**ỦY BAN NHÂN DÂN THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC SÀI GÒN**



**Lê Quang Vinh**

**Vũ Thị Thanh Vân**

**FACE RECOGNITION USING**

**CONVOLUTIONAL NEURAL NETWORKS**

**ĐỀ CƯƠNG LUẬN VĂN CỬ NHÂN**

NGÀNH: CÔNG NGHỆ THÔNG TIN

**Thành phố Hồ Chí Minh, năm 2025**

**ỦY BAN NHÂN DÂN THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC SÀI GÒN**

**Lê Quang Vinh**

**Vũ Thị Thanh Vân**

**FACE RECOGNITION USING**

**CONVOLUTIONAL NEURAL NETWORKS**

**ĐỀ CƯƠNG LUẬN VĂN CỬ NHÂN**

NGÀNH: CÔNG NGHỆ THÔNG TIN

**Người hướng dẫn Chủ nhiệm chuyên ngành**

**GV A GV B**

**Thành phố Hồ Chí Minh, năm 2025**

# ĐỀ CƯƠNG LUẬN VĂN

## **1. Lý do chọn đề tài**

## Trong thời đại công nghệ phát triển mạnh mẽ, Mạng Nơ-ron Tích Chập (CNN) đã trở thành một công cụ quan trọng trong lĩnh vực xử lý và nhận dạng hình ảnh. Các tập đoàn lớn như Facebook, Google, Amazon đang ứng dụng CNN để giải quyết nhiều bài toán quan trọng như nhận diện khuôn mặt, tìm kiếm hình ảnh và đề xuất sản phẩm, cho thấy tiềm năng rộng lớn của nó trong nhiều lĩnh vực.

## Không giống như các mạng nơ-ron truyền thống như Feedforward Neural Network, CNN được thiết kế để xử lý thông tin hình ảnh một cách hiệu quả hơn bằng cách chỉ kết nối cục bộ giữa các điểm ảnh và nút. Cách tiếp cận này giúp giảm đáng kể khối lượng tính toán khi kích thước hình ảnh tăng lên, đồng thời tối ưu hóa tài nguyên và nâng cao hiệu suất xử lý.

## Trong nhận diện khuôn mặt, CNN thực hiện một chuỗi các bước xử lý từ phát hiện, phân đoạn, trích xuất đặc trưng đến phân loại, mang lại độ chính xác cao và ứng dụng rộng rãi trong thực tế. Điều này mở ra cơ hội nghiên cứu sâu hơn nhằm tinh chỉnh và tối ưu hóa các bước xử lý để hệ thống ngày càng hoạt động hiệu quả hơn.

## Chính vì vậy, nghiên cứu về CNN không chỉ mang giá trị khoa học mà còn đóng góp quan trọng vào sự phát triển của các hệ thống thông minh hiện đại, giúp nâng cao khả năng xử lý hình ảnh và hỗ trợ nhiều ứng dụng thực tiễn.

## **2. Tổng quan vấn đề nghiên cứu**

## **2.1. Tình hình nghiên cứu hiện tại**

**Huấn luyện mạng Nơ-ron Inception V3:**

Đạt được độ chính xác cao nhất là 78,0% và top 5 là 93,9% trên ImageNet kiểm tra dữ liệu bao gồm 1000 lớp hình ảnh. [1]

### **Sử dụng tập dữ liệu Labeled Faces in the Wild (LFW)**

Tác giả chọn tập dữ liệu LFW, bao gồm hơn 13.000 hình ảnh của 5.749 cá nhân. Tuy nhiên, để đảm bảo đủ dữ liệu cho mỗi lớp, chỉ chọn 10 cá nhân có nhiều hình ảnh nhất (ít nhất 50 ảnh mỗi người). Điều này giúp mô hình học tốt hơn và giảm thiểu sự mất cân bằng lớp.

## **2.2. Hướng tiếp cận của đề tài**

Chứng minh khả năng tái sử dụng mô hình CNN tiền huấn luyện (Inception V3) cho nhiệm vụ nhận diện khuôn mặt, từ đó tiết kiệm thời gian và tài nguyên so với việc huấn luyện mô hình từ đầu.

Đề tài sử dụng **chuyển giao học sâu (transfer learning)** với mô hình **Inception V3** đã được huấn luyện trước trên ImageNet để giải quyết bài toán nhận diện khuôn mặt. Thay vì huấn luyện toàn bộ mạng, chỉ có **lớp cuối cùng** được thay đổi và huấn luyện lại để phân biệt các khuôn mặt từ bộ dữ liệu LFW. Kỹ thuật **tăng cường dữ liệu** và **căn chỉnh khuôn mặt** cũng được sử dụng để cải thiện độ chính xác. Hướng tiếp cận này giúp tiết kiệm thời gian huấn luyện và tài nguyên tính toán.

## **3. Mục đích và nhiệm vụ nghiên cứu**

**Mục đích nghiên cứu:** Mục đích nghiên cứu của bài báo là minh họa cách áp dụng chuyển giao học sâu (transfer learning) để phát triển hệ thống nhận diện khuôn mặt bằng mô hình Inception V3 đã được huấn luyện trước. Bài báo tập trung vào việc tinh chỉnh mô hình và tăng cường dữ liệu nhằm tiết kiệm thời gian huấn luyện và cải thiện độ chính xác trong phân loại khuôn mặt.

**Nhiệm vụ nghiên cứu:** Huấn luyện lại mô hình Inception-v3 của Google cho một nhiệm vụ mới là nhận diện khuôn mặt (facial recognition), thay vì phân loại các đối tượng chung trong ImageNet.

### Ứng dụng chuyển giao học sâu (Transfer Learning) : lý do sử dụng chuyển giao học sâu để tiết kiệm thời gian huấn luyện.

### Tinh chỉnh mô hình (Fine-tuning): quá trình chỉ huấn luyện lại lớp cuối cùng của mô hình, giữ nguyên các lớp trước đó, và chỉ thay đổi lớp đầu ra để nhận diện khuôn mặt.

### Áp dụng kỹ thuật tăng cường dữ liệu (Data Augmentation): mô tả các kỹ thuật như lật ảnh ngang và dịch chuyển ảnh để tăng cường bộ dữ liệu huấn luyện, nhằm giúp mô hình học được nhiều biến thể của các khuôn mặt.

### Đánh giá hiệu quả mô hình: việc sử dụng mô hình đã huấn luyện trước giúp tiết kiệm thời gian huấn luyện và tài nguyên tính toán.

## **4. Đối tượng và phạm vi nghiên cứu**

**Đối tượng nghiên cứu:** Đối tượng nghiên cứu của bài báo là hệ thống nhận diện khuôn mặt. Cụ thể, tác giả nghiên cứu việc sử dụng mô hình Convolutional Neural Network (CNN) đã được huấn luyện trước, cụ thể là mô hình Inception V3, để thực hiện tác vụ phân loại khuôn mặt người. Mô hình này được áp dụng để nhận diện và phân biệt các khuôn mặt từ một tập dữ liệu cụ thể, bao gồm những người trong bộ dữ liệu Labeled Faces in the Wild (LFW).

**Phạm vi nghiên cứu**

Phạm vi nghiên cứu của bài báo giới hạn ở việc **ứng dụng chuyển giao học sâu (transfer learning)** và **tinh chỉnh mô hình đã được huấn luyện trước** để giải quyết bài toán nhận diện khuôn mặt

Cụ thể, phạm vi nghiên cứu bao gồm:

* Không huấn luyện toàn bộ mạng nơ-ron (vì quá tốn tài nguyên), chỉ huấn luyện lại lớp cuối cùng (last layer).
* Sử dụng mô hình đã được huấn luyện trước (pre-trained) để tận dụng khả năng nhận diện các đặc trưng cơ bản như hình dạng, màu sắc, kết cấu…
* Dataset sử dụng: "Labeled Faces in the Wild" (LFW), giới hạn lại chỉ 10 người có nhiều hình ảnh nhất (mỗi người ít nhất 50 ảnh).
* Sử dụng phiên bản hình ảnh đã được căn chỉnh bằng kỹ thuật deep funneling nhằm giảm sự khác biệt nội bộ trong cùng một lớp*(intra-class variability)***.**

## **5. Phương pháp nghiên cứu**

Luận văn áp dụng phương pháp nghiên cứu kết hợp giữa phương pháp lý thuyết, phương pháp thực nghiệm, và phương pháp chuyên gia, nhằm đảm bảo tính hệ thống, khoa học và hiệu quả trong quá trình thực hiện.

Phương pháp lý thuyết:

* Các nghiên cứu liên quan đến các cơ chế trong nhận diện khuôn mặt sử dụng **Convolutional Neural Networks (CNN)** bao gồm các cơ chế chính như **chuyển giao học sâu (transfer learning)**, **tăng cường dữ liệu (data augmentation)** và **cơ chế Attention**. Các cơ chế này giúp cải thiện hiệu suất của mô hình CNN trong các tác vụ nhận diện khuôn mặt và hình ảnh.
* Bài báo sử dụng ba cơ chế chính: chuyển giao học sâu bằng mô hình **Inception V3**, **tăng cường dữ liệu** bằng các biến đổi ảnh như xoay, lật, thay đổi ánh sáng, và trực quan hóa quá trình huấn luyện qua **TensorBoard**. Các cơ chế này giúp rút ngắn thời gian huấn luyện và cải thiện độ chính xác.
* Lựa chọn nguồn tài liệu uy tín từ các tạp chí khoa học hàng đầu như: **IEEE Transactions on Neural Networks and Learning Systems, Journal of Machine Learning Research (JMLR), Neural Information Processing Systems (NeurIPS), International Conference on Computer Vision (ICCV), IEEE Conference on Computer Vision and Pattern Recognition (CVPR).**

Phương pháp thực nghiệm:

* **Huấn luyện lại mô hình Inception V3 đã được huấn luyện trước**, bằng cách thay thế lớp cuối cùng để phân loại khuôn mặt, sau đó huấn luyện lớp này trên một tập dữ liệu ảnh khuôn mặt được chuẩn bị riêng. Trong quá trình huấn luyện, tác giả sử dụng **kỹ thuật tăng cường dữ liệu** để cải thiện hiệu suất và **theo dõi kết quả bằng TensorBoard**, nhằm đánh giá độ chính xác và hàm mất mát của mô hình theo thời gian.
* Cài đặt mô hình và thuật toán bằng ngôn ngữ Python, sử dụng các thư viện: **TensorFlow, Keras (API của TensorFlow), TensorBoard, NumPy, OS, glob, Matplotlib**
* Kiểm thử trên tập dữ liệu, đánh giá hiệu quả dữ liệu dựa trên chỉ số như Accuracy (độ chính xác), Loss (hàm mất mát).

Phương pháp chuyên gia:

* Trao đổi thường xuyên với giảng viên hướng dẫn để làm rõ các vấn đề khoa học liên quan, nhận phản hồi và định hướng cho từng giai đoạn nghiên cứu.
* Dựa trên ý kiến của giảng viên, điều chỉnh cách tiếp cận, tối ưu hóa quá trình triển khai thực nghiệm, và đảm bảo tiến độ thực hiện luận văn.

Phương pháp nghiên cứu này kết hợp chặt chẽ giữa lý thuyết và thực nghiệm, cùng với sự hỗ trợ từ chuyên gia, nhằm đảm bảo kết quả nghiên cứu có tính khoa học, khả thi và giá trị ứng dụng cao.

## **6. Giả thuyết khoa học**

Việc áp dụng **chuyển giao học sâu** (transfer learning) với mô hình **Inception V3** đã được huấn luyện trước trên bộ dữ liệu ImageNet, kết hợp với **tăng cường dữ liệu** (data augmentation), sẽ cải thiện hiệu suất nhận diện khuôn mặt mà không cần phải huấn luyện mô hình từ đầu, đồng thời tiết kiệm thời gian và tài nguyên tính toán.

## **7. Những đóng góp mới của đề tài**

Đề tài **“Facial Recognition Using Google’s Convolutional Neural Network”** có những đóng góp mới đáng chú ý trong lĩnh vực nhận diện khuôn mặt. Đầu tiên, bài báo chứng minh hiệu quả của việc **áp dụng chuyển giao học sâu (transfer learning)** với mô hình **Inception V3** đã được huấn luyện sẵn trên bộ dữ liệu ImageNet, giúp tiết kiệm thời gian và tài nguyên huấn luyện mà vẫn đạt được độ chính xác cao trong bài toán nhận diện khuôn mặt. Thứ hai, tác giả sử dụng **kỹ thuật tăng cường dữ liệu** (data augmentation) như xoay, lật, thay đổi độ sáng của ảnh để tạo thêm mẫu dữ liệu huấn luyện, từ đó giúp mô hình học được nhiều đặc trưng hơn và giảm thiểu hiện tượng overfitting khi dữ liệu huấn luyện hạn chế. Cuối cùng, bài báo cũng đóng góp bằng cách **trực quan hóa quá trình huấn luyện** thông qua **TensorBoard**, giúp theo dõi các chỉ số như accuracy, loss và các đường cong huấn luyện, qua đó hỗ trợ đánh giá và điều chỉnh mô hình kịp thời. Những đóng góp này không chỉ cải thiện hiệu suất nhận diện khuôn mặt mà còn thúc đẩy ứng dụng của các kỹ thuật học sâu vào các bài toán thực tế trong lĩnh vực nhận diện ảnh.

## **8. Dự kiến kế hoạch nghiên cứu**

Bảng 8.1 Dự kiến kế hoạch nghiên cứu

| **STT** | **Nội dung** | **Dự kiến thời gian thực hiện** |
| --- | --- | --- |
| 1 | Nghiên cứu, chọn đề tài, xây dựng đề cương luận văn | 1 tháng |
| 3 | Nộp đề cương, sửa chữa hoàn thiện đề cương. | 0.5 tháng |
| 4 | Nghiên cứu, viết, hoàn thiện luận văn |  |
| Chương 1: Giới thiệu. | 0.5 tháng |
| Chương 2: Các nghiên cứu liên quan. | 0.5 tháng |
| Chương 3: Thực nghiệm | 2 tháng |
| Chương 4: Kết luận và hướng phát triển | 0.5 tháng |
| 5 | Chỉnh sửa, hoàn thiện luận văn và hồ sơ bảo vệ luận văn. | 1 tháng |

## **9. Dự kiến nội dung của luận văn**

**Chương 1: Giới thiệu** 1.1. Lý do chọn đề tài  
 1.2. Mục tiêu nghiên cứu  
 1.3. Đối tượng và phạm vi nghiên cứu  
 1.4. Phương pháp nghiên cứu  
 1.5. Những đóng góp mới của đề tài  
 1.6. Cấu trúc luận văn

**Chương 2: Cơ sở lý thuyết** 2.1. Tổng quan về mạng nơ-ron tích chập (CNN)  
 2.2. Giới thiệu về mô hình Inception V3  
 2.3. Chuyển giao học sâu (Transfer Learning)  
 2.4. Tăng cường dữ liệu (Data Augmentation)  
 2.5. Các phương pháp đánh giá hiệu suất mô hình: Accuracy, Loss

**Chương 3: Phương pháp nghiên cứu và mô hình đề xuất** 3.1. Quy trình nghiên cứu và các bước thực hiện  
 3.2. Tích hợp chuyển giao học sâu vào mô hình  
 3.3. Chuẩn bị dữ liệu và cài đặt mô hình  
 3.4. Đề xuất quy trình tối ưu hóa và huấn luyện mô hình

**Chương 4: Thực nghiệm và kết quả** 4.1. Thực hiện thí nghiệm trên bộ dữ liệu khuôn mặt  
 4.2. Đánh giá hiệu suất mô hình  
 4.3. Phân tích kết quả huấn luyện và kiểm thử  
 4.4. Biểu diễn kết quả qua bảng biểu và đồ thị

**Chương 5: Kết luận và hướng phát triển** 5.1. Tóm tắt kết quả đạt được  
 5.2. Đánh giá hiệu quả của mô hình  
 5.3. Hạn chế của nghiên cứu  
 5.4. Đề xuất hướng nghiên cứu và ứng dụng trong tương lai

## **10. Danh mục tài liệu tham khảo**

[1] **Koehrsen, W.** *Facial Recognition using Google’s Convolutional Neural Network*.