IEEE International Workshop on
Information Forensics and Security
WIFS 2018
December 11-13, Hong Kong

IEEE Signal Processing Society
IEEE Biometrics Council

in ictu oculi: Exposing DeepFake Videos by Detecting Eye Blinking

Yuezun Li, Ming-Ching Chang and Siwei Lyu
Computer Science Department
University at Albany, State University of New York

UNIVERSITY AT ALBANY
State University of New York

CVML@UA

What is DeepFake?

UNIVERSITY AT ALBANY
State University of New York

CVML@UA

The DeepFakes Horror

MOTHERBOARD

DEEPFAKES | By Samantha Cole | Dec 11 2017, 2:18pm

AI-Assisted Fake Porn Is Here and We're All F ***ed

Someone used an algorithm to paste the face of 'Wonder Woman' star Gal Gadot onto a porn video, and the implications are terrifying.

SHARE TWEET

CNN tech BUSINESS CULTURE GADGETS FUTURE STARTUPS CNNMONEY

Here's why it's so hard to spot deepfakes

by Lisa Fischer @CNNTech

Recommend 63

The National

38°c | 20:08 Isha

Abu Dhabi, UAE | Friday 24 August 2018

SUBSCRIBE

OPINION

Comment Editorial Feedback Cartoon

Deepfake technology could create huge potential for social unrest and even trigger wars

The only way to counteract the threat of deepfakes is to rely on the evidence of our own direct experience or authoritative proven sources, writes Rashmee Roshan Lall

Rashmee Roshan Lall

July 31, 2018

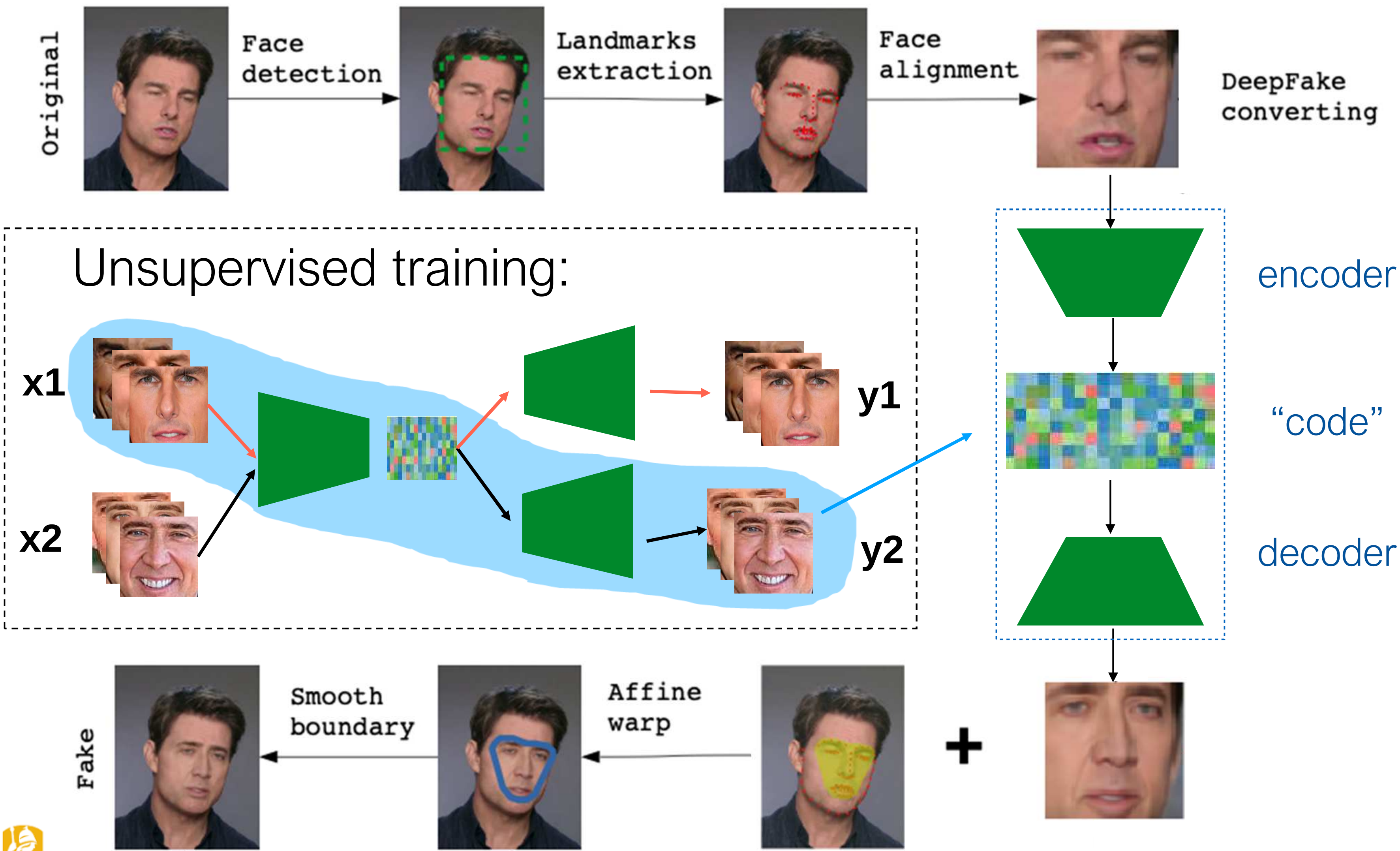
Updated: July 31, 2018 08:35 PM

765 shares

UNIVERSITY AT ALBANY State University of New York

CVML@UA

How DeepFake works?

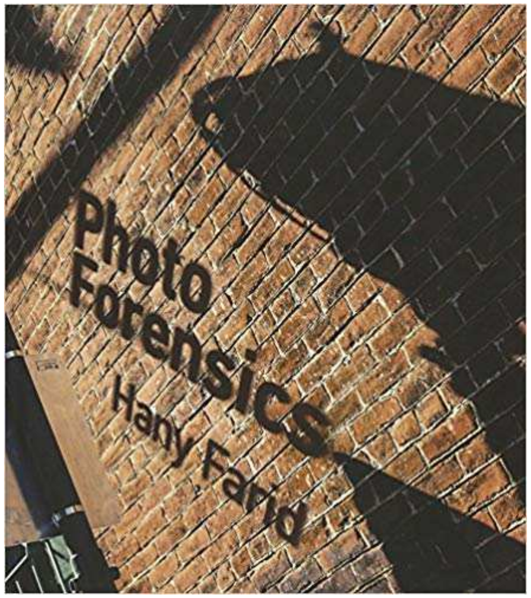


UNIVERSITY AT ALBANY
State University of New York

CVML@UA

Why not use traditional forensic methods?

- Traditional forensic methods
  - Signal based: JPEG, CFA, PRNU
  - Physics based: lighting, shadow, reflection
  - Semantic based: where, when, who
- These methods may not be the best solution for detecting AI generated fake videos.



MIT Press, 2016

Other works

- [1] exploited the color disparity between GAN generated images and real images.
- [2] trained a convolutional neural networks to directly classify real faces and fake faces.

[1] Haodong Li, Bin Li, Shunquan Tan, and Jiwu Huang, “Detection of deep network generated images using disparities in color components,” arXiv preprint arXiv:1808.07276, 2018.
[2] Darius Afchar, Vincent Nozick, Junichi Yamagishi, and Isao Echizen, “Mesonet: a compact facial video forgery detection network,” in IEEE International Workshop on Information Forensics and Security (WIFS), 2018

How to spot a DeepFake?

in ictu oculi

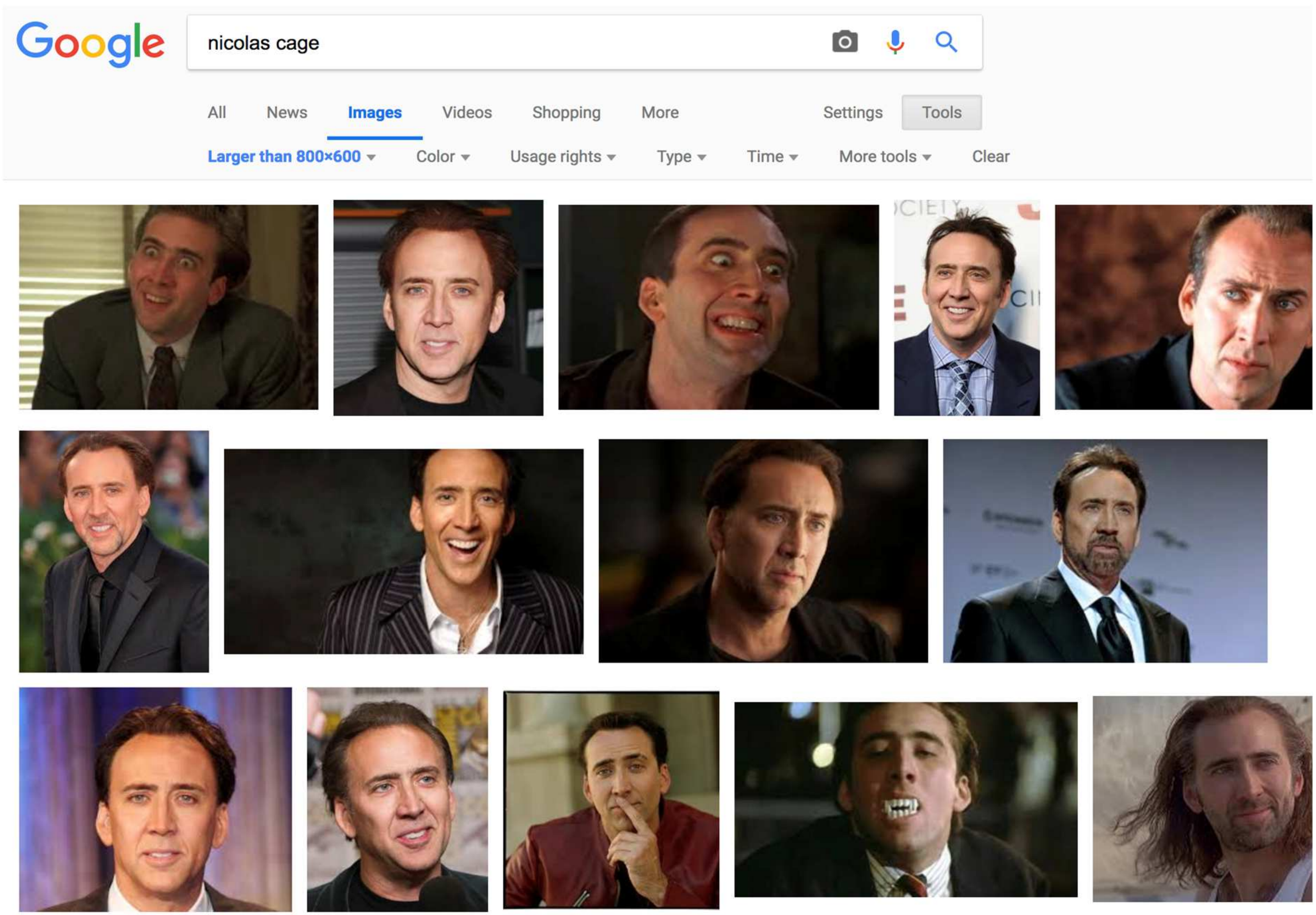
Real videos

DeepFake synthesized Faces do not blink!

DeepFake videos

Why no blinking?

reason: training data may not include closed eyes



UNIVERSITY AT ALBANY
State University of New York

CVML@UA

Detecting blinking with ML algorithm

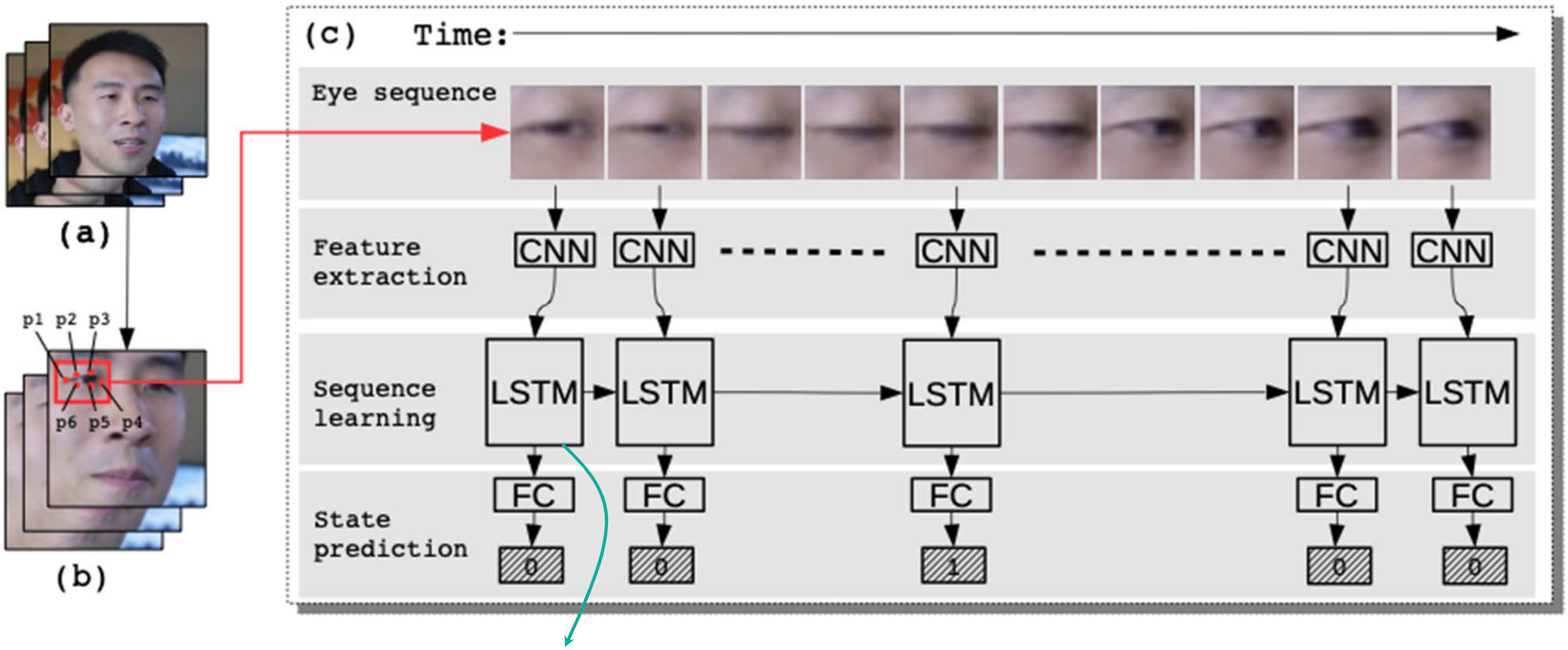
- Train a Long term Recurrent CNN (LRCN) [3] model to detect open/closed eye
- Apply this model to estimate blinking rate in a video to determine its authenticity

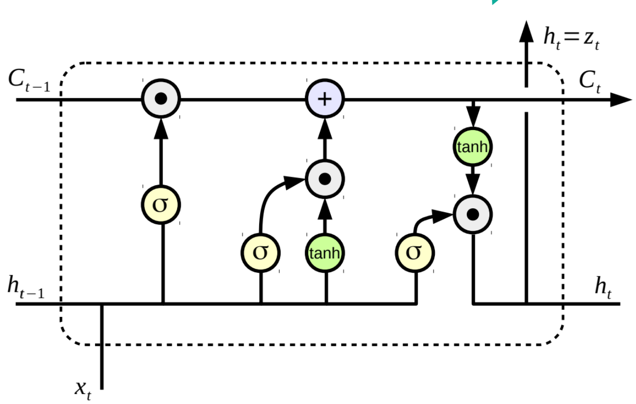
[3] J. Donahue, L. Anne Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, and T. Darrell. Long-term recurrent convolutional networks for visual recognition and description. In CVPR, pages 2625–2634, 2015

UNIVERSITY AT ALBANY
State University of New York

CVML@UA

The pipeline of our method





$$
\begin{aligned}
f_t &= \sigma(W_{fh}h_{t-1} + W_{fx}x_t + b_f) \\
i_t &= \sigma(W_{ih}h_{t-1} + W_{ix}x_t + b_i) \\
g_t &= tanh(W_{ch}h_{t-1} + W_{cx}x_t + b_c) \\
C_t &= f_t \odot C_{t-1} + i_t \odot g_t \\
o_t &= \sigma(W_{oh}h_{t-1} + W_{ox}x_t + b_o) \\
h_t &= o_t \odot tanh(C_t)
\end{aligned}
$$

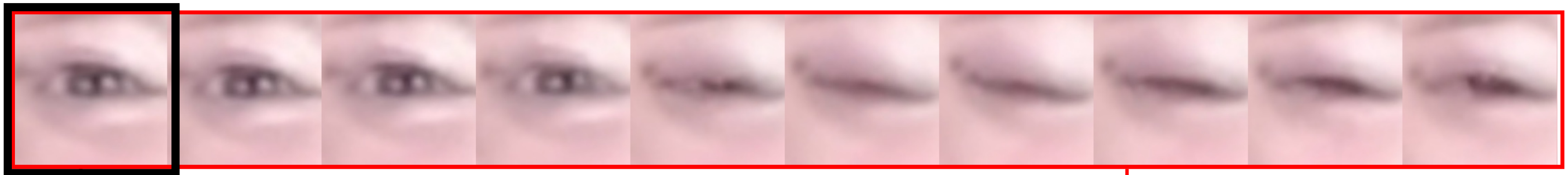
$$
\sigma(x) = \frac{1}{1+e^{-x}} \qquad tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}
$$

[4] S. Hochreiter and J. Schmidhuber. Long short-term memory. Neural Comput., 9(8):1735–1780, 1997

UNIVERSITY AT ALBANY
State University of New York
CVML@UA

Training LRCN

1. Data preparation

- 50 videos downloaded from Internet and annotate eye state
- 

2. Training CNN

Input size: 224x224
Batch size: 16
Learning rate: 0.01
Decay: 0.9
Optimizer: SGD
Epoch: 100

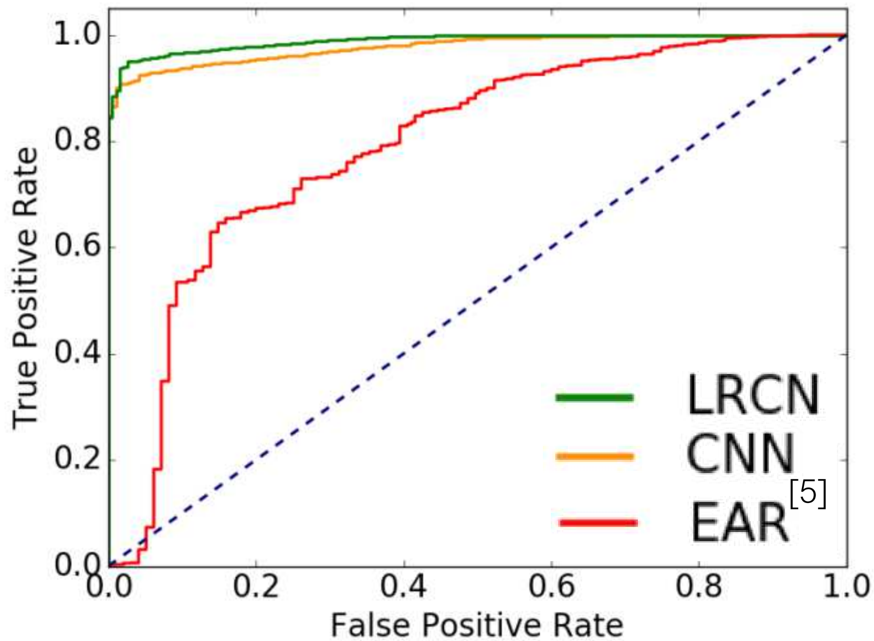
3. Training LRCN jointly

Input size: 224x224xN
Batch size: 4
Learning rate: 0.01
Decay: 0.9
Optimizer: ADAM
Epoch: 100

UNIVERSITY AT ALBANY
State University of New York

CVML@UA

The performance of LRCN



Video	Average video length	FPS	Rate of blinks
Origin	10 seconds	30	**34.1** / min
Fake	10 seconds	30	**3.4** / min

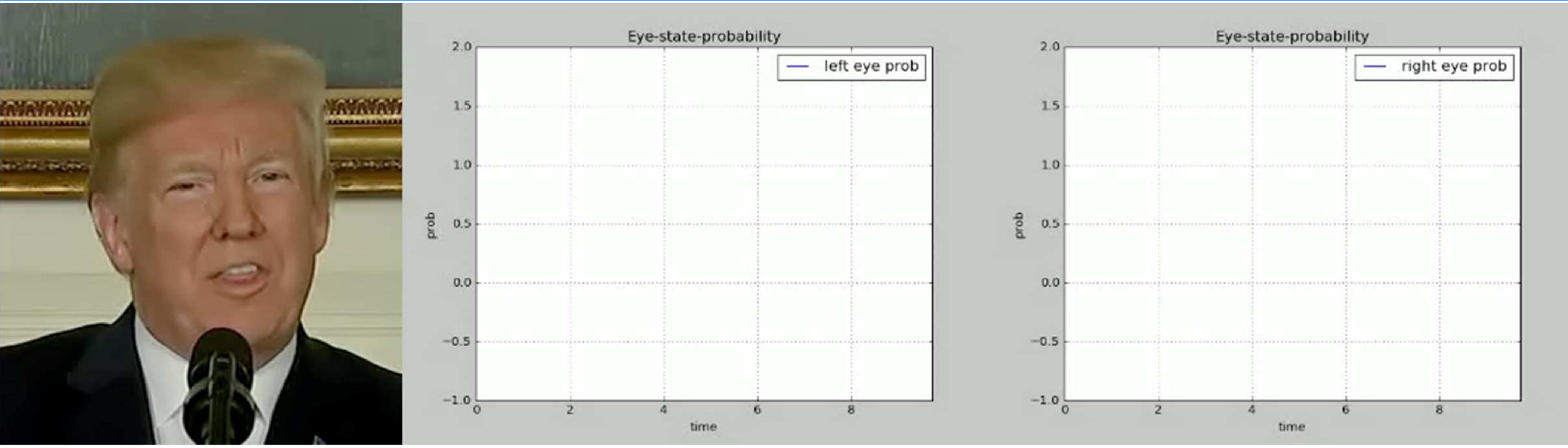
The blinking rate of normal human is set to 10/min [6]

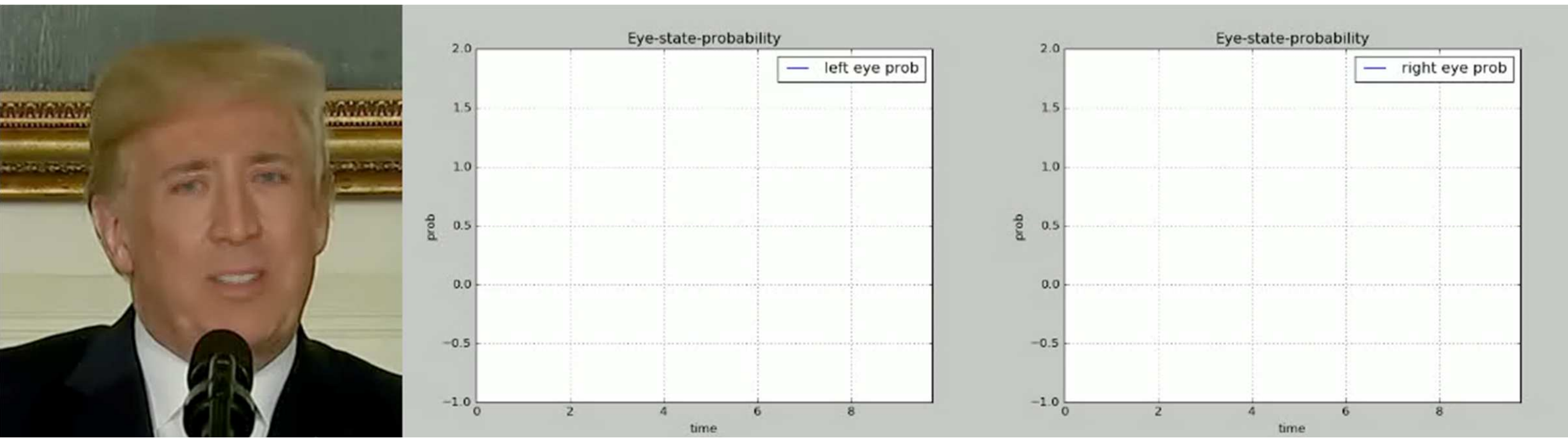
[5] T. Soukupova and J. Cech. Real-time eye blink detection using facial landmarks. In 21st Computer Vision Winter Workshop, pages 1–8, 2016
[6] A. R. Bentivoglio, S. B. Bressman, E. Cassetta, D. Carretta, P. Tonali, and A. Albanese. Analysis of blink rate patterns in normal subjects. Movement Disorders, 12(6):1028–1034, 1997

UNIVERSITY AT ALBANY State University of New York

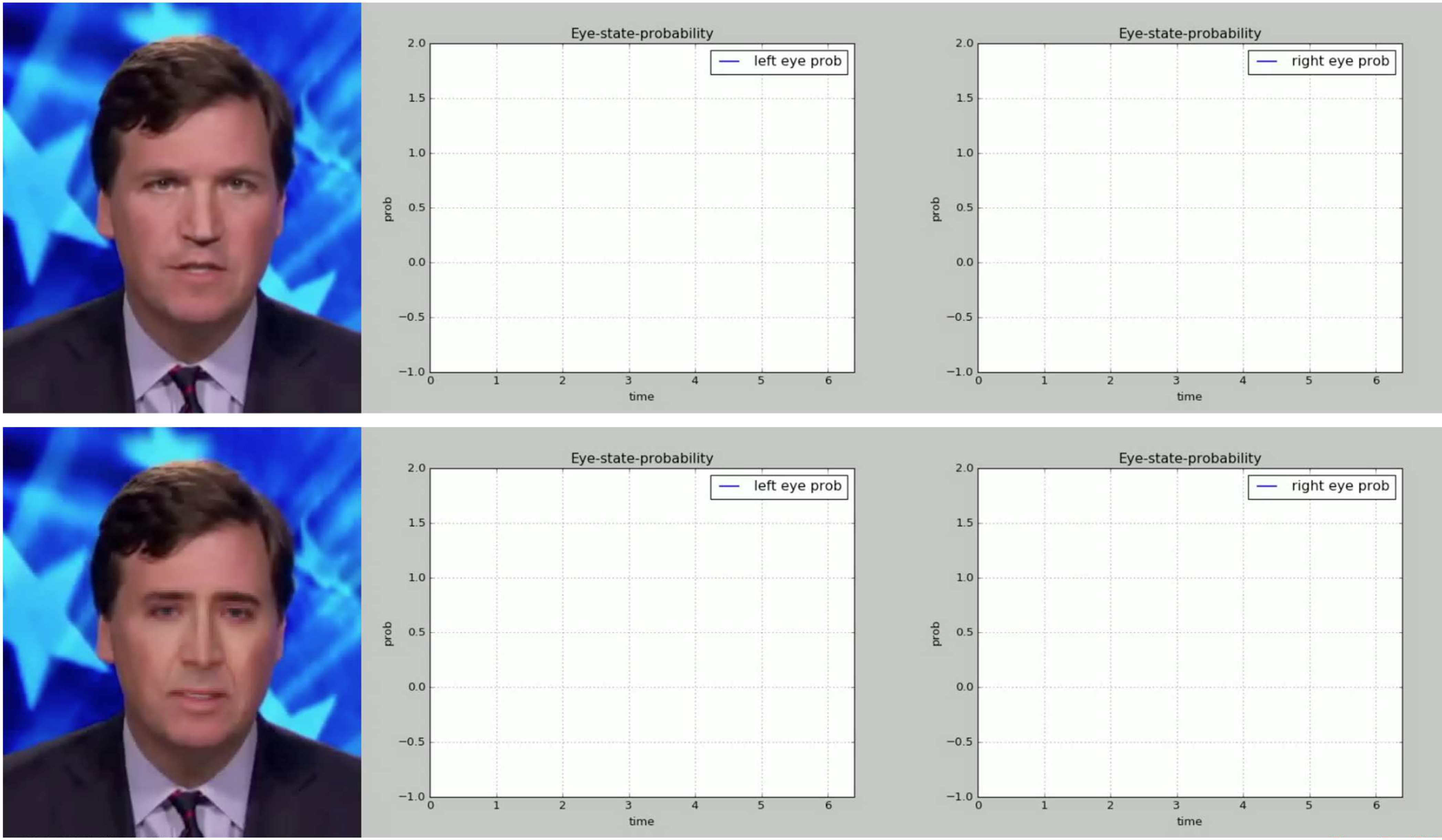
CVML@UA

Spot a DeepFake *in ictu oculi*





Spot a DeepFake *in ictu oculi*



Is this the end of DeepFake?

Forgery technology catches up quickly

- e.g., blinking can be fixed with using video frames as training data

- Despite this, our discovery can still increase the difficulties of DeepFake video generation. Now we are developing more effective method to expose the fake videos

Summary

- Technologies for creating DeepFake videos advance rapidly and can cause serious impacts to society
- Digital media forensics are catching up to control the negative effects of DeepFake videos
- The rivalry between forgeries and forensics will continue for coming years

UNIVERSITY AT ALBANY
State University of New York

CVML@UA

Thank you for your attention!

All non-public domain images/videos used in this talk belong to the creators.