THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo (tối đa 5 phút): *https://youtu.be/tJD46XZ1egY*
- Link slides (dạng .pdf đặt trên Github của nhóm): *https://github.com/trthminh/CS519.O11/blob/main/slide.pdf*
- *Mỗi thành viên của nhóm điền thông tin vào một dòng theo mẫu bên dưới*
- *Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in*

• Họ và Tên: Trương Thanh Minh
• MSSV: 21520064
 
	• Lớp: CS519.O11
	• Tự đánh giá (điểm tổng kết môn): 9.5/10
	• Số buổi vắng: 0
	• Link Github: https://github.com/trthminh/CS519.O11
	• Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm:
	  ○ Lên ý tưởng đề tài
	  ○ Viết phần Tóm tắt, Mục tiêu, Nội dung và Phương pháp
	  ○ Làm slide
	  ○ Làm poster
	  ○ Làm video YouTube

• Họ và Tên: Tô Anh Phát • MSSV: 21520085	• Lớp: CS519.O11 • Tự đánh giá (điểm tổng kết môn): 9.5/10 • Số buổi vắng: 0 • Link Github: https://github.com/trthminh/CS519.O11 • Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm: ○ Lên ý tưởng đề tài

	○ Viết phần Giới thiệu, Nội dung và Phương pháp, Kết quả mong đợi ○ Làm slide ○ Làm poster ○ Làm video YouTube

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

SỬ DỤNG TÌM KIẾM CÁC MẠNG KẾT HỢP KHÁC BIỆT TRUNG TÂM ĐỂ GIẢI QUYẾT BÀI TOÁN CHỐNG GIẢ MẠO KHUÔN MẶT

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

SEARCHING CENTRAL DIFFERENCE CONVOLUTIONAL NETWORKS FOR FACE ANTI-SPOOFING

TÓM TẮT *(Tối đa 400 từ)*

Đề tài này nhằm áp dụng các phương pháp tiên tiến trong bài toán chống gian lận xác thực khuôn mặt (face anti-spoofing), một bài toán quan trọng trong lĩnh vực nhận dạng sinh trắc học. Các phương pháp được sử dụng trong đề tài bao gồm Central Difference Convolution (CDC), một toán tử tích chập mới có khả năng mô tả chi tiết các thông tin bất biến trong ảnh khuôn mặt, và CDCN, CDCN++, hai mô hình mạng nơ-ron tích chập sử dụng CDC để trích xuất các đặc trưng phân biệt khuôn mặt thật và khuôn mặt giả. Các phương pháp này đã được chứng minh hiệu quả trên bộ dữ liệu OULU-NPU, một bộ dữ liệu tiêu chuẩn cho bài toán face anti-spoofing. Đề tài cũng nhằm kiểm tra khả năng của các phương pháp trên bộ dữ liệu của cuộc thi Zalo AI Challenge 2022, một bộ dữ liệu mới có nhiều biến thể thực tế. Đề tài mong muốn đạt được kết quả chính xác và ổn định trên cả hai bộ dữ liệu OULU-NPU và Zalo AI Challenge 2022.

GIỚI THIỆU *(Tối đa 1 trang A4)*

Các hệ thống nhận diện khuôn mặt hiện nay đã không còn quá xa lạ đối với mọi người. Ta có thể dễ dàng bắt gặp các hệ thống nhận diện khuôn mặt trong khi đăng nhập vào các thiết bị điện tử như điện thoại, laptop, . . ., hoặc các hệ thống chấm công trong các công ty. Có thể thấy, nhận dạng khuôn mặt đã phát triển thành một phương pháp xác thực sinh trắc học nổi bật, phổ biến. Tuy nhiên, nó có thể dễ bị tấn công bằng cách giả mạo khuôn mặt như in ảnh (print), quay video mặt người đó (video

replay) và thậm chí là đeo mặt nạ 3D (3D masks), những điều đó làm hạn chế đi sự đáng tin cậy của các hệ thống nhận diện khuôn mặt. Từ đó, việc phát hiện tự động các cuộc tấn công giả mạo là cần thiết để sử dụng các hệ thống nhận diện khuôn mặt một cách an toàn trong các tình huống không có sự giám sát.

Bài toán nhận đầu vào là 1 ảnh chứa hình người và đầu ra là 0/1, trong đó, 0 tương ứng với spoofing, tức là ảnh giả mạo, còn 1 tương ứng là ảnh real, tức là ảnh người thật.

Hình 1: Minh họa input, output của bài toán

Hiện nay, một số hướng tiếp cận để giải quyết bài toán face-anti-spoofing như sau:

- Phương pháp dựa trên hand-crafted features: Những hand-crafted descriptors cổ điển (như Local Binary Pattern (LBP) [7]) tận dụng những mối quan hệ cục bộ giữa các "hàng xóm" để làm đặc trưng phân biệt. Từ đó, mạnh mẽ hơn trong việc mô tả chi tiết các thông tin bất biến (như kết cấu màu sắc, . . .)
- Phương pháp dựa trên Deep learning: Nhờ sự xếp chồng các convolution operations với hàm kích hoạt phi tuyến, CNN (Convolutional neural networks) nắm giữ khả năng đại diện mạnh mẽ để phân biệt mặt thật và mặt giả.

Những phương pháp đó còn mang một số hạn chế như:

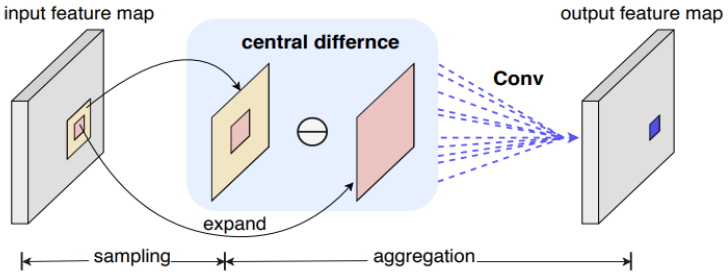
- Những phương pháp dựa trên hand-crafted features hạn chế trong việc nắm bắt những khác biệt phức tạp giữa khuôn mặt sống (living) và khuôn mặt giả mạo, nhất là các cuộc tấn công giả mạo phức tạp và thực tế hơn.
- Những phương pháp dựa trên CNN tập trung quá nhiều tới các đặc trưng deeper semantic, do đó nó yếu trong việc mô tả chi tiết thông tin fine-grained

giữa mặt thật và mặt giả. Ngoài ra, nó còn không hiệu quả khi môi trường thay đổi (ví dụ như điều kiện chiếu sáng khác nhau, . . .).

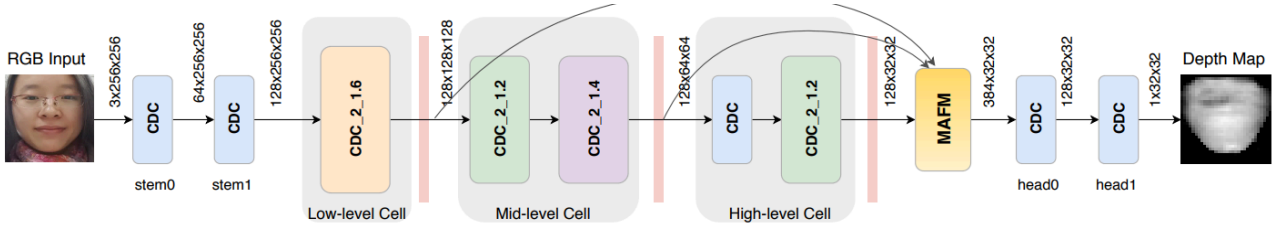
- Thông thường, các cấu trúc của mạng thường được thiết kế bởi chuyên gia, chính vì điều đó, nó có thể không tối ưu cho bài toán FAS (face-anti-spoofing).
- Những phương pháp state-of-the-art (SOTA) gần đây thường cần nhiều frames liên tục làm input để có thể trích xuất các đặc trưng dynamic spatio temporal (như chuyển động [4], rPPG [3], . . .). Tuy nhiên, các chuỗi video dài có thể không phù hợp cho điều kiện triển khai cụ thể mà cần thời gian nhanh. Do đó, hướng tiếp cận dựa trên frame thuận lợi hơn mặc dù có thể hiệu suất sẽ không cao như các phương pháp ở mức độ video.

Chính vì những điểm mạnh, điểm yếu của các phương pháp đó, ta mong muốn rằng sẽ tích hợp local descriptors với convolution operation lại với nhau để có thể tận dụng điểm mạnh của các phương pháp, từ đó ta sẽ có những feature representation mạnh mẽ. Một trong các phương pháp đó là Searching Central Difference Convolutional Networks for Face Anti-Spoofing.

Để xử lý các hạn chế đã đề cập ở trên, bài báo giới thiệu thêm 1 convolution operator là Central Difference Convolution (CDC), cái này rất tốt trong việc mô tả chi tiết các thông tin bất biến trong môi trường đa dạng. Ngoài ra, bài báo còn đề xuất CDCN++ để tổng hợp các CDC features ở nhiều level 1 cách hiệu quả.



Hình 2: Central difference convolution



Hình 3: Kiến trúc của CDCN++

Hướng tiếp cận này đã vượt qua các hướng tiếp cận khác state-of-the-art khác trên cả 6 benchmark datasets trong cả intra cũng như trên nhiều tập dữ liệu (cross-dataset testing).

Do đó, để kiểm chứng tính tổng quát của phương pháp này, nhóm chúng em sẽ thử nghiệm nó thêm trên bộ dataset của task Liveness Detection của cuộc thi Zalo AI Challenge 2022.

MỤC TIÊU

(Viết trong vòng 3 mục tiêu, lưu ý về tính khả thi và có thể đánh giá được)

1. Tìm hiểu tổng quan bài toán
2. Tìm hiểu và cài đặt CDCN, CDCN++
3. Đánh giá và phân tích phương pháp CDCN, CDCN++ trên bộ dữ liệu chuẩn OULU-NPU và trên tập dữ liệu cuộc thi Zalo AI Challenge 2022.

NỘI DUNG VÀ PHƯƠNG PHÁP

(Viết nội dung và phương pháp thực hiện để đạt được các mục tiêu đã nêu)

Trong đề tài nghiên cứu này, chúng tôi dự kiến sẽ nghiên cứu và tìm hiểu các nội dung chính như sau:

- **Nội dung 1: Tìm hiểu tổng quan bài toán**
    - **Phương pháp thực hiện**: Tìm hiểu, khảo sát và tổng hợp các tài liệu liên quan đến bài toán face-anti-spoofing, các phương pháp tiếp cận hiện có, các độ đo đánh giá và các bộ dữ liệu chuẩn.
    - **Kết quả dự kiến**: Tài liệu mô tả chi tiết về bài toán face-anti-spoofing, các thách thức, các ứng dụng và các xu hướng nghiên cứu.
- **Nội dung 2: Tìm hiểu và cài đặt CDCN, CDCN++**
    - **Phương pháp thực hiện**: Tìm hiểu, phân tích và cài đặt lại mô hình CDCN, CDCN++ dựa trên bài báo và mã nguồn của nhóm tác giả, sử dụng bộ dữ liệu OULU-NPU để huấn luyện và kiểm tra mô hình.
    - **Kết quả dự kiến**: Mã nguồn được chú thích chi tiết và tài liệu hướng

dẫn sử dụng mô hình CDCN, CDCN++, bảng thông tin về các thông số kỹ thuật trong quá trình huấn luyện và kiểm tra mô hình.

- **Nội dung 3: Đánh giá và phân tích phương pháp CDCN, CDCN++ trên bộ dữ liệu chuẩn OULU-NPU và trên tập dữ liệu cuộc thi Zalo AI Challenge 2022**
    - **Phương pháp thực hiện**: Sử dụng mô hình CDCN, CDCN++ đã huấn luyện để đánh giá trên bộ dữ liệu chuẩn OULU-NPU, so sánh và phân tích kết quả với các phương pháp state-of-the-art khác, áp dụng mô hình CDCN, CDCN++ trên bộ dữ liệu của cuộc thi Zalo AI Challenge 2022, đánh giá và phân tích kết quả trên tập dữ liệu đó.
    - **Kết quả dự kiến**: Bảng đánh giá mô hình CDCN, CDCN++ và các phương pháp khác trên bộ dữ liệu chuẩn, tài liệu phân tích và đánh giá mô hình CDCN, CDCN++ trên bộ dữ liệu của cuộc thi Zalo AI Challenge 2022, các trường hợp khó và hiệu suất của mô hình.

KẾT QUẢ MONG ĐỢI

(Viết kết quả phù hợp với mục tiêu đặt ra, trên cơ sở nội dung nghiên cứu ở trên)

- Tài liệu mô tả chi tiết về bài toán face-anti-spoofing, các thách thức, các ứng dụng và các xu hướng nghiên cứu.
- Mã nguồn được chú thích chi tiết và tài liệu hướng dẫn sử dụng mô hình CDCN, CDCN++.
- Bảng thông tin về các thông số kỹ thuật trong quá trình huấn luyện và kiểm tra mô hình CDCN, CDCN++.
- Bảng đánh giá mô hình CDCN, CDCN++ và các phương pháp khác trên bộ dữ liệu chuẩn, tài liệu phân tích và đánh giá mô hình CDCN, CDCN++ trên bộ dữ liệu của cuộc thi Zalo AI Challenge 2022, các trường hợp khó và hiệu suất của mô hình.

TÀI LIỆU THAM KHẢO *(Định dạng DBLP)*

[1] Amin Jourabloo, Yaojie Liu, and Xiaoming Liu. “Face despoofing: Anti-spoofing

via noise modeling". In: Proceedings of the European Conference on Computer Vision (ECCV) (2018).
[2] Anjith George and Sebastien Marcel. "Deep pixel-wise binary ´ supervision for face presentation attack detection". In: International Conference on Biometrics (2019).
[3] Xiao Yang, Wenhan Luo, Linchao Bao, Yuan Gao, Dihong Gong, Shibao Zheng, Zhifeng Li, and Wei Liu. "Face anti spoofing: Model matters, so does data". In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2019).
[4] Yaojie Liu, Joel Stehouwer, Amin Jourabloo, and Xiaoming Liu. "Deep tree learning for zero-shot face anti-spoofing". In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2019).
[5] Yuhui Xu, Lingxi Xie, Xiaopeng Zhang, Xin Chen, Guo-Jun Qi, Qi Tian, and Hongkai Xiong. "Pc-darts: Partial channel connections for memory-efficient differentiable architecture search". In: arXiv preprint arXiv:1907.05737. 2019.
[6] Zezheng Wang, Chenxu Zhao, Yunxiao Qin, Qiusheng Zhou, and Zhen Lei. "Exploiting temporal and depth information for multi-frame face anti-spoofing". In: arXiv preprint arXiv:1811.05118. 2018.
[7] Zinelabidine Boulkenafet, Jukka Komulainen, and Abdenour Hadid. "An image is worth 16x16 words: Transformers for image recognition at scale". In: IEEE international conference on image processing (ICIP) (2015).
[8] Zinelabinde Boulkenafet, Jukka Komulainen, Lei Li, Xiaoyi Feng, and Abdenour Hadid. "Oulunpu: A mobile face presentation attack database with real-world variations." In: FGR (2017).
[9] Zitong Yu, Chenxu Zhao, Zezheng Wang, Yunxiao Qin, Zhuo Su, Xiaobai Li, Feng Zhou, Guoying Zhao. "Searching Central Difference Convolutional Networks for Face Anti-Spoofing". In: CVPR 2020 (2020).