BÁO CÁO ĐỒ ÁN CUỐI KỲ

Môn học: CS519 - PHƯƠNG PHÁP LUẬN NCKH

Lớp: CS519.O11

GV: PGS.TS. Lê Đình Duy

Trường ĐH Công Nghệ Thông Tin, ĐHQG-HCM

ÁP DỤNG CƠ CHẾ ĐA CHÚ Ý VÀO NHẬN DẠNG ẢNH KHUÔN MẶT DEEPFAKE

Phan Tiến Quân - 21522502

Nguyễn Khánh Trình - 21522717

Tóm tắt

- Lớp: CS519.O11
- Link Github của nhóm: https://github.com/NgKhTr/CS519.O11.git
- Link YouTube video: https://youtu.be/VWJKArE2rW0

Phan Tiến Quân - 21522502

Nguyễn Khánh Trình - 21522717

Giới thiệu

CHÍNH PHỦ NƯỚC CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM

Báo Điện tử Chính phủ

Thứ tư, 31/1/2024 | MỚI NHẤT

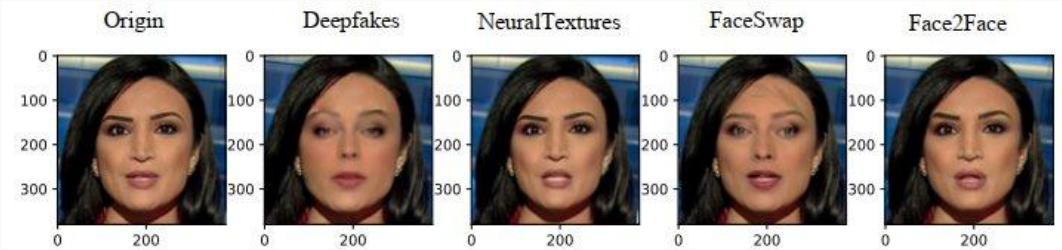
Xã hội > Pháp luật | Y tế | Đời sống | An sinh

Cảnh giác trước cuộc gọi lừa đảo "Deepfake"

11/07/2023 • 16:33

(Chinhphu.vn) - Thời gian gần đây, một số đối tượng lợi dụng công nghệ "Deepfake" để thực hiện các cuộc gọi video với hình ảnh, khuôn mặt giả mạo để lừa đảo "nạn nhân" nhằm thực hiện các hành vi vi phạm pháp luật. Bộ Công an khuyến cáo người dân những thông tin để nhận diện và phòng ngừa đối với loại tội phạm này.

Cần phải có phương pháp phát hiện deepfake



Bài toán phân loại ảnh nhị phân (binary image classification)

Bài toán phân loại ảnh chi tiết (fine-grained image classification)

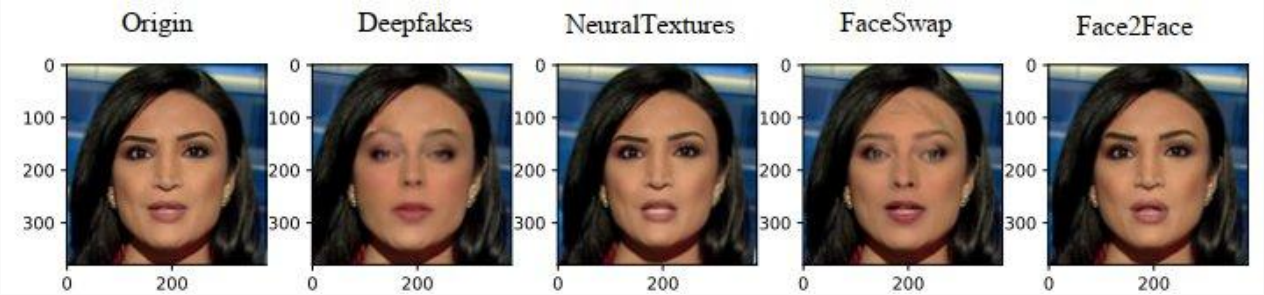
Mô hình như thế nào là tối ưu cho bài toán nhận dạng ảnh khuôn mặt deepfake khi ta xem xét nó dưới dạng bài toán phân loại ảnh chi tiết ?

Mục tiêu

- Thực hiện nghiên cứu tìm kiếm kiến trúc phù hợp cho 3 khối được đề ra.
- Nghiên cứu phương pháp huấn luyện và tăng cường dữ liệu tối ưu để các bản đồ chú ý độc lập và chứa các thông tin khác nhau.
- Thực hiện thử nghiệm mô hình được phát triển với các bộ dữ liệu nhận dạng deepfake hiện tại. Chứng minh được các bản đồ chú ý độc lập nhau.

Nội dung và Phương pháp

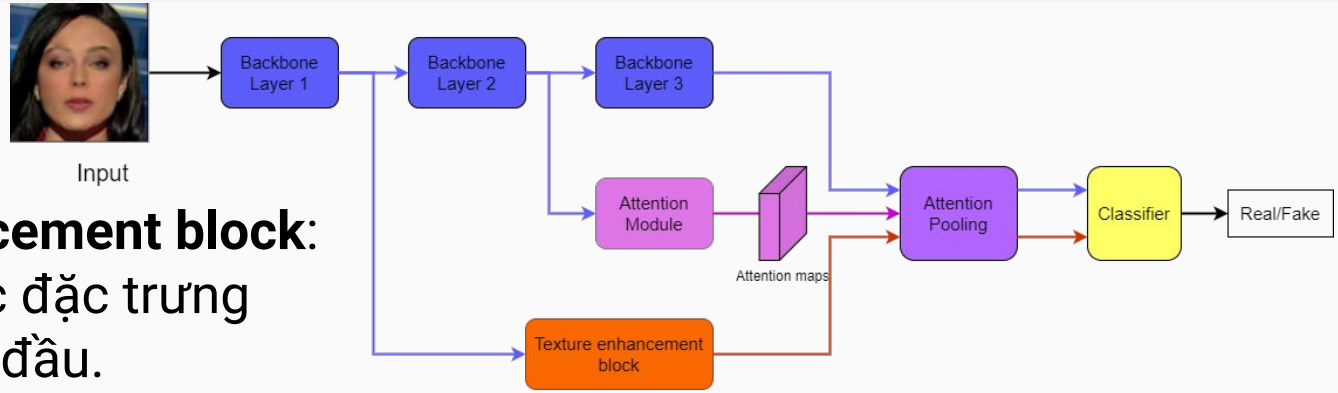
Nội dung



- Phân loại ảnh chi tiết đối với ảnh khuôn mặt: mắt, mũi, miệng, kết cấu khuôn mặt, … chứa thông tin quyết định đến kết quả phân loại.
- Cơ chế đa chú ý (multi-attention): nhận dạng và chú ý vào đa dạng (multi) vùng cục bộ chứa trong ảnh chứa thông tin để phân biệt.

Nội dung và Phương pháp

Nội dung



- **Texture enhancement block**: tăng cường các đặc trưng nông ở các lớp đầu.
- **Module attention**: tạo các bản đồ chú ý (attention map).
- **Module attention pooling**: gộp thông tin từ bản đồ chú ý với:
    - đặc trưng cấp thấp được tăng cường.
    - đặc trưng ngữ nghĩa cấp cao.
- **Backbone layer**: lớp trích xuất đặc trưng.

Cần có phương pháp huấn luyện mới để các attention map độc lập nhau

Nội dung và Phương pháp

Phương pháp

- Khảo sát các lớp của các mạng phân loại ảnh tốt hiện nay (EfficientNet, XceptionNet, MobileNet, …) để đưa vào backbone layer.
- Nghiên cứu, tìm kiếm
  - Kiến trúc phù hợp cho 3 khối được đề ra.
  - Phương pháp huấn luyện phù hợp cho mô hình.
- Chứng minh hiệu quả của mô hình chúng tôi đề xuất bằng cách:
  - Thực nghiệm trên các bộ dữ liệu FaceForencis++, DFDC dataset, Celeb-DF và so sánh kết quả với các mô hình hiện tại.
  - Minh hoạ bản đồ chú ý để kiểm chứng mô hình.

Kết quả dự kiến

- Tìm ra được
  - Kiến trúc mạng phù hợp cho 3 khối được đề ra.
  - Phương pháp huấn luyện phù hợp.
- Kết quả của mô hình tìm được sẽ tốt hơn so với các mô hình hiện tại bài toán này.
- Các bản đồ chú ý thu được sẽ cho ra các kết quả độc lập → nắm bắt được các đặc trưng cục bộ từ nhiều vùng trên khuôn mặt.

Tài liệu tham khảo

[1]. Justus Thies, Michael Zollhöfer, Marc Stamminger, Christian Theobalt, Matthias Nießner: Face2Face: Real-time Face Capture and Reenactment of RGB Videos. CoRR abs/2007.14808 (2020)
[2]. Supasorn Suwajanakorn, Steven M. Seitz, Ira Kemelmacher-Shlizerman: Synthesizing Obama: learning lip sync from audio. ACM Trans. Graph. 36(4): 95:1-95:13 (2017)
[3]. Albert Pumarola, Antonio Agudo, Aleix M. Martínez, Alberto Sanfeliu, Francesc Moreno-Noguer: GANimation: Anatomically-Aware Facial Animation from a Single Image. ECCV (10) 2018: 835-851
[4]. Yuval Nirkin, Yosi Keller, Tal Hassner: FSGAN: Subject Agnostic Face Swapping and Reenactment. ICCV 2019: 7183-7192
[5]. Wayne Wu, Yunxuan Zhang, Cheng Li, Chen Qian, Chen Change Loy: ReenactGAN: Learning to Reenact Faces via Boundary Transfer. ECCV (1) 2018: 622-638
[6]. Andreas Rössler, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, Matthias Nießner: FaceForensics++: Learning to Detect Manipulated Facial Images. ICCV 2019: 1-11
[7]. Feng Wang, Weiyang Liu, Hanjun Dai, Haijun Liu, Jian Cheng: Additive Margin Softmax for Face Verification. ICLR (Workshop) 2018
[8]. Xin Yang, Yuezun Li, Siwei Lyu: Exposing Deep Fakes Using Inconsistent Head Poses. ICASSP 2019: 8261-8265
[9]. Yuezun Li, Ming-Ching Chang, Siwei Lyu: In Ictu Oculi: Exposing AI Created Fake Videos by Detecting Eye Blinking. WIFS 2018: 1-7
[10]. Lingzhi Li, Jianmin Bao, Ting Zhang, Hao Yang, Dong Chen, Fang Wen, Baining Guo: Face X-Ray for More General Face Forgery Detection. CVPR 2020: 5000-5009
[11]. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin: Attention is All you Need. NIPS 2017: 5998-6008
[12]. Kai Han, Yunhe Wang, Hanting Chen, Xinghao Chen, Jianyuan Guo, Zhenhua Liu, Yehui Tang, An Xiao, Chunjing Xu, Yixing Xu, Zhaohui Yang, Yiman Zhang, Dacheng Tao: A Survey on Visual Transformer. CoRR abs/2012.12556 (2020)
[13]. Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL-HLT (1) 2019: 4171-4186
[14]. Heliang Zheng, Jianlong Fu, Tao Mei, Jiebo Luo: Learning Multi-attention Convolutional Neural Network for Fine-Grained Image Recognition. ICCV 2017: 5219-5227
[15]. Brian Dolhansky, Russ Howes, Ben Pflaum, Nicole Baram, Cristian Canton-Ferrer: The Deepfake Detection Challenge (DFDC) Preview Dataset. CoRR abs/1910.08854 (2019)
[16]. Yuezun Li, Xin Yang, Pu Sun, Honggang Qi, Siwei Lyu: Celeb-DF: A Large-Scale Challenging Dataset for DeepFake Forensics. CVPR 2020: 3204-3213
[17]. Andreas Rössler, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, Matthias Nießner: FaceForensics++: Learning to Detect Manipulated Facial Images. ICCV 2019: 1-11