

TÌM HIỂU VỀ MÔ HÌNH CNN TRONG THỊ GIÁC MÁY TÍNH VÀ ỨNG DỤNG TRONG NHẬN DIỆN KHUÔN MẶT

GVHD: Đoàn Phước Miền

Sinh viên: Trần Quốc Thắng
Mssv:110123048
Lớp: DA23TTA

MỤC LỤC

- 1.MỤC ĐÍCH NGHIÊN CỨU
- 2.CƠ SỞ LÍ THUYẾT KIẾN TRÚC CNN
- 3.QUY TRÌNH THỰC HIỆN HÓA HỆ THỐNG
- 4.TIỀN XỬ LÝ DỮ LIỆU
- 5.TĂNG CƯỜNG DỮ LIỆU
- 6.HUẤN LUYỆN MÔ HÌNH CNN
- 7.TRIỂN KHAI NHẬN DIỆN QUA WEBCAM
- 8.HIỆU NĂNG HỆ THỐNG
- 9.KẾT QUẢ ĐẠT ĐƯỢC
- 10.KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

1. MỤC ĐÍCH NGHIÊN CỨU



NGHIÊN CỨU LÍ THUYẾT

Tìm hiểu chuyên sâu về kiến trúc, nguyên lý vận hành của mạng nơ-ron tích chập (CNN) và cách trích xuất đặc trưng.

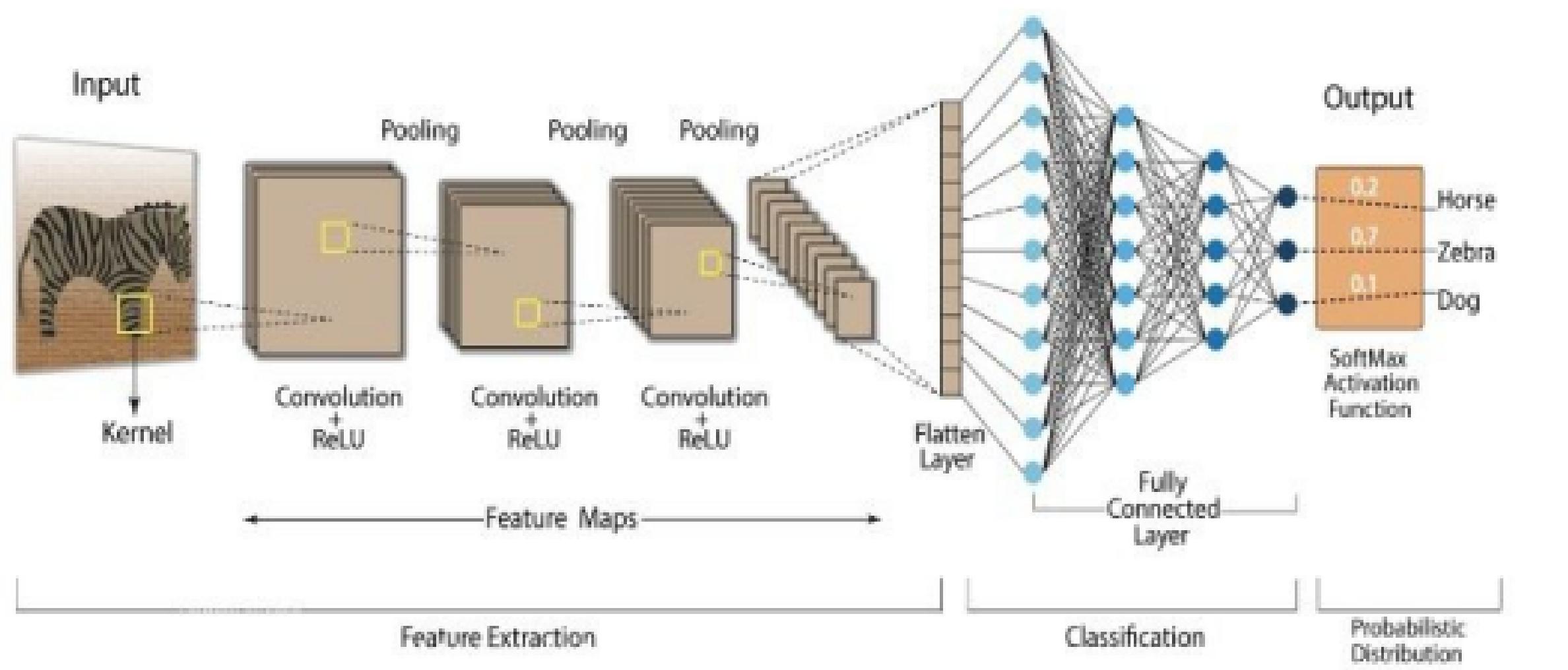
PHÂN TÍCH KIẾN TRÚC VÀ CƠ CHẾ HOẠT ĐỘNG CỦA CNN

Phân tích cấu trúc các tầng: Tích chập (Convolution), Kích hoạt (Activation), Gộp (Pooling) và Kết nối đầy đủ (Fully Connected).
Làm rõ cơ chế tự động trích xuất đặc trưng (feature extraction) thay vì thiết kế thủ công.

TRIỂN KHAI ỨNG DỤNG TRONG THỰC TẾ

Hiện thực hóa mô hình đã huấn luyện vào hệ thống nhận diện khuôn mặt.
Xây dựng chức năng nhận diện thời gian thực (real-time) thông qua webcam.

2. CƠ SỞ LÍ THUYẾT-KIẾN TRÚC CNN



Convolution (tích chập) Activation (ReLU):

Dùng bộ lọc (kernel) quét qua ảnh để tìm đặc trưng.

Activation (ReLU):

Loại bỏ giá trị âm, giữ lại thông tin quan trọng.

Pooling (Gộp):

Giảm kích thước dữ liệu nhưng giữ lại đặc trưng chính (Max Pooling).

Fully Connected:

Tổng hợp thông tin để phân loại.

3. QUY TRÌNH THỰC HIỆN HÓA HỆ THỐNG



Thu thập dữ liệu --> Tiền xử lý --> Tăng
cường dữ liệu --> Huấn luyện CNN -->
Kiểm thử --> Triển khai Webcam

4. TIỀN XỬ LÍ DỮ LIỆU

Phát hiện khuôn mặt: Sử dụng thuật toán Haar Cascade để xác định vị trí khuôn mặt trong ảnh gốc.

Cắt & Chuẩn hóa:

- Cắt vùng chứa khuôn mặt để loại bỏ nhiễu nền.
- Resize ảnh về kích thước cố định 100×100 pixel.
- Chuyển đổi sang ảnh xám (Grayscale) để giảm khối lượng tính toán.

Chuẩn hóa giá trị: Đưa giá trị pixel về khoảng $[0, 1]$ (chia cho 255) giúp mô hình hội tụ nhanh hơn

5. TĂNG CƯỜNG DỮ LIỆU

Số lượng ảnh thu thập thực tế thường ít, dễ dẫn đến hiện tượng học vẹt (Overfitting).

Giải pháp: Sử dụng kỹ thuật sinh dữ liệu mới từ ảnh gốc thông qua thư viện ImageDataGenerator.

Các kỹ thuật áp dụng:

- Xoay ảnh: Góc xoay ngẫu nhiên 25 độ
- Phóng to/Thu nhỏ: Tỷ lệ 30%
- Dịch chuyển: Dịch ngang/dọc 20%.
- Lật ảnh: Lật theo trục ngang

6. HUẤN LUYỆN MÔ HÌNH CNN



MÔI TRƯỜNG HUẤN LUYỆN

- Nền tảng: Google Colab.
- Phần cứng: Tận dụng GPU miễn phí của Google để tăng tốc độ tính toán các phép tích chập.
- Thư viện chính: TensorFlow, Keras, OpenCV.

THIẾT LẬP THAM SỐ

- Kích thước lô (Batch size): 32 (Số ảnh được xử lý trong một lần cập nhật trọng số).
- Số vòng lặp (Epochs): 30 – 50 vòng (Đảm bảo mô hình đủ thời gian hội tụ).
- Phân chia dữ liệu (Validation Split):
 - + 80% dữ liệu dùng để huấn luyện (Training set).
 - + 20% dữ liệu dùng để kiểm thử chéo (Validation set) nhằm đánh giá khách quan ngay trong quá trình train.

7.TRIỂN KHAI NHẬN DIỆN QUA WEBCAM(1)

Bước 1 - Thu nhận:

Đọc luồng video liên tục từ Webcam thông qua thư viện OpenCV

Bước 2 - Phát hiện:

Sử dụng Haar Cascade để tìm vị trí khuôn mặt (nhanh, nhẹ) trước khi đưa vào mô hình.

Bước 3 - Chuẩn hóa:

Cắt vùng mặt (Crop), đưa về kích thước 100x100, và chia cho 255 để khớp với đầu vào của CNN.

Bước 4 - Dự đoán:

Mô hình CNN trả về vector xác suất cho từng nhãn

7. TRIỂN KHAI NHẬN DIỆN QUA WEBCAM(2)

Để tránh nhận nhầm người không có trong dữ liệu, hệ thống áp dụng bộ lọc kép:

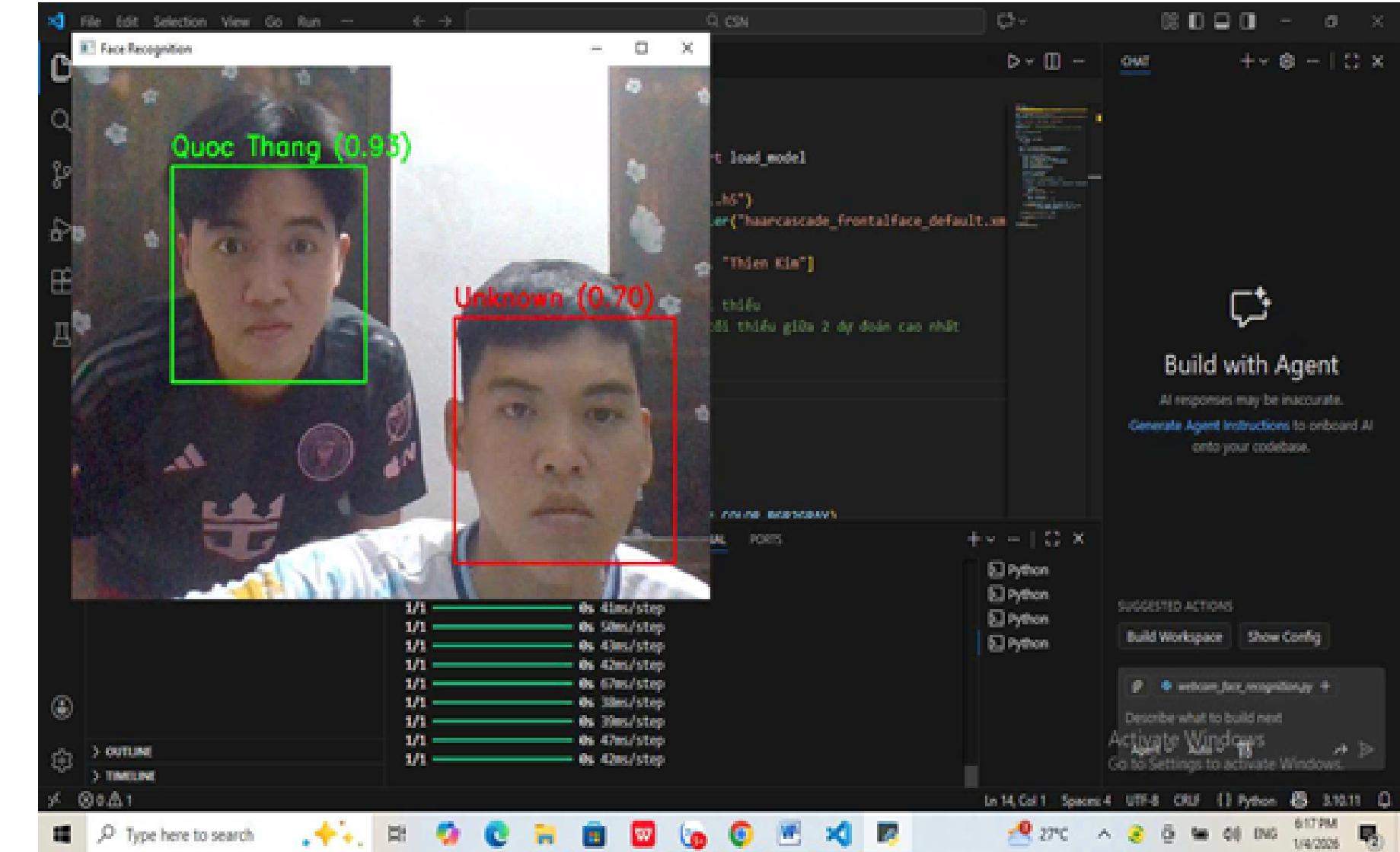
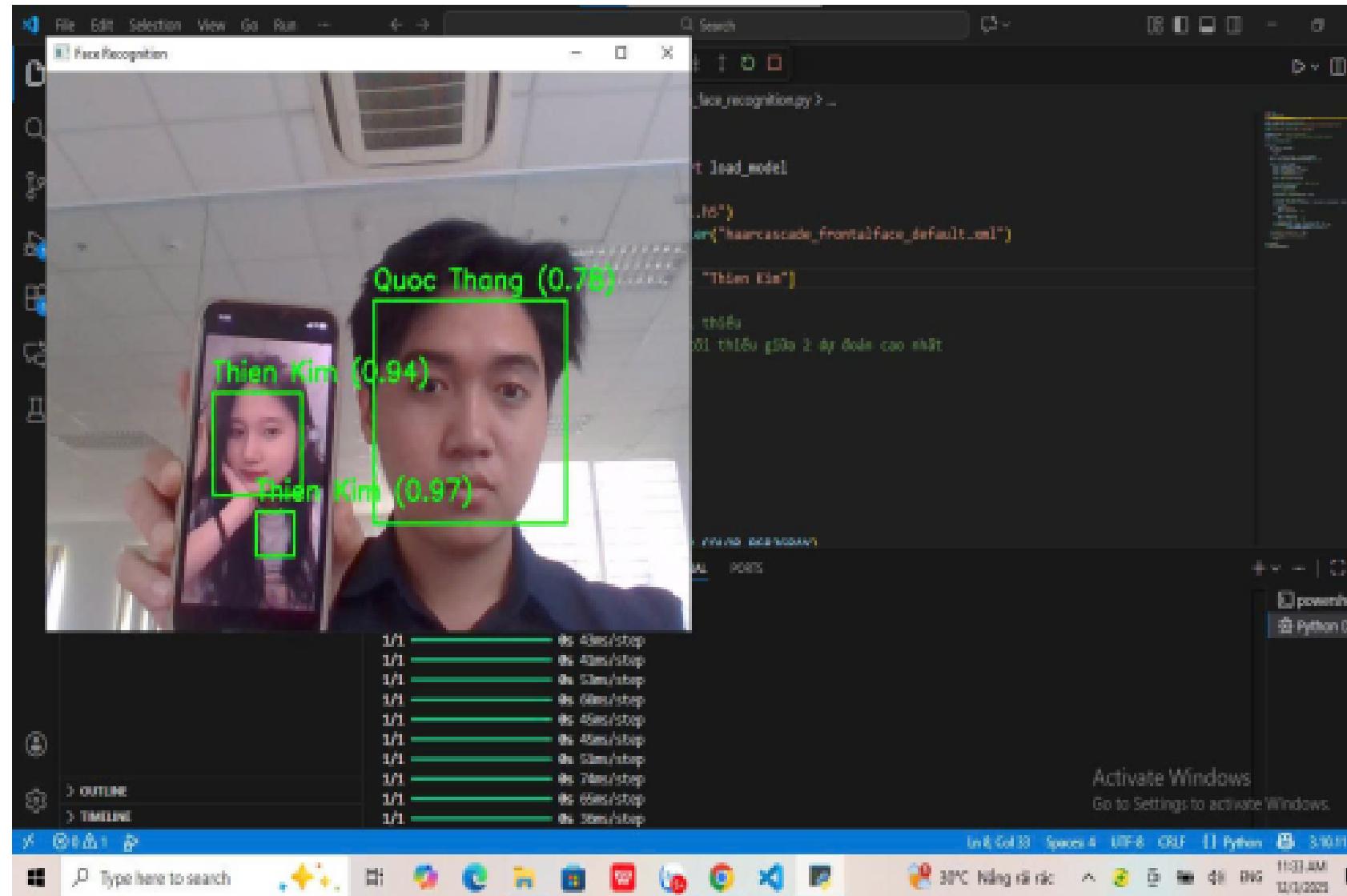
- Ngưỡng tin cậy (Threshold > 0.75): Xác suất dự đoán cao nhất phải lớn hơn 75%.
 - Độ tách biệt (Margin > 0.15): Chênh lệch giữa xác suất cao nhất và người thứ nhì phải lớn hơn 15% (để tránh sự lưỡng lự).
- > Nếu không thỏa mãn cả hai điều kiện: Kết luận là "Unknown".

8.HIỆU NĂNG HỆ THỐNG

- Tốc độ: Đạt 20 – 30 FPS (Khung hình/giây).
- Độ trễ: Khoảng 40 – 70 ms cho mỗi khuôn mặt được xử lý.

| | | |
|-----|----|------------|
| 1/1 | 85 | 58ms/step |
| 1/1 | 85 | 55ms/step |
| 1/1 | 85 | 45ms/step |
| 1/1 | 85 | 72ms/step |
| 1/1 | 85 | 156ms/step |
| 1/1 | 85 | 54ms/step |
| 1/1 | 85 | 66ms/step |
| 1/1 | 85 | 90ms/step |
| 1/1 | 85 | 93ms/step |
| 1/1 | 85 | 77ms/step |
| 1/1 | 85 | 71ms/step |

9. KẾT QUẢ ĐẠT ĐƯỢC



Kết quả cho thấy:

- + Hệ thống hạn chế hiện tượng nhận nhầm người lạ thành người quen.
- + Khi xuất hiện khuôn mặt không thuộc tập huấn luyện, hệ thống phản hồi đúng bằng cách hiển thị nhãn “Unknown”.

10. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN (1)

- Hệ thống hoàn chỉnh: Xây dựng thành công quy trình từ thu thập dữ liệu đến nhận diện thời gian thực.
- Hiệu năng: Nhận diện chính xác, phân biệt tốt Người lạ (Unknown) và đạt tốc độ xử lý nhanh (Real-time).
- Tính khả thi: Chứng minh hiệu quả vượt trội của CNN so với các phương pháp truyền thống.

10. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN (2)

- Bảo mật: Tích hợp Liveness Detection (Chống giả mạo bằng ảnh/video).
- Tối ưu hóa: Triển khai trên thiết bị nhúng/di động cấu hình thấp.
- Mở rộng: Phát triển cơ chế học liên tục (thêm người mới không cần train lại từ đầu).

Champ
Year

