Introduction to Deepfakes and Deepfake detection

Thomas Marcoux, January 2020

Image: Facebook [1]

Free Solutions

DeepFaceLab(DFL) - Russian

Faceswap - American

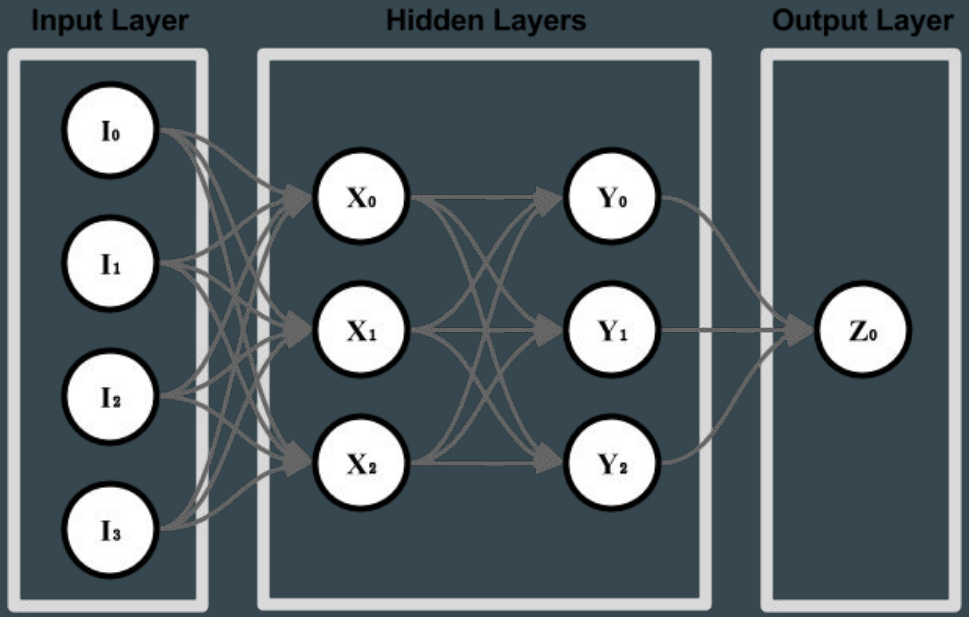
Fakeapp - Unknown, free but proprietary

Method

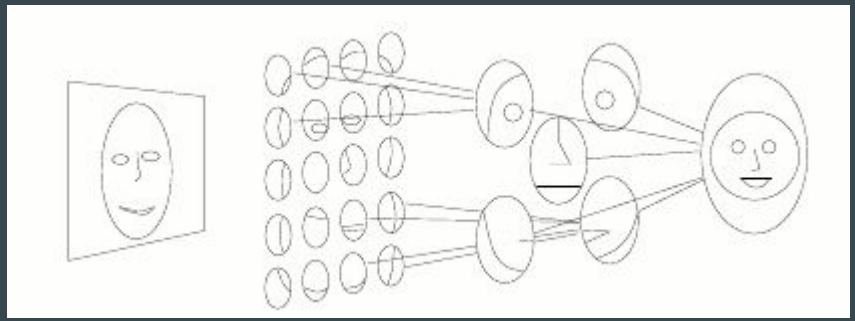
- Steps:
  - Extraction
  - Training - Neural Network
  - Conversion
- Challenges:
  - Multiple faces
  - Blurry videos
  - Obstructed faces

More on Neural Networks

Convolutional Neural Network



Features detection



Images: alanzucconi.com[7]

Research & Potential for Slander

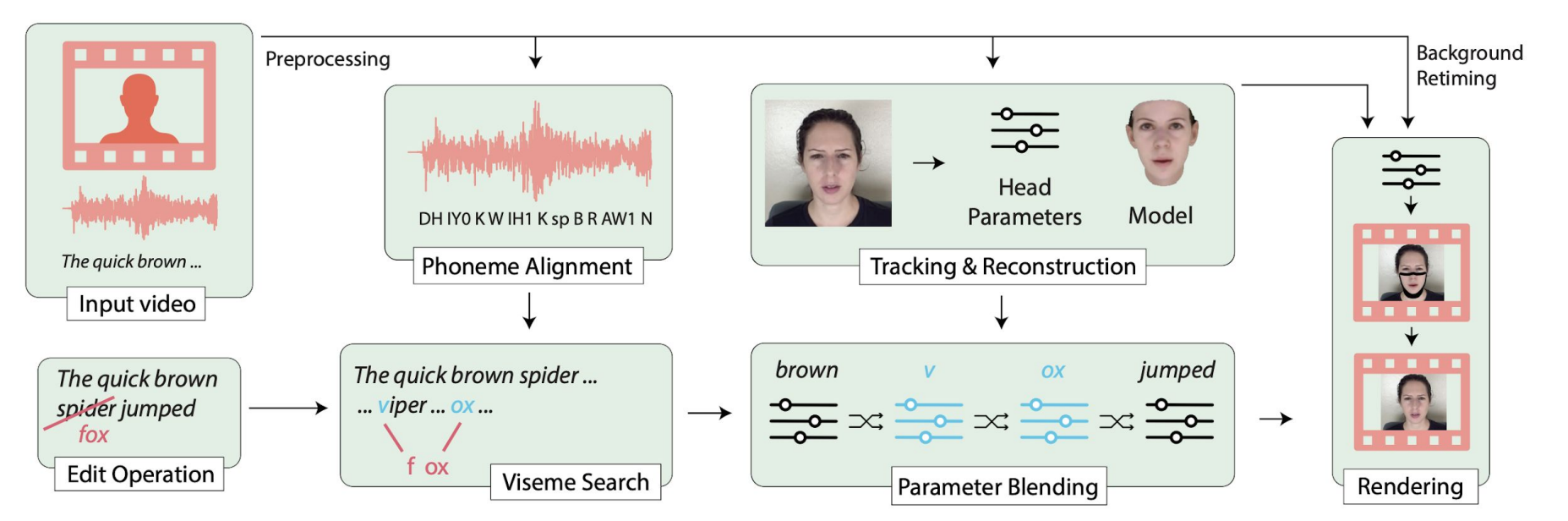


Image: Fried, O. et al [2]

Research on Deepfake Detection

Binary classification with supervised training [10].

Researchers look for digital integrity, physical integrity, or semantic integrity [1].

Deepfakes leave distinctive artifacts in end-product [3].

Some research uses a convolutional neural network - as is the case to generate deepfakes - to extract features. They then train a recurrent neural network to detect image manipulation [4,10].

Reality Defender to come out in early 2020 [1].

References

1. https://spectrum.ieee.org/tech-talk/robotics/artificial-intelligence/facebook-ai-launches-its-deepfake-detection-challenge
2. Fried, Ohad, Maneesh Agrawala, Ayush Tewari, Michael Zollhöfer, Adam Finkelstein, Eli Shechtman, Dan B Goldman, Kyle Genova, Zeyu Jin, and Christian Theobalt. “Text-Based Editing of Talking-Head Video.” ACM Transactions on Graphics 38, no. 4 (July 12, 2019): 1–14. https://doi.org/10.1145/3306346.3323028.
3. Li, Yuezun, and Siwei Lyu. “Exposing DeepFake Videos By Detecting Face Warping Artifacts,” n.d., 7.
4. Guera, David, and Edward J. Delp. “Deepfake Video Detection Using Recurrent Neural Networks.” In 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 1–6. Auckland, New Zealand: IEEE, 2018. https://doi.org/10.1109/AVSS.2018.8639163.
5. https://medium.com/@jsoverson/from-zero-to-deepfake-310551e59aa3
6. https://www.alanzucconi.com/2018/03/14/introduction-to-deepfakes/
7. https://www.alanzucconi.com/2018/03/14/an-introduction-to-autoencoders/
8. https://github.com/iperov/DeepFaceLab
9. https://faceswap.dev/
10. Nguyen, Thanh Thi, Cuong M. Nguyen, Dung Tien Nguyen, Duc Thanh Nguyen, and Saeid Nahavandi. “Deep Learning for Deepfakes Creation and Detection.” ArXiv:1909.11573 [Cs, Eess], September 25, 2019. http://arxiv.org/abs/1909.11573.