HELLO!

I am Alessandro Ferrari

Founder & CEO - ARGO Vision | www.argo.vision

- 2001 - Finalista premio «**Net Economy Ideas**» - Uni. Bocconi
- 2002 - **Diploma in Elettronica**, ITIS Galilei Milano
- 2006 - **Laureato in Comunicazione Digitale** @UniMI
- 2008 - **Laureato in Informatica** Magistrale @UniMI
- 2009 - R&D in **Computer Vision e Machine Learning**
- 2016 - **ARGO Vision** (**www.argo.vision**), startup #AI
- 2020 - ARGO Vision entra nel gruppo **SEA Vision / Marchesini**
- 2022 - Docente @UniPV all'interno del corso di "**Persuasive Design**"
- 2023 - Docente @UniPV, corso "**Comunicazione Tech & AI**"

NEURAL AGEs

The proto-age

The ‘40s to 2012

First wave

The 2012 to present

«Conscious» age

The 2030: what’s next?

#history

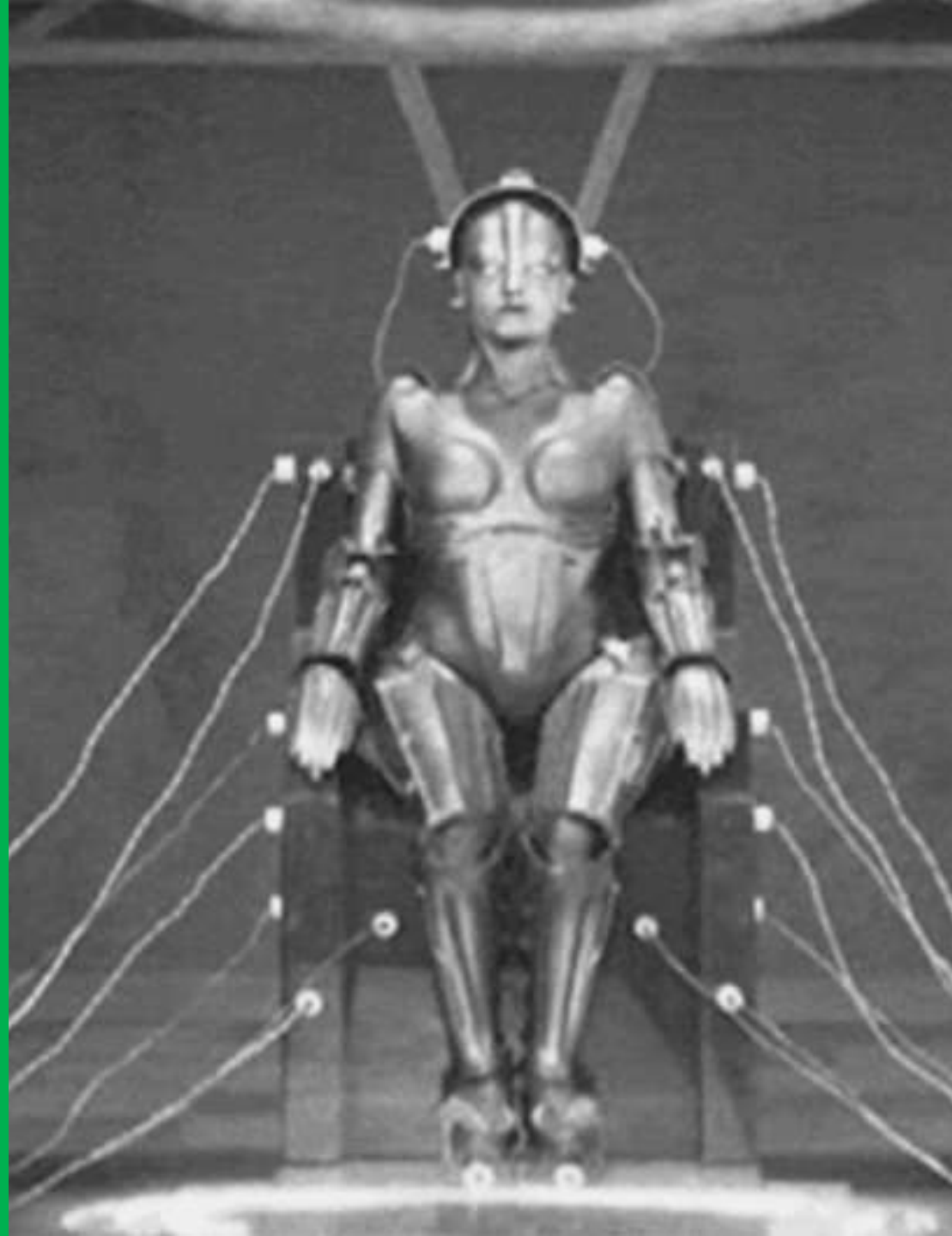
Late ‘40s: McCulloch–Pitts, the artificial neuron.

The ‘50s: Rosenblatt, a successful use case.

Late ‘60s: Minsky and huge limitations to AI.

The ‘90s: Hinton & LeCun, the fathers of "deep".

more



2012

#2012

"ImageNet Classification with deep convolutional neural networks" **blow away** any previous old school method.

A mix of standard concepts (convolutional layers, pooling) with new insights (GPU, ReLU, Dropout).

The CNN error rate was **%15.3**, whereas the second closest was **%26.2**. Whoa!

more

Six Facts

1. #Rockstar

AI Scientists are media rock stars. A clip on youtube can reach for Millions views.

2. #Leadership

AI pioneers left academia and entered the business world (LeCun, Andrew Ng, etc.).c

3. #Dataset

A better (not only bigger) dataset could be stronger than a good algorithm.

4. #DeepLearning

Vision, Language and a mix of them (Deep Learning based) are the stars of AI revolution

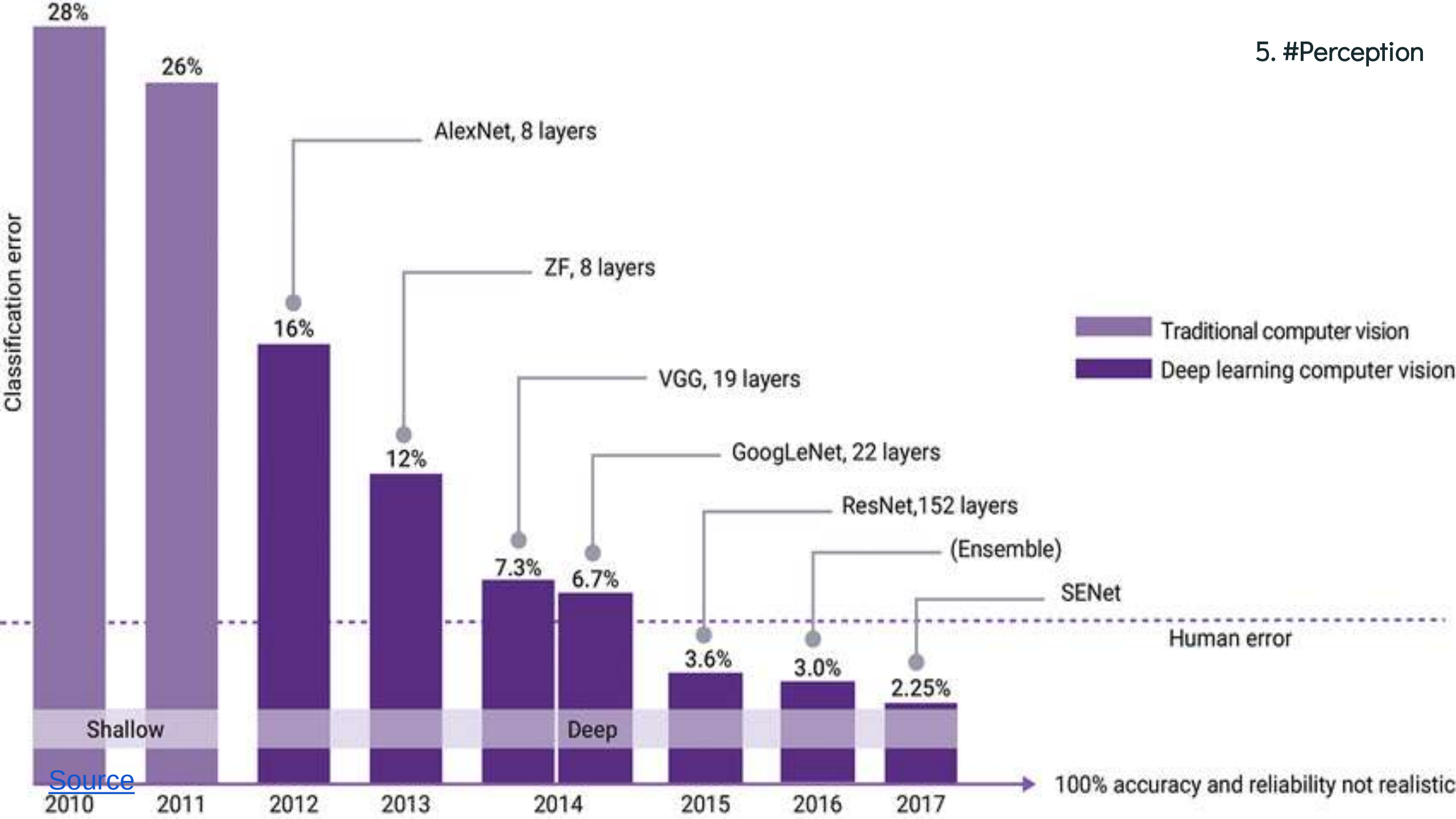
5. #Perception

Neural approach offers the higher accuracy ever. Often better than humans. Magic?

6. #Technology

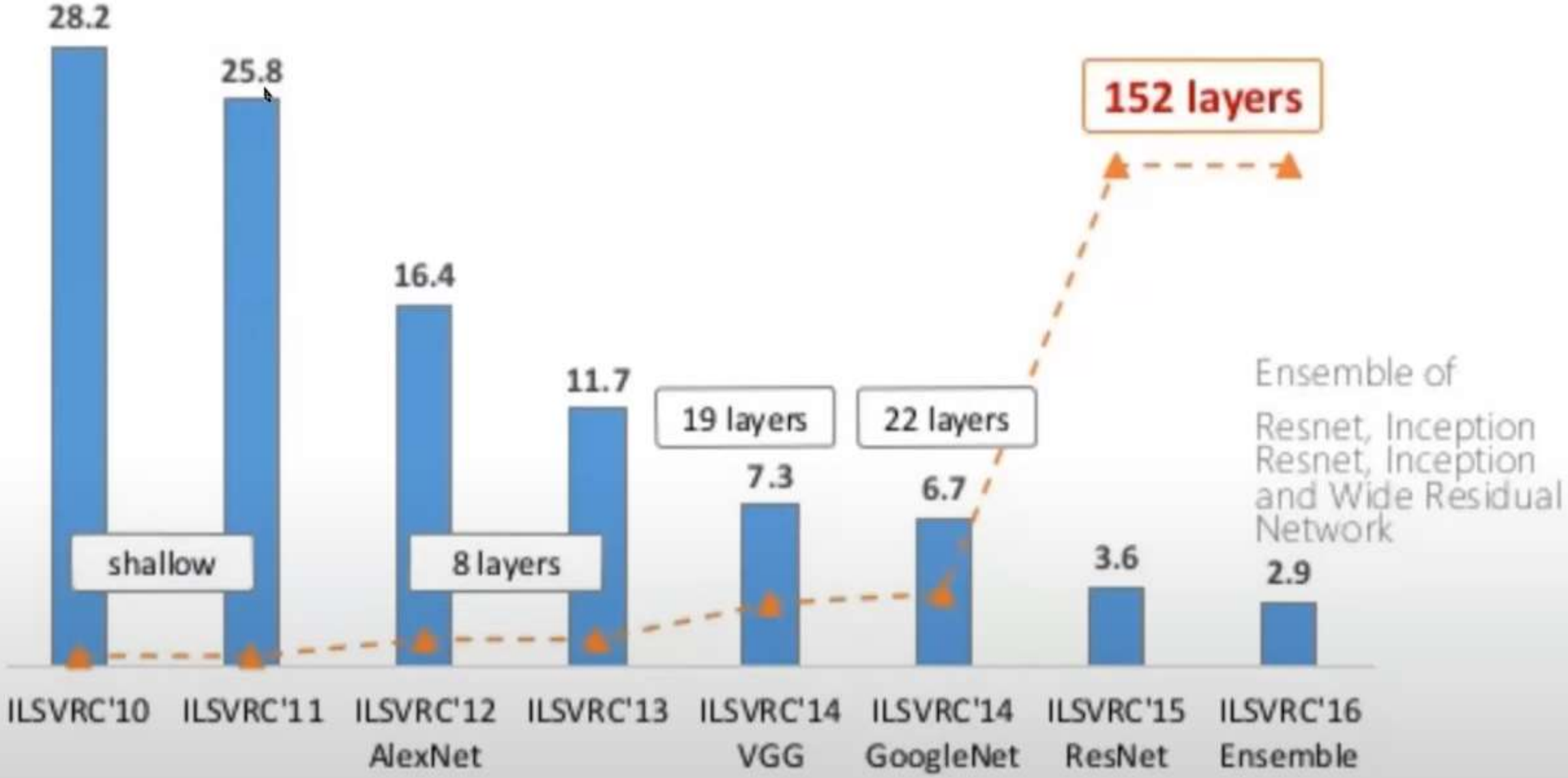
GPUs , many open frameworks, GitHub projects, GB of pre-trained models, etc.

more


Classification error
28%
26%
16%
12%
7.3%
6.7%
3.6%
3.0%
2.25%
AlexNet, 8 layers
ZF, 8 layers
VGG, 19 layers
GoogLeNet, 22 layers
ResNet,152 layers
(Ensemble)
SENet
Human error
Traditional computer vision
Deep learning computer vision
Shallow
Deep
2010
2011
2012
2013
2014
2015
2016
2017
100% accuracy and reliability not realistic

Source

Depth inflation

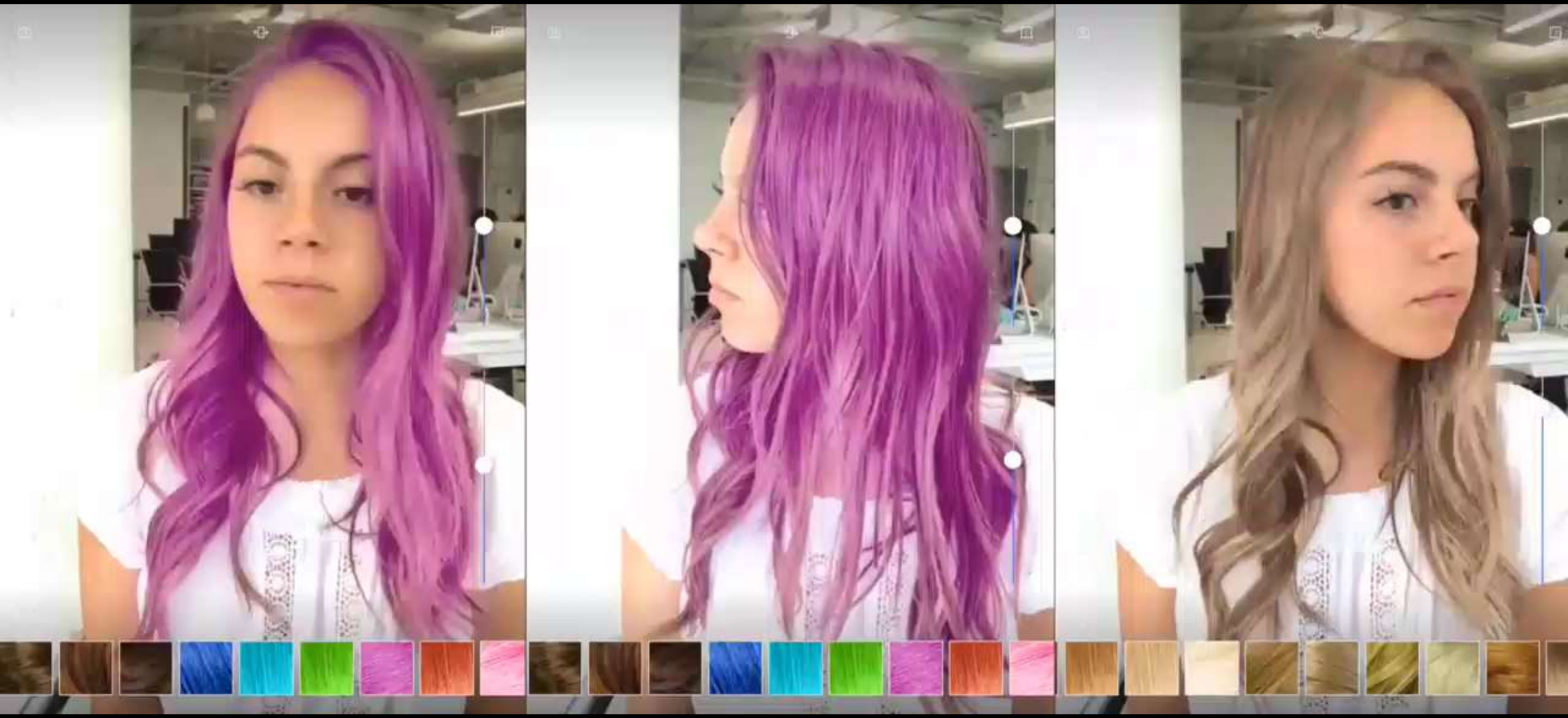


Source

(Figure: Anirudh Koul)

EVOLUTION

the DTC paradigm



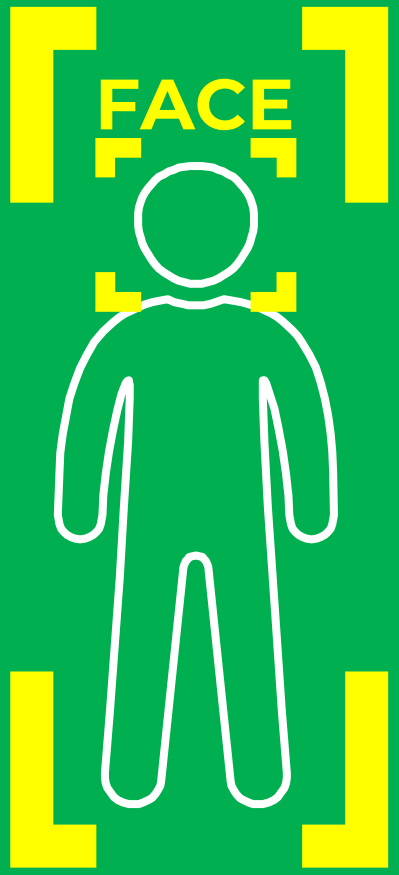
more


gucci

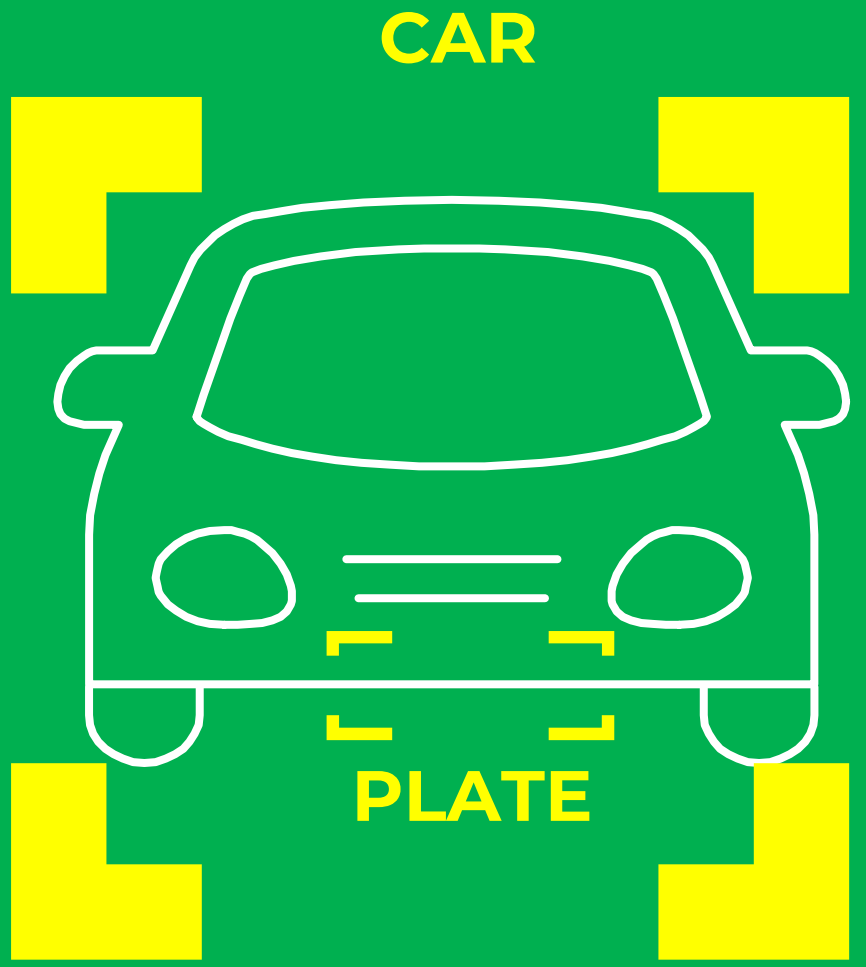

Latent
44
-
+
Seed
0.002
Step Size
Reset
w
w+
Drag
Drag
Add point
Reset point
Start
Stop
Steps: 0
Mask
Flexible area
Fixed area
Reset mask
Show mask
50
-
+
Radius
20
-
+
Lambda
Capture
Capture
_screenshots
Save image
Save GUI

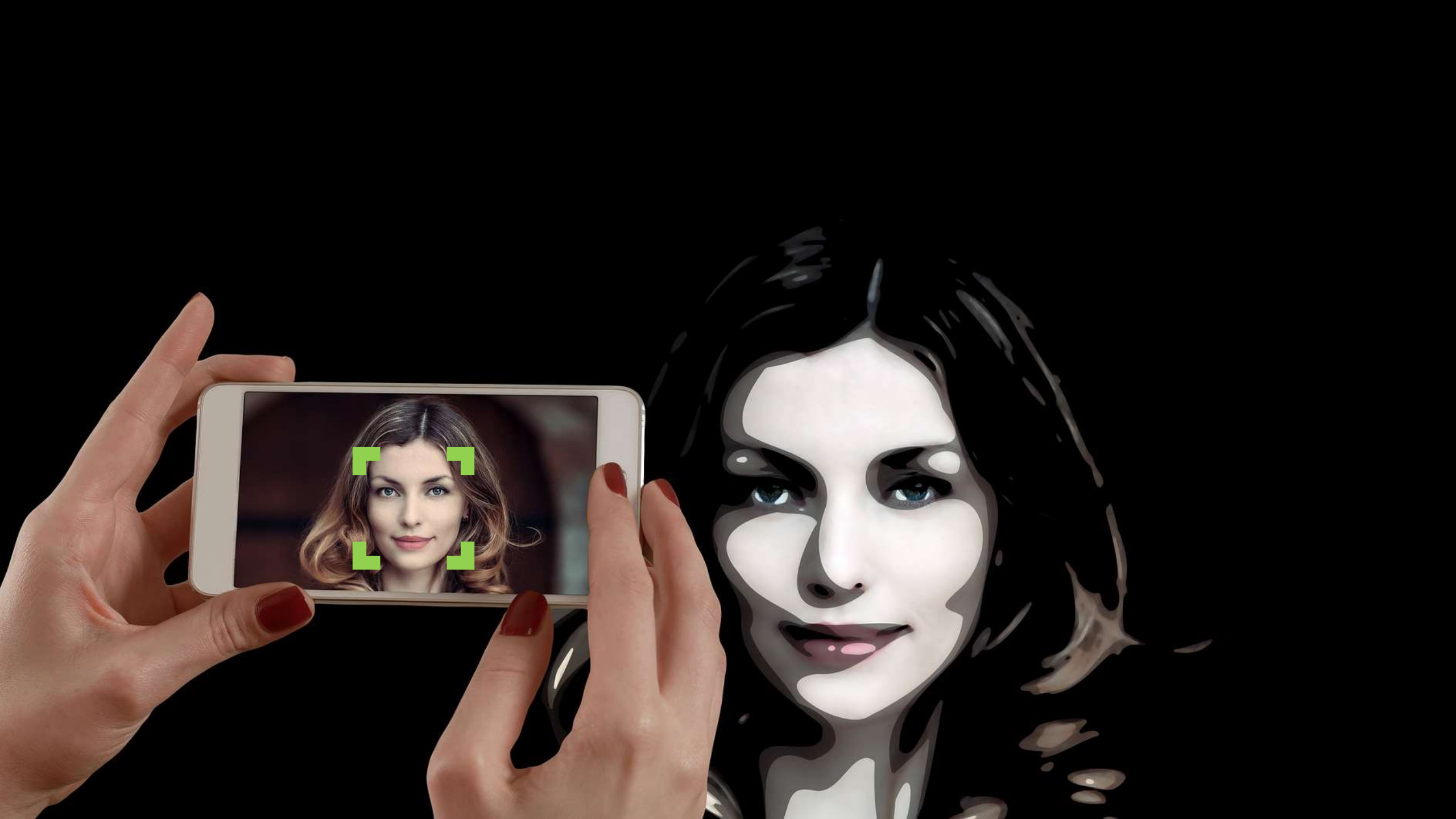
D-T-C paradigm

Detection, Tracking and Classification of instances of semantic objects of a certain class (such as faces, humans, cars, etc.) in digital images and videos

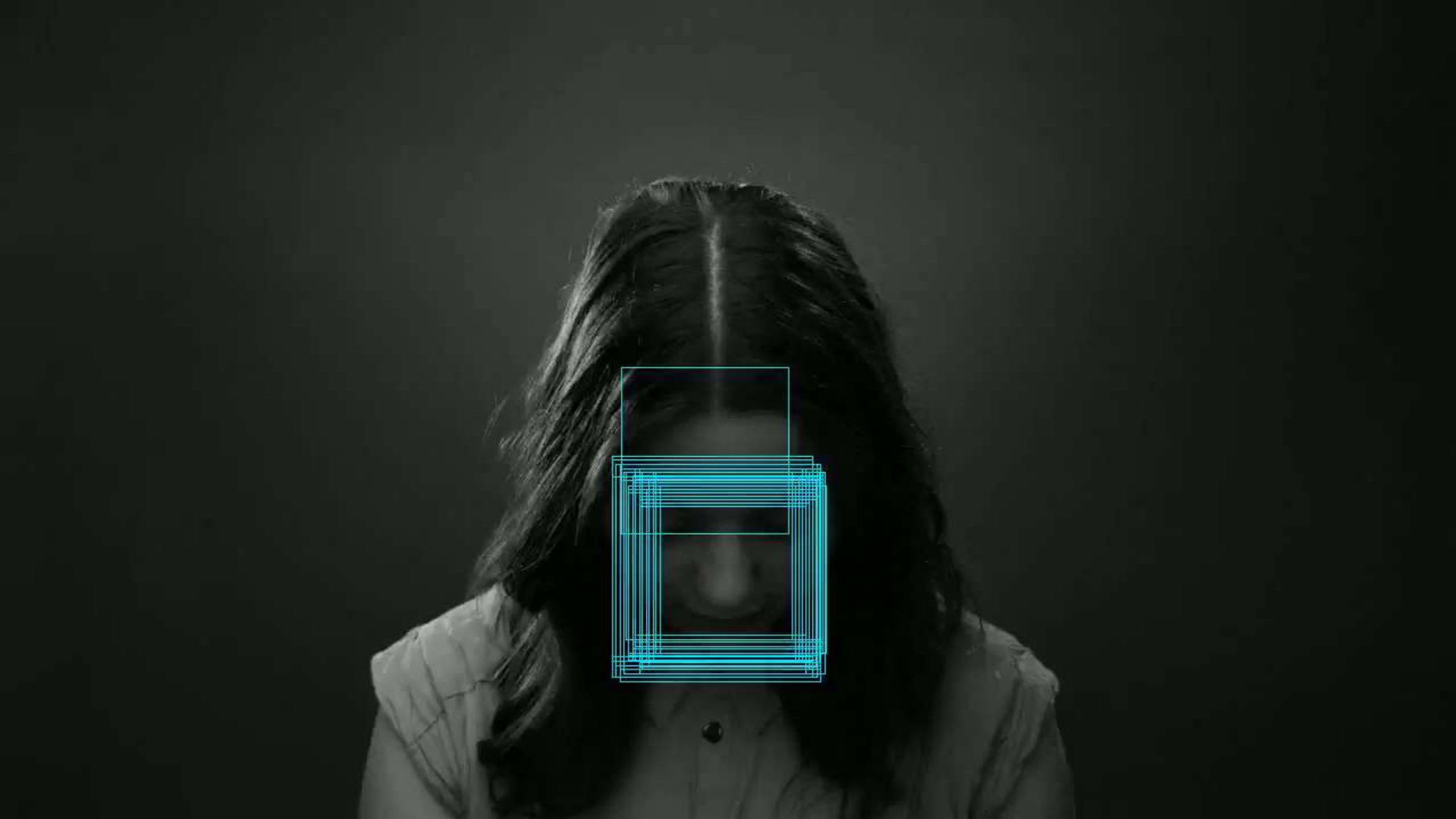


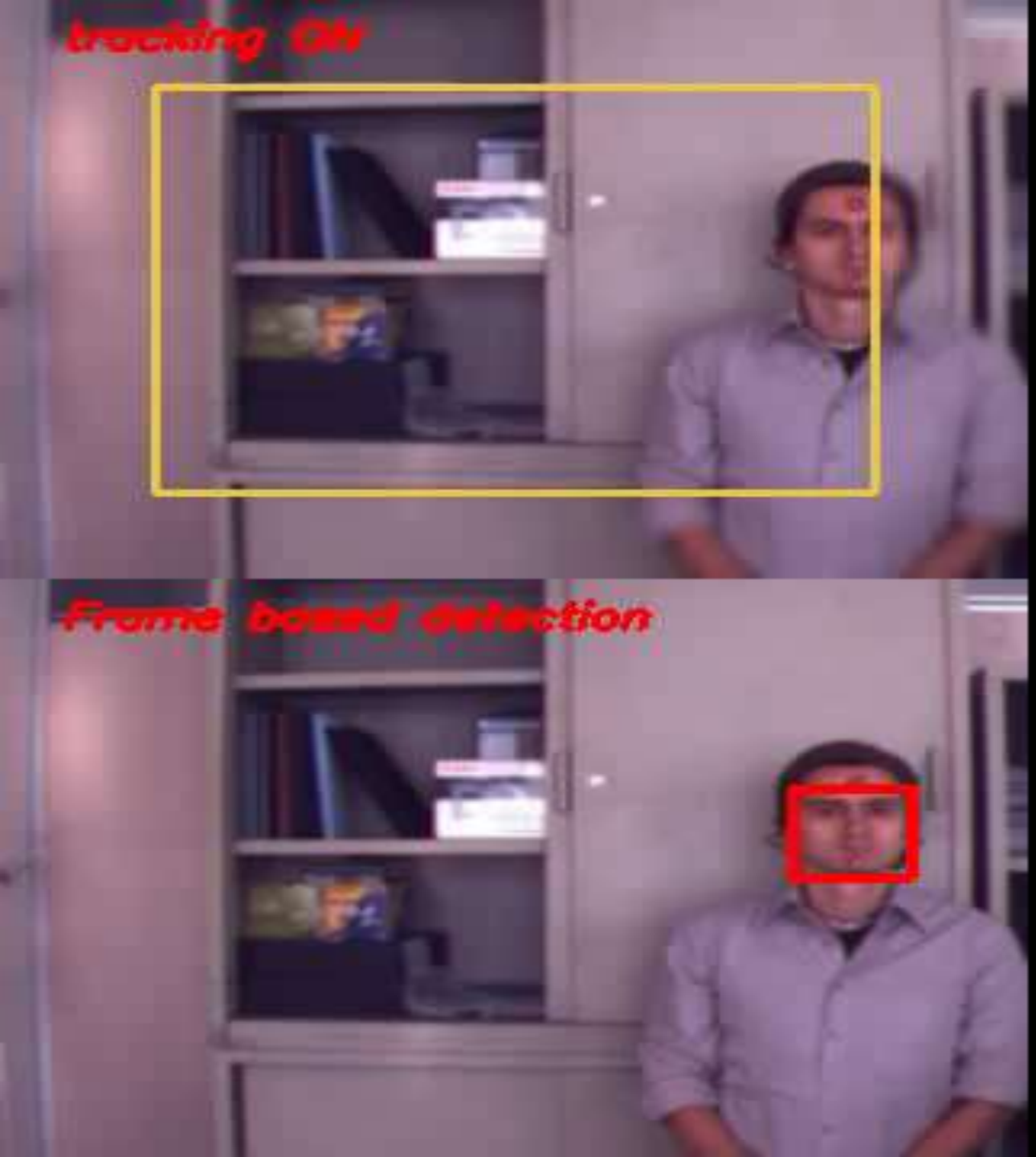
HUMAN



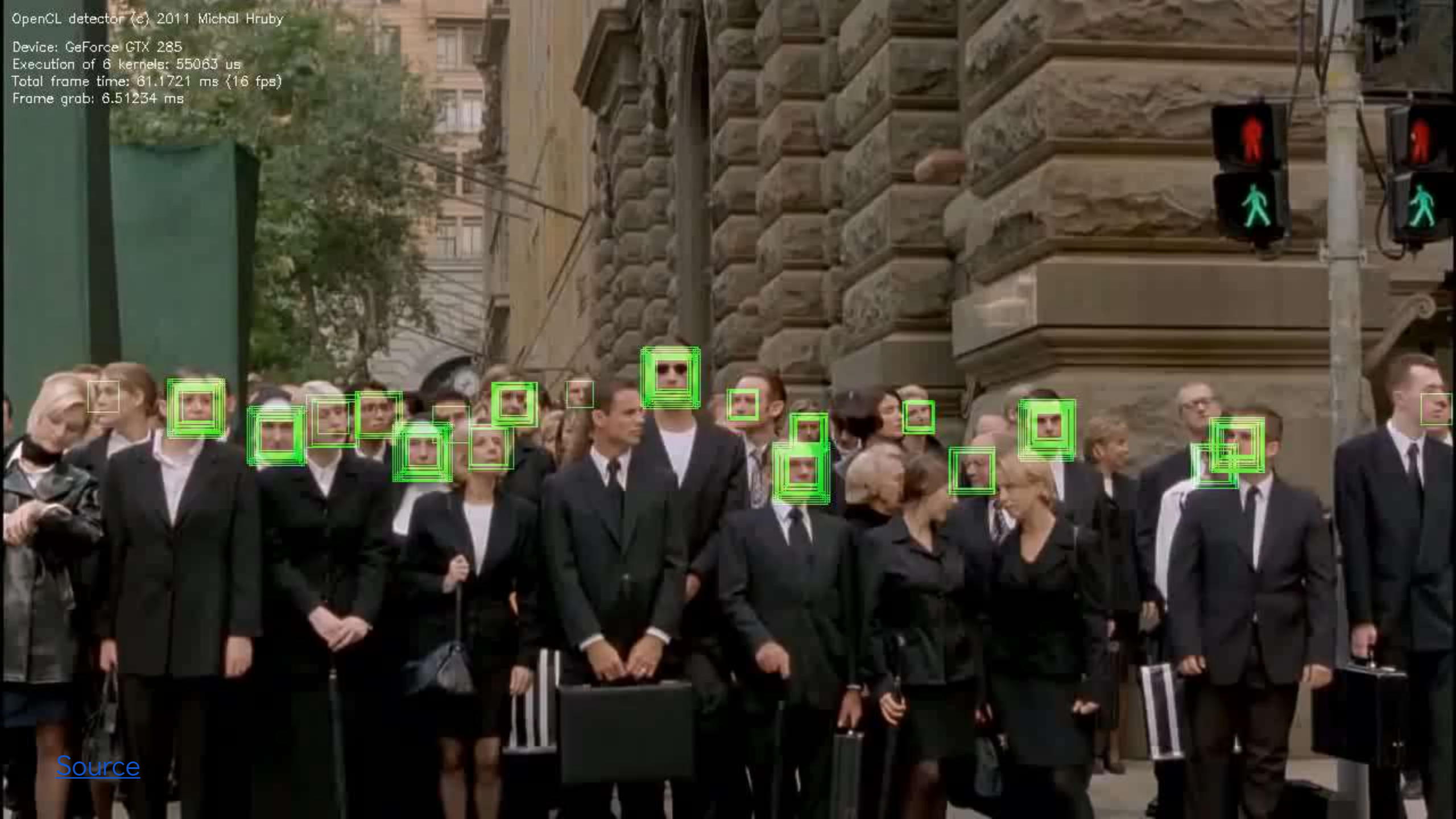


2000




tracking ON
Frame based detection

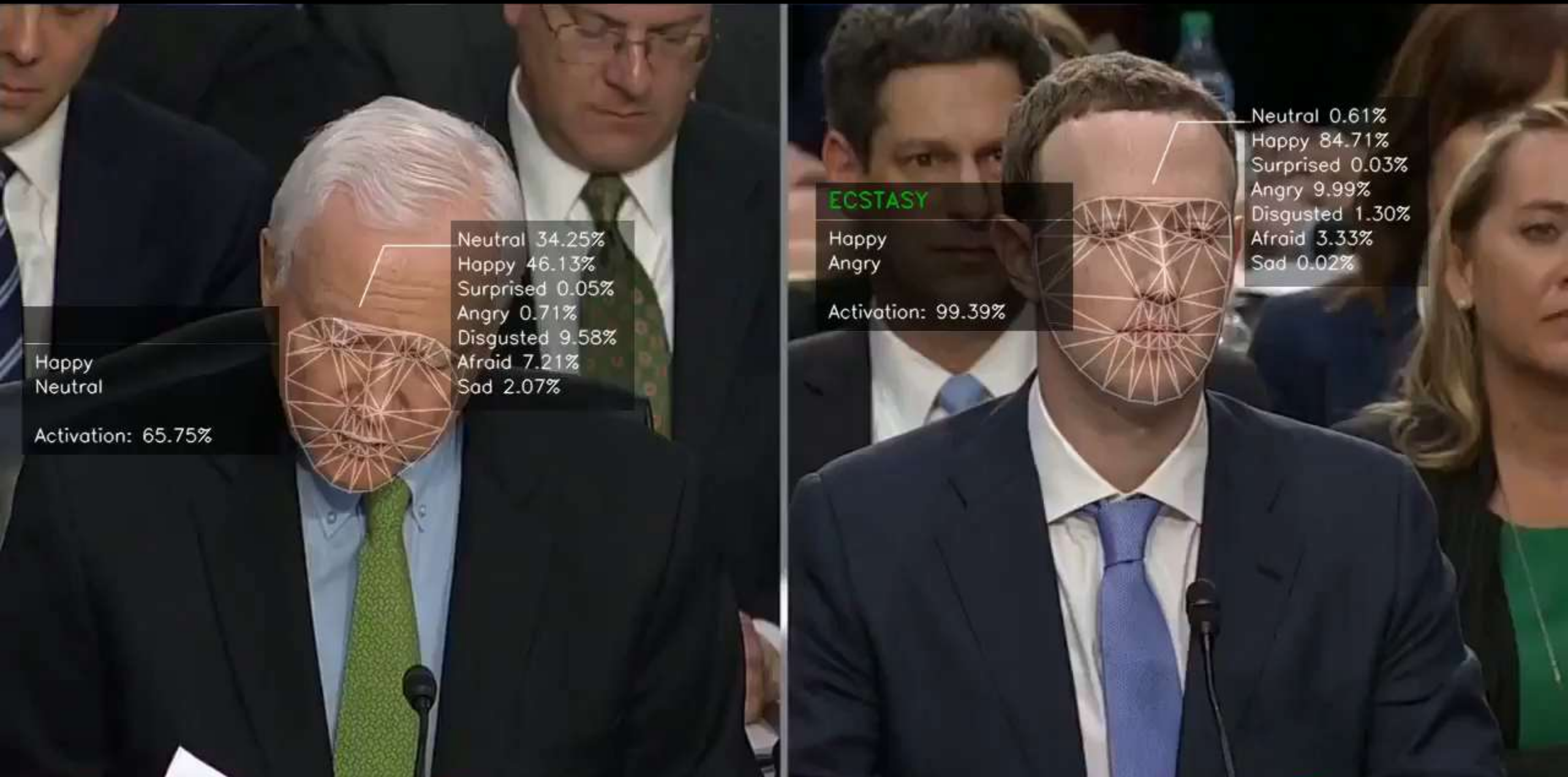
2010


OpenCL detector (c) 2011 Michal Hruby
Device: GeForce GTX 285
Execution of 6 kernels: 55063 us
Total frame time: 61.1721 ms (16 fps)
Frame grab: 6.51234 ms

Source

Then...


WH
.GOV
more

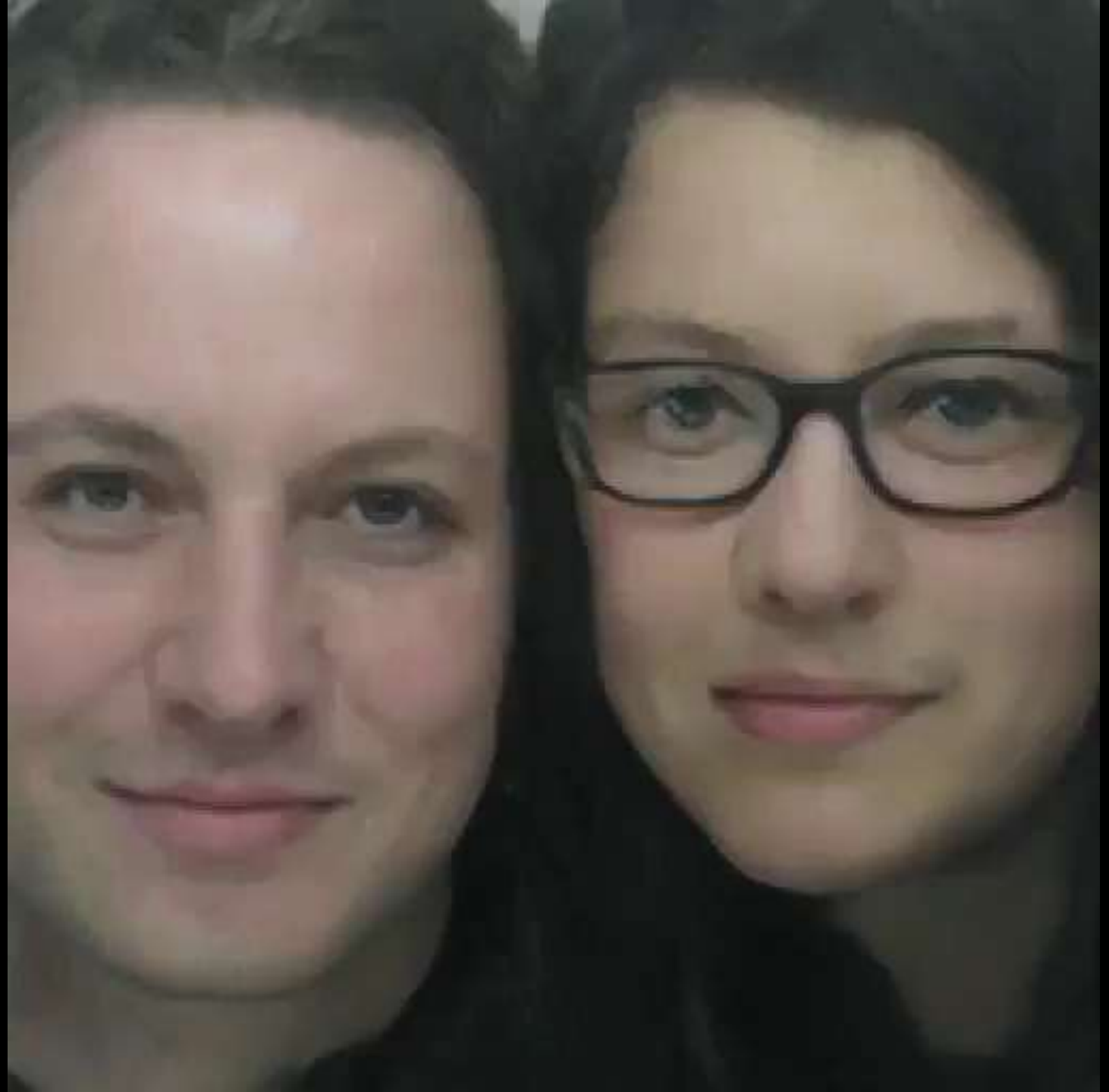

Neutral 34.25%
Happy 46.13%
Surprised 0.05%
Angry 0.71%
Disgusted 9.58%
Afraid 7.21%
Sad 2.07%
Happy
Neutral
Activation: 65.75%
Neutral 0.61%
Happy 84.71%
Surprised 0.03%
Angry 9.99%
Disgusted 1.30%
Afraid 3.33%
Sad 0.02%
ECSTASY
Happy
Angry
Activation: 99.39%

Source

PREVEDO CHE SARÀ LA COLONNA
SONORA DI TUTTI I RAVE D'ITALIA 😂


ОФІС
ПРЕЗИДЕНТА
УКРАЇНИ
OFFICE
OF THE PRESIDENT
OF UKRAINE

Source







ARGO VISION
PRESENTS

Then...


Imagen Video

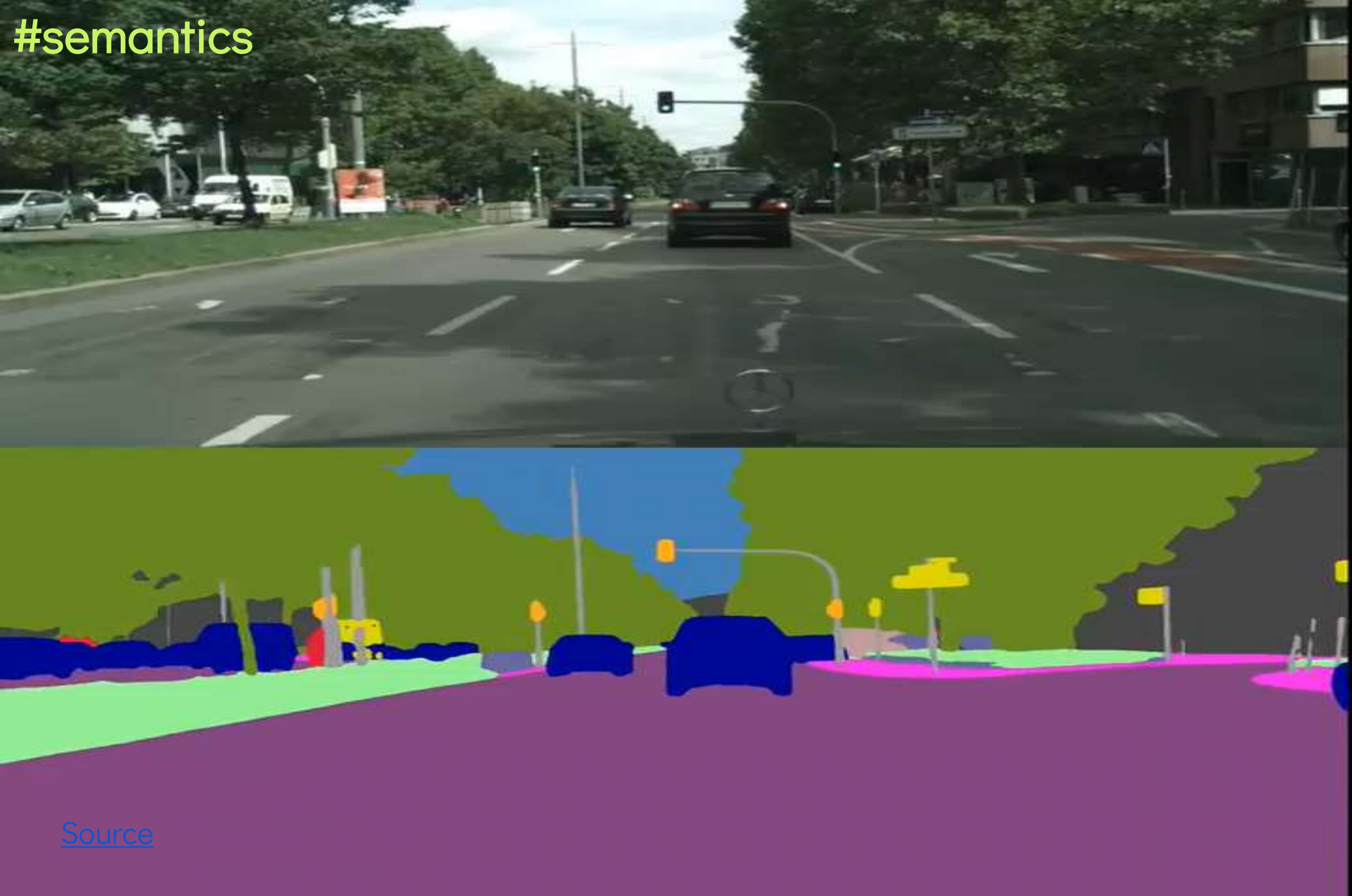
And more

++ click me please ++

Cars

Source

#semantics



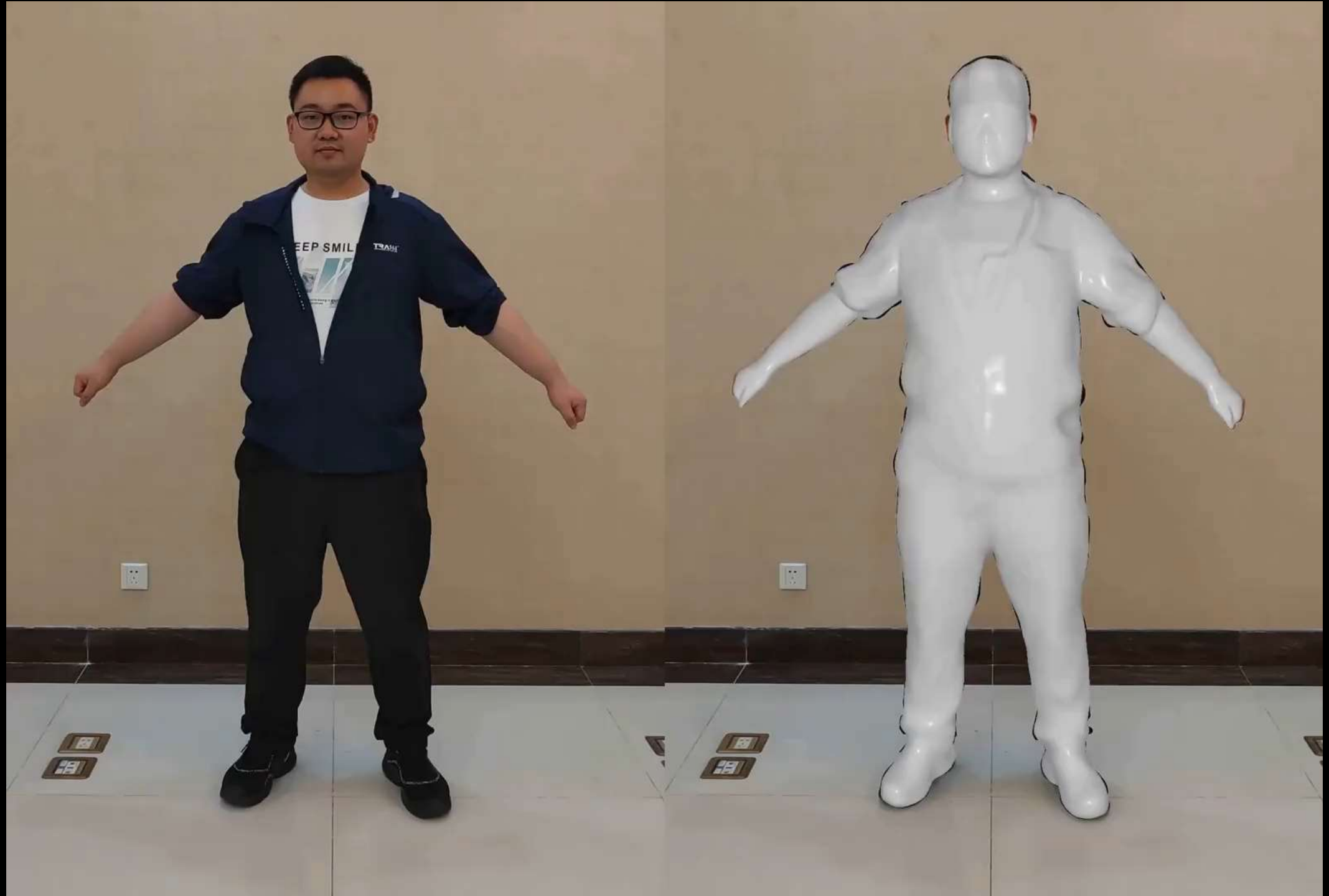
road
sidewalk
building
wall
fence
pole
traffic light
traffic sign
vegetation
terrain
sky
person
rider
car
truck
bus
train
motorcycle
bicycle

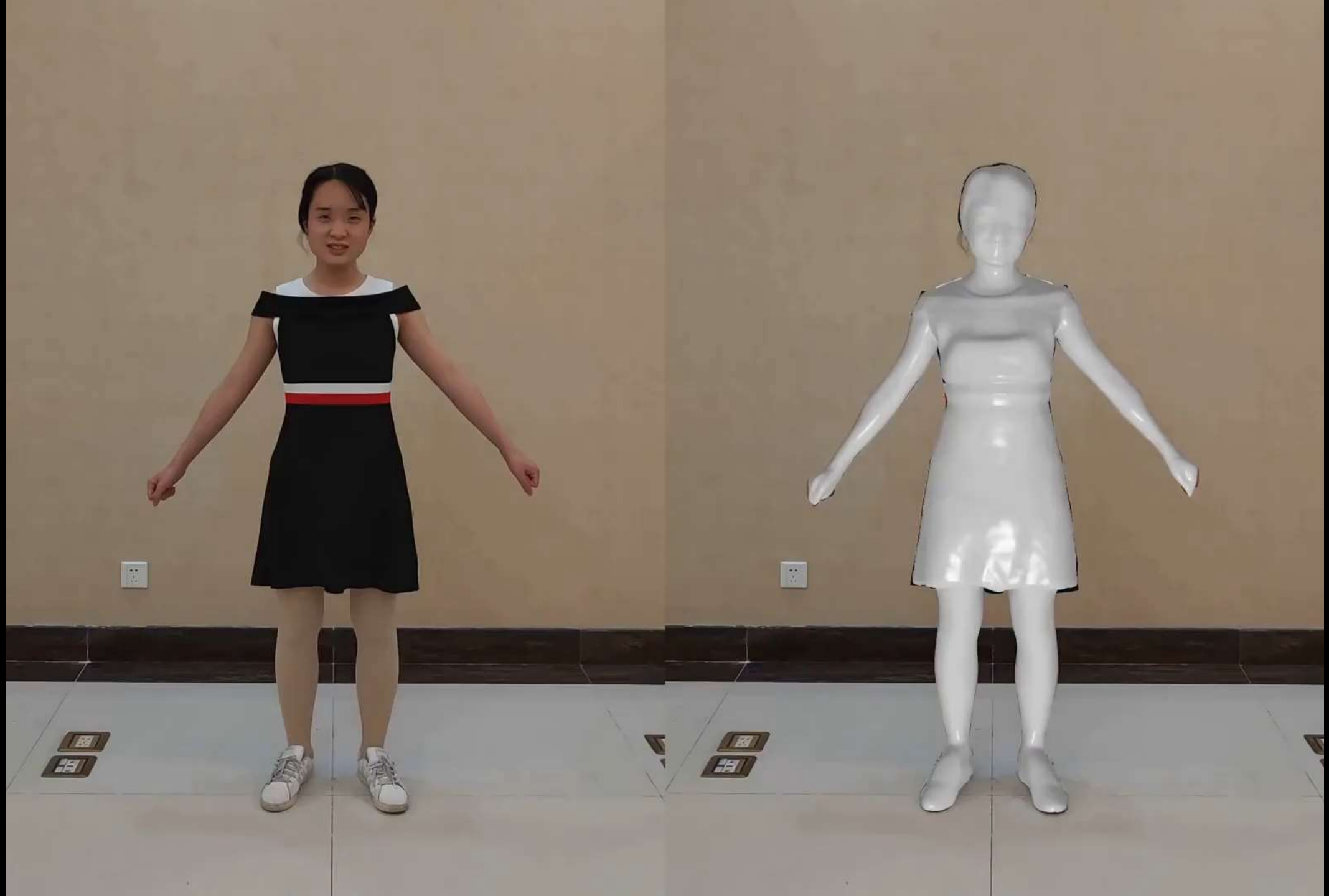
Source

Humans

Source


MIXTURE






FIFA
9
FIFA

Visualization with estimated contacts

LET'S RECAP

Evolution

From bounding box to more sophisticated descriptors. The goal is the semantic comprehension of the world.

Framework

Detection, tracking and classification not separated anymore. One problem, one framework (E2E)

Big Data Generator

From descriptive models to generative frameworks. Virtual/Generative data is the new frontier of the AI.

Trends

Many of the most important problems remain open, both in terms of theory and in terms of applications.

Accuracy

highest accuracy ever in DTC. In many challenges, better than humans (face recognition, ImageNet, etc.).

GPU

GPU is mandatory. All the "amazing" clips you have seen are accelerated by GPUs (Nvidia)


VOLUNTEERS
ARE AWESOME!
memegenerator.net

LET'S PLAY!

HOW FAR FROM ARTIFICIAL INTELLIGENCE?

PERFORMANCE?

Human: 52%

LipNet: 93%


BBC
LG CLOi™
LG's home robot
has made a dreadful debut
more

GRAZIE!

- **LinkedIn** (40,000+)
  https://www.linkedin.com/in/visionarynet/
- **AI with Papers** (7,000+)
  https://t.me/AI_DeepLearning