

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI

ĐỒ ÁN TỐT NGHIỆP

Nhận diện khuôn mặt từ một ảnh mẫu đơn dùng mô hình CNN kết hợp kỹ thuật ảnh 3D

Phạm Vũ Dũng

dung.pv183508@sis.hust.edu.vn

Ngành: Khoa học máy tính

Giảng viên hướng dẫn: PGS.TS. Nguyễn Thị Hoàng Lan

Chữ kí GVHD

Khoa: Khoa học máy tính

Trường: Công nghệ thông tin và Truyền thông

HÀ NỘI, 07/2023

LỜI CAM KẾT

Họ và tên sinh viên: Phạm Vũ Dũng
Điện thoại liên lạc: 0346909604
Email: dung.pv183508@sis.hust.edu.vn
Lớp: Khoa học máy tính - 03
Hệ đào tạo: Đại học chính quy

Tôi – *Phạm Vũ Dũng* – cam kết Đồ án Tốt nghiệp (ĐATN) là công trình nghiên cứu của bản thân tôi dưới sự hướng dẫn của *PGS.TS. Nguyễn Thị Hoàng Lan*. Các kết quả nêu trong ĐATN là trung thực, là thành quả của riêng tôi, không sao chép theo bất kỳ công trình nào khác. Tất cả những tham khảo trong ĐATN – bao gồm hình ảnh, bảng biểu, số liệu, và các câu từ trích dẫn – đều được ghi rõ ràng và đầy đủ nguồn gốc trong danh mục tài liệu tham khảo. Tôi xin hoàn toàn chịu trách nhiệm với dù chỉ một sao chép vi phạm quy chế của nhà trường.

Hà Nội, ngày tháng năm

Tác giả ĐATN

Họ và tên sinh viên

LỜI CẢM ƠN

Trước hết, tôi muốn gửi lời cảm ơn sâu sắc đến Đại học Bách Khoa Hà Nội, nơi đã cung cấp một môi trường học thuật xuất sắc, nơi tôi được đào tạo và phát triển kỹ năng chuyên môn của mình trong suốt quãng thời gian đại học.

Tôi xin chân thành cảm ơn các thầy cô trong Trường Công Nghệ Thông Tin Truyền Thông, đã tận tình hướng dẫn và giúp đỡ tôi trong quá trình học tập và thực hiện đồ án tốt nghiệp này. Đặc biệt, tôi xin bày tỏ lòng biết ơn sâu sắc đối với cô Hoàng Lan, người đã dành rất nhiều thời gian và công sức để chỉ bảo, hỗ trợ và khích lệ tôi trong suốt quá trình thực hiện đồ án.

Tôi cũng xin gửi lời cảm ơn tới gia đình, bạn bè, và những người bạn học của tôi, đã luôn ủng hộ và khích lệ tôi, cung cấp sự giúp đỡ và hỗ trợ khi tôi gặp khó khăn.

Cuối cùng, tôi xin cảm ơn tất cả mọi người đã ủng hộ và đồng hành cùng tôi trong suốt quá trình học tập và thực hiện đồ án tốt nghiệp này. Sự thành công này không chỉ là của tôi, mà còn là của tất cả mọi người đã đóng góp vào nó.

Tôi rất trân trọng mọi đóng góp và sự giúp đỡ mà tôi đã nhận được. Tôi hy vọng rằng những gì tôi đã học được và những kinh nghiệm tôi đã tích lũy sẽ giúp tôi tạo ra những đóng góp có giá trị trong lĩnh vực của mình trong tương lai.

TÓM TẮT NỘI DUNG ĐỒ ÁN

Đồ án "Nhận diện khuôn mặt từ một ảnh mẫu đơn dùng mô hình CNN (Convolutional Neural Networks) kết hợp kỹ thuật ảnh 3D" khám phá việc áp dụng mạng học sâu (Deep Learning), đặc biệt là mạng CNN vào bài toán nhận diện khuôn mặt từ một ảnh mẫu đơn (Single Sample Per Person - SSPP).

Bằng cách sử dụng phương pháp học sâu, đồ án này giới thiệu một cách tiếp cận mới cho bài toán SSPP, kết hợp với kỹ thuật tái tạo khuôn mặt 3D từ ảnh 2D để nâng cao khả năng nhận diện khuôn mặt. Mô hình CNN được huấn luyện để trích xuất đặc trưng khuôn mặt từ ảnh đầu vào, trong khi các phương pháp tái tạo khuôn mặt 3D được sử dụng để tạo ra một tập dữ liệu mẫu mở rộng, giúp cải thiện độ chính xác của việc nhận diện khuôn mặt từ một ảnh mẫu đơn.

Kết quả thử nghiệm cho thấy phương pháp này cung cấp một phương pháp hiệu quả cho bài toán nhận diện khuôn mặt từ một ảnh đơn mẫu, đồng thời mở ra hướng nghiên cứu mới trong việc kết hợp các mô hình học sâu với kỹ thuật ảnh 3D để giải quyết các vấn đề phức tạp trong lĩnh vực nhận diện khuôn mặt.

Sinh viên thực hiện
(Ký và ghi rõ họ tên)

MỤC LỤC

CHƯƠNG 1. GIỚI THIỆU ĐỀ TÀI.....	1
1.1 Đặt vấn đề.....	1
1.2 Các giải pháp hiện tại và hạn chế	3
1.2.1 Giải pháp hiện tại.....	3
1.2.2 Hạn chế.....	4
1.3 Mục tiêu và định hướng giải pháp	4
1.3.1 Mục tiêu.....	4
1.3.2 Định hướng giải pháp	5
1.4 Đóng góp của đề án	5
1.5 Bố cục đề án	5
CHƯƠNG 2. CƠ SỞ LÝ THUYẾT	7
2.1 Ngữ cảnh của bài toán.....	7
2.2 Các kết quả nghiên cứu liên quan	8
2.2.1 Các nghiên cứu liên quan tới tăng cường dữ liệu	8
2.2.2 Các nghiên cứu liên quan tới bài toán SSPP.....	10
2.3 Dựng khuôn mặt 3D.....	12
2.3.1 Giới thiệu	12
2.3.2 Hiểu về dựng khuôn mặt 3D và học sâu.....	12
2.3.3 Các hướng giải bài toán dựng khuôn mặt 3D với học sâu	12
2.3.4 Dựng khuôn mặt 3D với CNN [4]	13
2.3.5 Học có giám sát yếu	14
2.3.6 Học có giám sát yếu cho một ảnh.....	15
2.3.7 Tổng hợp neuron cho nhiều ảnh.....	16
2.3.8 Huấn luyện không nhãn	17

2.4 Nhận diện khuôn mặt	17
2.4.1 Giới thiệu	17
2.4.2 Hiểu về nhận diện khuôn mặt	17
2.4.3 Các hướng giải bài toán nhận diện khuôn mặt với học sâu.....	18
2.4.4 Nhận diện khuôn mặt với ResNet và ArcFace	18
CHƯƠNG 3. GIẢI PHÁP ĐỀ XUẤT	20
3.1 Tổng quan giải pháp.....	20
3.2 Tăng cường dữ liệu với khuôn mặt 3D	20
3.3 Trích xuất đặc trưng khuôn mặt với ResNet và ArcFace.....	21
3.4 Nhận diện khuôn mặt với hướng tìm kiếm vector hoạt động như một bộ phân lớp	21
CHƯƠNG 4. XÂY DỰNG HỆ THỐNG NHẬN DIỆN KHUÔN MẶT	23
4.1 Kiến trúc hệ thống nhận diện khuôn mặt.....	23
4.2 Pha huấn luyện	23
4.2.1 Chuẩn bị dữ liệu.....	23
4.2.2 Huấn luyện mô hình	26
4.3 Pha nhận diện.....	27
4.3.1 Lưu trữ đặc trưng khuôn mặt trong Cơ sở dữ liệu Vector.....	27
4.3.2 Bộ phân lớp sử dụng ElasticSearch và độ đo khoảng cách nhỏ nhất	27
CHƯƠNG 5. ĐÁNH GIÁ THỰC NGHIỆM.....	29
5.1 Các tham số đánh giá	29
5.1.1 Tham số đánh giá một-một	29
5.1.2 Tham số đánh giá một-nhiều	29
5.2 Phương pháp thực nghiệm	30
5.3 Phương pháp đánh giá một-một.....	30
5.4 Phương pháp đánh giá một-nhiều	31

CHƯƠNG 6. KẾT LUẬN	33
6.1 Kết luận.....	33
6.2 Hướng phát triển trong tương lai	33
TÀI LIỆU THAM KHẢO.....	36

DANH MỤC HÌNH VẼ

Hình 3.1	Luồng hoạt động của hệ thống	20
Hình 3.2	Kết quả tăng cường dữ liệu	20
Hình 4.1	Kiến trúc hệ thống nhận diện khuôn mặt	23
Hình 4.2	Sơ đồ khối pha huấn luyện	24
Hình 4.3	Phát hiện khuôn mặt với RetinaFace	25
Hình 4.4	Cấu trúc thư mục chứa dữ liệu huấn luyện	26
Hình 4.5	Sơ đồ khối pha nhận diện	27
Hình 5.1	Dữ liệu đánh giá một-một	30
Hình 5.2	Dữ liệu thực tế	30
Hình 5.3	Kết quả đánh giá một-một	31
Hình 5.4	Kết quả đánh giá bộ dữ liệu thực tế	31
Hình 5.5	Kết quả đánh giá bộ dữ liệu người nổi tiếng	32

DANH MỤC THUẬT NGỮ VÀ TỪ VIẾT TẮT

Thuật ngữ	Ý nghĩa
3DMM	3D Morphable Model
CNN	Mạng nơron tích chập (Convolutional Neural Network)
DA	Chuyển miền dữ liệu (Domain Adaptation)
KNN	K-Hàng xóm gần nhất (K-Nearest Neighbors)
RHDA	Robust Heterogeneous Discriminative Analysis
SSPP	Mỗi người một mẫu ảnh (Single Sample Per Person)

CHƯƠNG 1. GIỚI THIỆU ĐỀ TÀI

1.1 Đặt vấn đề

Nhận diện khuôn mặt với phương pháp học sâu (Deep Learning) đang được rộng rãi ứng dụng trong nhiều lĩnh vực như an ninh, quảng cáo hay công nghệ thông tin. Tuy nhiên, việc áp dụng học sâu vào nhận diện khuôn mặt đối mặt với nhiều thách thức, đặc biệt khi chỉ có ít dữ liệu.

Các thách thức khi xây dựng hệ thống nhận diện khuôn mặt khi có ít dữ liệu:

Đào tạo mô hình: Mô hình học sâu thường yêu cầu một lượng lớn dữ liệu để đào tạo hiệu quả. Khi có ít dữ liệu, mô hình có thể không học được các đặc điểm quan trọng cần thiết để phân biệt giữa các khuôn mặt khác nhau. Điều này dẫn đến hiệu suất nhận diện kém.

Quá khớp (Overfitting): Khi dữ liệu huấn luyện không đủ, mô hình có thể gặp phải vấn đề quá khớp. Điều này có nghĩa là mô hình học quá tốt trên dữ liệu huấn luyện nhưng không thể tổng quát hóa tốt trên dữ liệu mới hoặc chưa từng gặp.

Biến đổi và nhiễu: Khi có ít dữ liệu, mô hình có thể không được đào tạo để đối phó với các biến đổi như góc nhìn, ánh sáng, biểu cảm khuôn mặt, hoặc nhiễu trong quá trình chụp ảnh. Điều này làm giảm độ chính xác của việc nhận diện khuôn mặt.

Thiếu đa dạng: Nếu dữ liệu huấn luyện không đa dạng (ví dụ, chủ yếu là khuôn mặt của người da trắng hoặc chỉ có khuôn mặt nam giới), mô hình có thể không nhận diện tốt trên những nhóm không được đại diện trong dữ liệu huấn luyện.

Thiếu dữ liệu gán nhãn: Trong những trường hợp cụ thể, chúng ta có thể có nhiều dữ liệu không gán nhãn. Trong tình huống này, việc huấn luyện một mô hình học sâu đòi hỏi sự can thiệp của con người để gán nhãn, hoặc sử dụng các phương pháp học không giám sát hoặc bán giám sát, nhưng chúng có thể không mang lại hiệu suất tốt như học giám sát với dữ liệu gán nhãn đầy đủ.

Nhận diện khuôn mặt với một mẫu duy nhất cho mỗi người là một vấn đề thách thức do thông tin cung cấp cho mô hình để học các đặc điểm đặc trưng của từng khuôn mặt cá nhân bị giới hạn. Các mô hình học sâu truyền thống yêu cầu lượng dữ liệu lớn để tổng quát hóa tốt. Tuy nhiên, khi chỉ có một ví dụ được cung cấp cho mỗi lớp, như là trường hợp của nhận diện khuôn mặt với một mẫu duy nhất cho mỗi người, các mô hình có thể gặp khó khăn để học và tổng quát hóa.

Các giải pháp cho vấn đề này thường liên quan đến các kỹ thuật có thể tìm ra hoặc tạo ra thêm thông tin hữu ích từ số lượng dữ liệu hạn chế có sẵn. Qua khảo