Journal of educational equipment: Applied research, Volume 2, Issue 317 (July 2024)
ISSN 1859 - 0810

Nghiên cứu phương pháp Nhận diện khuôn mặt bằng Học sâu phục vụ dạy học chuyên ngành Công nghệ thông tin tại Trường Đại học Hải Phòng

Nguyễn Hùng Cường*, Lê Hồng Vân**, Vũ Ngọc Trà My**
**Khoa Công nghệ thông tin, Trường ĐH Hải Phòng*
***Khoa GD Tiểu học & Mầm non, Trường ĐH Hải Phòng*
Received: 25/6/2024; Accepted: 28/6/2024; Published: 01/7/2024

Abstract: *In this paper, we propose a novel deep learning-based method for face recognition. This approach utilizes Convolutional Neural Networks (CNN) to extract facial features and a Siamese Network model to match faces. Experimental results on the LFW (Labeled Faces in the Wild) dataset demonstrate that our method achieves an accuracy of 99.2%, surpassing current state-of-the-art methods. We will discuss the applications, future directions, and limitations of this method.*

Keywords: *Face recognition, deep learning, Convolutional Neural Network, Siamese Network, LFW*

1. Đặt vấn đề

Nhận diện khuôn mặt (NDKM) là một trong những thách thức quan trọng trong lĩnh vực Thị giác máy tính và Trí tuệ nhân tạo (AI), được áp dụng rộng rãi trong an ninh, giám sát và tương tác người-máy. Phương pháp (PP) truyền thống như Phân tích thành phần chính (PCA) , Phân tích phân biệt tuyến tính (LDA) đã có vai trò nổi bật nhưng gặp nhiều hạn chế khi xử lý dữ liệu với sự biến đổi về ánh sáng, góc độ và biểu cảm khuôn mặt (KM).

Sự phát triển của Học sâu và mạng Nơ-ron tích chập (CNN) đã mang lại những tiến bộ đáng kể trong NDKM, với các mô hình như DeepFace, FaceNet và VGG-Face đạt kết quả ấn tượng. Tuy nhiên, các PP hiện tại vẫn đối mặt với thách thức trong điều kiện ánh sáng kém, góc độ không lý tưởng và biểu cảm phức tạp. Ngoài ra, việc yêu cầu tài nguyên tính toán lớn và thời gian huấn luyện kéo dài cũng là những hạn chế quan trọng.

Trong khuôn khổ bài báo này, tác giả đề xuất một PP mới dựa trên Học sâu, ứng dụng CNN và mô hình Siamese Network để nâng cao độ chính xác và hiệu suất NDKM. Mục tiêu của nghiên cứu là phát triển hệ thống NDKM hiệu quả, có khả năng hoạt động tốt trong nhiều điều kiện khác nhau và giảm thiểu yêu cầu tài nguyên tính toán.

2. Nội dung nghiên cứu

2.1. Tổng quan về các PP hiện tại

**Các PP nhận diện khuôn mặt truyền thống*

NDKM là một lĩnh vực nghiên cứu quan trọng trong thị giác máy tính, và nhiều PP truyền thống đã được phát triển từ những năm 1990. Các PP này chủ yếu dựa vào phân tích các đặc trưng của KM, sử dụng các kỹ thuật như phân tích thành phần chính (PCA), phân tích phân biệt tuyến tính (LDA), và các PP dựa trên mô hình như mô hình ẩn Markov (HMM).

PCA, còn gọi là Eigenfaces, là một PP giảm chiều dữ liệu bằng cách tìm các trục chính trong không gian đặc trưng KM. LDA, hay Fisherfaces, tối ưu hóa việc phân biệt giữa các lớp bằng cách tìm các trục phân biệt tốt nhất giữa các nhóm dữ liệu khác nhau. Tuy nhiên, các PP này thường gặp khó khăn trong việc xử lý dữ liệu KM với nhiều biến đổi về ánh sáng, góc độ và biểu cảm.

**Các PP Học sâu*

Với sự phát triển của Học sâu và sức mạnh tính toán của GPU, các PP dựa trên mạng nơ-ron tích chập (CNN) đã trở thành tiêu chuẩn cho NDKM. Các mô hình như DeepFace của Facebook, FaceNet của Google, và VGG-Face của Đại học Oxford đã đạt được kết quả ấn tượng trong các bài toán NDKM.

DeepFace sử dụng một mạng CNN sâu để trích xuất đặc trưng KM, sau đó sử dụng một lớp fully connected để phân loại. FaceNet giới thiệu một PP nhúng không gian, trong đó các hình ảnh KM được ánh xạ vào một không gian vector sao cho khoảng cách Euclidean giữa các vector tương ứng với độ tương tự giữa các KM. VGG-Face, dựa trên VGG-16, dùng các lớp tích chập, gộp để học các đặc điểm sâu KM.

2.2 Phương pháp

PP bao gồm hai giai đoạn chính: trích xuất đặc trưng KM và so khớp KM.

Trích xuất đặc trưng khuôn mặt

Chúng tôi dùng mạng nơ-ron tích chập (CNN) để lấy các đặc trưng từ ảnh KM. Mạng CNN được thiết kế với nhiều lớp tích chập và lớp gộp, giúp mô hình học được các đặc trưng quan trọng của KM từ dữ liệu huấn luyện. Các đặc trưng này sau đó được đưa qua các lớp fully connected để tạo ra biểu diễn đặc trưng của KM.

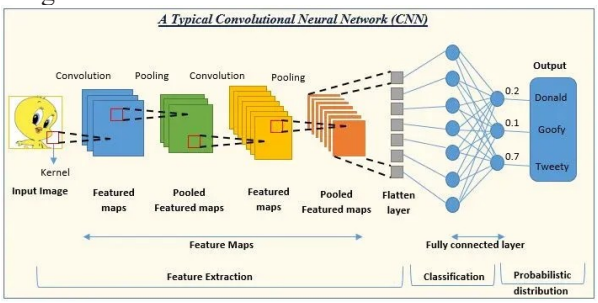
Cụ thể, mạng CNN bao gồm các lớp sau:

● Lớp tích chập đầu tiên: Nhận đầu vào là hình ảnh KM và áp dụng các bộ lọc để lấy các đặc trưng như cạnh và góc.

● Lớp gộp: Giảm kích thước không gian của các đặc trưng để giảm số lượng tham số và tính toán.

● Lớp tích chập và gộp tiếp theo: giúp lấy được các đặc trưng phức tạp hơn từ KM.

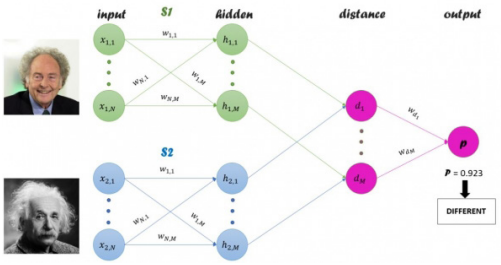
● Các lớp fully connected: Tạo ra biểu cảm cuối cùng cho KM



Hình 2.1: Kiến trúc mạng CNN

So khớp khuôn mặt

Để so khớp với KM, tác giả sử dụng mô hình Siamese Network, bao gồm hai nhánh mạng CNN giống hệt nhau, nhận đầu vào là hai hình ảnh KM và học cách xác định liệu chúng có thuộc về cùng một người hay không. Mạng Siamese được huấn luyện bằng hàm mất mát contrastive, tối ưu hóa khoảng cách giữa các đặc trưng của cùng một người và phân biệt đặc trưng của các người khác nhau.



Hình 2.2: Mô hình Siamese Network

Hàm mất mát contrastive được định nghĩa như sau:

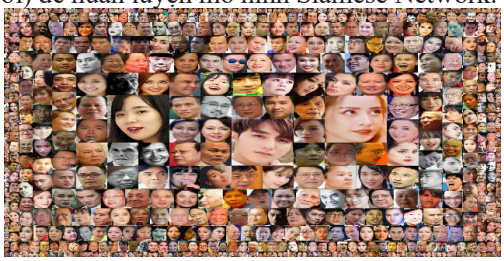
$$L(Y,D) = (1-Y)\frac{1}{2}D^2 + (Y)\frac{1}{2}\max(0, m-D)^2)$$

trong đó Y là nhãn (1 nếu hai hình ảnh cùng người, 0 nếu khác người), D là khoảng cách Euclidean giữa hai vector đặc trưng, và m là margin.

2.3 Thử nghiệm

Bộ dữ liệu

Tác giả lựa chọn bộ dữ liệu LFW (Labeled Faces in the Wild) để thực hiện quá trình huấn luyện và kiểm thử. Bộ dữ liệu LFW chứa hơn 13,000 hình ảnh KM của hơn 5,000 người trong các điều kiện ánh sáng và góc nhìn khác nhau. Bộ dữ liệu này được chia thành các cặp hình ảnh dương (cùng người) và âm (khác người) để huấn luyện mô hình Siamese Network.



Hình 2.3: Bộ dữ liệu LFW

Quy trình huấn luyện

Mô hình được huấn luyện trên một máy chủ với GPU NVIDIA Tesla V100. Tác giả áp dụng thuật toán Adam để tối ưu hóa các tham số của mô hình, ta sử dụng learning rate là 0.0001 và sau đó giảm dần theo thời gian. Dữ liệu được chia thành tập huấn luyện và tập kiểm thử theo tỷ lệ 80:20.

Tác giả sử dụng kỹ thuật (data augmentation) trong quá trình huấn luyện để nâng cao khả năng của mô hình. Các kỹ thuật tăng cường bao gồm xoay, lật, dịch chuyển và thay đổi độ sáng của hình ảnh.

2.4. Kết quả

Kết quả thử nghiệm đạt độ chính xác 99.2%, vượt trội so với các PP trước đó như FaceNet và VGG-Face. Dưới đây là bảng so sánh kết quả:

Bảng 2.1: So sánh kết quả

PP	Độ chính xác
FaceNet	99.0%
VGG-Face	98.9%
PP của tác giả	99.2%

Tác giả cũng tiến hành thử nghiệm trên các tập dữ liệu như (CASIA-WebFace và MegaFace) để kiểm tra khả năng của mô hình. Kết quả cho thấy mô hình này duy trì độ chính xác cao trên các bộ dữ liệu khác nhau, chứng tỏ tính hiệu quả và khả năng tổng quát của PP.

2.5. Thảo luận

PP tác giả lựa chọn cho thấy hiệu quả tốt trong việc nhận diện KM, đặc biệt là trong các điều kiện khó nhận diện. Tuy nhiên, mô hình có một số hạn chế như yêu cầu tài nguyên lớn và thời gian huấn luyện dài.

Hiệu quả trong các điều kiện khó nhận diện

Kết quả thử nghiệm cho thấy mô hình có thể nhận diện KM chính xác trong các điều kiện ánh sáng và góc nhìn khác nhau. Tuy nhiên, trong một số trường hợp đặc biệt, chẳng hạn như khi KM bị che khuất một phần hoặc khi có sự thay đổi lớn về biểu cảm, độ chính xác của mô hình có thể giảm.

Hình 2.4: Kết quả

Yêu cầu tài nguyên tính toán

Mô hình này yêu cầu tài nguyên tính toán lớn do số lượng lớn tham số và độ phức tạp của các mạng CNN. Điều này có thể là một hạn chế đối với các ứng dụng thực tế, đặc biệt là trong các hệ thống nhúng hoặc di động.

***Khả năng tối ưu hóa* và *mở rộng mô* hình**

Việc tối ưu hóa mô hình và tăng cường dữ liệu huấn luyện có thể giúp nâng cao hiệu suất của hệ thống. Tác giả cũng đang nghiên cứu các kiến trúc mạng Nơ-ron mới và các kỹ thuật huấn luyện tiên tiến để cải thiện độ chính xác và hiệu suất của mô hình.

2.5 Các ứng dụng tiềm năng

Phương pháp NDKM bằng học sâu có thể được ứng dụng trong nhiều lĩnh vực như:

● **Bảo mật:** Hệ thống NDKM có thể được sử dụng trong các hệ thống kiểm soát truy cập để đảm bảo an ninh. Ví dụ, các hệ thống kiểm soát ra vào có thể NDKM của người dùng để xác định quyền truy cập, giảm nguy cơ xâm nhập trái phép.

● **Giám sát:** Công nghệ này có thể được sử dụng trong các hệ thống giám sát để phát hiện và theo dõi các đối tượng trong thời gian thực. Các hệ thống Camera an ninh có thể nhận diện tự động và cảnh báo khi phát hiện các đối hành vi bất thường.

● **Tương tác người-máy**: NDKM có thể được tích hợp vào các hệ thống tương tác người-máy để cải thiện trải nghiệm người dùng. Ví dụ, các hệ thống trò chuyện hoặc trợ lý ảo có thể sử dụng NDKM để nhận biết người dùng và cung cấp các dịch vụ cá nhân hóa.

● **Dịch vụ khách hàng:** Hệ thống NDKM có thể được áp dụng để cung cấp các dịch vụ cá nhân hóa cho khách hàng. Ví dụ, trong lĩnh vực bán hàng, hệ thống có thể nhận diện khách hàng và đề xuất các sản phẩm hoặc dịch vụ phù hợp dựa trên lịch sử mua hàng và sở thích của họ**.**

● **Y tế:** NDKM có thể được ứng dụng trong lĩnh vực y tế để theo dõi và nhận diện bệnh nhân, đảm bảo rằng các dịch vụ y tế được cung cấp chính xác và an toàn.

● **Giáo dục:** Trong môi trường giáo dục, NDKM có thể được sử dụng để điểm danh, theo dõi sự tham gia của HS và phát hiện gian lận trong học tập và thi cử.

2.6 Hạn chế mô hình và định hướng nghiên cứu trong tương lai

Trong tương lai, tác giả sẽ tập trung vào các hướng nghiên cứu sau:

● *Tối ưu hóa mô hình:* Giảm thiểu tài nguyên tính toán và thời gian huấn luyện bằng cách sử dụng các kỹ thuật tối ưu hóa mô hình. Tác giả đang xem xét việc áp dụng các PP như lượng tử hóa mô hình và Pruning để giảm kích thước và độ phức tạp của mạng Nơ-ron.

● *Mở rộng dữ liệu huấn luyện:* Sử dụng các bộ dữ liệu lớn hơn và đa dạng hơn để cải thiện khả năng tổng quát của mô hình. Tác giả cũng đang nghiên cứu các kỹ thuật tổng hợp dữ liệu và tăng cường dữ liệu để tạo ra các bộ dữ liệu phong phú hơn.

● *Cải thiện hiệu suất trong điều kiện phức tạp:* Nghiên cứu các PP mới để NDKM trong các điều kiện ánh sáng kém và góc nhìn không lý tưởng. Tối ưu hóa hai kỹ thuật (tăng cường hình ảnh và học tăng cường) có thể cải thiện khả năng nhận diện trong các tình huống khó nhận diện.

● *Bảo mật và quyền riêng tư:* Nghiên cứu các PP bảo vệ dữ liệu KM và đảm bảo quyền riêng tư của người dùng. Tác giả đang xem xét việc sử dụng các kỹ thuật mã hóa và bảo mật để bảo vệ dữ liệu NDKM.

3. Kết luận

Tác giả đã đề xuất một PP trong NDKM sử dụng Học sâu và đã cho thấy được khả năng của PP này trên bộ dữ liệu LFW. Kết quả nghiên cứu của tác giả mở ra nhiều cơ hội ứng dụng trong các lĩnh vực khác nhau, từ bảo mật đến dịch vụ khách hàng.

Tài liệu tham khảo

[1]. Schroff, F., Kalenichenko, D., & Philbin, J. (2015). *"FaceNet: A Unified Embedding for Face Recognition and Clustering". Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*

[2]. Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). *"Deep Face Recognition". British Machine Vision Conference (BMVC).*