

**ĐẠI HỌC QUỐC GIA TP.HCM**  
**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

**Báo cáo đồ án môn: Nhập môn học máy**

**Đề tài: Face Recognition**

**GIÁO VIÊN HƯỚNG DẪN: Bùi Tiến Lên**  
**Nguyễn Ngọc Đức**

**THÀNH VIÊN NHÓM: Nguyễn Hoàng Việt**  
**Nguyễn Vương Đạt Vũ**

**TP.HCM - Năm 2020**

- **Bố cục bài báo cáo**

Nội dung của bài gồm các nội dung chính sau đây:

**Chương 1.** Giới thiệu tổng quan

Nội dung chương này giới thiệu khái quát về nhận diện khuôn mặt.

**Chương 2.** Nhận dạng khuôn mặt bằng LBP

Nội dung chương này trình bày việc xây dựng một hệ thống nhận dạng bao gồm việc phát hiện ra khuôn mặt,.

**Chương 3.** Xây dựng và cho ra kết quả

Nội dung chương 3 tập trung phân tích yêu cầu của hệ thống, xác định các chức năng chính.

## CHƯƠNG 1

### GIỚI THIỆU TỔNG QUAN

#### 1.1. TỔNG QUAN SINH TRẮC HỌC

Sinh trắc học là công nghệ sử dụng những thuộc tính vật lý, đặc điểm sinh học riêng của mỗi cá nhân như vân tay, móng mắt, khuôn mặt, giọng nói... để nhận diện.

#### 1.2. BÀI TOÁN NHẬN DIỆN KHUÔN MẶT

Nhận diện khuôn mặt thuộc lĩnh vực computer vision, là một phương pháp sinh trắc để xác minh một cá nhân nào đó bằng cách so sánh dữ liệu hình ảnh thu được trực tiếp hoặc hình ảnh với data đã lưu lại trước đó. Đa số nó được áp dụng vào các hệ thống an ninh như camera chống trộm, xác thực danh tính,...

Bên cạnh những thành công đã được ghi nhận thì nhận dạng khuôn mặt cũng còn gặp nhiều khó khăn như về độ sáng của môi trường, hướng nghiêng khuôn mặt, kích thước hình ảnh đưa vào, diện mạo, biểu hiện cảm xúc của khuôn mặt .

Để xây dựng một hệ thống nhận dạng khuôn mặt có đầu vào của hệ thống là một hình ảnh kỹ thuật số hay một khung hình video từ một nguồn video. Đầu ra là xác định hoặc xác minh người ở trong bức hình hoặc trong video đó là ai. Hướng tới mục tiêu này chúng ta thường chia thủ tục nhận dạng khuôn mặt gồm ba bước: Phát hiện khuôn mặt, trích rút đặc trưng và nhận dạng khuôn mặt.



*Hình 1.1.Hệ thống nhận dạng khuôn mặt*

Phát hiện khuôn mặt (Face Detection): Chức năng chính của bước này là định vị khuôn mặt, tìm kiếm tất cả khuôn mặt trong hình ảnh hoặc video. Phát hiện khuôn mặt phụ thuộc vào độ sáng, hướng nghiêng khuôn mặt, v.v. biểu hiện cảm xúc trên khuôn mặt hay các yếu tố môi trường khác. Để hệ thống nhận dạng hoạt động đạt hiệu quả cao thì hình ảnh khuôn mặt sau khi được phát hiện cần chuẩn hóa về

kích thước, ánh sáng.

Trích rút đặc trưng (Feature Extraction): Sau khi phát hiện ra khuôn mặt trong bức ảnh, chúng ta tiến hành trích rút những đặc trưng của khuôn mặt. Bước này trích xuất ra một vector đặc trưng đại diện cho một khuôn mặt. Nó phải đảm bảo được tính duy nhất của một khuôn mặt.

Nhận dạng khuôn mặt (Face Recognition): Với hình ảnh đầu vào sau khi phát hiện ra khuôn mặt, trích rút các đặc trưng của khuôn mặt và đem so sánh các đặc trưng này với cơ sở dữ liệu khuôn mặt.

Bài toán nhận dạng khuôn mặt được ứng dụng nhiều trong các lĩnh vực đời sống đặc biệt ở những lĩnh vực công nghệ cao, yêu cầu về an ninh, bảo mật. Do đó để hệ thống nhận dạng khuôn mặt hoạt động mạnh mẽ với tốc độ và độ tin cậy thì có rất nhiều các phương pháp về nhận dạng khuôn mặt được đưa ra. Các phương pháp có thể được phân loại theo các tiêu chí khác nhau như nhận dạng với dữ liệu ảnh đầu vào là ảnh tĩnh 2D. Phương pháp này là phổ biến nhất và tương lai sẽ là 3D. Tuy nhiên trên thực tế người ta hay chia phương pháp nhận dạng khuôn mặt ra thành 2 loại:

- Nhận dạng dựa trên các đặc trưng của các phần tử trên khuôn mặt (Feature Base Face Recognition)
- Nhận dạng dựa trên xét tổng thể toàn khuôn mặt (Appearance Based Face Recognition)

### **1.2.1. Nhận dạng dựa trên thu thập bộ dữ liệu khuôn mặt**

Để tiếp thực hiện nhận dạng thành công khuôn mặt, phải tiến hành giải quyết theo 3 giai đoạn riêng biệt:

- Phát hiện khuôn mặt và thu thập dữ liệu.
- Huấn luyện bộ dữ liệu vừa thu nhận được.
- Nhận dạng khuôn mặt.

Phương pháp được nhóm sử dụng trong hướng tiếp cận này là:

- LBP (Local Binary Pattern)

### **1.2.2. Điểm mạnh và hạn chế của nhận dạng khuôn mặt**

#### **a. Điểm mạnh**

- Phạm vi kiểm soát rộng

- Không đòi hỏi sự hợp tác của các đối tượng thử nghiệm.
- Có thể phát hiện các cá nhân giữa đám đông mà không bỏ sót.

***b. Hạn chế***

- Để có được hình ảnh đối chiếu với hệ thống thì yêu cầu người dân hoặc khách hàng phải quay ít nhất 35 độ về phía camera và không sử dụng khẩu trang, mũ, nón...
- Cần phải đáp ứng đủ điều kiện môi trường (ánh sáng, độ phân giải của hình ảnh...)

## CHƯƠNG 2

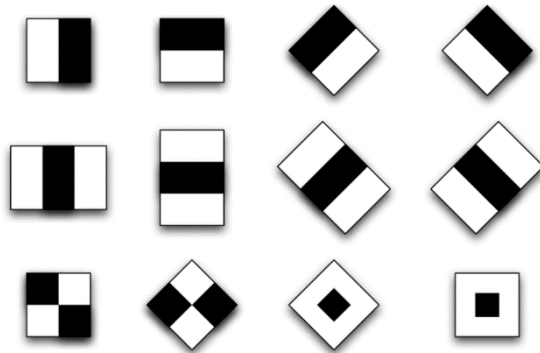
### KỸ THUẬT NHẬN DẠNG KHUÔN MẶT

Các pha trong hệ thống nhận dạng khuôn mặt bao gồm: phát hiện khuôn mặt (face detection); trích chọn đặc trưng (feature extraction), nhận dạng khuôn mặt (face recognition).

#### 2.1. PHÁT HIỆN KHUÔN MẶT

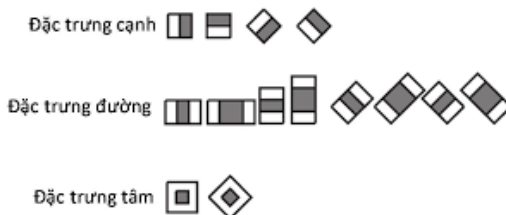
##### 2.1.1. Đặc trưng Haar-like

Đặc trưng Haar Like được tạo thành bằng việc kết hợp các hình chữ nhật đen, trắng với nhau theo một trật tự, một kích thước nào đó. Hình dưới đây mô tả 4 đặc trưng Haar Like cơ bản như sau:



*Hình 2.1 Các đặc trưng haar like cơ bản.*

Để sử dụng các đặc trưng này vào việc xác định khuôn mặt người, 4 đặc trưng Haar Like cơ bản được mở rộng ra và được chia làm 3 tập đặc trưng như sau:



*Hình 2.2 Các đặc trưng haar like mở rộng.*

### 2.1.2. Thuật toán tăng tốc Adaboost

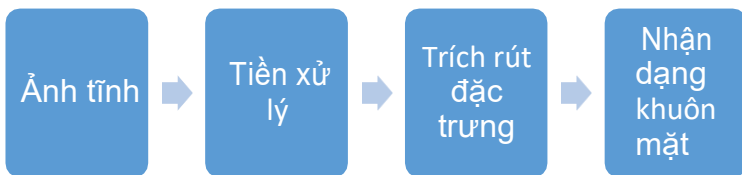
AdaBoost (Freund & Schapire, 1995) là một bộ phân loại mạnh phi tuyến phức, hoạt động trên nguyên tắc kết hợp tuyến tính các bộ phân loại yếu để tạo nên một bộ phân loại mạnh. AdaBoost sử dụng trọng số để đánh dấu các mẫu khó nhận dạng. Trong quá trình huấn luyện cứ mỗi bộ phân loại yếu được xây dựng thì thuật toán sẽ tiến hành cập nhật lại trọng số để chuẩn bị cho việc xây dựng bộ phân loại tiếp theo. Cập nhật bằng cách tăng trọng số của các mẫu nhận dạng sai và giảm trọng số của các mẫu được nhận dạng đúng bởi bộ phân loại yếu vừa xây dựng. Bằng cách này thì bộ phân loại sau có thể tập trung vào các mẫu mà bộ phân loại trước nó làm chưa tốt. Cuối cùng các bộ phân loại yếu sẽ được kết hợp lại tùy theo mức độ tốt của chúng để tạo nên một bộ phân loại mạnh.

Bộ phân loại yếu  $h_k$  được biểu diễn như sau:

$$h_k(x) = \begin{cases} 1 & \text{nếu } p_k f_k(x) < p_k \theta_k \\ 0 & \text{nếu ngược lại} \end{cases}$$

Với  $x$  là cửa sổ con cần xét,  $\theta_k$  là ngưỡng,  $f_k$  là giá trị đặc trưng Haar Like và  $p_k$  là hệ số quyết định chiều của phương trình.

## 2.2. NHẬN DẠNG KHUÔN MẶT



Hình 2.3. Hệ thống nhận dạng khuôn mặt cơ bản

### 2.2.1. Tiền xử lý

Quá trình này giúp nâng cao chất lượng hình ảnh, chuẩn hóa về mặt dữ liệu, kích thước hình ảnh. Giúp cho việc trích rút đặc trưng được chính xác hơn.

Các công việc thường gặp trong tiền xử lý hình ảnh:

- Crop ảnh, resize ảnh.
- Biến đổi ảnh thành dạng ảnh khác: ảnh màu sang ảnh xám, ảnh xám sang nhị phân, RGB→HSV...
- Cân bằng sáng
- Lọc nhiễu
- Xoay ảnh, warp ảnh

...

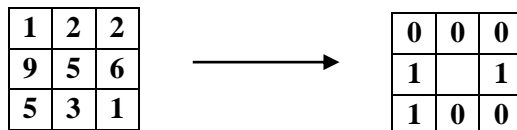
### 2.2.2. Trích rút đặc trưng

Trích rút đặc trưng là kỹ thuật sử dụng các thuật toán để lấy ra những thông tin mang những đặc điểm riêng biệt của một người. **Local Binary Pattern**- LBP là phương pháp trích chọn kết cấu của ảnh thành vector đặc trưng gọi là đặc trưng LBP. LBP là toán tử đơn từ hiệu quả gắn nhãn các pixel của ảnh bằng cách đánh giá các vùng lân cận của mỗi pixel và xem xét kết quả là một số nhị phân. LBP được dùng để đo độ tương phản cục bộ ảnh.

Với một điểm ảnh xét 8 điểm lân cận và sử dụng chính giá trị của điểm ảnh tại trung tâm để phân ngưỡng 8 giá trị lân cận. Kết quả tìm được một chuỗi bit nhị phân có chiều dài bằng 8 tương ứng với 8 điểm lân cận được định nghĩa theo thứ tự nhất định. Chuỗi bit sẽ được chuyển sang hệ thập phân và giá trị thập phân này thay thế giá trị ban đầu của điểm ảnh đang xét. Số điểm lân cận có thể được thay đổi bằng đại lượng bán kính

#### a. Thuật toán LBP

Bước đầu tiên để xây dựng một LBP descriptor là chuyển đổi input image về grayscale. Với mỗi pixel trong grayscale image, chúng ta chọn ra vùng lân cận (neighborhood) bao quanh pixel trung tâm đó, tất nhiên sẽ là  $3 \times 3$ . Một giá trị LBP sẽ được tính toán cho pixel trung tâm này và lưu vào một ma trận 2D có kích thước giống với input image.



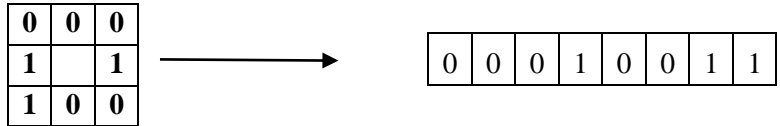


Công thức LBP :

$$LBP_{r,p} = \sum_{n=0}^{p-1} S(X_{r,p,n} - X_p) 2^n$$

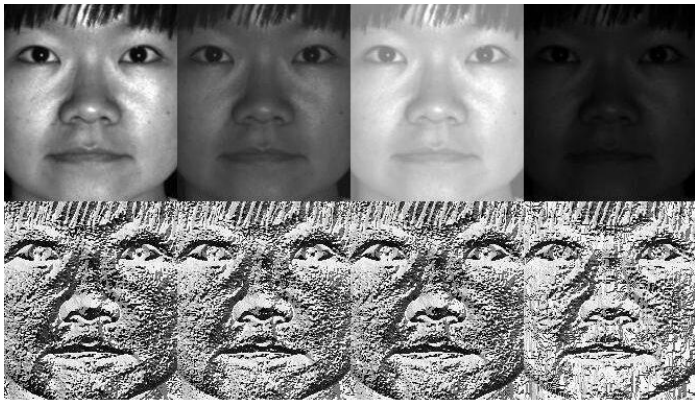
Trong đó:  $S(x) = \begin{cases} 1, & \text{khi } x \geq 0 \\ 0, & \text{ngược lại} \end{cases}$

Ví dụ: trong ảnh ta xét pixel 5, các vùng lân cận của nó là: 1 – 2 – 2 – 9 – 5 – 3 – 1 – 6 . Ta sẽ đi tìm trọng số (gọi là X) của mỗi điểm ảnh để tính toán giá trị nhị phân cần thiết. Với các vùng lân cận (tạm gọi là A), ta so sánh mỗi pixel với pixel đang xét (tạm gọi là B) (ở đây là pixel mang label = 5), nếu:  $A < B \Rightarrow X = 1$  ngược lại nếu  $A > B \Rightarrow X = 0$ . Ta có ma trận trọng số như sau:



Sau khi có ma trận trọng số, ta sẽ có dãy nhị phân: 00010111  $\Rightarrow$  số thập phân: 19.

Đây là kết quả sau khi sử dụng LBP, với cùng 1 gương mặt sử dụng các độ sáng tối, khác nhau ta sẽ nhận được kết quả giống nhau.



Hình 2.3. Kết quả

***b. Ưu điểm của thuật toán LBP***

Ưu điểm của phương pháp LBP là có chi phí tính toán thấp, ổn định khi cường độ thay đổi đơn điệu và dễ mở rộng lên không gian nhiều chiều như ảnh màu trong hệ RGB. Tuy nhiên, nhược điểm của đặc trưng dựa trên gradient là, đầu tiên, gradient phác họa cường độ phân phối xung quanh điểm ảnh một cách khá thô do chỉ biểu thị bởi hai giá trị 0 và 1. Hai mức gradient giống nhau có thể thuộc hai đối tượng cục bộ khác nhau, do đó không rõ ràng. Thứ hai, thông tin về gradient không đủ biểu diễn cho các đối tượng trong trường hợp gặp môi trường phức tạp gây ra nhiễu.

## CHƯƠNG 3

### XÂY DỰNG HỆ THỐNG

#### 3.1. PHÂN TÍCH THIẾT KẾ HỆ THỐNG

##### 3.1.1. Xác định yêu cầu

Đầu vào: Nguồn video thu được từ camera theo thời gian thực.

Đầu ra: Chương trình giúp ta có thể quan sát, ghi lại dữ liệu, phát hiện ra khuôn mặt, theo vết chuyển động của khuôn mặt, nhận dạng khuôn mặt, đưa ra cảnh báo khi người lạ xuất hiện trong vùng quan sát

##### *a. Yêu cầu chức năng*

- Hiện thị hình ảnh trên màn hình quan sát theo thời gian thực.
- Chức năng phát hiện khuôn mặt và theo vết khuôn mặt khi có người xuất hiện trong khung hình.

##### *b. Yêu cầu phi chức năng*

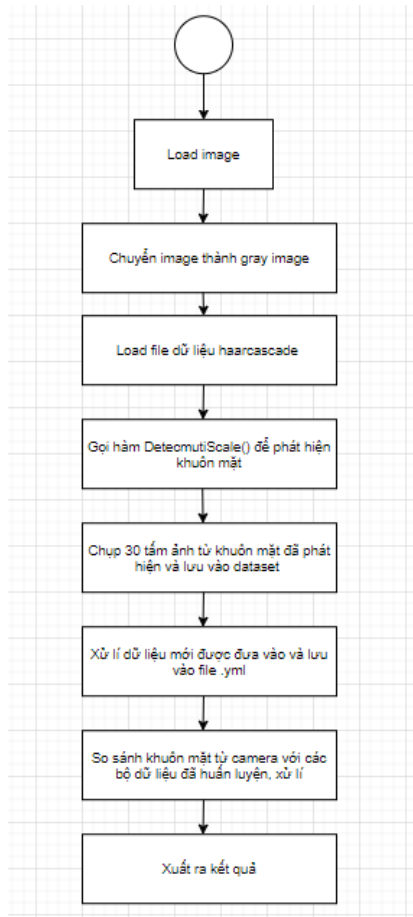
- Thiết kế giao diện thân thiện, dễ sử dụng.
- Quản lý dữ liệu hợp lý.
- Đầy đủ các chức năng theo yêu cầu.

### ***c. Môi trường và công cụ sử dụng***

- Sử dụng trên hệ điều hành Window
- Sử dụng thư viện OpenCV
- Điều kiện ánh sáng tốt, điều chỉnh hướng đặt camera để thu được khuôn mặt chính diện, khoảng cách đặt camera không quá xa để dễ dàng nhận diện khuôn mặt.

### **Chức năng phát hiện khuôn mặt người**

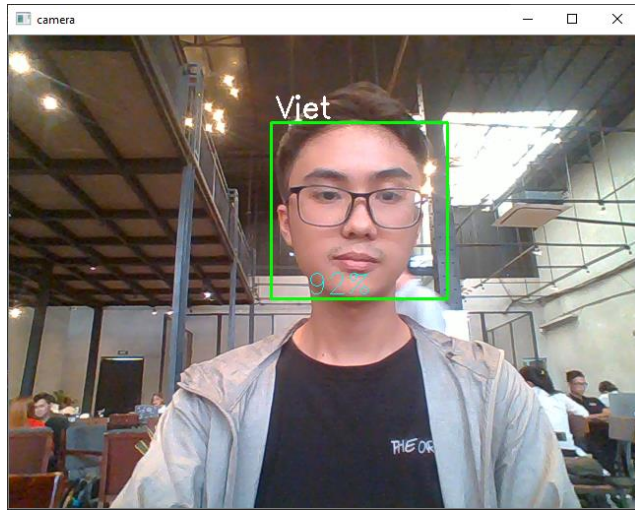
Khi phát hiện ra khuôn mặt trong video thì tiến hành tô viền xung quanh khuôn mặt để tiện cho việc theo vết khuôn mặt.



*Hình 3.1. Sơ đồ thuật toán phát hiện khuôn mặt*

### 3.2. KẾT QUẢ CHẠY CHƯƠNG TRÌNH

Giao diện chương trình chính



Hình 3.2: Giao diện chương trình khi nhận dạng khuôn mặt

### 3.3. KẾT QUẢ THỰC NGHIỆM

#### 3.3.1. Chức năng phát hiện khuôn mặt

Chức năng này được thực nghiệm trong các điều kiện về ánh sáng, hướng của khuôn mặt và che khuất khuôn mặt khác nhau.

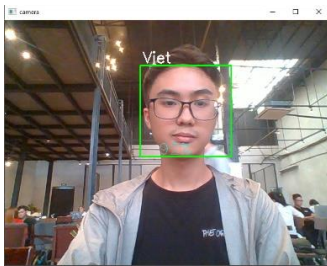
	Điều kiện thử nghiệm	Số khuôn mặt	Khuôn mặt phát hiện	Tỉ lệ phát hiện được
Ánh sáng	Tốt	6	6	100%
	Trung bình	6	5	83,3%
Hướng khuôn mặt	0 <sup>0</sup>	6	6	100%
	15 <sup>0</sup>	6	4	66%
	90 <sup>0</sup>	6	0	0%
Che khuất	Che hơn 50%	6	0	0%
	Che dưới 50%	6	3	50%

Khi chương trình chạy ở điều kiện ánh sáng tốt, hướng người nhìn chính diện vào camera, khoảng cách không qua xa so với camera thì tỉ lệ nhận diện đúng khuôn mặt cao. Tuy nhiên nếu bị ảnh hưởng bởi các yếu tố như bị che mặt hoặc môi trường không tốt thì tỷ lệ nhận diện sẽ bị giảm và đôi khi không thể nhận diện được.

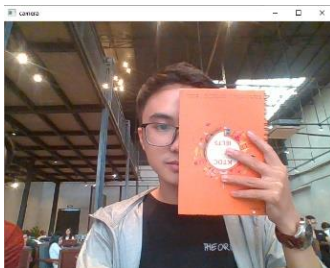
Hình ảnh dưới là một ví dụ cho trường hợp với cùng một điều kiện ánh sáng khi khuôn mặt nhìn thẳng vào camera và khuôn mặt bị che khuất một phần.

- Khi nhìn thẳng vào camera hệ thống phát hiện khuôn mặt chính xác.

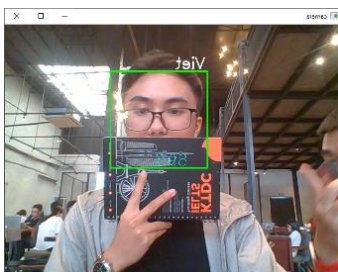
- Khi khuôn mặt bị che khuất hơn 50% thì hệ thống không phát hiện ra được khuôn mặt.



Khi nhìn trực diện vào camera



Che hơn 50% khuôn mặt

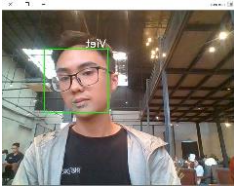


Che dưới 50% khuôn mặt

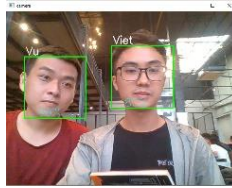
Quá trình phát hiện khuôn mặt đóng 1 vai trò quan trọng để việc nhận dạng được thành công. Hình ảnh khuôn mặt được phát hiện là dữ liệu chính cho qua trình nhận dạng, kết quả nhận dạng tỉ lệ thuận vào việc phát hiện khuôn mặt.

### 3.3.2. Chức năng nhận diện được khuôn mặt

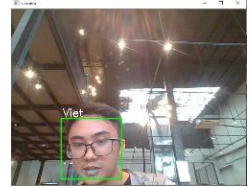
Sau khi phát hiện ra khuôn mặt hệ thống theo vết khuôn mặt khi đối tượng chuyển động. Nếu khuôn mặt thay đổi nhiều về hướng, điều kiện ánh sáng thì hệ thống đưa ra kết quả không chính xác. Đồng thời chương trình còn tính được độ chính xác của khuôn mặt nhận diện so với mẫu đã huấn luyện.



Vị trí 1



Vị trí 2



Vị trí 3

## KẾT LUẬN

Bài báo cáo của nhóm đã trình bày tổng quan về phương pháp nhận diện khuôn mặt người, phát hiện khuôn mặt, tính độ chính xác; xây dựng chương trình sử dụng theo phương pháp nhận diện khuôn mặt người bằng thuật toán LBP

Tuy nhiên việc nhận dạng mặt người cũng có hạn chế so với những phương pháp nhận diện khác. Kết quả của nhận diện khuôn mặt phải phụ thuộc vào các điều kiện cụ thể như điều kiện môi trường, hướng khuôn mặt,...