

BÁO CÁO TÌM HIỂU ĐỀ TÀI

# NHẬN DIỆN KHUÔN MẶT

Face Detection & Face Recognition



Thực hiện: Phạm Đình Thục

**Link Github:** <https://github.com/phamdinhtthuc020100/Project-FaceRecognizer.git>

Tháng 11 năm 2021

## Mục lục

<b>A. LEVEL 01</b>	<b>3</b>
<b>I. Đặt vấn đề</b>	<b>3</b>
<b>II. Mục đích tìm hiểu</b>	<b>3</b>
<b>III. Quá trình tìm hiểu</b>	<b>3</b>
<b>IV. Nội dung tìm hiểu</b>	<b>3</b>
<b>B. LEVEL 02</b>	<b>5</b>
<b>I. Đặt vấn đề</b>	<b>5</b>
<b>II. Mục đích tìm hiểu</b>	<b>6</b>
<b>III. Quá trình tìm hiểu</b>	<b>6</b>
<b>IV. Nội dung tìm hiểu</b>	<b>6</b>
<b>C. Tài liệu tham khảo</b>	<b>11</b>

## A. LEVEL 01

### I. Đặt vấn đề

Vấn đề nhận dạng khuôn mặt hiện nay đang rất được quan tâm, nó được sử dụng cho nhiều mục đích như ngăn ngừa tội phạm, bảo vệ dữ liệu và thông tin cá nhân thông qua sinh trắc học

Hiện nay, thuật toán nhận dạng đã được các nền tảng lớn sử dụng thường xuyên hơn, như Facebook đã dùng thuật toán nhận dạng khuôn mặt để gắn thẻ bạn bè, người thân trên các tấm ảnh người dùng upload với độ chính xác 98% [1]

### II. Mục đích tìm hiểu

- Nắm được các lý thuyết cơ bản của nhận diện khuôn mặt
- Tìm hiểu mức cơ bản quá trình nhận diện khuôn mặt bằng webcam
- Demo mức cơ bản

### III. Quá trình tìm hiểu

- Đọc, hiểu tài liệu liên quan đến Nhận diện khuôn mặt
- Áp dụng thư viện OpenCV để chạy mô hình nhận diện khuôn mặt
- Demo cơ bản bằng webcam trên laptop

### IV. Nội dung tìm hiểu

Hệ thống nhận dạng khuôn mặt là một ứng dụng máy tính tự động xác định hoặc nhận dạng một người nào đó từ một bức hình ảnh kỹ thuật số hoặc một khung hình video từ một nguồn video. Một trong những cách để thực hiện điều này là so sánh các đặc điểm khuôn mặt chọn trước từ hình ảnh và một cơ sở dữ liệu về khuôn mặt

Hiện nay có nhiều thuật toán, công cụ hỗ trợ cho huấn luyện, thực thi các mô hình nhận diện khuôn mặt. Có thể kể đến như: OpenCV, MTCNN, FaceNet, ...

Nhận diện khuôn mặt gồm có 2 công việc chính:

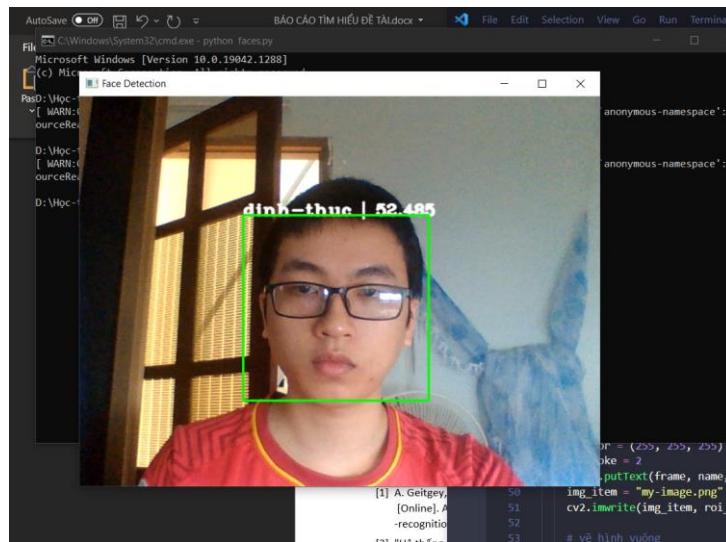
- Phát hiện khuôn mặt
- Nhận diện khuôn mặt đó

Các bước chính của nhận diện khuôn mặt

- Bước 1. Chuẩn bị một cơ sở dữ liệu để lưu trữ thông tin của mỗi người như: tên, tuổi, chức danh...
- Bước 2. Thực hiện chụp ảnh hoặc tải về ảnh có sẵn
- Bước 3. Train model để nhận diện
- Bước 4. Tiến hành nhận dạng khuôn mặt

### Ảnh Demo

- Ta demo trên tập dữ liệu nhỏ (5 người) cho ra kết quả

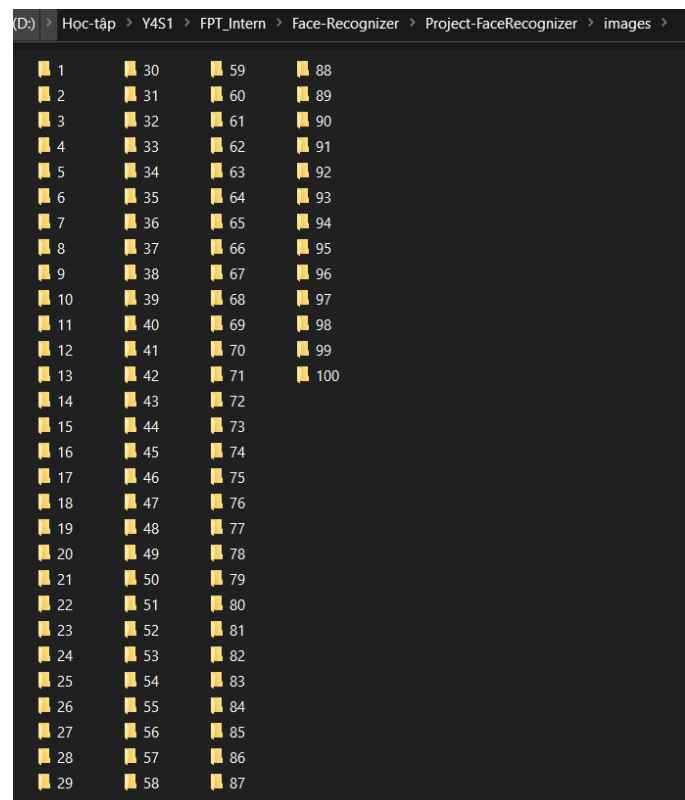


- Tiếp theo ta test trên bộ dữ liệu lớn hơn (khoảng 100 người)
  - o Dataset được chọn là: VN-celeb dataset với khoảng 23k khuôn mặt của hơn 1000 người Việt [2]. Ở đây ta dùng 99 tập của 99 người và 1 tập của cá nhân. Tổng cộng 100 tập

## Cấu trúc file

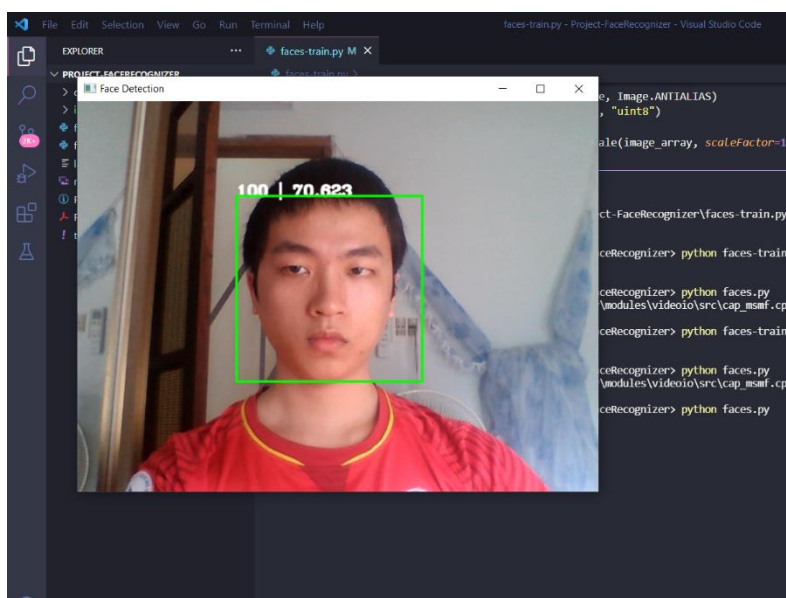
Folder 1 – 99: Người Việt nổi tiếng

Folder 100: đình thực

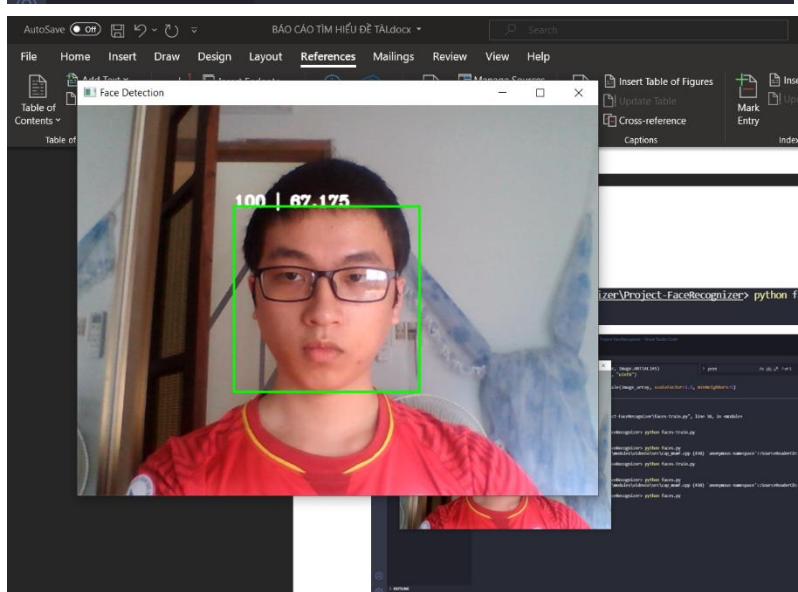


## Kết quả test

```
PS D:\Hoc-tap\Y4S1\FPT_Intern\Face-Recognizer\Project-FaceRecognizer> python faces-train.py
Loading ...
finished
```



(Không đeo kính)



(Có đeo kính)

## B. LEVEL 02

### I. Đặt vấn đề

Với tình hình dịch bệnh COVID-19 trên toàn cầu hiện nay, cùng với đó việc áp dụng các thuật toán nhận diện khuôn mặt truyền thống đang gặp các khó khăn như: người dân đeo khẩu trang, đeo kính và tuân thủ khoảng cách an toàn đang là tiêu chuẩn mới nhằm ngăn ngừa dịch bệnh. Điều này khiến khuôn mặt của họ bị che đi, không còn nhiều các đặc điểm đặc trưng để có thể nhận dạng nhanh chóng. NIST đã tiến hành các bài kiểm tra độ chính xác của nhận dạng khuôn mặt gần đây. Báo cáo của họ tiết lộ rằng các thuật toán FR chính thống đã tăng tỷ lệ lỗi từ 5% đến 50% trên khuôn mặt có đeo khẩu trang. [3]

Vì thế cần có một phương pháp mới để có thể nhận diện một người nào đó dựa vào một số ít các đặc trưng tiêu biểu

## II. Mục đích tìm hiểu

- Hiểu được các lý thuyết cơ bản về nhận dạng khuôn mặt cải tiến (khi đeo khẩu trang)
- Hiểu được cách các mô hình học máy, học sâu được áp dụng

## III. Quá trình tìm hiểu

- Đọc hiểu đề tài
- Tìm kiếm thông tin đề tài
- Remake source code (nếu có) để demo

## IV. Nội dung tìm hiểu

### a. Lý thuyết

- Các bước cho quá trình nhận dạng khuôn mặt cải tiến (khi đeo khẩu trang)
  - **Thu thập dữ liệu**  
Đầu tiên, ta cần thu thập hình ảnh liên quan đến khuôn mặt người khi đeo và không đeo khẩu trang  
Có thể thu thập bằng nhiều cách: tải các bộ dataset có sẵn, hoặc crawl trên web bằng các thư viện như Selenium hoặc BeautifulSoup
  - **Tiền xử lý**  
Ở bước này, ta tiến hành các thao tác xử lý ảnh (rotate, crop, ...) để đưa ảnh về kích thước phù hợp cho quá trình huấn luyện  
Để làm giàu dữ liệu ta có thể sử dụng các phương pháp tăng số lượng và chất lượng ảnh như: Mask the face, ...

### Phương pháp Mask the face [4]:

- Nó sử dụng một máy dò các cạnh khuôn mặt dựa trên dlib để xác định độ nghiêng của khuôn mặt và các đặc điểm chính của khuôn mặt cần thiết để đeo khẩu trang vào (khẩu trang ảo). Khẩu trang làm mẫu sẽ được biến đổi theo các hướng của khuôn mặt sao cho khớp nhất có thể. Tham khảo mô hình MaskTheFace

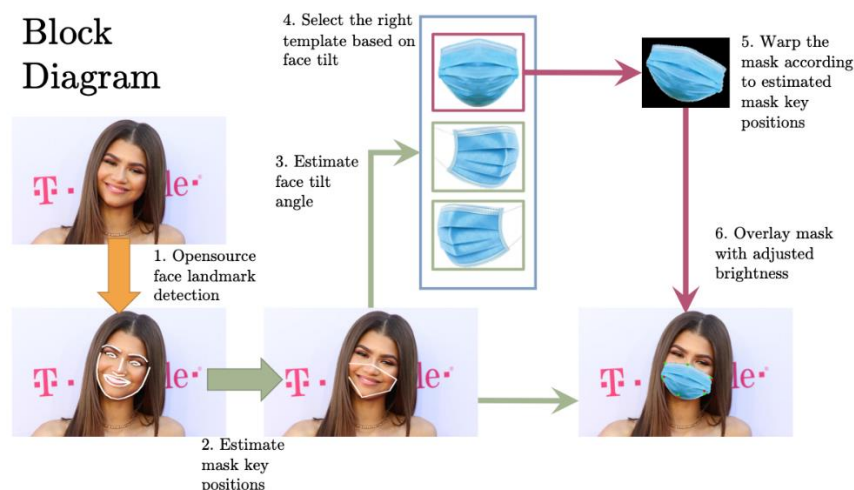






Figure 1: Dataset Generation: Original => Masked => Cropped

- Trong khi sử dụng dlib để tạo mặt nạ, chúng ta ghi lại vị trí của mặt nạ và tạo một binary segmentation maps ứng với từng ảnh để loại bỏ khẩu trang



Figure 2: Mask Segmentation

- Ở đây, tác giả dùng nhiều mẫu khẩu trang để làm đa dạng bộ dữ liệu, và thuật toán được điều chỉnh để có thể gắn khẩu trang cho nhiều người trên một khung hình

Ngoài ra ta có thể sử dụng bộ dữ liệu có sẵn: Masked faces in real world for face recognition (MFR2) với 53 danh tính của những người nổi tiếng và chính trị gia với tổng số 269 hình ảnh được thu thập từ internet, mỗi người có trung bình 5 ảnh.



### ○ Xây dựng mô hình

- Tác giả thử nghiệm trên 2 mô hình: **VGG** and **ResNet**. Kết quả cho ra VGG không thể hội tụ trên tập dữ liệu với các siêu tham số mà nhóm tác giả đã thử nghiệm. ResNet đã hoạt động tốt hơn. Vì thế, **ResNet50** đã được lựa chọn làm phương thức chính cho mô hình này
- Cụ thể, sẽ mở rộng ResNet50 với 1 lớp Dropout và dùng 1 lớp tuyến tính mới để tạo các bản nhúng 512 bit
- Lớp FC cuối, được cấu hình như bản sau

ID	Layers	In	Out	Kernel Size
0	Dropout			
1	AdaptiveAvgPool2d	2048	2048	$1 \times 1$
2	Flatten			
3	Linear	2048	512	
4	BatchNorm1d			

Table 1: Extension layers

### ○ Loss function

- Các hệ thống FR truyền thống như FaceNet sử dụng chức năng Triplet Loss. Ở đây tác giả đề xuất một Loss function mới được gọi là **Quintuplet Loss** để xử lý các hình ảnh có đeo và không đeo khẩu trang.
- Quintuplet Loss gồm 5 ảnh: Anchor (A), Positive (P), Negative (N), Masked Positive (Pm) and Masked Negative (Nm).
  - **Anchor** là hình ảnh được nhận dạng
  - **Positive** là hình ảnh không đeo khẩu trang khác của cùng một người
  - **Negative** là hình ảnh không đeo khẩu trang của một người khác
  - **Masked Positive** là hình ảnh đeo khẩu trang của cùng một người
  - **Masked Negative** là hình ảnh đeo khẩu trang của một người khác
  - $\alpha$  là margin

The Loss Function is:

$$\mathcal{L} = ||f(A) - f(P)||^2 + ||f(A) - f(P_m)||^2 - ||f(A) - f(N)||^2 - ||f(A) - f(N_m)||^2 + \alpha$$

Để tránh mô hình bị model collapsing, ta cần chọn Quintuplets phù hợp (không quá khó, không quá dễ). Nếu quá dễ dàng, chúng sẽ không giúp train neural network. Nếu quá khó, chúng có thể dẫn đến cực tiểu cục bộ hoặc model collapse. Để tối ưu, ta cần chọn Quintuplets với điều kiện:

$$0 < ||f(A) - f(N)||^2 - ||f(A) - f(P)||^2 < \alpha$$

$$0 < ||f(A) - f(N_m)||^2 - ||f(A) - f(P_m)||^2 < \alpha$$

### ○ Xóa khẩu trang và tái tạo hình ảnh (MRII)

- Thực hiện xóa khẩu trang và tái tạo hình ảnh trước khi đưa hình ảnh vào mô hình FR.
- Một pretrained generative model được tạo từ Edgeconnect để tháo khẩu trang của hình ảnh.



- Binary mask segmentation map sẽ được tạo cùng với cùng với hình ảnh trong quá trình mask the face. Ngoài ra ta cũng có thể dùng các thuật toán object detection và localization algorithm.

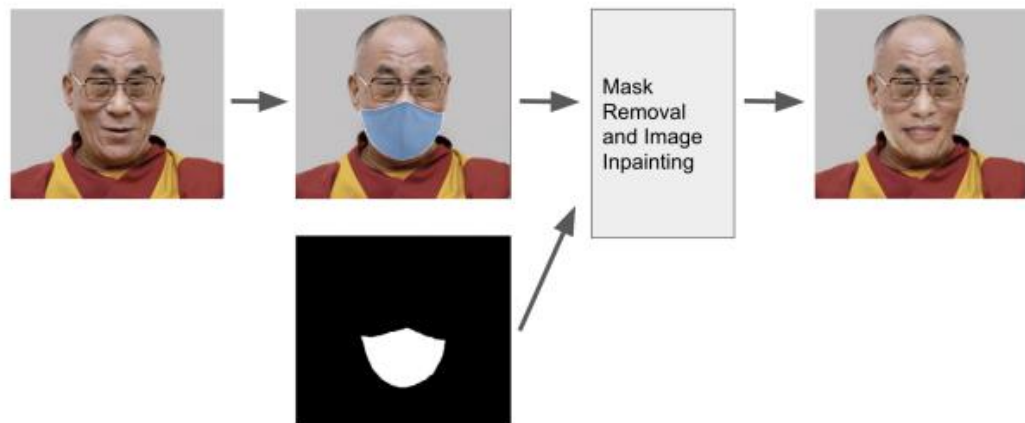


Figure 3: Mask removal and image inpainting

- Quá trình được diễn ra theo các bước
  - o Từ ảnh không gốc đeo (hoặc không) khẩu trang
  - o Thông qua quá trình mask the face (đối với ảnh không đeo khẩu trang) để tạo Binary mask segmentation map
  - o Đối với các ảnh đã đeo khẩu trang, dùng các thuật toán object detection và localization algorithm để tạo Binary mask segmentation map
  - o Tổng hợp hình ảnh đeo khẩu trang cùng với Binary mask segmentation map để gỡ bỏ khẩu trang và tái tạo lại khuôn mặt
  - o Cuối cùng, ta đưa kết quả đến FaceNet để dự đoán

#### o Kết quả

Model	FaceNet	MaskFaceNet	MRII
Accuracy	67.28%	75.06%	82.39%

Table 2: Evaluation

- Mô hình **FaceNet** hoạt động không tốt lắm vì nó chưa được đào tạo với hình ảnh khuôn mặt bị che.
- Mô hình **MaskFaceNet** hoạt động tốt hơn FaceNet, cho thấy quintuplet loss function giúp mô hình nhận ra hình ảnh khuôn mặt bị che. Độ chính xác tổng thể của mô hình không cao lắm vì phải sử dụng batch size nhỏ, bị giới hạn bởi bộ nhớ GPU và thời gian đào tạo cũng ít hơn nhiều so với các mô hình như FaceNet. Với nhiều tài nguyên máy tính hơn, MaskFaceNet sẽ có hiệu suất tốt hơn
- Mô hình **Mask removal and image inpainting (MRII)** cải thiện đáng kể hiệu suất của FaceNet trên tập dữ liệu khuôn mặt bị che. Nó cho thấy rằng một số đặc điểm trên

khuôn mặt được phục hồi bằng kỹ thuật inpainting có thể giúp hệ thống FR nhận ra hình ảnh khuôn mặt bị che

### ○ **Kết luận**

Trong đề tài này, em đã nghiên cứu các cách tiếp cận để cải thiện hiệu suất của hệ thống FR trên hình ảnh khuôn mặt của những người đeo khẩu trang. Kết quả đánh giá cho thấy các cách tiếp cận đều có thể cải thiện độ chính xác của hệ thống FR trên hình ảnh khuôn mặt được che mặt. Ta có thể đạt được sự cải thiện hơn nữa nếu có nhiều tài nguyên máy tính hơn hoặc đào tạo lại một số mô hình.

### **b. Thực hành**

- Soure FaceTheMask: <https://github.com/aeelanwar/MaskTheFace>
- Soure MaskFaceNet: <https://github.com/zhnli/MaskFaceNet>

## C. Tài liệu tham khảo

- [1] A. Geitgey, "Machine Learning is Fun! Part 4: Modern Face Recognition with Deep Learning," [Online]. Available: <https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78>.
- [2] "VN-celeb: Dữ liệu khuôn mặt người nổi tiếng Việt Nam và bài toán Face recognition," [Online]. Available: <https://viblo.asia/p/vn-celeb-du-lieu-khuon-mat-nguoi-noi-tieng-viet-nam-va-bai-toan-face-recognition>.
- [3] "Facial Recognition for People Wearing Masks," [Online]. Available: [http://cs230.stanford.edu/projects\\_fall\\_2020/reports/55790723.pdf](http://cs230.stanford.edu/projects_fall_2020/reports/55790723.pdf).
- [4] "MaskTheFace," [Online]. Available: <https://github.com/aaqelanzar/MaskTheFace>.
- [5] "Hệ thống nhận dạng khuôn mặt," [Online]. Available: [https://vi.wikipedia.org/wiki/H%E1%BB%87\\_th%E1%BB%91ng\\_nh%E1%BA%ADn\\_d%E1%BA%A1nh\\_khuon\\_m%E1%BA%B7t](https://vi.wikipedia.org/wiki/H%E1%BB%87_th%E1%BB%91ng_nh%E1%BA%ADn_d%E1%BA%A1nh_khuon_m%E1%BA%B7t).
- [6] "[Face Recognize] Thử làm hệ thống chấm công bằng nhận dạng khuôn mặt," [Online]. Available: <https://www.miai.vn/2019/08/13/face-recognize-thu-lam-he-thong-cham-cong-bang-nhan-dang-khuon-mat>.
- [7] "OpenCV Python TUTORIAL #4 for Face Recognition and Identification," [Online]. Available: [https://www.youtube.com/watch?v=PmZ29Vta7Vc&ab\\_channel=CodingEntrepreneurs](https://www.youtube.com/watch?v=PmZ29Vta7Vc&ab_channel=CodingEntrepreneurs).