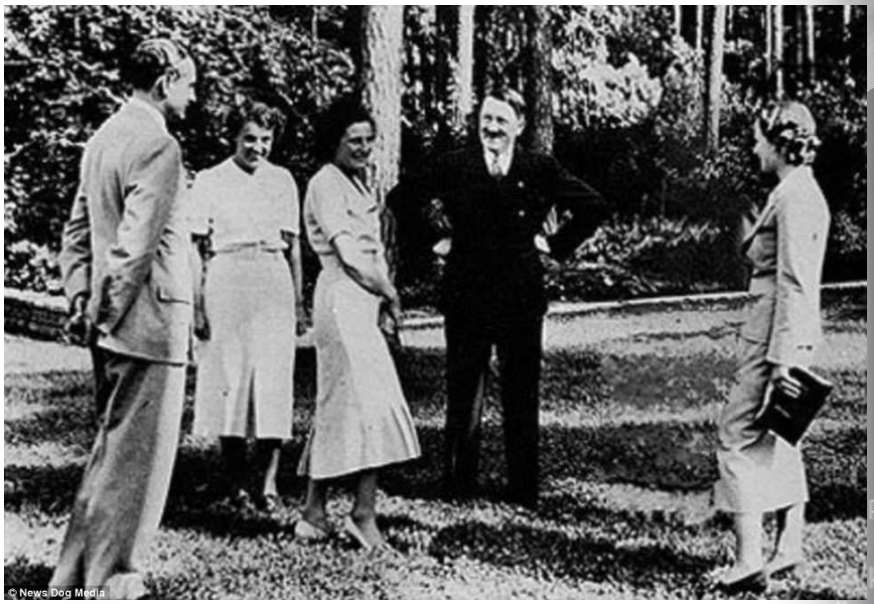
Cottingley Fairies hoax

Retouching history

И. В. Сталин и С. М. Киров в Ленинграде (1926 г.)
Слева — И. Е. Аптинков, справа — П. М. Шверник

Retouching history





Different head / different body

Fake news

https://en.unesco.org/fightfakenews

Strategies to detect digitally manipulated images

- Image reverse search
- Get image metadata
- Check the light/shadows
- Use photo-manipulation detection tools, e.g. *https://29a.ch/photo-forensics/*

Forensically βeta Open File Help

Magnifier

Magnification 4

Enhancement
Histogram Equalization

Clone Detection

Error Level Analysis

Noise Analysis

Level Sweep

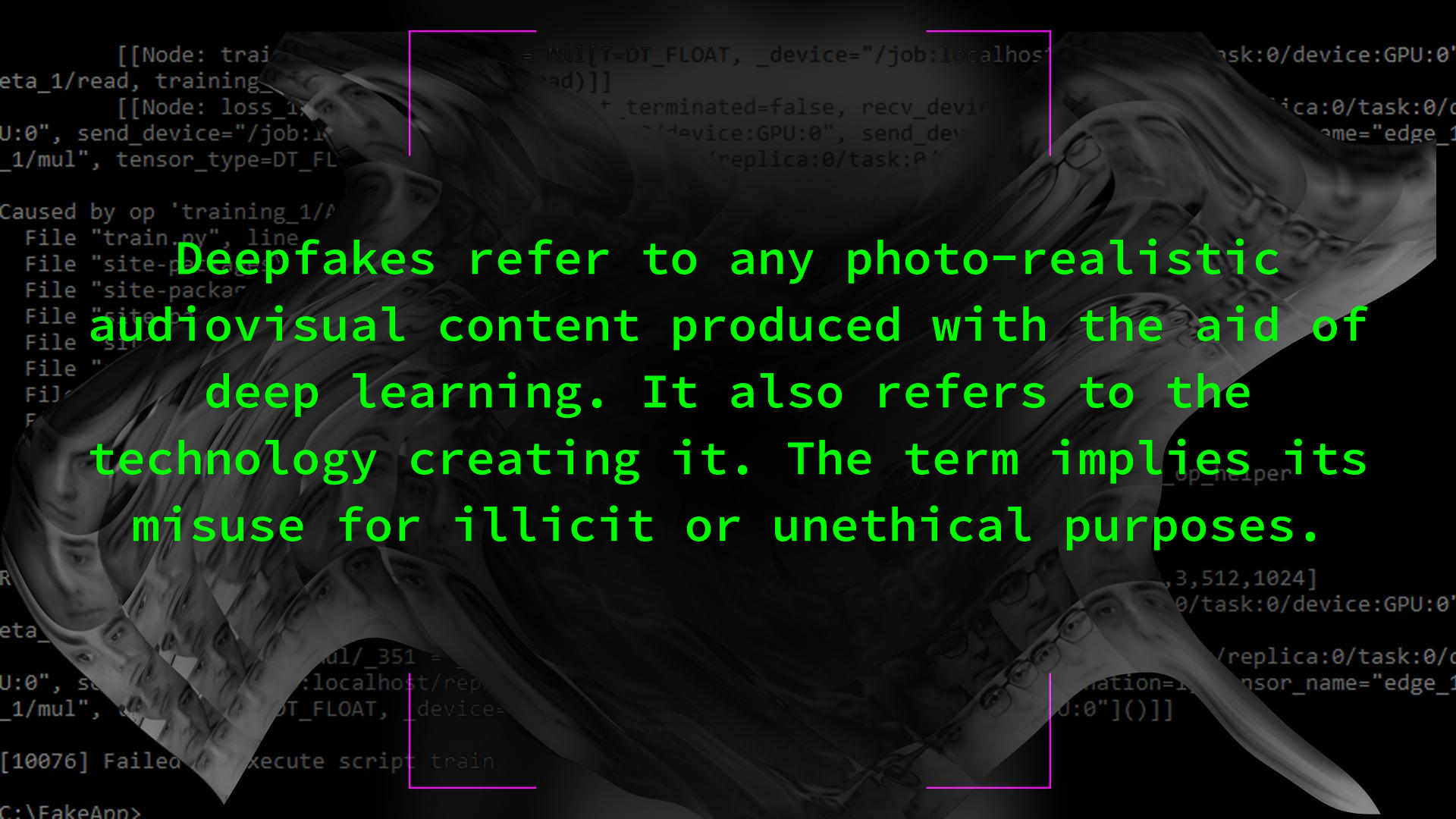
Luminance Gradient

Principal Component

"Synthesizing Obama: Learning Lip Sync from Audio" (The University of Washington 2017)

Researchers explaining the algorithm: https://youtu.be/9Yq67CjDqvw


2. DEEP FAKE VIDEOS


Deepfakes refer to any photo-realistic audiovisual content produced with the aid of deep learning. It also refers to the technology creating it. The term implies its misuse for illicit or unethical purposes.

FakeApp available on Reddit

reddit

r/GifFakes

Posts

Posted by u/GifFakes 1 year ago

55

FakeApp download links and How-To Guide

FakeApp was created to give everyday people access to realistic faceswapping technology for creative and exploratory use.

If you use this tool, please use it responsibly.

The FakeApp website will always have up-to-date links.

FakeApp v2.1:
https://mega.nz/#!z0dxTTJS!9azjngbENwU2WSL7kfJPAIyilMG94bHbhARt4073s6Q

Previous How-To thread:
https://www.reddit.com/r/GifFakes/comments/7w1pp9/how_to_use_fakeapp_to_make_your_own_faceswap_gifs/

FakeApp v1.1 (Google Drive):
https://drive.google.com/file/d/1d9jCoPmPzjtBRtRyZhf2I1GYzyNBDgiR/view

FakeApp v1.1 (Mega Mirror):
https://mega.nz/#!hTgA2b6b!mI6k9dFt_w__jIEUQO2ZePhzFMg6JWUpBZWiV2TDgs4

22 Comments Share Save 100% Upvoted

COMMUNITY DETAILS

r/GifFakes

10.5k Subscribers 73 Online

Deepfake face swapped gifs and videos. "I'm a dude playin a dude disguised as another dude." (Also known as deepfakes, faceswaps, gif fakes, etc.)

JOIN

CREATE POST

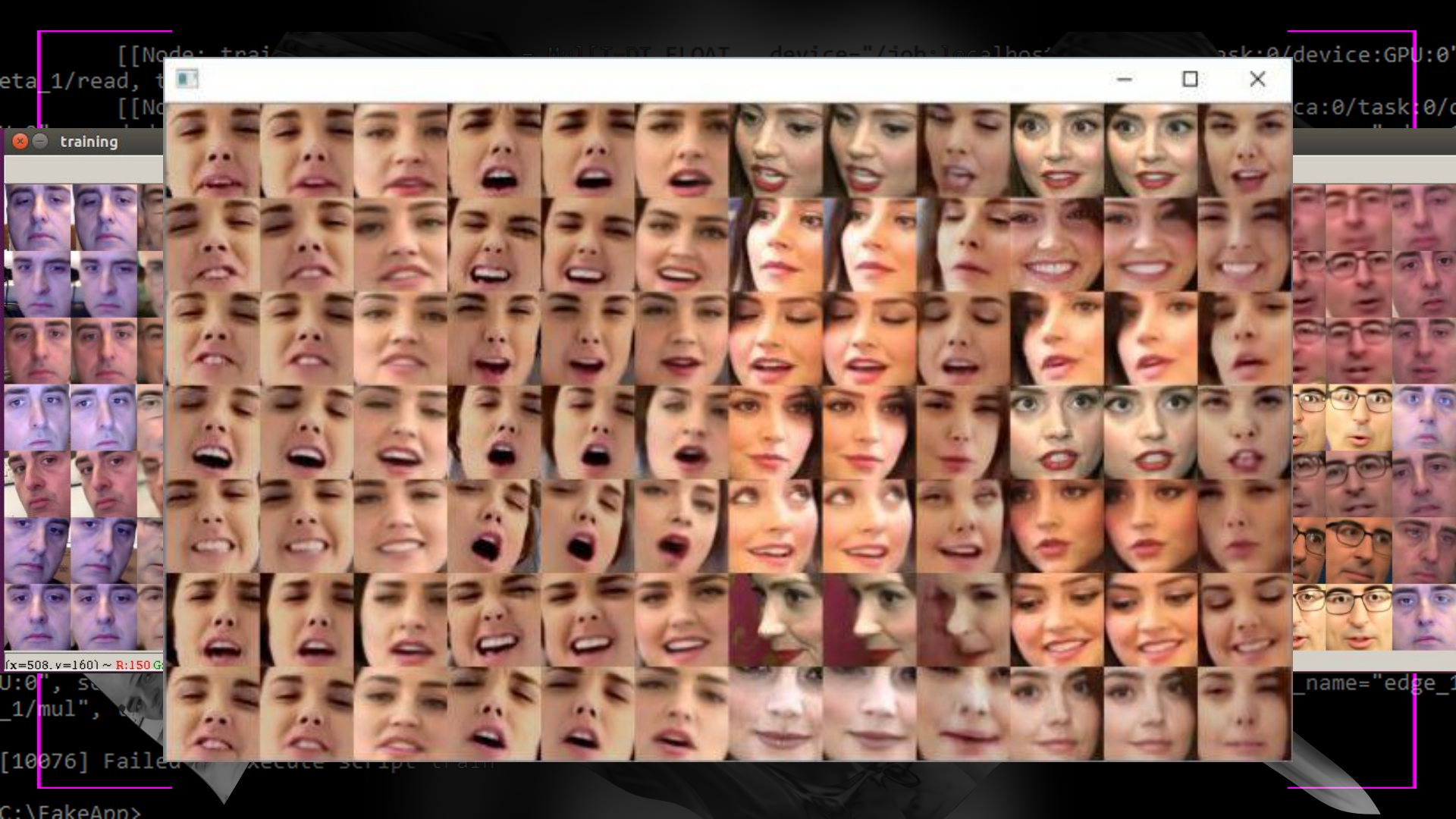
MODERATORS

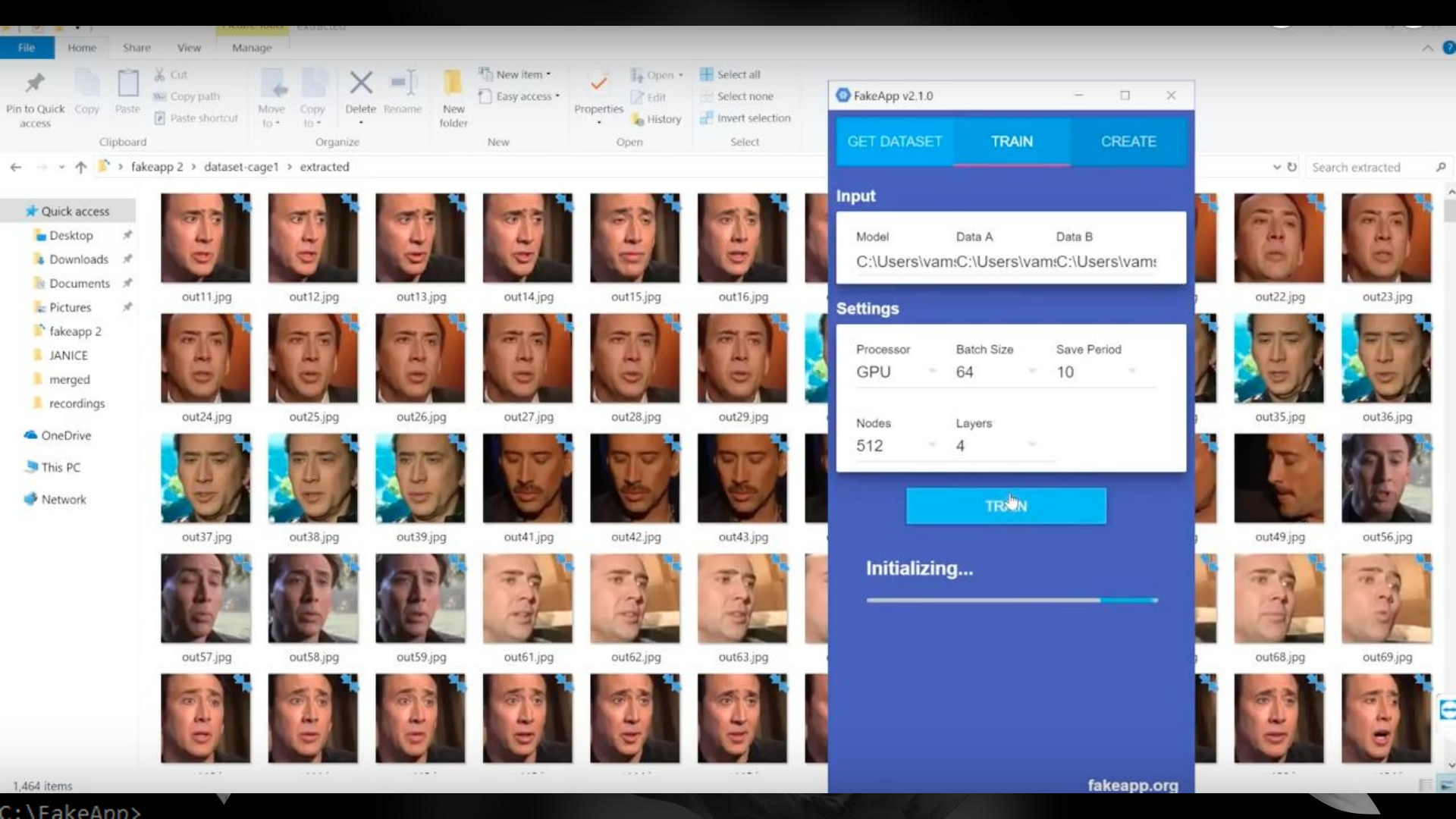
u/POVOH

u/GifFakes

u/derpfakes Content Creator

u/deepfakeapp FakeApp Developer


training


File
Home
Share
View
Manage
Pin to Quick access
Copy
Paste
Cut
Copy path
Paste shortcut
Clipboard
Move to
Copy to
Delete
Rename
New folder
New item
Easy access
Organize
New
Properties
Open
Edit
History
Open
Select all
Select none
Invert selection
Select
fakeapp 2 › dataset-cage1 › extracted
Search extracted
Quick access
Desktop
Downloads
Documents
Pictures
fakeapp 2
JANICE
merged
recordings
OneDrive
This PC
Network
out11.jpg
out12.jpg
out13.jpg
out14.jpg
out15.jpg
out16.jpg
out22.jpg
out23.jpg
out24.jpg
out25.jpg
out26.jpg
out27.jpg
out28.jpg
out29.jpg
out35.jpg
out36.jpg
out37.jpg
out38.jpg
out39.jpg
out41.jpg
out42.jpg
out43.jpg
out49.jpg
out56.jpg
out57.jpg
out58.jpg
out59.jpg
out61.jpg
out62.jpg
out63.jpg
out68.jpg
out69.jpg
1,464 items
FakeApp v2.1.0
GET DATASET
TRAIN
CREATE
Input
Model
Data A
Data B
C:\Users\vam
C:\Users\vam
C:\Users\vam
Settings
Processor
GPU
Batch Size
64
Save Period
10
Nodes
512
Layers
4
TRAIN
Initializing...
fakeapp.org
C:\FakeApp>

Belgian socialist party circulates ‘deep fake’ Donald Trump video


Deep Learning

Deep Learning

machine learning

we teach computer systems to effectively perform a specific task
without explicit instructions

computer vision, speech recognition, natural language processing, audio recognition,
social network filtering, machine translation, bioinformatics, drug design, medical image
analysis, material inspection, board games, etc.

using deep neural networks (DNNs)

Deep Neural Network (DNN)

deep

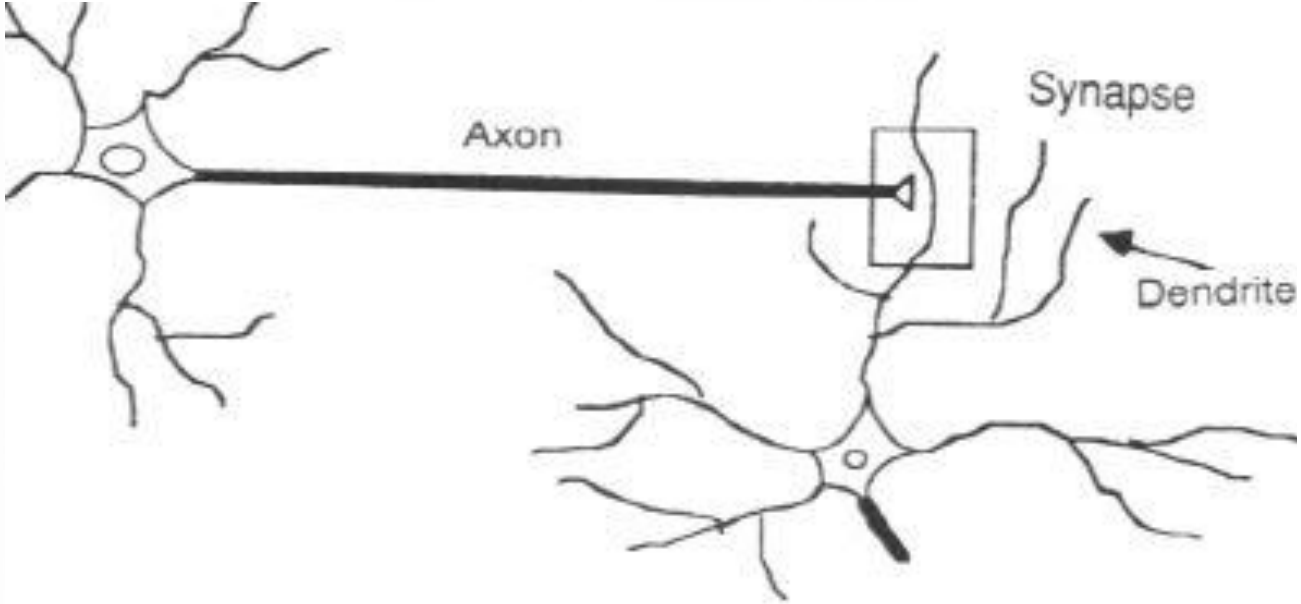
network

of neurons

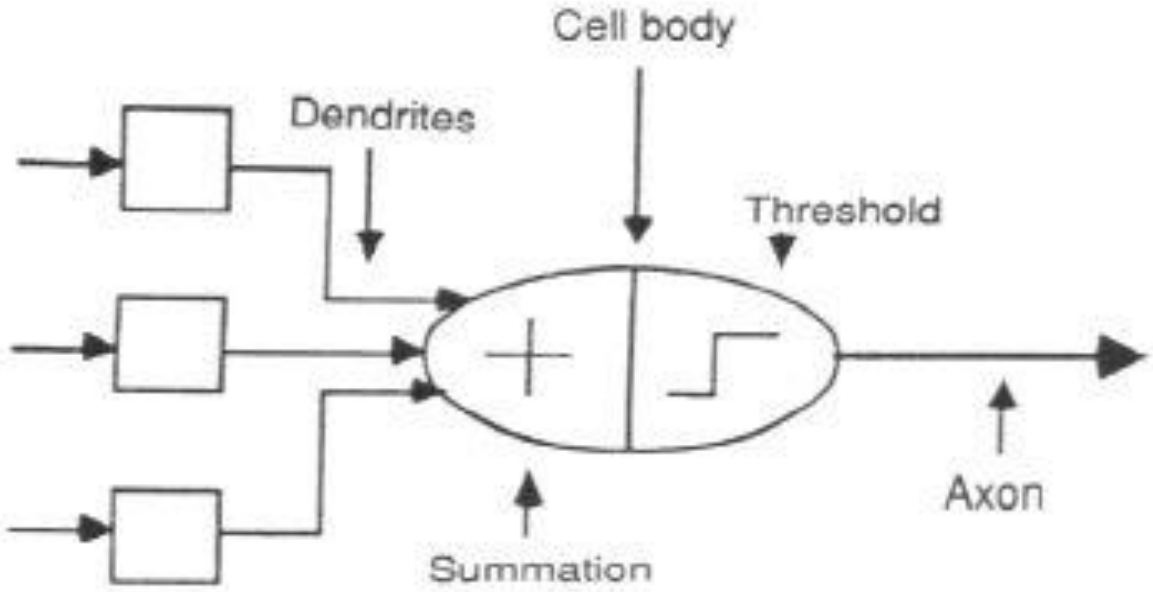
inspired by information processing and communication patterns

in biological nervous systems

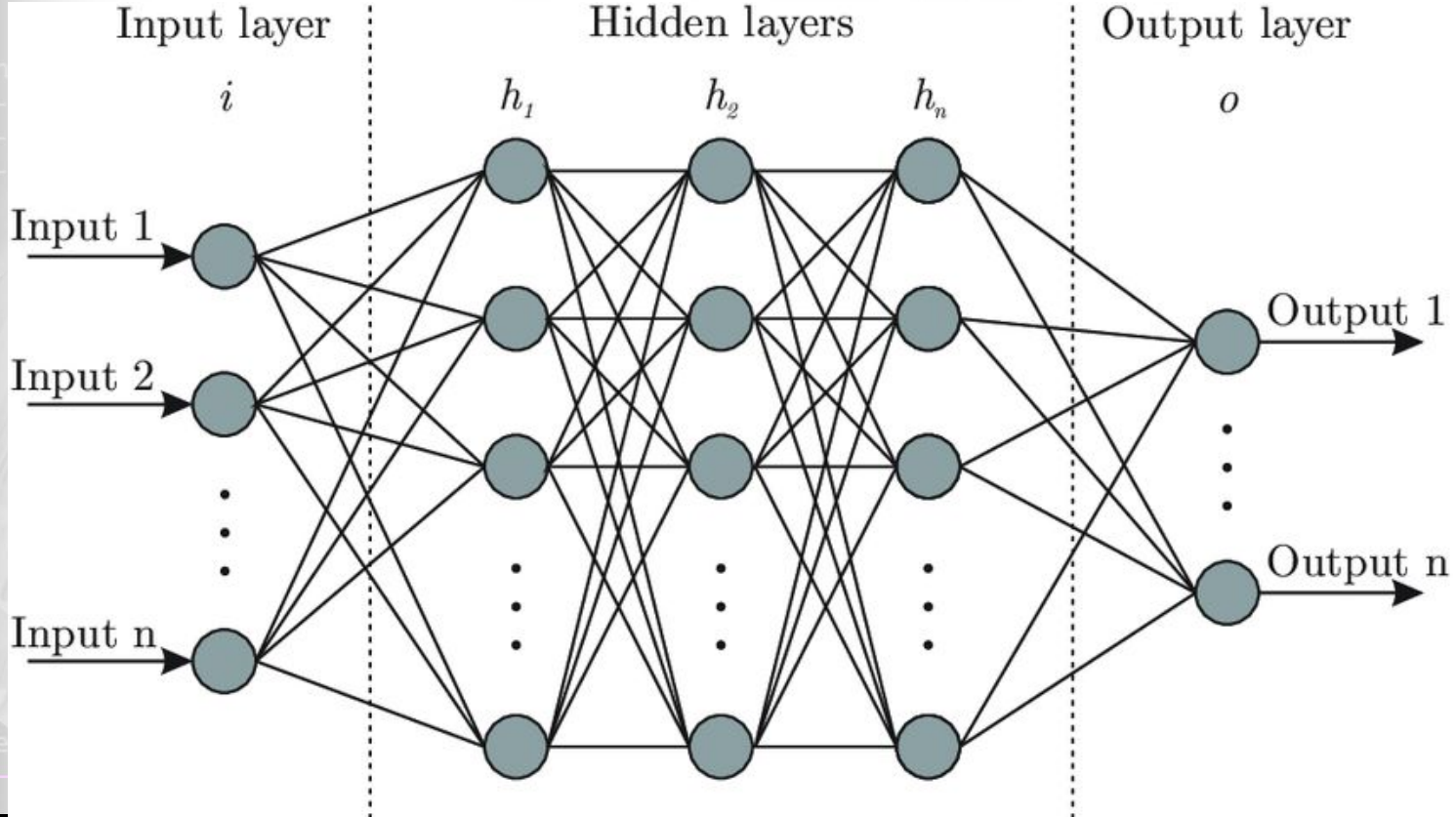
Neuron



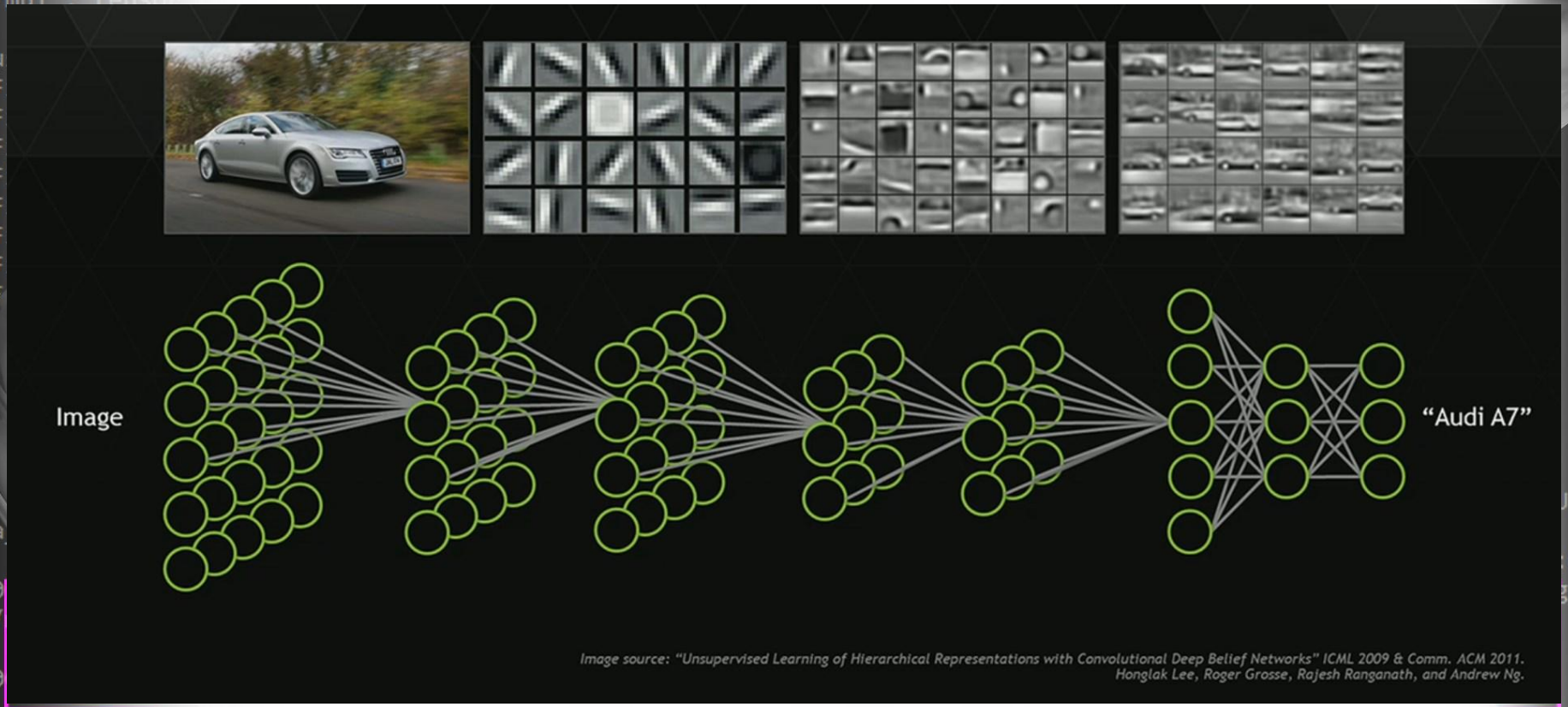
Artificial Neuron



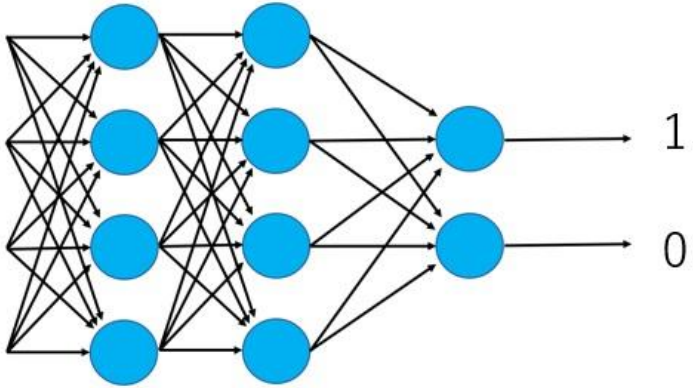
Artificial Neural Network (ANN)

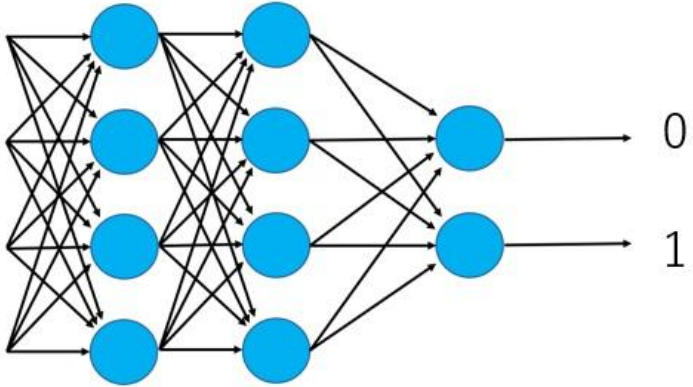


Deep Neural Network (DNN)

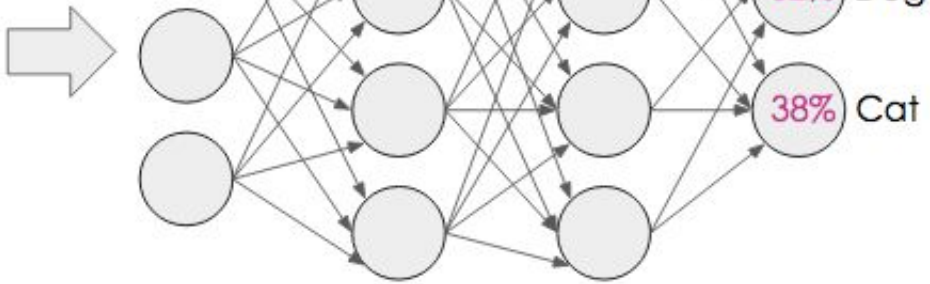


Deep Neural Network (DNN)



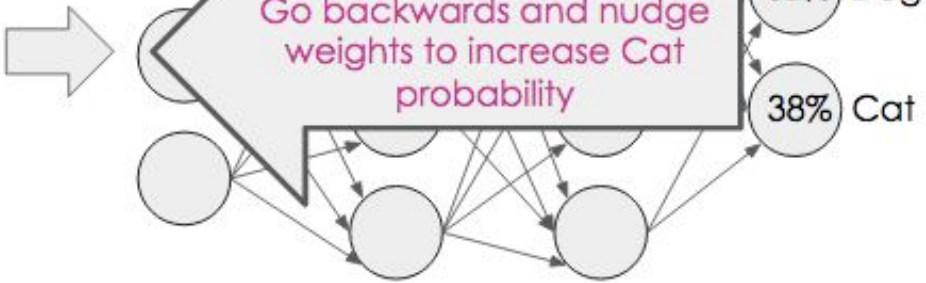


Deep Neural Network (DNN)



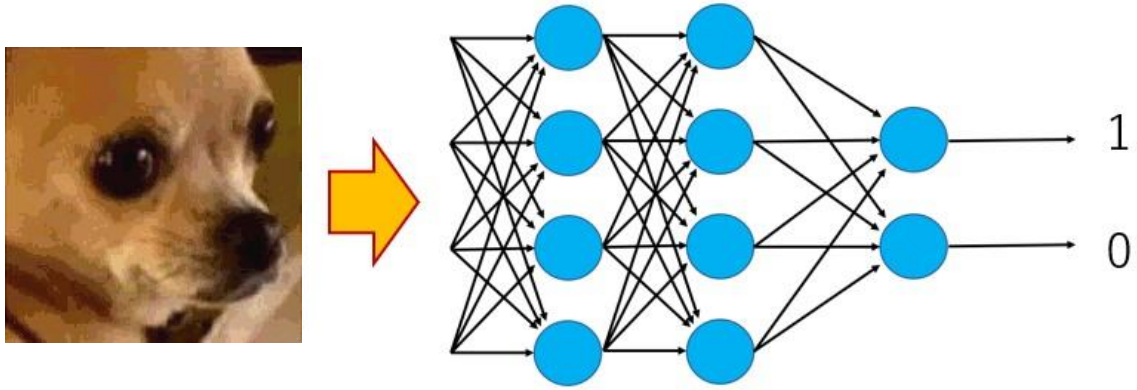
It should be
100% Cat :(

Deep Neural Network (DNN)

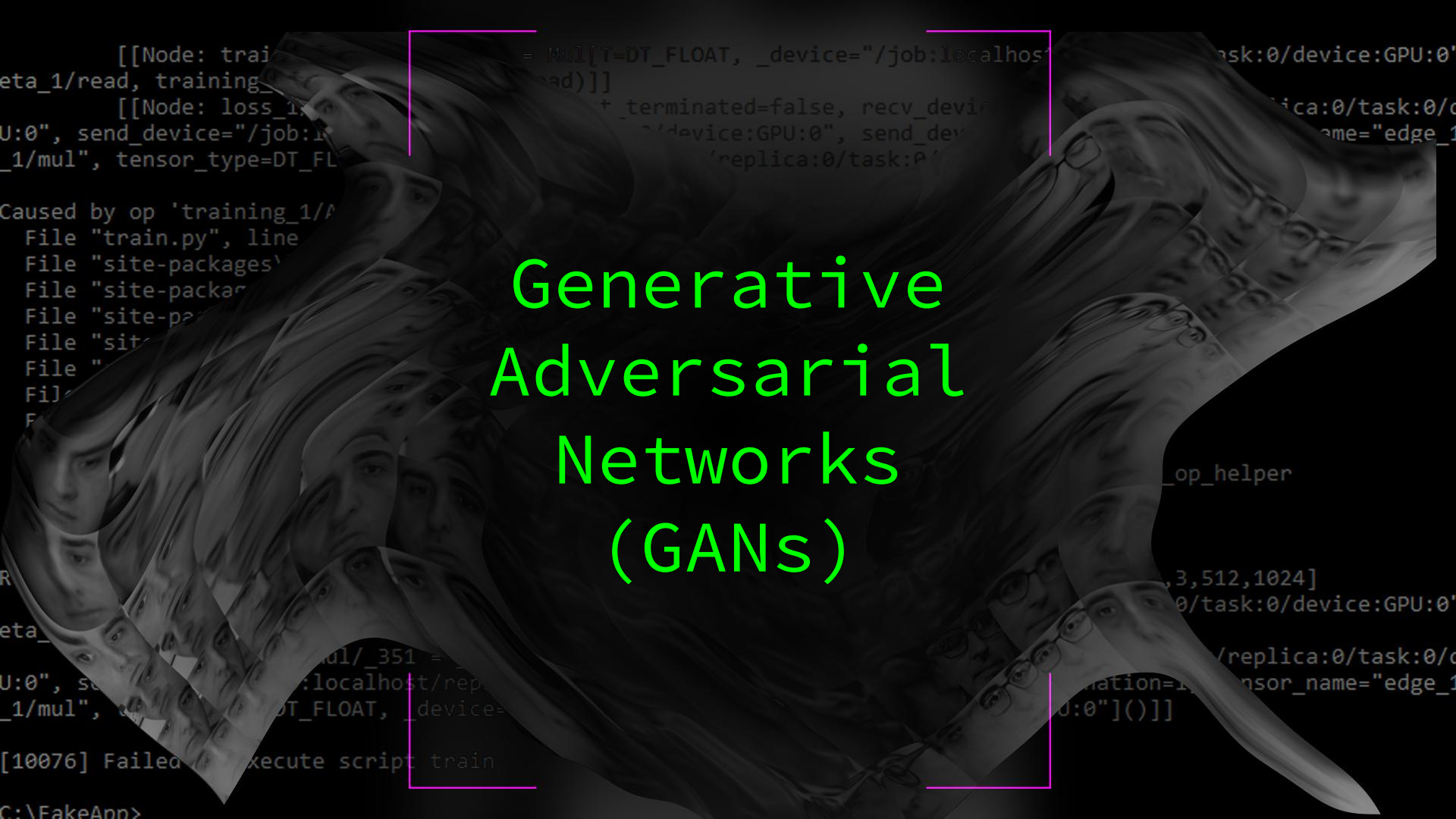


It should be 100% Cat :(

Deep Neural Network (DNN)



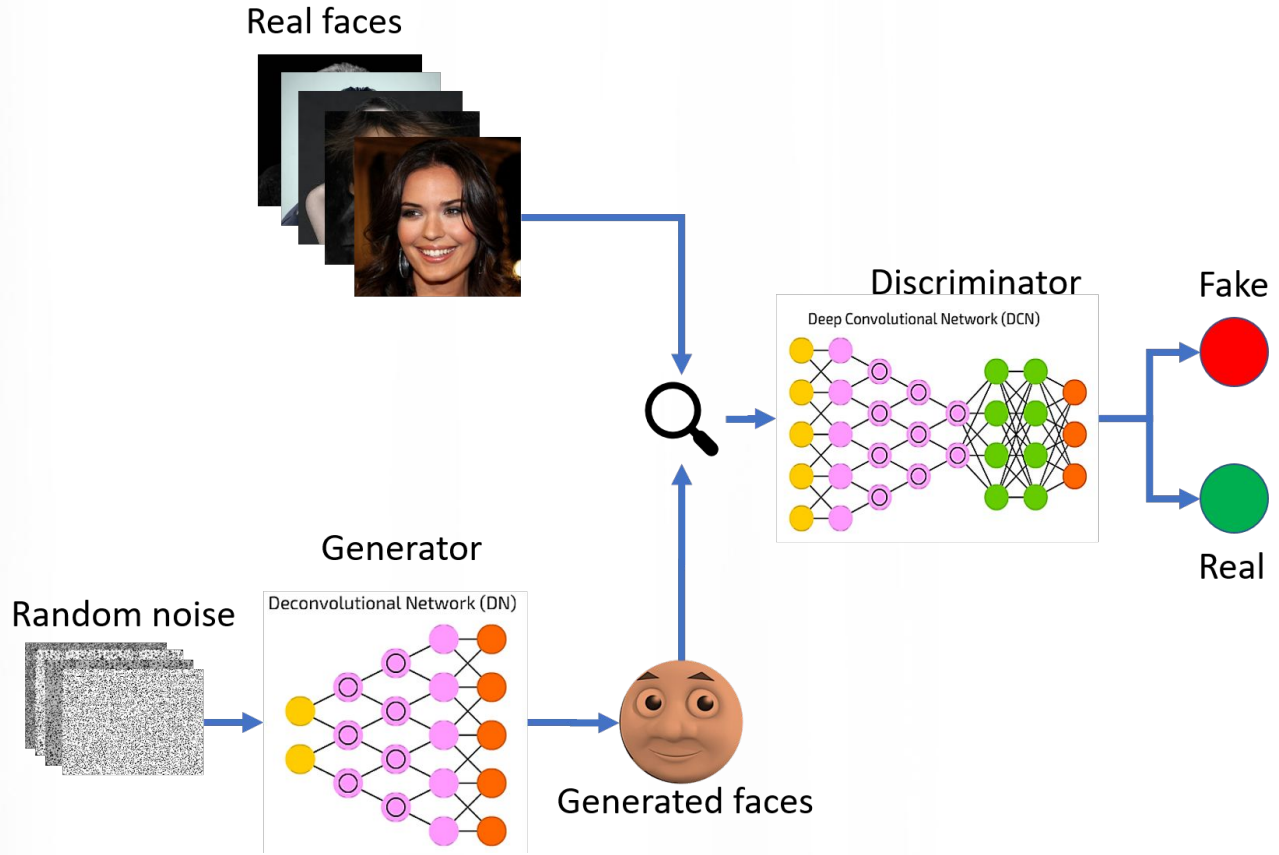

3. THE AFTER


Generative
Adversarial
Networks
(GANs)

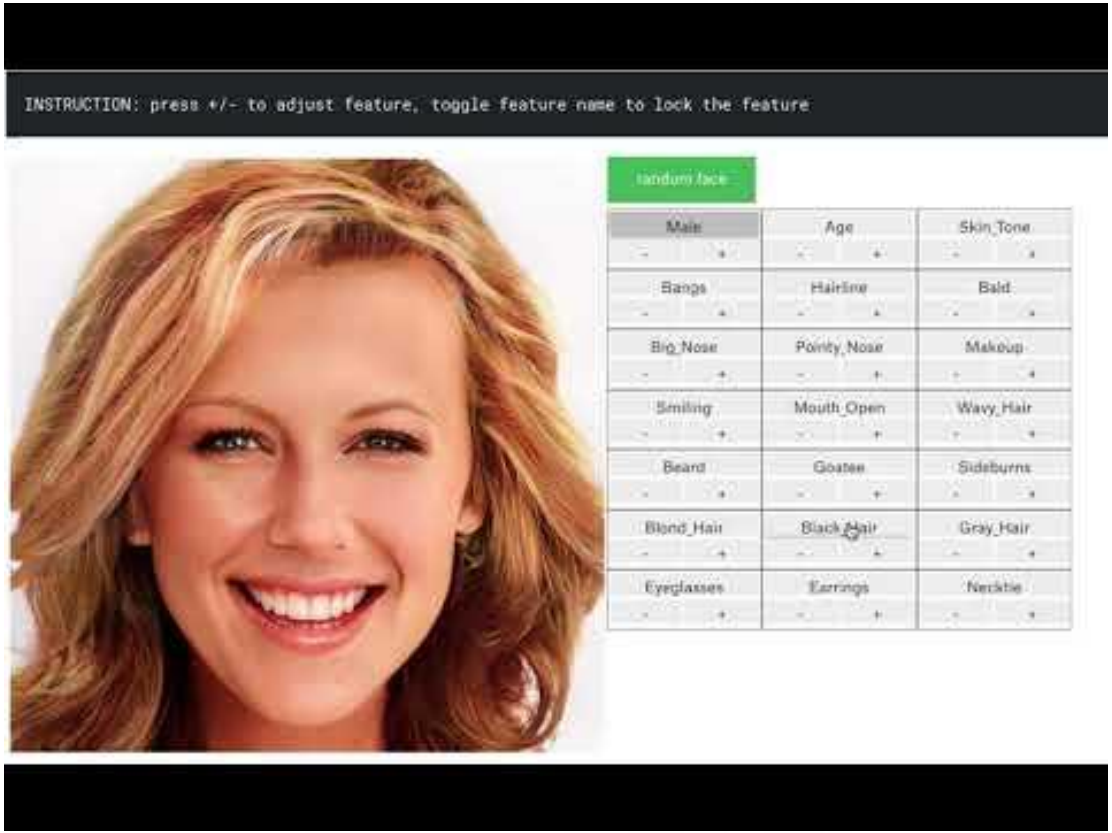
Generative Adversarial Networks (GANs)

- introduced in 2014 by Ian Goodfellow and other researchers at the University of Montreal
- *https://arxiv.org/abs/1406.2661*
- two neural networks contesting each other
  - discriminative network
  - generative network

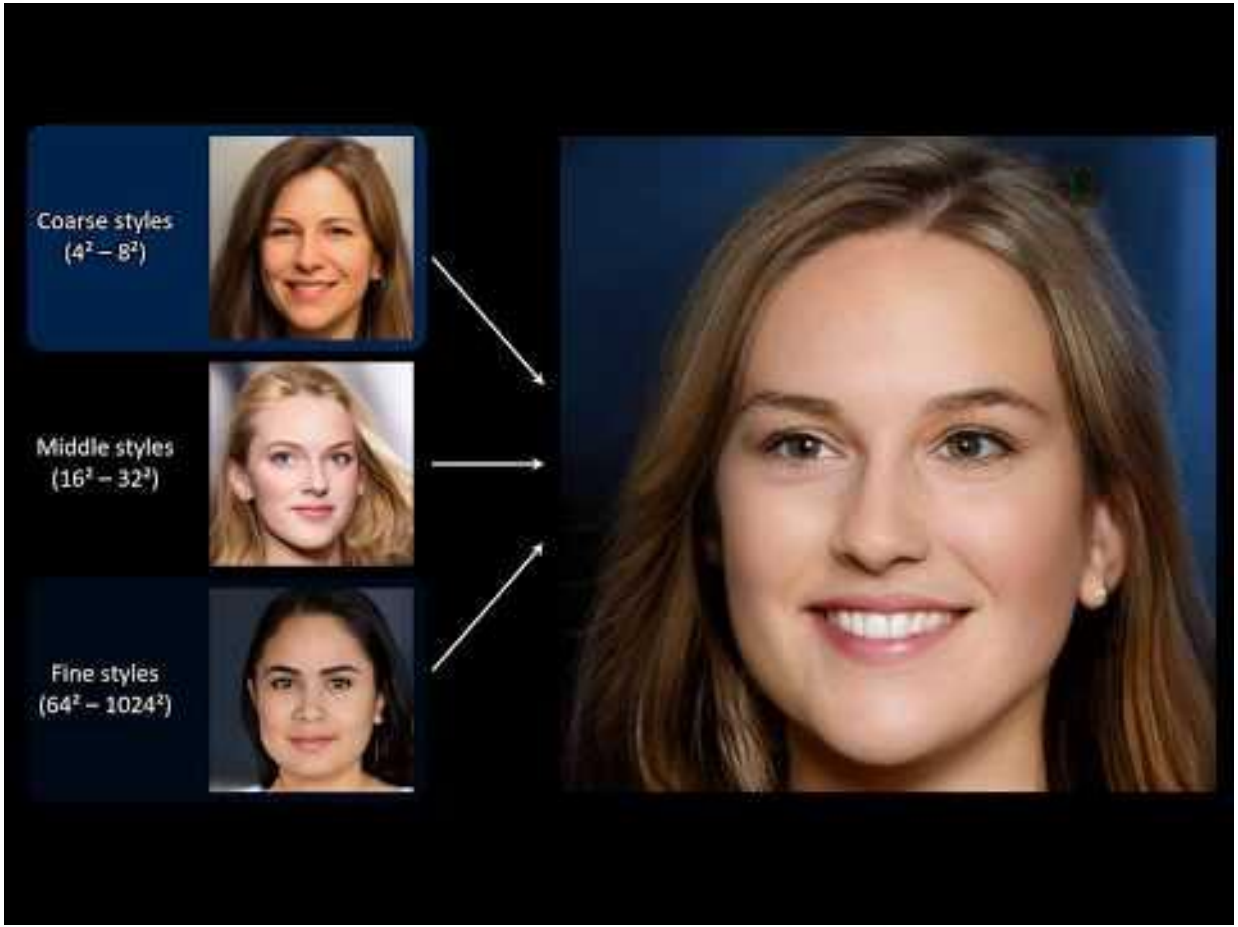
Generative Adversarial Networks (GANs)



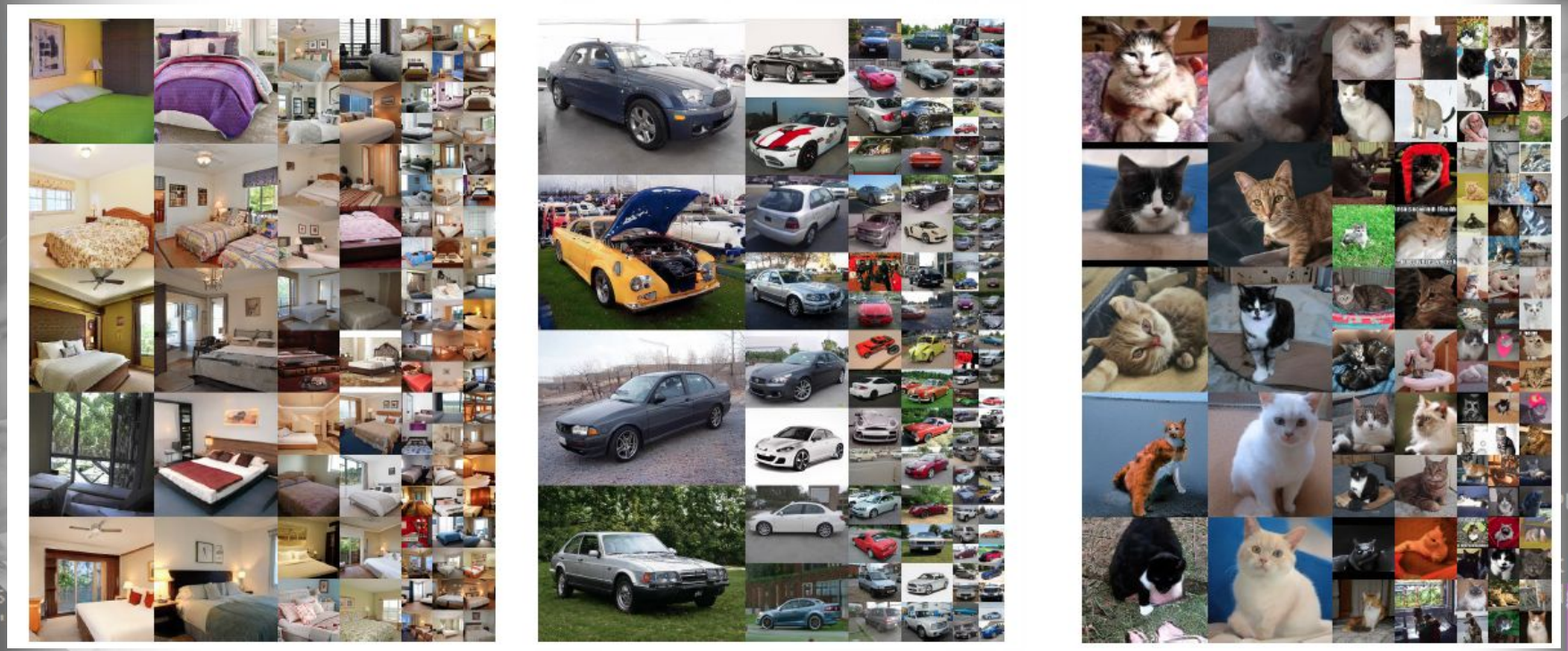
StyleGAN (NVIDIA)



StyleGAN (NVIDIA)

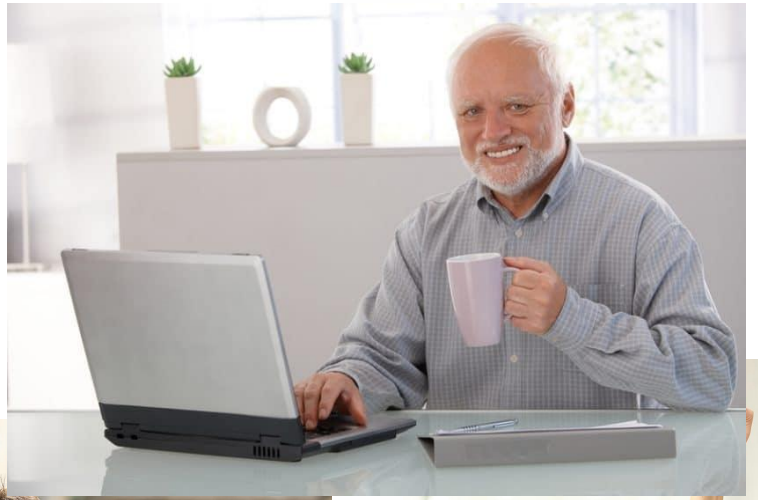


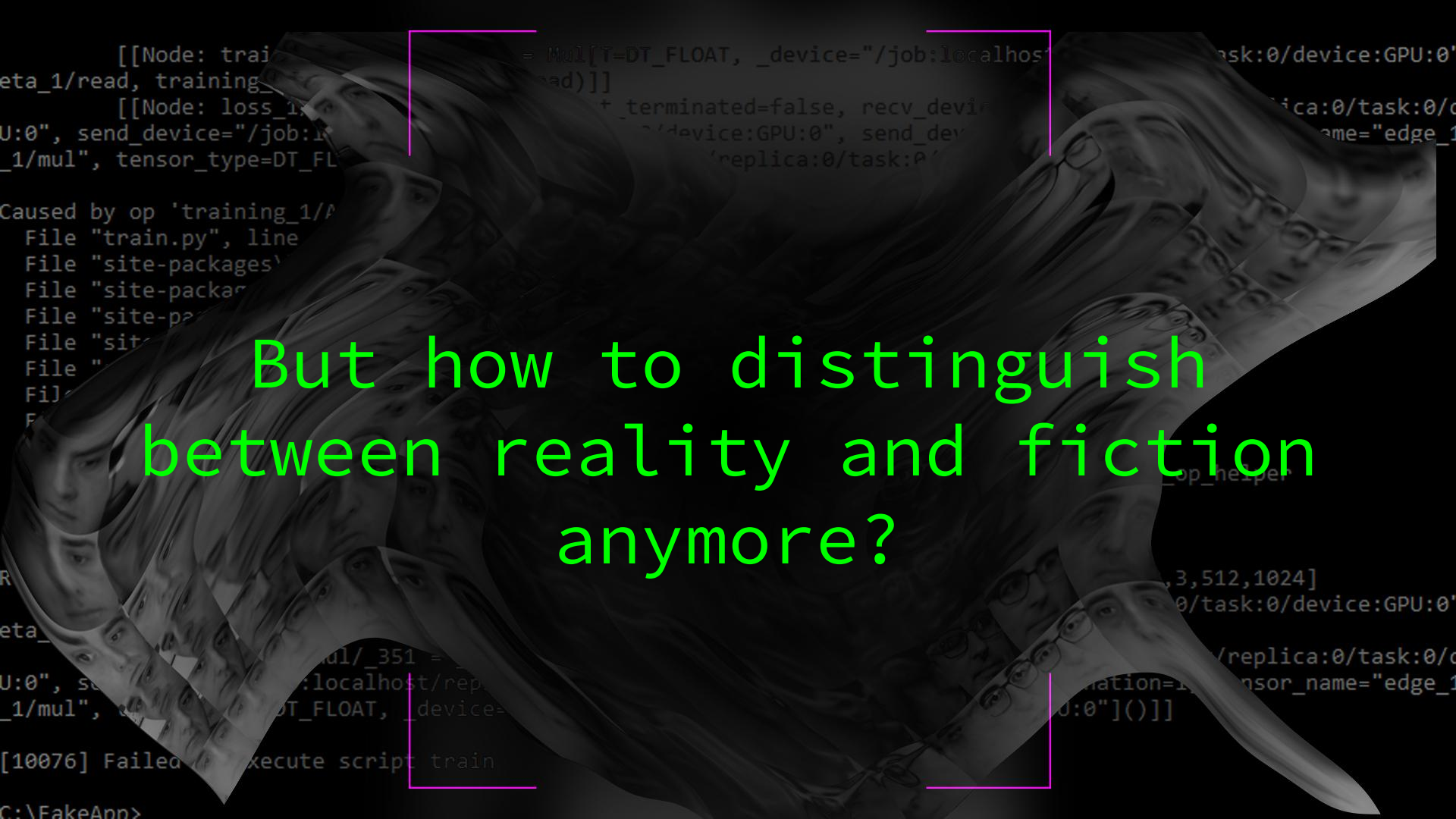
StyleGAN (NVIDIA)



thispersondoesnotexist.com

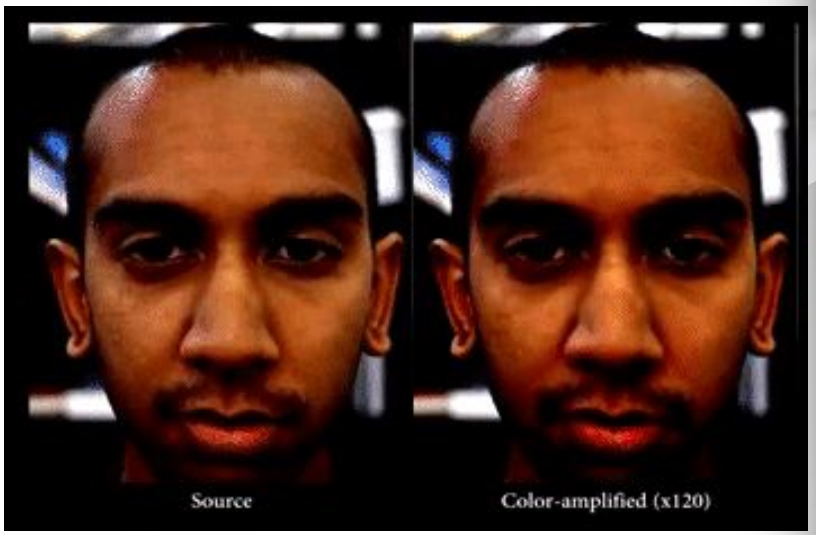
NVIDIA Research: “GauGAN”



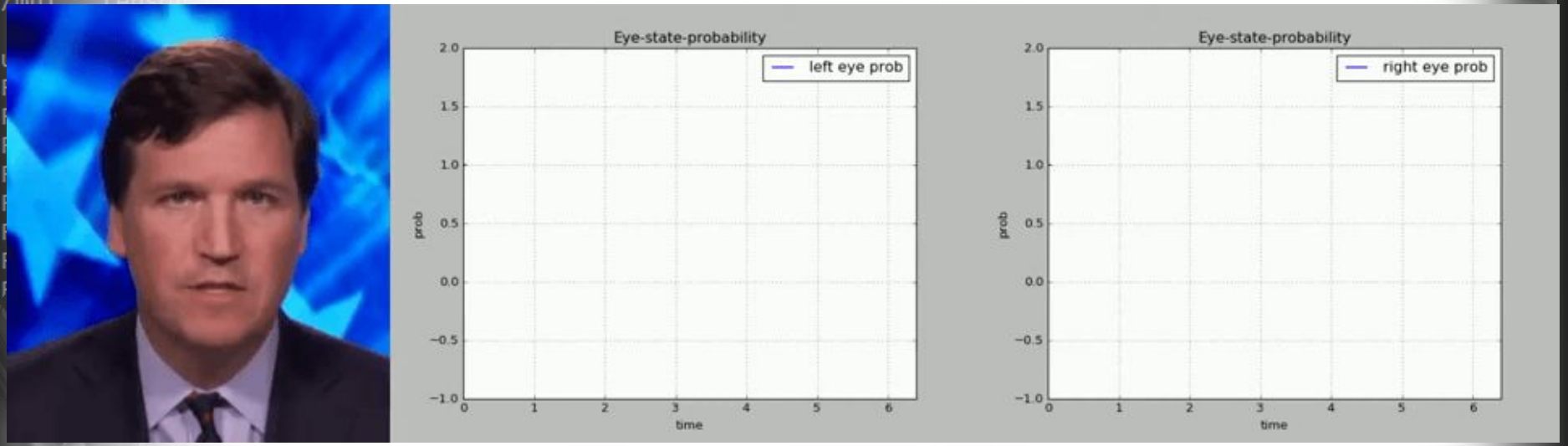

But how to distinguish between reality and fiction anymore?

THIS LOOKS SHOPPED

I CAN TELL FROM SOME OF THE PIXELS AND FROM SEEING QUITE A FEW SHOPS IN MY TIME



MIT Video Magnification: *https://people.csail.mit.edu/mrub/vidmag/*

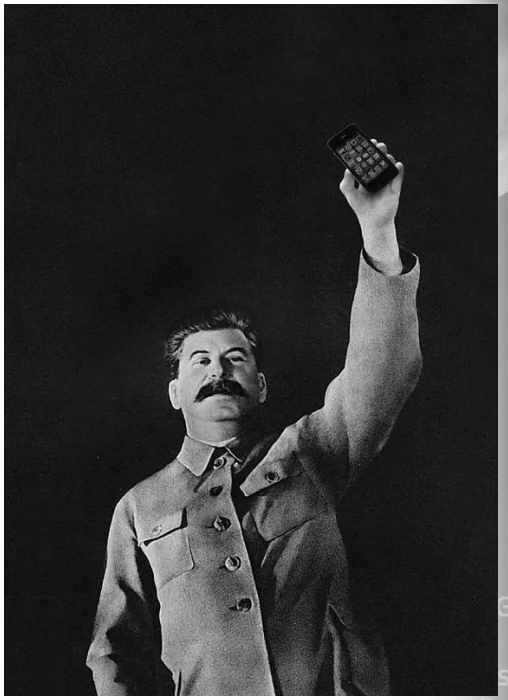


In Ictu Oculi: Exposing AI Generated Fake Face Videos by Detecting Eye Blinking: *https://arxiv.org/abs/1806.02877*


4. CRYPTOGRAPHY

Authenticity of digitally signed image

- Canon OSK-E3 (Original Data Security Kit), 2007
- encryption/decryption, **verification** function for **authenticating** image **originality**
- not only image data (pixels)
- but also metadata (GPS coordinates + time)
- only top DSLR models

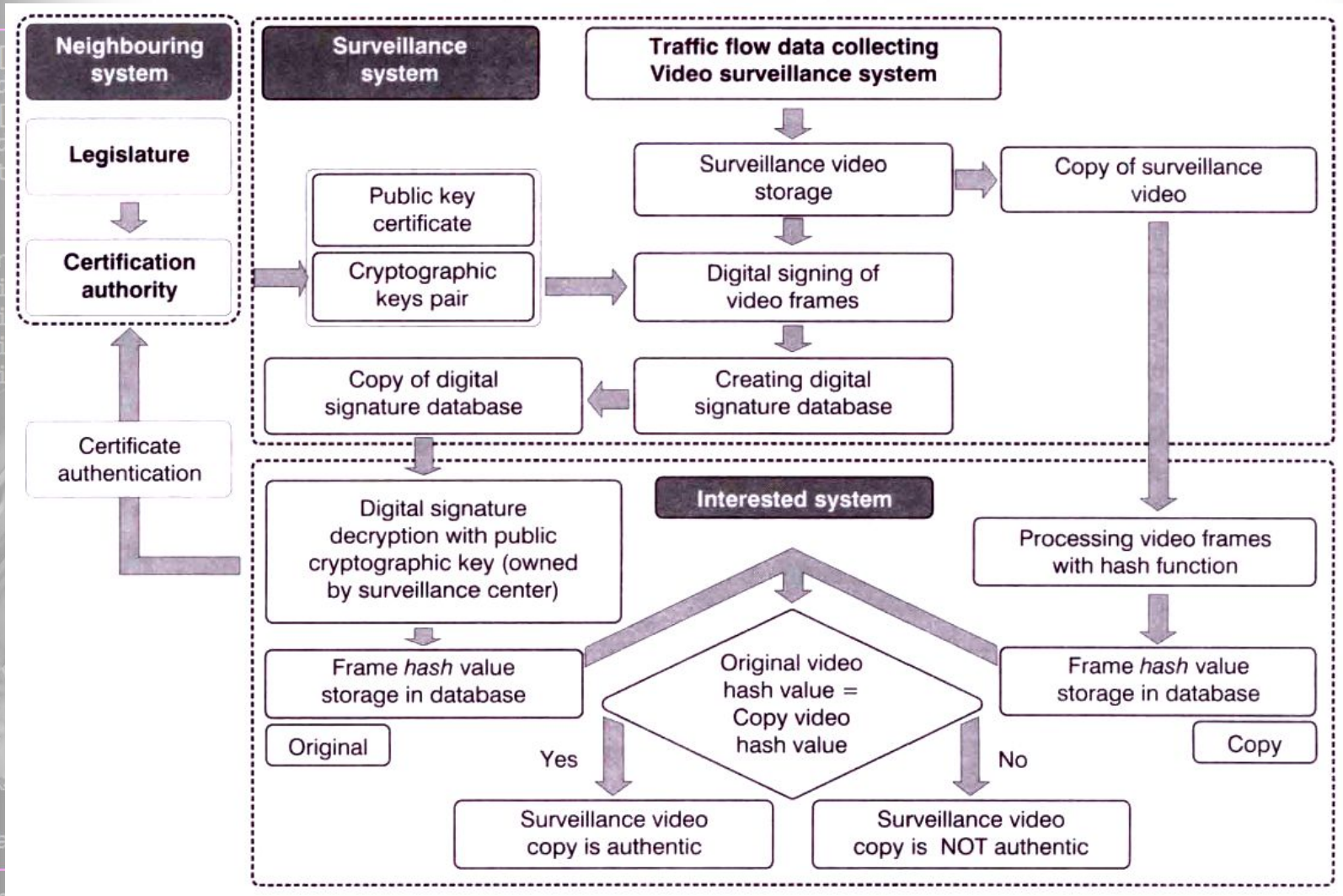


hacked in 2010 by Dmitry Sklyarov

Authenticity of digitally signed video

Credibility and Authenticity of Digitally Signed Videos in Traffic (2008)

by Ivan Grgurević, Adam Stančić, Pero Škorput


Neighbouring system
Legislature
Certification authority
Certificate authentication
Surveillance system
Traffic flow data collecting
Video surveillance system
Surveillance video storage
Copy of surveillance video
Public key certificate
Cryptographic keys pair
Digital signing of video frames
Copy of digital signature database
Creating digital signature database
Interested system
Digital signature decryption with public cryptographic key (owned by surveillance center)
Processing video frames with hash function
Frame hash value storage in database
Original video hash value = Copy video hash value
Frame hash value storage in database
Original
Copy
Yes
No
Surveillance video copy is authentic
Surveillance video copy is NOT authentic

Problems

Centralized

Not universal

Algorithms can be hacked

Insufficient proof of creation

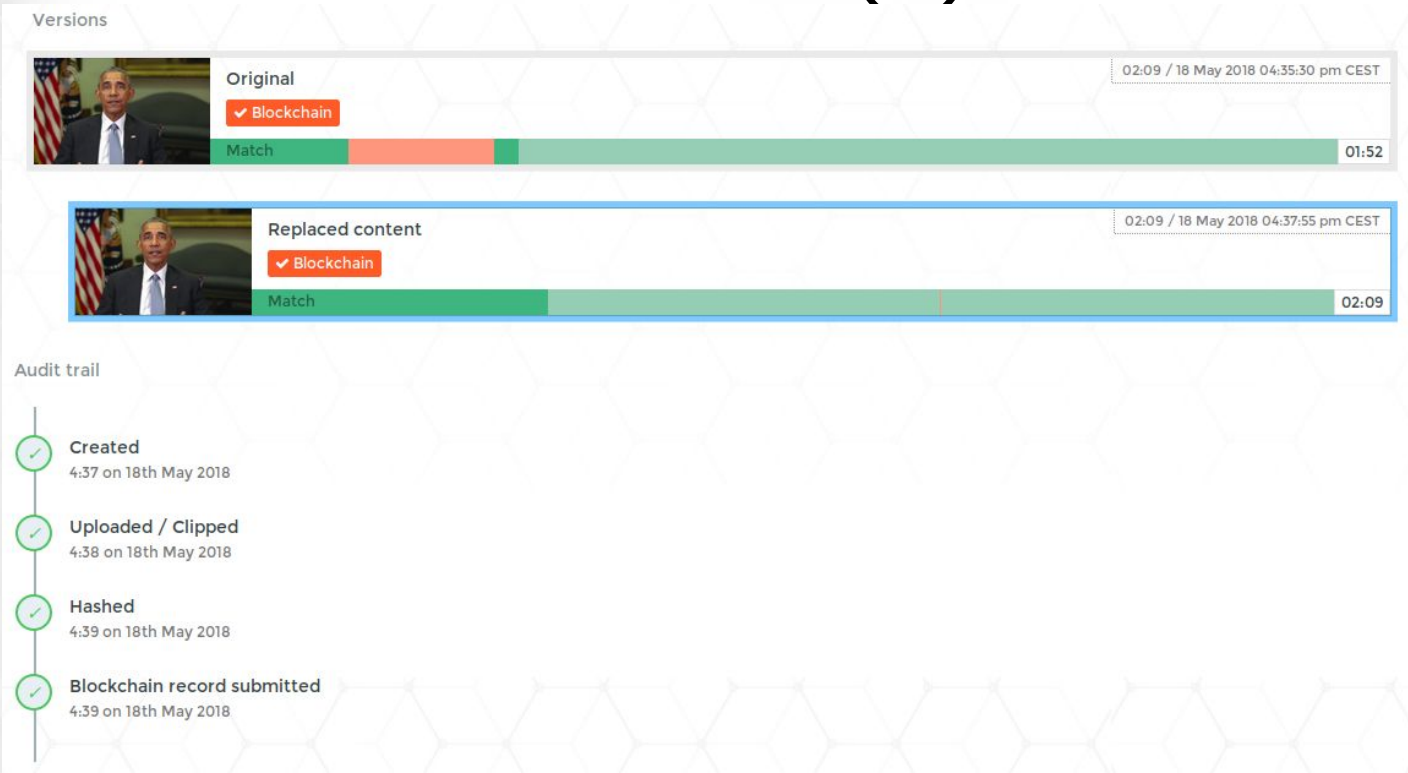
Solution

1. record 30s of video
2. compute hash of recorded chunk
3. submit hash to public blockchain
4. go to step 1

Solution (?)

https://app.ambervideo.co/public/

Solution (?)



https://app.ambervideo.co/public/

For more information check YouTube's video: https://youtu.be/1PGm8LslEb4


5. THE FUTURE

Thank you!

Contact us:

@lenkahamosova
@pavolrusnak

lenka@hamosova.com
pavol@rusnak.io

Check also: ALT.TAB (FB)
Critical & Speculative
Design Platform