**TRƯỜNG ĐẠI HỌC CÔNG THƯƠNG TP. HỒ CHÍ MINH**

**KHOA CÔNG NGHỆ THÔNG TIN**

**------------------------------**

A blue and white logo with a book and a square cap

AI-generated content may be incorrect.

**BÁO CÁO HỌC PHẦN:**

**THỰC HÀNH DEEP LEARNING**

Đề tài:

**NHẬN DIỆN CẢM XÚC GƯƠNG MẶT THÔNG QUA HÌNH ẢNH BẰNG MÔ HÌNH CNN**

**Thành phố Hồ Chí Minh, tháng 5 năm 2025**

**TRƯỜNG ĐẠI HỌC CÔNG THƯƠNG TP. HỒ CHÍ MINH**

**KHOA CÔNG NGHỆ THÔNG TIN**

**------------------------------**

A blue and white logo with a book and a square cap

AI-generated content may be incorrect.

**BÁO CÁO HỌC PHẦN:**

**THỰC HÀNH DEEP LEARNING**

Đề tài:

**NHẬN DIỆN CẢM XÚC GƯƠNG MẶT THÔNG QUA HÌNH ẢNH BẰNG MÔ HÌNH CNN**

**Giảng viên hướng dẫn :** Nguyễn Đình Toàn

**Sinh viên thực hiện :**

2001221088 - Nguyễn Hồng Hà

2001225449 – Trịnh Võ Thùy Trâm

**Thành phố Hồ Chí Minh, tháng 5 năm 2025**

**LỜI MỞ ĐẦU**

Trong bối cảnh thế giới đang bước vào thời kỳ chuyển đổi số mạnh mẽ, công nghệ trí tuệ nhân tạo (Artificial Intelligence – AI) không chỉ còn là một xu hướng nhất thời mà đã trở thành nền tảng cốt lõi trong việc tái cấu trúc lại toàn bộ các lĩnh vực từ kinh tế, giáo dục, y tế, đến giao tiếp và đời sống hàng ngày. Trong số các nhánh nghiên cứu quan trọng của AI, học sâu (Deep Learning) – đặc biệt là các mô hình mạng nơ-ron tích chập (CNN – Convolutional Neural Network) – đang đóng vai trò trung tâm trong các ứng dụng về thị giác máy tính (Computer Vision), mở ra nhiều tiềm năng ứng dụng to lớn trong thực tiễn.

Một trong những hướng phát triển nổi bật trong lĩnh vực thị giác máy tính chính là khả năng nhận diện cảm xúc của con người thông qua biểu cảm gương mặt. Đây không chỉ là một bài toán kỹ thuật mang tính thách thức cao, mà còn có ý nghĩa lớn trong việc nâng cao khả năng tương tác giữa con người và máy móc, giúp máy móc hiểu được trạng thái cảm xúc của người dùng, từ đó điều chỉnh hành vi phản hồi một cách phù hợp, thông minh và “nhân văn” hơn.

Với đề tài "Nhận diện cảm xúc gương mặt thông qua hình ảnh bằng mô hình CNN", nhóm chúng em mong muốn tiếp cận và khai thác tiềm năng của học sâu trong lĩnh vực nhận diện cảm xúc. Đề tài không chỉ dừng lại ở việc xây dựng một mô hình có khả năng phân loại cảm xúc như vui vẻ, buồn bã, giận dữ, ngạc nhiên... mà còn hướng đến việc triển khai hệ thống có thể hoạt động trong môi trường thực tế, sử dụng webcam để phát hiện và phản hồi cảm xúc theo thời gian thực.

Mục tiêu của đề tài là nghiên cứu và xây dựng một pipeline hoàn chỉnh gồm các bước: thu thập và xử lý dữ liệu gương mặt, huấn luyện mô hình CNN dựa trên tập dữ liệu chuẩn (FER2013), kiểm thử độ chính xác, triển khai demo trực quan, và đánh giá hiệu quả thực tiễn của mô hình. Bên cạnh đó, nhóm còn đề xuất một số hướng phát triển mở rộng trong tương lai, như tích hợp mô hình vào ứng dụng chăm sóc khách hàng, giáo dục từ xa, hoặc hệ thống cảnh báo tâm lý.

Trong quá trình thực hiện, nhóm đã gặp không ít khó khăn, từ việc xử lý dữ liệu hình ảnh chất lượng thấp, đến việc tinh chỉnh mô hình sao cho đạt độ chính xác cao mà vẫn tối ưu tốc độ xử lý. Tuy nhiên, nhờ có sự hướng dẫn tận tình từ giảng viên, sự hỗ trợ về kiến thức từ các tài liệu chuyên ngành và sự đoàn kết, cố gắng không ngừng của các thành viên trong nhóm, đề tài đã được hoàn thiện theo đúng tiến độ và mục tiêu đề ra.

Chúng em hy vọng rằng đề tài này không chỉ là sản phẩm học thuật phục vụ cho việc đánh giá kết quả học tập, mà còn là bước khởi đầu cho những nghiên cứu chuyên sâu hơn về AI trong tương lai. Rất mong nhận được sự đánh giá, góp ý và phản biện từ quý Thầy/Cô để nhóm có thể hoàn thiện hơn và tiếp tục phát triển kỹ năng nghiên cứu của mình trong những bước đường tiếp theo.

Xin trân trọng cảm ơn!

**BẢNG PHÂN CÔNG CÔNG VIỆC**

|  |  |  |  |
| --- | --- | --- | --- |
| **MSSV** | **Họ và tên** | **Nội dung** | **Đánh giá** |
| 2001221088 | Nguyễn Hồng Hà | Soạn nội dung, code | 100 |
| 2001225449 | Trịnh Võ Thùy Trâm | Soạn nội dung, code | 100 |

**MỤC LỤC**

[CHƯƠNG 1. GIỚI THIỆU CHUNG 8](#_Toc199427928)

[1.1. Giới thiệu đề tài 8](#_Toc199427929)

[1.2. Mục tiêu đề tài 8](#_Toc199427930)

[1.3. Đối tượng và phạm vi nghiên cứu 8](#_Toc199427931)

[1.4. Phương pháp nghiên cứu 9](#_Toc199427932)

[1.5. Ý nghĩa của đề tài 9](#_Toc199427933)

[CHƯƠNG 2. NỘI DUNG NGUYÊN CỨU 10](#_Toc199427934)

[2.1. Cảm xúc và nhận diện cảm xúc qua khuôn mặt 10](#_Toc199427935)

[2.2. Môi trường thực nghiệm và thư viện sử dụng 10](#_Toc199427936)

[2.3. Cài đặt và huấn luyện mô hình nhận diện cảm xúc 11](#_Toc199427937)

[2.3.1. Xử lý dữ liệu 11](#_Toc199427938)

[2.3.2. Mô hình CNN 11](#_Toc199427939)

[2.4. Kết quả 13](#_Toc199427940)

[2.5. Biểu đồ Loss theo Epoch 13](#_Toc199427941)

[2.6. Giao diện và chức năng ứng dụng nhận diện cảm xúc 14](#_Toc199427942)

[CHƯƠNG 3. KẾT LUẬN 16](#_Toc199427943)

[TÀI LIỆU THAM KHẢO 17](#_Toc199427944)

**DANH MỤC HÌNH ẢNH**

[Hình 1. Biều đồ huấn luyện Accurary / Loss 14](#_Toc199427876)

[Hình 2. Màn hình kết quả dự đoán 15](#_Toc199427877)

# GIỚI THIỆU CHUNG

Cảm xúc đóng vai trò quan trọng trong giao tiếp và hành vi con người. Việc nhận diện cảm xúc từ khuôn mặt đang trở thành một hướng nghiên cứu tiềm năng trong trí tuệ nhân tạo, đặc biệt nhờ sự phát triển của các mô hình học sâu như CNN. Chương này giới thiệu tổng quan về đề tài, mục tiêu nghiên cứu, phạm vi và phương pháp thực hiện.

## 1.1. Giới thiệu đề tài

Cảm xúc là một yếu tố cốt lõi trong giao tiếp con người, phản ánh trạng thái tinh thần và định hướng hành vi. Trong thời đại chuyển đổi số, việc hiểu và phân tích cảm xúc qua các phương tiện kỹ thuật số – đặc biệt là khuôn mặt – đóng vai trò quan trọng trong nhiều lĩnh vực như chăm sóc sức khỏe tâm thần, giáo dục thông minh, marketing, và hệ thống tương tác người–máy. Nhận diện cảm xúc từ khuôn mặt là một hướng nghiên cứu liên ngành kết hợp giữa tâm lý học, khoa học thần kinh và trí tuệ nhân tạo.

Nhu cầu tăng cao về các hệ thống nhận biết cảm xúc để nâng cao trải nghiệm người dùng trong các hệ thống tự động (ứng dụng học tập cá nhân hoá, hỗ trợ điều trị tâm lý…). Đề tài này tập trung vào việc xây dựng mô hình học sâu CNN để nhận diện cảm xúc khuôn mặt từ ảnh tĩnh, hướng đến mục tiêu tạo nên một mô hình có độ chính xác cao, có thể tích hợp vào các hệ thống thông minh thực tế.

Mô hình CNN hiện nay là một trong những phương pháp hiệu quả nhất trong thị giác máy tính, đặc biệt là trong nhận dạng khuôn mặt và phân tích biểu cảm. Tính khả thi về mặt kỹ thuật và tài nguyên: có thể tiếp cận các bộ dữ liệu chuẩn (FER-2013, CK+, AffectNet) và triển khai bằng các framework mã nguồn mở (TensorFlow, Keras…).

## 1.2. Mục tiêu đề tài

* Tìm hiểu các lý thuyết liên quan đến cảm xúc con người và cách biểu hiện qua khuôn mặt.
* Khảo sát và phân tích các kỹ thuật nhận diện cảm xúc hiện nay.
* Xây dựng hệ thống nhận diện cảm xúc dựa trên hình ảnh khuôn mặt sử dụng mô hình CNN.
* Thử nghiệm hệ thống trên tập dữ liệu chuẩn FER2013 để đánh giá độ chính xác của mô hình.
* Đề xuất các hướng cải tiến và ứng dụng thực tế của mô hình.

## 1.3. Đối tượng và phạm vi nghiên cứu

* Đối tượng nghiên cứu: Hình ảnh khuôn mặt.
* Phạm vi:
* Nghiên cứu lý thuyết về biểu hiện cảm xúc cơ bản (Ekman) và biểu hiện qua cơ mặt.
* Ứng dụng mạng CNN trong phân loại cảm xúc từ hình ảnh tĩnh.
* Sử dụng dữ liệu hình ảnh từ các bộ dữ liệu có sẵn như FER2013.

## 1.4. Phương pháp nghiên cứu

* Thu thập dữ liệu: sử dụng bộ dữ liệu FER-2013 – tập dữ liệu tiêu chuẩn bao gồm 35.887 ảnh xám kích thước 48x48 pixels, đã được gán nhãn với 7 cảm xúc cơ bản.
* Tiền xử lý dữ liệu: cân bằng dữ liệu, chuẩn hóa ảnh, tăng cường dữ liệu bằng xoay, lật, thay đổi độ sáng.
* Xây dựng mô hình CNN: thử nghiệm với mô hình CNN xác định hiệu năng tối ưu.
* Huấn luyện và đánh giá: dùng cross-validation, theo dõi các chỉ số như accuracy, F1-score, confusion matrix.
* Công cụ triển khai: Python, TensorFlow/Keras, Google Colab.

## 1.5. Ý nghĩa của đề tài

* Về lý thuyết: góp phần nghiên cứu khả năng áp dụng CNN trong bài toán phân tích cảm xúc khuôn mặt – một trong những bài toán quan trọng của thị giác máy tính.
* Về thực tiễn**:** mô hình xây dựng có thể mở rộng tích hợp vào các hệ thống giao tiếp người–máy, chẩn đoán tâm lý, hoặc hỗ trợ giáo dục cảm xúc.
* Về kỹ năng học thuật**:** giúp người thực hiện nắm vững kỹ thuật học sâu hiện đại và kỹ năng triển khai mô hình AI thực tế.

# NỘI DUNG NGUYÊN CỨU

Để xây dựng một hệ thống nhận diện cảm xúc hiệu quả từ khuôn mặt, việc hiểu rõ bản chất của cảm xúc con người cũng như các nguyên lý hoạt động của mô hình học sâu là điều cần thiết. Chương này trình bày các kiến thức nền tảng phục vụ cho việc phát triển mô hình, bao gồm khái niệm về cảm xúc, cách biểu hiện cảm xúc qua khuôn mặt, và các cảm xúc cơ bản theo lý thuyết của Paul Ekman. Đồng thời, chương cũng giới thiệu tổng quan về mạng nơ-ron tích chập (CNN) – công nghệ cốt lõi trong bài toán nhận diện hình ảnh, từ đó làm nền tảng cho các chương tiếp theo trong quá trình thiết kế và triển khai mô hình nhận diện cảm xúc.

## 2.1. Cảm xúc và nhận diện cảm xúc qua khuôn mặt

Cảm xúc là phản ứng tâm lý sinh học của con người, thường thể hiện qua nét mặt. Nhận diện cảm xúc qua khuôn mặt là một bài toán quan trọng trong lĩnh vực trí tuệ nhân tạo (AI), giúp máy móc tương tác tự nhiên hơn với con người.

Theo lý thuyết của Paul Ekman, có 7 cảm xúc cơ bản có thể biểu hiện rõ qua khuôn mặt:

* Angry (Giận dữ)
* Disgust (Ghê tởm)
* Fear (Sợ hãi)
* Happy (Hạnh phúc)
* Sad (Buồn)
* Surprise (Ngạc nhiên)
* Neutral (Bình thường)

Để mô hình AI nhận diện cảm xúc hiệu quả, cần xử lý chính xác các biểu hiện khuôn mặt, đặc biệt ở các vùng mắt, lông mày, miệng. Một số thách thức trong bài toán này bao gồm biểu hiện mờ nhạt, micro-expression, điều kiện ảnh không lý tưởng và dữ liệu mất cân bằng.

## 2.2. Môi trường thực nghiệm và thư viện sử dụng

* Hệ điều hành: Windows 11 (64-bit)
* Python: Phiên bản 3.8
* Phần cứng: CPU Intel Core i5 hoặc AMD Ryzen 5, RAM tối thiểu 8GB
* Framework: TensorFlow 2.12 + Keras
* IDE: Visual Studio Code hoặc Visual Studio 2022

**Các thư viện chính:** Cài đặt nhanh qua pip: pip install tensorflow opencv-python scikit-learn matplotlib seaborn.

|  |  |
| --- | --- |
| **Thư viện** | **Vai trò** |
| tensorflow | Xây dựng mô hình học sâu |
| opencv-python | Đọc và xử lý ảnh đầu vào |
| Sklearn | Đánh giá mô hình, chia dữ liệu |
| matplotlib, seaborn | Vẽ biểu đồ và biểu diễn kết quả |

## 2.3. Cài đặt và huấn luyện mô hình nhận diện cảm xúc

### Xử lý dữ liệu

* Ảnh được đọc từ thư mục theo từng cảm xúc.
* Resize về kích thước chuẩn 48x48, chuyển sang grayscale và chuẩn hóa giá trị pixel về [0,1].
* Dữ liệu được chia: Train (90%), Validation (10%), và Test riêng biệt.

### Mô hình CNN

Mô hình gồm 3 khối tích chập (Conv2D), mỗi khối gồm:

* Convolution → BatchNormalization → MaxPooling → Dropout
* Cuối cùng là các lớp fully-connected và softmax để phân loại 7 cảm xúc.

#### Khối tích chập (Convolutional Block)

Mỗi khối tích chập bao gồm bốn thành phần chính:

* Lớp tích chập (Convolution): Đây là bước quan trọng nhất, nơi mà mô hình dùng nhiều bộ lọc nhỏ để “quét” qua ảnh và phát hiện các đặc trưng như cạnh, góc, hoặc các mẫu phức tạp hơn. Số lượng bộ lọc tăng dần qua các khối, giúp mô hình học được từ các đặc trưng cơ bản đến phức tạp hơn.
* Chuẩn hóa theo batch (Batch Normalization): Sau khi có được kết quả từ lớp tích chập, dữ liệu được chuẩn hóa để ổn định quá trình huấn luyện, giúp mô hình học nhanh hơn và tránh bị lệ thuộc quá mức vào giá trị đầu ra ở từng bước.
* Phân vùng lấy mẫu lớn nhất (Max Pooling): Giảm kích thước không gian của ảnh đầu ra bằng cách lấy giá trị lớn nhất trong vùng nhỏ, giúp giảm số lượng tham số, tăng khả năng tổng quát và giảm chi phí tính toán.
* Giảm quá khớp bằng Dropout: Một số neuron trong lớp được tắt ngẫu nhiên khi huấn luyện nhằm tránh việc mô hình học quá kỹ các đặc trưng trong dữ liệu huấn luyện, giúp mô hình tổng quát tốt hơn với dữ liệu mới.

#### Các khối tích chập xếp chồng

Ba khối tích chập được xếp chồng lên nhau, mỗi khối tăng dần số bộ lọc để có thể nhận diện các đặc trưng ngày càng phức tạp. Khối đầu tiên học các đặc trưng cơ bản như cạnh và hình dạng đơn giản, trong khi các khối tiếp theo học các mẫu phức tạp hơn, giúp mô hình phân biệt tốt các biểu cảm khuôn mặt khác nhau.

#### Lớp kết nối đầy đủ (Fully Connected Layers)

Sau khi các khối tích chập hoàn tất việc trích xuất đặc trưng, các dữ liệu đa chiều được chuyển thành một vector một chiều (flatten). Vector này được đưa vào các lớp kết nối đầy đủ, nơi mô hình học được các mối quan hệ phi tuyến tính giữa các đặc trưng đã trích xuất.

* Một lớp ẩn với một số lượng neuron nhất định giúp mô hình học các kết hợp phức tạp hơn của đặc trưng.
* Tiếp theo là một lớp chuẩn hóa để ổn định kết quả và một lớp Dropout mạnh hơn nhằm tiếp tục giảm nguy cơ quá khớp.

#### Lớp đầu ra với Softmax

Cuối cùng, mô hình có một lớp đầu ra với số lượng neuron bằng số lớp cần phân loại (trong trường hợp này là 7 cảm xúc). Lớp này sử dụng hàm Softmax để chuyển đổi kết quả thành xác suất cho từng lớp, giúp mô hình đưa ra dự đoán xác suất mỗi cảm xúc đúng nhất.

#### Tăng cường dữ liệu (Augmentation)

Sử dụng ImageDataGenerator để xoay, dịch chuyển, zoom, và lật ảnh, giúp mô hình học tốt hơn và tránh overfitting.

#### Lịch trình giảm learning rate (Cosine Annealing)

Hàm học được điều chỉnh theo lịch trình cosine annealing nhằm tối ưu tốc độ hội tụ.

#### Callback

* ModelCheckpoint: lưu mô hình tốt nhất
* EarlyStopping: dừng sớm nếu không cải thiện
* ReduceLROnPlateau: giảm tốc độ học nếu không cải thiện
* LearningRateScheduler: sử dụng hàm cosine\_annealing

## 2.4. Kết quả

Sau khi huấn luyện, mô hình đạt độ chính xác Validation Accuracy được in ra cùng với biểu đồ loss/accuracy qua từng epoch. Mô hình được lưu tại best\_deep\_model\_2.h5 và có thể dùng để nhận diện cảm xúc từ ảnh mới.

* Đường màu xanh (Train Accuracy): thể hiện độ chính xác của mô hình trên tập huấn luyện theo từng epoch.
* Đường màu cam (Validation Accuracy): thể hiện độ chính xác trên tập kiểm tra/validation.

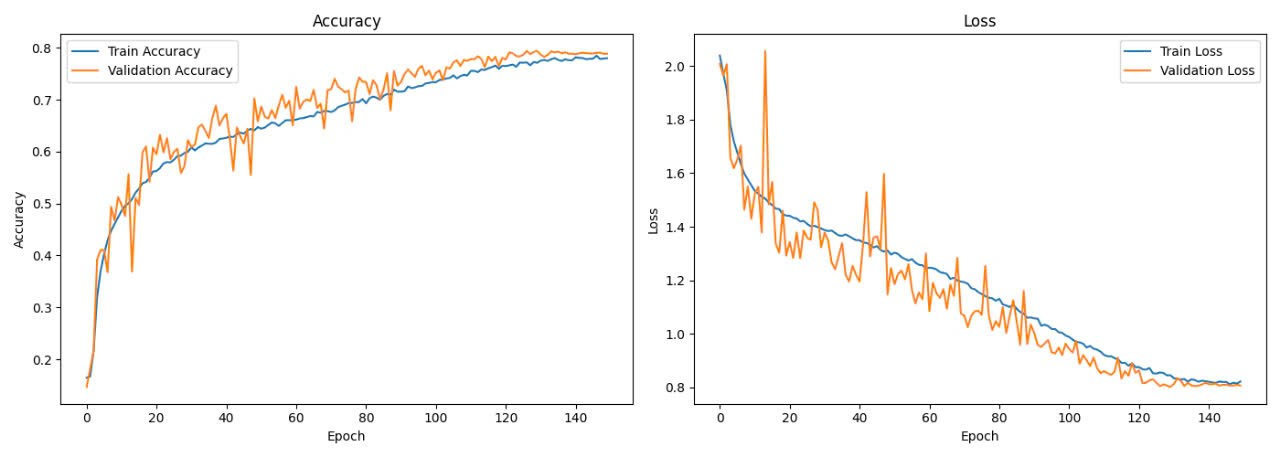
Nhận xét:

* Cả hai đường đều tăng dần → mô hình học được thông tin tốt hơn theo thời gian.
* Đường validation accuracy dao động mạnh hơn và có những điểm rớt đột ngột → cho thấy tập validation có độ biến thiên lớn hoặc dữ liệu không ổn định (có thể do augmentation, dữ liệu không đều, hoặc batch size nhỏ).
* Không có dấu hiệu rõ ràng của overfitting (khi đường train accuracy tăng mà validation accuracy giảm).

## 2.5. Biểu đồ Loss theo Epoch

Đường màu xanh (Train Loss): tổn thất của mô hình trên tập huấn luyện.

Đường màu cam (Validation Loss): tổn thất trên tập kiểm tra/validation.



Hình 1. Biều đồ huấn luyện Accurary / Loss

Nhận xét:

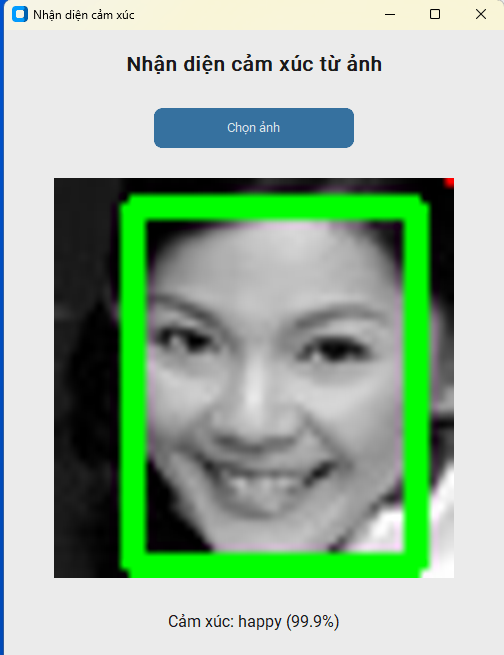
* Train loss giảm đều và ổn định → mô hình đang học hiệu quả.
* Validation loss dao động nhiều, có nhiều spike (đột biến tăng mạnh) → điều này thường do:
  + Dữ liệu kiểm tra có nhiều biến động hoặc không đồng nhất.
  + Batch size validation nhỏ → ảnh hưởng đến độ ổn định.
  + Regularization yếu hoặc learning rate chưa tối ưu.
* Mặc dù dao động nhưng nhìn chung validation loss có xu hướng giảm → dấu hiệu tích cực.

Tổng kết:

* Mô hình đang học tốt trên tập huấn luyện.
* Tập validation có dấu hiệu nhiễu, nhưng không gây overfitting rõ rệt.
* Để cải thiện:
  + Tăng batch size của validation.
  + Cân bằng hoặc lọc lại dữ liệu validation.
  + Xem xét giảm learning rate hoặc thêm regularization (dropout, weight decay).
  + Dùng kỹ thuật như EarlyStopping để tránh train quá dài khi validation không cải thiện.

## 2.6. Giao diện và chức năng ứng dụng nhận diện cảm xúc

* Nhằm nâng cao tính ứng dụng thực tiễn và trải nghiệm người dùng, nhóm đã xây dựng một giao diện đồ họa (GUI) đơn giản, hỗ trợ tải ảnh tĩnh từ máy tính và nhận diện cảm xúc.



Hình 2. Màn hình kết quả dự đoán

#### Chức năng chính của ứng dụng:

* Giao diện web dễ sử dụng (không cần dòng lệnh)
* Cho phép chọn ảnh từ thiết bị cá nhân
* Dự đoán và hiển thị cảm xúc tương ứng bằng mô hình CNN đã huấn luyện

# KẾT LUẬN

Hệ thống nhận diện cảm xúc gương mặt dựa trên mô hình mạng nơ-ron tích chập (CNN) đã được phát triển và thử nghiệm. Qua quá trình huấn luyện và đánh giá, mô hình đạt được độ chính xác cao trong việc phân loại các trạng thái cảm xúc cơ bản như vui, buồn, giận dữ, sợ hãi, ngạc nhiên và trung tính.

Các kỹ thuật tăng cường dữ liệu (data augmentation) như xoay, dịch chuyển, lật ảnh, điều chỉnh độ sáng… đã giúp làm phong phú bộ dữ liệu huấn luyện, từ đó cải thiện khả năng tổng quát hóa của mô hình khi gặp các hình ảnh thực tế khác nhau. Đồng thời, việc sử dụng các phương pháp tối ưu như dropout, batch normalization, cùng với điều chỉnh tham số học giúp giảm thiểu hiện tượng overfitting, tăng độ tin cậy của hệ thống khi hoạt động trong môi trường thực tế.

Mô hình CNN chứng tỏ tính khả thi cao trong việc nhận diện cảm xúc gương mặt với độ chính xác ổn định, kể cả trong điều kiện ánh sáng và góc quay không hoàn hảo. Điều này mở ra nhiều khả năng ứng dụng trong các lĩnh vực thực tế đa dạng.

Trong lĩnh vực chăm sóc sức khỏe tâm thần, hệ thống có thể hỗ trợ theo dõi trạng thái cảm xúc của bệnh nhân, góp phần vào việc chẩn đoán và điều trị hiệu quả hơn. Trong giáo dục, công nghệ này giúp giáo viên nhận biết cảm xúc học sinh, từ đó điều chỉnh phương pháp giảng dạy phù hợp. Trong ngành marketing và bán lẻ, việc phân tích cảm xúc khách hàng có thể nâng cao trải nghiệm người dùng và hiệu quả quảng cáo. Ngoài ra, các ứng dụng trong an ninh, giao tiếp người-máy, hoặc các thiết bị thông minh cũng có thể tận dụng công nghệ nhận diện cảm xúc để tương tác tự nhiên và linh hoạt hơn.

Mặc dù đã đạt được kết quả tích cực, hệ thống vẫn còn một số hạn chế như độ chính xác giảm khi gặp biểu cảm phức tạp, ánh sáng yếu hoặc khuôn mặt bị che khuất. Vì vậy, cần tiếp tục cải tiến mô hình bằng cách mở rộng dữ liệu huấn luyện, áp dụng các kiến trúc mạng sâu hơn, hoặc kết hợp thêm các cảm biến khác như âm thanh, chuyển động để nâng cao khả năng nhận diện toàn diện hơn.

Tóm lại, công nghệ nhận diện cảm xúc gương mặt sử dụng mô hình CNN không chỉ có tính khả thi cao mà còn có tiềm năng ứng dụng thực tiễn rộng lớn, góp phần tạo nên các sản phẩm, dịch vụ thông minh và thân thiện hơn với người dùng trong nhiều lĩnh vực của cuộc sống.

# TÀI LIỆU THAM KHẢO

1. Ekman, P. (1992). An Argument for Basic Emotions. Cognition & Emotion.
2. LeCun, Y., et al. (1998). Gradient-Based Learning Applied to Document Recognition. Proceedings of the IEEE.
3. 3. Goodfellow, I., Bengio, Y., Courville, A. (2016). Deep Learning. MIT Press.
4. Li, S. Z., Jain, A. (2011). Handbook of Face Recognition. Springer.
5. Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2017). AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. IEEE Transactions on Affective Computing.
6. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. CVPR.
7. Giáo trình Học sâu (Deep Learning), PGS. TS. Trần Minh Triết - Đại học Quốc gia TP.HCM.
8. Priyanka, V. T., Reddy, Y. R., Vajja, D., Ramesh, G., & Gomathy, S. (2023, May). A novel emotion based music recommendation system using CNN. In 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 592-596). IEEE.
9. Sana, S. K., Sruthi, G., Suresh, D., Rajesh, G., & Reddy, G. S. (2022). Facial emotion recognition based music system using convolutional neural networks. Materials Today: Proceedings, 62, 4699-4706.
10. Akhand, M. A. H., Roy, S., Siddique, N., Kamal, M. A. S., & Shimamura, T. (2021). Facial emotion recognition using transfer learning in the deep CNN. Electronics, 10(9), 1036.
11. Altaher, A., Salekshahrezaee, Z., Abdollah Zadeh, A., Rafieipour, H., & Altaher, A. (2020). Using multi-inception CNN for face emotion recognition. Journal of Bioengineering Research, 3(1), 1-12.