

**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

**KHOA ĐIỆN TỬ - VIỄN THÔNG
BỘ MÔN MÁY TÍNH – HỆ THỐNG NHÚNG**



**TRẦN TIỀN THÀNH
MAI VĂN TRÃI**

Đề tài:

**HỆ THỐNG ĐIỂM DANH
ỨNG DỤNG NHẬN DIỆN KHUÔN MẶT**

Chuyên ngành Máy Tính - Hệ Thống Nhúng

TP. Hồ Chí Minh, tháng 8 năm 2020



**ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

**KHOA ĐIỆN TỬ - VIỄN THÔNG
BỘ MÔN MÁY TÍNH - HỆ THỐNG NHÚNG**



TRẦN TIẾN THÀNH – MSSV: 1620234

MAI VĂN TRÃI – MSSV: 1620261

Đề tài:

**HỆ THỐNG ĐIỂM DANH
ỨNG DỤNG NHẬN DIỆN KHUÔN MẶT**

**KHÓA LUẬN TỐT NGHIỆP CỦ NHÂN
NGÀNH KỸ THUẬT ĐIỆN TỬ - VIỄN THÔNG
CHUYÊN NGÀNH MÁY TÍNH - HỆ THỐNG NHÚNG**

**NGƯỜI HƯỚNG DẪN KHOA HỌC
Th.S Lê Đức Trị**

TP. Hồ Chí Minh, tháng 8 năm 2020

LỜI CẢM ƠN

Lời đầu tiên, chúng em xin gửi lời cảm ơn đến quý Thầy Cô khoa Điện Tử - Viễn Thông, trường Đại học Khoa Học Tự Nhiên – Đại học Quốc gia Thành phố Hồ Chí Minh đã truyền đạt cho chúng em nhiều kiến thức quý báu trong suốt thời gian qua.

Chúng em cũng xin bày tỏ lòng biết ơn sâu sắc tới Thầy Lê Đức Trị. Thầy đã tận tình hướng dẫn, giúp đỡ, định hướng và chỉ bảo chúng em trong suốt thời gian thực hiện khóa luận tốt nghiệp này.

Sau cùng, chúng em xin chân thành bày tỏ tình cảm, lòng biết ơn và sự kính trọng sâu sắc đến Cha Mẹ đã luôn luôn động viên, ủng hộ. Những người bạn trong lớp K16-DTV đã gắn bó, chia sẻ rất nhiều kiến thức và kinh nghiệm trong thời gian thực hiện đề tài, để đề tài có thể thực hiện một cách thành công nhất.

Kiến thức là vô hạn, và sự tiếp thu của mỗi người thì luôn có giới hạn nhất định. Vì vậy, mặc dù đã đầu tư thời gian và công sức trong việc nghiên cứu và thực hiện đề tài này tuy nhiên không thể tránh khỏi những thiếu sót. Vì thế, chúng em rất mong nhận được sự chỉ bảo và đóng góp ý kiến của quý Thầy Cô để có thể bổ thêm sung kiến thức, nâng cao chuyên môn phục vụ tốt cho công việc sau này.

TP. Hồ Chí Minh, ngày... tháng 8, năm 2020

Sinh viên

Trần Tiến Thành

Mai Văn Trãi

MỞ ĐẦU

Trong những năm gần đây, máy học (machine learning) là một lĩnh vực đang được phát triển mạnh mẽ và được ứng dụng trong nhiều lĩnh vực khác như xe tự hành, chatbot tự động, xử lý hình ảnh,... Bằng máy học, máy tính đã thực hiện được những công việc vượt xa khả năng của một thiết bị điện tử thông thường như tự tạo chủ thích cho ảnh, nhận diện giọng nói, nhận diện chữ viết tay, và rất nhiều ứng dụng thực tế khác. Xử lý ảnh là một trong những ứng dụng phát triển mạnh mẽ nhất của máy học. Trong khóa luận này, sinh viên tập chung vào nghiên cứu và xây dựng ứng dụng nhận diện đối tượng dựa trên máy học. Đối tượng được chọn ở đây là khuôn mặt người. Và ứng dụng nhận diện khuôn mặt để xây dựng “Hệ thống điểm danh sinh viên dựa trên nhận diện khuôn mặt” trên nền tảng Web Application.

Cùng với sự phát triển của công nghệ, Internet phát triển như một điều tất yếu để phục vụ cho cuộc sống của con người. Internet là một hệ thống thông tin toàn cầu có thể được truy cập công cộng gồm các mạng máy tính được liên kết với nhau. Hệ thống này truyền thông tin theo kiểu nối chuyển gói dữ liệu (*packet switching*) dựa trên một giao thức liên kết mạng đã được chuẩn hóa (*giao thức IP*). Hệ thống này bao gồm hàng ngàn mạng máy tính nhỏ hơn của các doanh nghiệp, của các viện nghiên cứu và các trường đại học, của các mạng cá nhân và các chính phủ trên toàn cầu. Các ứng dụng trên nền tảng web ngày càng phổ biến vì tính đa nền tảng, dễ dàng tiếp cận và bảo trì.

Nội dung khóa luận bao gồm những nội dung chính sau:

Chương 1: Cơ sở lý thuyết.

Chương 2: Tổng quan về nhận diện khuôn mặt.

Chương 3: Tổng quan về C#, Microsoft SQL, ReactJS, thư viện EmguCV và ứng dụng Web.

Chương 4: Hệ thống điểm danh tự động bằng nhận diện khuôn mặt.

Chương 5: Kết luận và hướng phát triển.

NHẬN XÉT

(Của giảng viên hướng dẫn)

NHẬN XÉT

(Của giảng viên phản biện)

DANH MỤC CÁC CHỮ VIẾT TẮT

A	
API	Application Programming Interface - giao diện lập trình ứng dụng
H	
HOG	Histogram of oriented gradients – Biểu đồ của dốc định hướng
I	
IoT	Internet of Things – Internet kết nối vạn vật
K	
k-NN	k-Nearest Neighbors – k láng giềng gần nhất
L	
LDA	Linear Discriminant Analysis - Phân tích phân biệt tuyến tính
LBP	Local Binary Pattern – Mẫu nhị phân cục bộ
P	
PCA	Principal Component Analysis – Phân tích thành phần chính
R	
RDBMS	Relational Database Management System – Hệ quản trị dữ liệu quan hệ
S	
SVM	Support Vector Machine – Máy Vector Hỗ trợ
SQL	Structured Query Language - Ngôn ngữ truy vấn có cấu trúc

DANH MỤC CÁC HÌNH

Hình 1.1 Giảng viên đọc tên điểm danh sinh viên	16
Hình 1.2 Sơ đồ phương pháp điểm danh truyền thống	17
Hình 1.3 Máy quét vân tay.....	18
Hình 1.4 Sơ đồ hệ thống điểm danh bằng vân tay.	18
Hình 1.5 Hệ thống điểm danh bằng vân tay trong thực tế.....	19
Hình 1.6 Nhận diện khuôn mặt.....	21
Hình 1.7 Ứng dụng nhận diện khuôn mặt vào điểm danh.	21
Hình 2.1 Nhận diện khuôn mặt trong ảnh	23
Hình 2.2 Hệ thống nhận diện khuôn mặt.....	25
Hình 2.3 Các điểm mứt đặc trưng trên một khuôn mặt.....	28
Hình 2.4 Công nghệ nhận diện khuôn mặt 3D	29
Hình 2.5 FaceID trên điện thoại iPhone	30
Hình 2.6 Sơ đồ hệ thống nhận diện khuôn mặt	31
Hình 2.7: Phát hiện các khuôn mặt trong ảnh.....	32
Hình 2.8 Các đặc hình trưng trong Haar-Like	34
Hình 2.9 Các đặc trưng Haar-Like cho phát hiện khuôn mặt.....	34
Hình 2.10 Hình ảnh tích hợp (Integral Image).....	35
Hình 2.11: Các bộ lọc Haar-Like trượt trên hình ảnh.....	36
Hình 2.12 Mô hình phân tầng Haar-Cascade	37
Hình 2.13: Kết hợp các bộ phân loại yếu thành bộ phân loại mạnh.....	38
Hình 2.14 Sơ đồ hệ thống phát hiện khuôn mặt Viola Jones	38
Hình 2.15 Mô hình phân tầng định vị khuôn mặt trong ảnh.....	39
Hình 2.16 Mô hình giảm chiều dữ liệu với PCA	41
Hình 2.17: Mô hình biểu đồ tính PCA	41
Hình 2.18 Khuôn mặt được giảm chiều dữ liệu với PCA	44
Hình 2.19: Tính toán giá trị LBP	46
Hình 2.20: Giá trị bán kính R và số pixel lân cận P	46

<i>Hình 2.21: Hình ảnh gốc và sau khi được biểu diễn bằng LBP</i>	47
<i>Hình 2.22: Hình ảnh khuôn mặt được mô tả với LBPH trong các điều kiện sáng khác nhau</i>	48
<i>Hình 2.23: Chia khuôn mặt thành nhiều block cục bộ</i>	48
<i>Hình 2.24 Các biểu đồ của từng khối (block)</i>	49
<i>Hình 2.25 Chọn khuôn mặt phù hợp bằng khoảng cách euclidean</i>	51
<i>Hình 2.26 Ảnh training gốc và ảnh training được lật gương</i>	53
<i>Hình 3.1 Sơ đồ các thành phần trong .NET Framework</i>	56
<i>Hình 3.2 Cấu trúc một ứng dụng Web</i>	60
<i>Hình 4.1 Sơ đồ hệ thống điểm danh sinh viên bằng nhận diện khuôn mặt</i>	62
<i>Hình 4.2 Cơ sở dữ liệu và ràng buộc quan hệ giữa các bảng</i>	64
<i>Hình 4.3 Giao diện đăng nhập của giảng viên</i>	67
<i>Hình 4.4 Giao diện lịch giảng dạy của giảng viên</i>	68
<i>Hình 4.5 Danh sách các môn học mà giảng viên quản lý giảng dạy</i>	69
<i>Hình 4.6 Giao diện chi tiết lớp học đang mở</i>	69
<i>Hình 4.7 Màn hình điểm danh sinh viên</i>	70
<i>Hình 4.8 Màn hình quản lý sinh viên của một lớp</i>	71
<i>Hình 4.9 Thông tin chi tiết sinh viên</i>	71
<i>Hình 4.10 Giao diện quản lý lịch học của sinh viên</i>	72
<i>Hình 4.11 Giao diện upload ảnh training cho sinh viên</i>	73

DANH MỤC CÁC BẢNG, BIỂU ĐỒ

Bảng

Bảng 2.1 Bảng chú thích các biến và kích thước của biến trong PCA.....42

Biểu đồ

Biểu đồ 2.1 Độ chính xác của các thuật toán với số lượng ảnh training.....52

MỤC LỤC

GIỚI THIỆU VỀ ĐỀ TÀI.....	12
CHƯƠNG 1 CƠ SỞ LÝ THUYẾT	15
1.1 Bài toán điểm danh.....	15
1.2 Các phương pháp điểm danh đã được triển khai và những bất cập... 	16
1.2.1 Phương pháp truyền thống.	16
1.2.2 Phương pháp điểm danh bằng vân tay.	17
1.2.3 Phương pháp điểm danh bằng quét thẻ ID.	20
1.3 Áp dụng nhận diện khuôn mặt vào bài toán điểm danh.	20
CHƯƠNG 2 TỔNG QUAN VỀ NHẬN DIỆN KHUÔN MẶT.....	23
2.1 Bài toán nhận diện khuôn mặt	23
2.2 Những khó khăn và thách thức trong bài toán nhận diện khuôn mặt	25
2.3 Hệ thống nhận diện khuôn mặt hoạt động như thế nào?	26
2.3.1 Cơ sở lý thuyết.....	26
2.3.1.1 Phát hiện khuôn mặt.....	32
2.3.1.1.1 Tìm hiểu về các đặc trưng Haar-Like.....	33
2.3.1.1.2 AdaBoost.....	37
2.3.2 Nhận diện khuôn mặt.	39
2.4 Các thuật toán phổ biến trong nhận diện khuôn mặt.....	40
2.4.1 Eigen Face Recognizer.....	40
2.4.1.1 Tìm hiểu về PCA.....	40
2.4.1.2 Tìm hiểu về EigenFace.	42
2.4.2 Local Binary Pattern Histogram	45

2.4.2.1 Local Binary Pattern (LBP) là gì?	45
2.4.2.2 LBPH trong nhận diện khuôn mặt.	47
2.5 So sánh độ chính xác giữa các thuật toán và bài toán tối ưu hóa.	51
CHƯƠNG 3 TỔNG QUAN VỀ C#, MICROSOFT SQL, REACTJS, THU'	
VIỆN EMGUCV VÀ ỨNG DỤNG WEB	54
3.1 Ngôn ngữ lập trình C# và .NET Framework.....	54
3.1.1 Giới thiệu ngôn ngữ lập trình C#	54
3.1.2 .NET Framework	55
3.2 Hệ quản trị dữ liệu Microsoft SQL Server	57
3.3 Ngôn ngữ lập trình JavaScript và ReactJs Library.....	57
3.3.1 JavaScript là gì?	57
3.3.2 ReactJs Library.....	58
3.4 Tổng quan về thư viện EmguCv.	58
3.5 Ứng dụng Web.	59
3.5.1 Ứng dụng web là gì và những tiện ích của ứng dụng web.	59
3.5.2 Cấu trúc của một ứng dụng web.....	60
3.5.3 Dịch vụ Web và Web API.....	61
CHƯƠNG 4 HỆ THỐNG ĐIỂM DANH TỰ ĐỘNG BẰNG NHẬN DIỆN	
KHUÔN MẶT	62
4.1 Thiết kế hệ thống.	62
4.1.1 Phân tích nghiệp vụ.....	62
4.1.2 Thiết kế cơ sở dữ liệu.....	63
4.1.3 Nền tảng công nghệ.....	66
4.2 Một số chức năng chính của hệ thống.	67

4.2.1 Giao diện đăng nhập giảng viên.	67
4.2.2 Quản lý lịch giảng dạy của giảng viên.....	68
4.2.3 Điểm danh sinh viên.....	70
4.2.4 Thống kê, quản lý các môn học, sinh viên.	71
4.2.5 Giao diện đăng nhập sinh viên.....	72
4.2.6 Quản lý lịch học sinh viên.	72
4.2.7 Upload ảnh training.....	72
CHƯƠNG 5 KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN.	74
DANH MỤC TÀI LIỆU THAM KHẢO	76

GIỚI THIỆU VỀ ĐỀ TÀI

A. Lý do chọn đề tài.

Trong kỷ nguyên cách mạng công nghiệp 4.0, việc ứng dụng công nghệ thông tin vào thực tế đời sống không còn là điều mới mẻ. Internet hóa (Internet of Things) và máy học (Machine Learning) đang trở thành một xu thế tất yếu trong nhiều lĩnh vực đời sống kinh tế, khoa học, giáo dục của nhân loại.

Sự ra đời của công nghệ thông tin là sự tích hợp đồng thời các tiến bộ về công nghệ và tổ chức thông tin, đem đến nhiều ảnh hưởng tích cực cho sự phát triển của giáo dục. Công nghệ thông tin và đặc biệt là sự phát triển của internet đã mở ra một kho kiến thức vô cùng đa dạng và phong phú cho người học và người dạy, giúp cho việc tìm hiểu kiến thức đơn giản hơn rất nhiều, cải thiện chất lượng học và dạy.

So với cách làm việc truyền thống, việc ứng dụng công nghệ mang lại tiềm năng lớn trong việc tự động hóa và thực hiện các nhiệm vụ quản trị cho các giáo viên. Ở các trường đại học, một trong những công việc gây tốn thời gian, công sức của cả giảng viên và sinh viên nhất đó là điểm danh.

Điểm danh theo phương pháp truyền thống gây khó khăn trong việc quản lý và thống kê sinh viên đi học. Hiện cũng có một số ứng dụng công nghệ như nhận dạng vân tay, quét thẻ ID vào điểm danh nhưng vẫn có nhiều bất cập.

Biết được những nhu cầu đó, trong khóa luận này, sinh viên đã ứng dụng những kiến thức đã học và tìm hiểu được để xây dựng ứng dụng điểm danh tự động dựa trên công nghệ nhận diện khuôn mặt. Song song với sự bùng nổ về Internet như hiện nay là sự phát triển mạnh mẽ của các ứng dụng trên nền tảng web. Miễn là có kết nối mạng (wifi, 3G, 4G) các ứng dụng web có thể được truy cập mọi lúc, mọi nơi và trên đa thiết bị. Với những lợi ích đó, sinh viên đã chọn nền tảng web để xây dựng “Hệ thống điểm danh sinh viên bằng nhận diện khuôn

mặt” để tạo sự linh hoạt cho quá trình sử dụng của người dùng - ở đây là sinh viên và giảng viên.

B. Mục đích, ý nghĩa chọn đề tài.

Việc ứng dụng công nghệ để thay thế các phương pháp truyền thông như một điều tất yếu trong quá trình phát triển của thế giới. Công nghệ nhận diện khuôn mặt mặc dù vẫn còn đang trong giai đoạn nghiên cứu và cải thiện tuy nhiên không thể phủ nhận sự hiệu quả của nó khi áp dụng thực tế như không tốn kém chi phí cơ sở hạ tầng, xử lý nhanh.

Nhận diện mặt người (Face recognition) là một lĩnh vực nghiên cứu của ngành Thị giác máy tính (Computer Vision), và cũng được xem là một lĩnh vực nghiên cứu của ngành Sinh trắc học (Biometrics - tương tự như nhận dạng vân tay – Fingerprint recognition, hay nhận dạng mống mắt – Iris recognition). Xét về nguyên tắc chung, nhận dạng mặt có sự tương đồng rất lớn với nhận dạng vân tay và nhận dạng mống mắt, tuy nhiên sự khác biệt nằm ở bước trích chọn đặc trưng (feature extraction) của mỗi lĩnh vực. Trong khi nhận dạng vân tay và mống mắt đã đạt tới độ chín, tức là có thể áp dụng trên thực tế một cách rộng rãi thì nhận dạng mặt người vẫn còn nhiều thách thức và vẫn là một lĩnh vực nghiên cứu thú vị với nhiều người. So với nhận dạng vân tay và mống mắt, nhận dạng mặt có nguồn dữ liệu phong phú hơn, chúng ta có thể nhìn thấy mặt người ở bất cứ tấm ảnh, video clip nào liên quan tới con người trên mạng và ít đòi hỏi sự tương tác có kiểm soát hơn (để thực hiện nhận dạng vân tay hay mống mắt, dữ liệu input lấy từ con người đòi hỏi có sự hợp tác trong môi trường có kiểm soát).

Nhận thấy quy trình điểm danh truyền thống và cả các phương pháp điểm danh hiện đại hơn đã và đang được triển khai vẫn còn những bất cập như tốn thời gian, khó khăn trong việc quản lý (phương pháp điểm danh truyền thống), setup hệ thống phức tạp, tốn kém chi phí (điểm danh bằng quét thẻ ID) hoặc thậm chí còn tiềm ẩn nguy cơ lây các bệnh truyền nhiễm (điểm danh bằng nhận dạng vân

tay). “Hệ thống điểm danh bằng nhận diện khuôn mặt” được sinh viên nghiên cứu và xây dựng để khắc phục những bất cập mà các phương pháp cũ đang gặp phải.

C. Phương pháp nghiên cứu.

- Về mặt lý thuyết: các thuật toán trong xử lý ảnh để phát hiện và nhận diện khuôn mặt, kỹ thuật lập trình, cách thức hoạt động của một website sử dụng ngôn ngữ lập trình C#
- Về mặt thực hành: Sử dụng ngôn ngữ C# và .NET Framework để xây dựng website hệ thống. Triển khai các thuật toán nhận diện khuôn mặt để ứng dụng để điểm danh các sinh viên đang tham gia lớp học.

D. Ý nghĩa thực tiễn của đề tài.

“Hệ thống điểm danh bằng nhận diện khuôn mặt” được sinh viên xây dựng trong khóa luận này hoàn toàn có thể triển khai vào thực tế. Vì được xây dựng trên nền tảng Web nên dễ dàng truy cập và sử dụng, đáp ứng được các nhu cầu cơ bản và thiết yếu mà quy trình điểm danh cần, cùng các chức năng về quản trị sinh viên, môn học. Bên cạnh đó không tốn quá nhiều chi phí để triển khai cũng như bảo trì dễ dàng. Hệ thống có thể thêm một số chức năng như cho phép giảng viên giao bài tập về nhà và chấm điểm, hay phát triển thêm chức năng học online và điểm danh online.

CHƯƠNG 1 CƠ SỞ LÝ THUYẾT

1.1 Bài toán điểm danh.

Với phương pháp đào tạo theo hệ thống tín chỉ hiện nay ở hầu hết các trường đại học ở Việt Nam và trên thế giới thì việc điểm danh trở nên không thể thiếu như một phương pháp giúp sinh viên đến giảng đường tham gia lớp học đều đặn.

Tùy vào quy mô lớn nhỏ của cơ sở đào tạo mà số lượng lớp mở trong mỗi học kỳ khác nhau. Tuy nhiên, con số này cũng dao động từ 200-500 lớp được mở mỗi học kỳ. Đầu mỗi học kỳ, giảng viên sẽ nhận từ Phòng Đào Tạo thời khóa biểu của sinh viên và bảng theo dõi tình hình môn học của những môn mà giảng viên đó tham gia giảng dạy.

Hàng ngày, giảng viên lên lớp phải mở bảng theo dõi tình hình môn học để điểm danh sinh viên vắng bằng cách “gọi tên” từng sinh viên, sinh viên nào vắng mặt thì giảng viên sẽ ghi số tiết vắng vào bảng theo dõi. Giảng viên tính số tiết vắng của những buổi đã học của sinh viên, nếu sinh viên nào có số tiết vắng đạt gần 30% tổng số tiết, giảng viên sẽ thông báo cho sinh viên đó số tiết vắng. Sinh viên sẽ tự kiểm tra số tiết vắng đó có đúng hay không, nếu không đúng sẽ thông báo lại cho giảng viên và giảng viên sẽ tính lại.

Cuối mỗi học kỳ, giảng viên sẽ mở lại bảng theo dõi để đếm số tiết vắng của từng sinh viên, từ đó tính điểm chuyên cần của sinh viên theo văn bản hướng dẫn thực hiện quy chế 25/2006/QĐ-BGDDT trong việc cho điểm quá trình. Giảng viên ghi điểm chuyên cần và đánh dấu những sinh viên mất tư cách về mặt thời gian của từng môn học và nộp lại cho Phòng Đào Tạo.

Cuối kỳ, khi lãnh đạo có yêu cầu, giảng viên phải dựa vào bảng theo dõi để lập báo cáo về số tiết vắng của sinh viên và tình hình sinh viên vắng của trường trong học kỳ.

Với việc quản lý học sinh sinh viên như trước đây thì đa phần do giáo viên trực tiếp đứng lớp điểm danh, việc vắng mặt hay không đều do giáo viên đảm nhiệm và việc này thường mất một khoảng thời gian trong giờ dạy học, nó cũng có những mặt tiêu cực xảy ra bởi giáo viên hoặc học sinh có thể điểm danh cho nhau để mà trốn học, việc này thô sơ cũng tạo nên sự lạc hậu, kèm theo đó những thông tin dữ liệu đó phải cập nhật trong sổ sách, quá nhiều giấy tờ lưu kho chiếm mất diện tích của trường học. Giải pháp nào để thoát khỏi việc áy?



Hình 1.1 Giảng viên đọc tên điểm danh sinh viên.

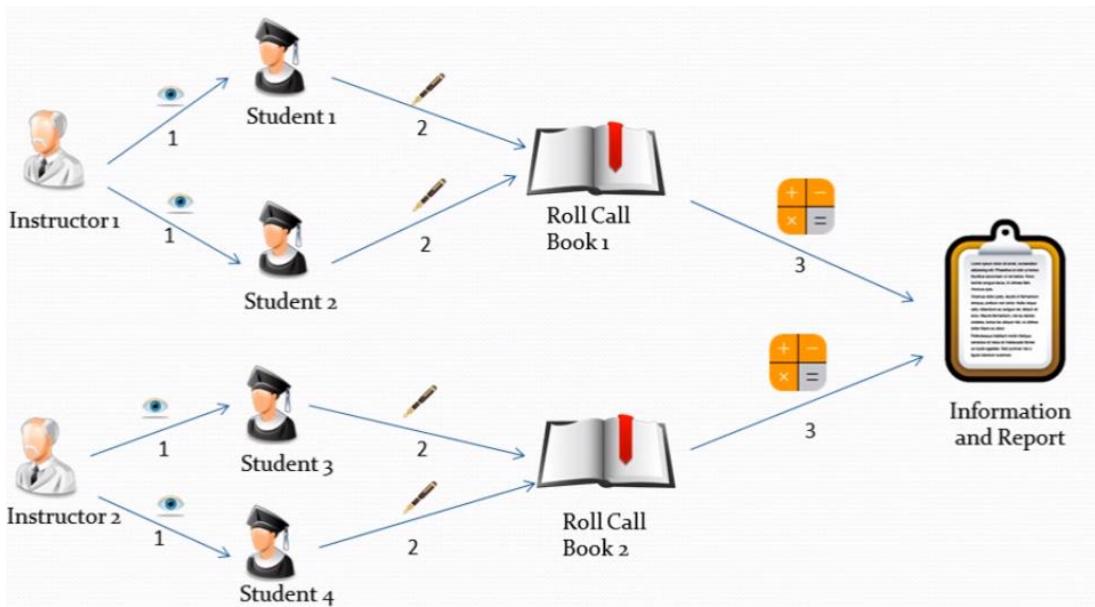
1.2 Các phương pháp điểm danh đã được triển khai và những bất cập.

1.2.1 Phương pháp truyền thống.

Hiện tại, hầu hết việc điểm danh ở các trường đại học ở Việt Nam được thực hiện bằng cách thủ công, là giảng viên sẽ lấy danh sách sinh viên lớp học, đọc tên từng sinh viên và đánh dấu vào danh sách. Công việc này rất mất thời gian và công sức của cả giảng viên lẫn sinh viên. Đặc biệt là một ngày giảng viên phải giảng dạy nhiều lớp, và mỗi lớp có cả trăm sinh viên thì phương pháp này thực sự tốn thời gian và công sức.

Với công nghệ phát triển hiện tại, việc điểm danh đã đơn giản hơn với việc sử dụng các công cụ thống kê như Excel, Google Sheet được lưu trữ dưới

dạng các file dữ liệu cũng như giao diện thân thiện và dễ thao tác. Tuy nhiên, việc này cùng đòi hỏi giảng viên phải có thời gian để đọc tên và đánh dấu điểm danh từng sinh viên, cùng các bước tính toán để ra các số liệu như các buổi đi học, nghỉ học.Thêm vào đó, việc thống kê và lưu trữ các file này cũng đòi giảng viên phải ghi nhớ vị trí các file trong ổ cứng, chưa kể các bất cập như giảng viên có thể không luôn đem theo máy tính cá nhân lên giảng đường hoặc xảy ra các sự cố về thất thoát dữ liệu hoặc hư hỏng máy tính.



Hình 1.2 Sơ đồ phương pháp điểm danh truyền thống.

Phương pháp điểm danh truyền thống này tuy phổ biến vì dễ dàng thực hiện, không tốn thời gian training nhưng lại có quá nhiều bất cập và tốn kém cả về thời gian và công sức. Với số lượng lớp giảng dạy và sinh viên theo học nhiều thì đôi khi còn xảy ra sai sót. Lúc này cả sinh viên và giảng viên sẽ phải ngồi thống kê lại các buổi học và buổi nghỉ để đánh giá lại.

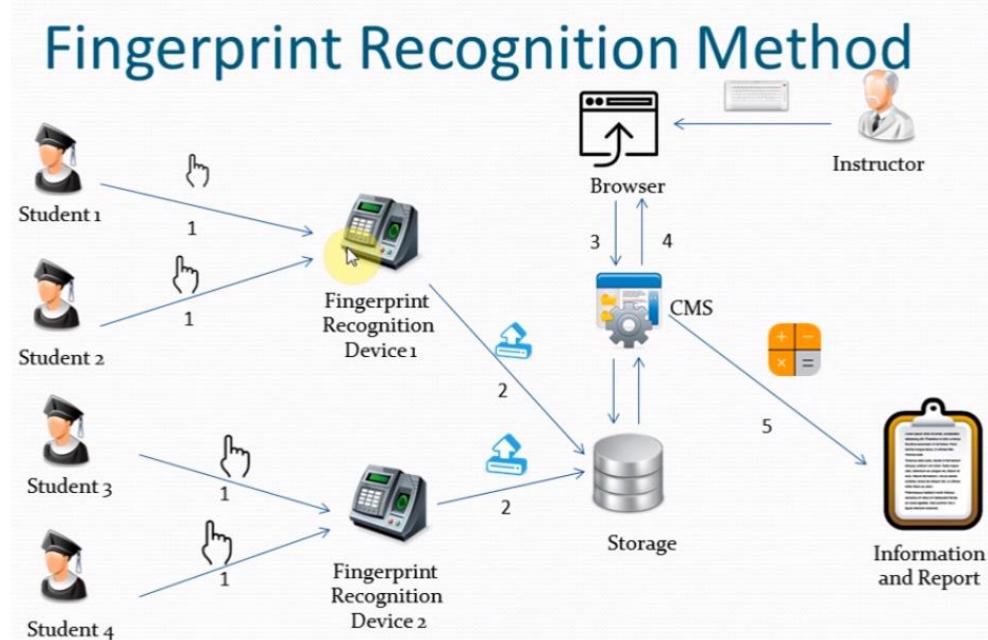
1.2.2 Phương pháp điểm danh bằng vân tay.

Nhận diện vân tay hiện nay đã đạt được một số thành tựu nhất định và được áp dụng rộng rãi trong thực tế. Nó thậm chí có thể được tìm thấy trên các tiện ích gia đình, chẳng hạn như một số máy tính xách tay và điện thoại thông minh mà chúng ta vẫn thường dùng, hay hệ thống quét vân tay ở nơi làm việc...



Hình 1.3 Máy quét vân tay.

Với công nghệ phát triển và có nguồn tài chính tốt thì hiện nay một số trường đã và đang triển khai hệ thống nhận dạng vân tay cho các phòng học, giảng đường. Phương pháp này có ưu điểm là làm việc nhanh, hoạt động hầu như chính xác tuyệt đối vì công nghệ nhận diện vân tay hiện nay gần như đạt tới độ chín về công nghệ. Các hệ thống nhận diện vân tay cũng được sử dụng rộng rãi trong các môi trường doanh nghiệp để chấm công cho số lượng lớn nhân viên. Hệ thống nhận diện vân tay cũng dễ dàng xuất các file báo cáo.



Hình 1.4 Sơ đồ hệ thống điểm danh bằng vân tay.

Tuy nhiên, nhược điểm của hệ thống này là chi phí không hề rẻ cùng với setup hệ thống phức tạp, chưa kể chi phí bảo trì cũng như vận hành. Nếu trường học setup cho tất cả các phòng học thì nó trở nên cực kỳ tốn kém. Nếu không setup đủ thì lại gây khó khăn cho việc chia phòng học. Sau mỗi buổi học, giảng viên lại phải cắm cáp từ máy nhận diện vân tay vào máy tính cá nhân để lấy file báo cáo sau đó lại ngồi tính toán thủ công các buổi học buổi nghỉ, điểm chuyên cần cho từng sinh viên. Điều này khiến quy trình điểm danh trở nên phức tạp hơn. Đặc biệt là một phòng học lại có thể sử dụng cho nhiều lớp dạy.

Cũng có giải pháp là cài thêm ứng dụng quản lý trực tiếp trên máy tính cá nhân của giảng viên để khi tham gia lớp học, giảng viên chỉ cần kết nối máy tính cá nhân với hệ thống quét vân tay của phòng học. Khi đó, hệ thống lại phải detect lớp học đang diễn ra là của môn nào, giảng viên nào, có những sinh viên nào theo học để lấy cơ sở dữ liệu cho hợp lý. Tuy nhiên điều này lại đòi hỏi máy tính của giảng viên phải cài đặt phần mềm hệ thống. Qua đó có thể thấy hệ thống setup cực kỳ phức tạp.

Một nhược điểm nữa của hệ thống điểm danh bằng vân tay là có thể là vật trung gian lây các bệnh truyền nhiễm. Điều này trở nên cực kỳ nguy hiểm trong môi trường đại học với cả ngàn sinh viên theo học trong trường.



Hình 1.5 Hệ thống điểm danh bằng vân tay trong thực tế.

1.2.3 Phương pháp điểm danh bằng quét thẻ ID.

Phương pháp này có một số ưu điểm và nhược điểm giống hệ thống điểm danh bằng vân tay. Ngoài ra, hệ thống này còn đòi hỏi tất cả sinh viên trong trường đều phải có thẻ sinh viên bằng thẻ từ. Điều này gây tốn kém hơn. Thêm vào đó còn xảy ra tình trạng quét thẻ hộ khiến việc quản lý điểm danh trở nên có kẽ hở.

1.3 Áp dụng nhận diện khuôn mặt vào bài toán điểm danh.

Khi nền công nghiệp 4.0 ra đời, rất cả những vấn đề trong cuộc sống đều được máy tính xử lý một cách hoàn toàn tự động và nhanh chóng. Hiện nay, công nghệ nhận diện khuôn mặt ngày càng trở nên phổ biến và được áp dụng nhiều trong thực tế cuộc sống. Tuy còn vẫn còn đang trong giai đoạn nghiên cứu, phát triển và đôi khi chưa đạt được độ chính xác tuyệt đối song công nghệ nhận diện khuôn mặt đang ngày càng được hoàn thiện cả về độ chính xác lẫn hiệu quả áp dụng trong thực tế.

Hiện nay, chúng ta có thể thấy công nghệ nhận diện khuôn mặt phổ biến khắp nơi, và được ứng trên nhiều thiết bị trong cuộc sống như một số điện thoại cho phép mở khóa bằng khuôn mặt, các mạng xã hội như facebook, instagram,... tự động tag các ảnh có khuôn mặt của bạn vào hay việc kiểm soát các người ra vào văn phòng tại các cơ quan, doanh nghiệp. Hệ thống này thường được sử dụng trong các hệ thống an ninh và có thể được so sánh với các dạng sinh trắc học khác như các hệ thống nhận dạng vân tay hay tròng mắt.

Trong số các kỹ thuật sinh trắc học, nhận dạng khuôn mặt có thể không đáng tin cậy và hiệu quả nhất. Tuy nhiên, một trong những lợi thế quan trọng là nó không đòi hỏi sự hợp tác của các đối tượng thử nghiệm. Các hệ thống thiết kế được lắp đặt tại các sân bay, khu chung cư, và những nơi công cộng khác có thể xác định các cá nhân giữa đám đông, mà không bỏ sót một ai. Sinh trắc học khác như dấu vân tay, quét móng mắt, và nhận dạng giọng nói không thể thực hiện được như vậy.



Hình 1.6 Nhận diện khuôn mặt.

Việc áp dụng công nghệ nhận diện khuôn mặt vào việc điểm danh giúp giảng viên thực hiện công việc này một cách nhanh chóng, không tốn quá nhiều thời gian và công sức vì hệ thống làm việc hoàn toàn tự động.



Hình 1.7 Ứng dụng nhận diện khuôn mặt vào điểm danh.

Nhận diện khuôn mặt 2D tuy đôi khi đem lại độ chính xác chưa hoàn hảo nhất và có thể bị đánh lừa bởi một bức ảnh chứa khuôn mặt. Nhưng trong bài toán điểm danh, kết hợp với sự quan sát trực quan của giảng viên đứng lớp thì những vấn đề này hoàn toàn có thể được giải quyết

Hệ thống này có thể đem lại một số lợi ích như:

- Không cần phải trực tiếp tiếp xúc với thiết bị để xác thực (các kỹ thuật xác thực sinh trắc học dựa trên tiếp xúc khác như máy quét dấu vân tay, có thể không hoạt động chính xác nếu có vết bẩn trên tay của một người).
- Cải thiện mức độ bảo mật.
- Yêu cầu xử lý ít hơn so với các kỹ thuật xác thực sinh trắc học khác.
- Độ chính xác được cải thiện theo thời gian.
- Có thể được sử dụng để giúp tự động hóa việc xác thực.

Có thể thấy việc áp dụng công nghệ nhận diện khuôn mặt kết hợp với ứng dụng trên nền tảng web vào bài toán điểm danh có thể khắc phục được hầu hết các nhược điểm của các phương pháp cũ đã được triển khai.

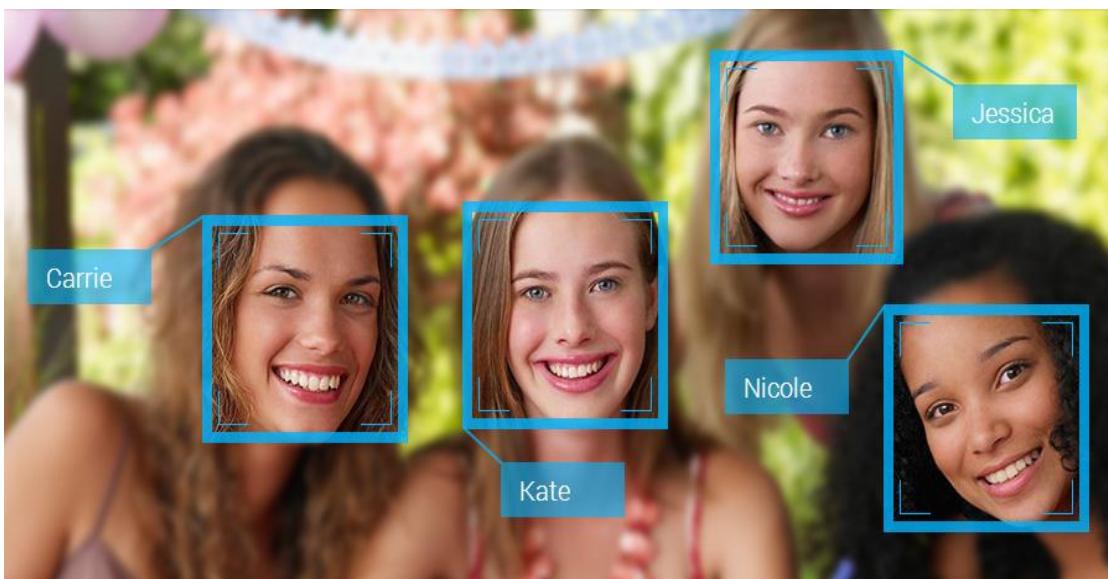
Tuy nhiên, hệ thống này không phải là không có nhược điểm như độ chính xác có thể thấp hơn hay vấn đề quyền riêng tư như dữ liệu khuôn mặt có thể bị rò rỉ ra bên ngoài. Nhưng với những lợi ích mà nó đem lại thì nó có thể giải quyết bài toán chúng ta đặt ra một cách hiệu quả nhất.

CHƯƠNG 2 TỔNG QUAN VỀ NHẬN DIỆN KHUÔN MẶT

2.1 Bài toán nhận diện khuôn mặt.

Nhận dạng hay nhận biết một đối tượng thông qua các giác quan của con người đặc biệt là thị giác là một đặc điểm hết sức quan trọng. Nó giúp chúng ta phân biệt được các sự vật hiện tượng và cảm nhận được thế giới thiên nhiên cũng như mọi người xung quanh.

Ngày nay, với sự phát triển của khoa học công nghệ, vấn đề nhận dạng đang được đặc biệt quan tâm và phát triển, nhất là trong lĩnh vực robot, tự động hóa, bảo mật... Con người rất giỏi việc nhận diện khuôn mặt và các hình mẫu phức tạp. Ngay cả khi một thời gian trôi qua vẫn không ảnh hưởng đến khả năng này và con người ngày càng mong ước tạo ra được một hệ thống nhận dạng có khả năng tương đương hoặc thậm chí là vượt hơn hệ thống nhận dạng sinh học.



Hình 2.1 Nhận diện khuôn mặt trong ảnh.

Nhận diện khuôn mặt người là một trong những vấn đề được khoa học thế giới nghiên cứu và phát triển, nó có rất nhiều ứng dụng trong thực tế cuộc sống cũng như trong khoa học công nghệ. Ngày nay, cùng với sự bùng nổ thông tin, sự phát triển của công nghệ cao, sự giao tiếp giữa con người và máy tính đang thay đổi rất nhanh, giờ đây, giao tiếp này không còn đơn thuần dùng những thiết bị cơ

học như chuột, bàn phím... mà còn có thể thông qua các biểu hiện của khuôn mặt. Bên cạnh đó, công nghệ ngày càng phát triển thì giá cả ngày càng giảm, thêm vào đó tốc độ xử lý của máy tính càng cao. Do đó, hệ thống xử lý khuôn mặt đang được phát triển rất mạnh mẽ.

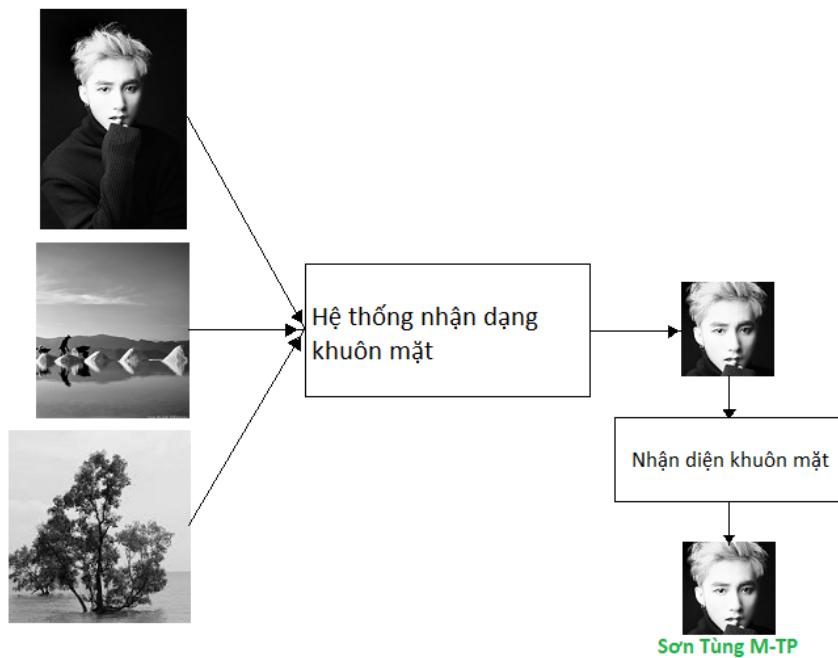
Hơn một thập kỷ qua, có rất nhiều công trình nghiên cứu về bài toán xác định khuôn mặt người từ ảnh đen trắng, ảnh xám, đến ảnh màu. Các nghiên cứu đi từ bài toán đơn giản là ảnh chỉ chứa một khuôn mặt người nhìn thẳng vào thiết bị thu hình và đầu ở tư thế thẳng đứng, cho đến ảnh màu với nhiều khuôn mặt người trong cùng ảnh, khuôn mặt có quay một góc nhỏ, hay bị khuất một phần, và với ảnh nền của ảnh phức tạp (ảnh chụp ngoài thiên nhiên) nhằm đáp ứng nhu cầu thật sự cần thiết của con người.

Bài toán xác định khuôn mặt người là một kỹ thuật máy tính để xác định các vị trí và các kích thước của khuôn mặt trong ảnh bất kỳ (ảnh kỹ thuật số). Kỹ thuật này nhận biết các đặc trưng khuôn mặt và bỏ qua những thứ khác như: tòa nhà, cây cối, cơ thể...

Xét về nguyên tắc chung, nhận dạng mặt có sự tương đồng rất lớn với nhận dạng vân tay và nhận dạng móng mắt, tuy nhiên sự khác biệt nằm ở bước trích chọn đặc trưng (feature extraction) của mỗi lĩnh vực. Trong khi nhận dạng vân tay và móng mắt đã đạt tới độ chín, tức là có thể áp dụng trên thực tế một cách rộng rãi thì nhận dạng mặt người vẫn còn nhiều thách thức và vẫn là một lĩnh vực nghiên cứu thú vị với nhiều người. So với nhận dạng vân tay và móng mắt, nhận dạng mặt có nguồn dữ liệu phong phú hơn (chúng ta có thể nhìn thấy mặt người ở bất cứ tấm ảnh, video clip nào liên quan tới con người trên mạng) và ít đòi hỏi sự tương tác có kiểm soát hơn (để thực hiện nhận dạng vân tay hay móng mắt, dữ liệu input lấy từ con người đòi hỏi có sự hợp tác trong môi trường có kiểm soát).

Hiện nay các phương pháp nhận dạng mặt được chia thành nhiều hướng theo các tiêu chí khác nhau: nhận dạng với dữ liệu đầu vào là ảnh tĩnh 2D (still image based FR) là phổ biến nhất, tuy nhiên tương lai có lẽ sẽ là 3D FR (vì việc bố trí nhiều camera 2D sẽ cho dữ liệu 3D và đem lại kết quả tốt hơn, đáng tin cậy

hơn), cũng có thể chia thành 2 hướng là: làm với dữ liệu ảnh và làm với dữ liệu video.



Hình 2.2 Hệ thống nhận diện khuôn mặt

Hệ thống nhận diện khuôn mặt có thể giúp:

- Tra cứu thông tin tội phạm.
- Giám sát bằng camera để phát hiện tội phạm tại các khu vực công cộng.
- Tìm trộm larc dựa vào các camera đặt ở các nơi công cộng.
- Phát hiện các nhân vật VIP đặt chân vào khách sạn.
- Phát hiện tội phạm tại các nơi công cộng.
- Có thể dùng trong khoa học để so sánh một thực thể với nhiều thực thể khác.
- Điểm danh, chấm công

2.2 Những khó khăn và thách thức trong bài toán nhận diện khuôn mặt.

Bài toán nhận diện khuôn mặt đã được đưa ra nghiên cứu từ những năm 70. Tuy nhiên, đây là một bài toán khó nên những nghiên cứu hiện tại vẫn chưa đạt

được kết quả mong muốn. Có thể kể đến những khó khăn của bài toán nhận diện khuôn mặt người như dưới đây:

- *Tư thế, góc chụp*: ảnh chụp khuôn mặt có thể thay đổi rất nhiều bởi vì góc chụp giữa camera và khuôn mặt. Chẳng hạn như: chụp thẳng, chụp chéo bên trái 45° , chụp từ trên xuống, chụp từ dưới lên,... Với các tư thế khác nhau, các thành phần trên khuôn mặt như mắt, mũi, miệng có thể bị khuất một phần hoặc thậm chí khuất hết.
- *Sự xuất hiện hoặc thiếu một số thành phần trên khuôn mặt*: các đặc trưng như râu mép, râu cằm, mắt kính... có thể xuất hiện hoặc không. Vấn đề này càng làm cho bài toán trở nên khó hơn rất nhiều.
- *Cảm xúc biểu hiện trên khuôn mặt*: cảm xúc có thể làm ảnh hưởng đáng kể lên các thông số của khuôn mặt. Chẳng hạn, cùng một khuôn mặt một người, nhưng có thể sẽ rất khác khi họ cười hoặc sợ hãi...
- *Sự che khuất*: khuôn mặt có thể bị che khuất một phần bởi các đối tượng khác hoặc một khuôn mặt khác.
- *Hướng của ảnh*: các ảnh của khuôn mặt có thể bị biến đổi rất nhiều với các góc quay khác nhau của trực camera. Chẳng hạn chụp với trực máy ảnh nghiêng làm cho khuôn mặt bị nghiêng so với trực thẳng của ảnh.
- *Điều kiện, chất lượng của ảnh*: ảnh được chụp trong các điều kiện khác nhau về: chiếu sáng, các loại máy ảnh (máy ảnh kỹ thuật số, máy hồng ngoại,...) độ phân giải của ảnh ảnh hưởng rất nhiều đến sự hiệu quả trong nhận diện khuôn mặt.

2.3 Hệ thống nhận diện khuôn mặt hoạt động như thế nào?

2.3.1 Cơ sở lý thuyết.

Nhận dạng khuôn mặt là một bài toán khá phức tạp, nó đòi hỏi một loạt các vấn đề cần thực hiện:

- Việc làm đầu tiên đó là cần phải tìm kiếm tất cả những khuôn mặt có trong bức hình.
- Focus vào từng khuôn mặt chắc chắn rằng chúng ta có thể nhận ra cùng một người từ các góc nhìn hoặc điều kiện sáng tối khác nhau.
- Lựa chọn những feature đặc trưng trên từng khuôn mặt - VD như đôi mắt to, khuôn mặt dài...
- So sánh những đặc trưng này với những người khác để chúng ta có thể biết được tên của họ.

Bộ não của con người có rất nhiều các nơron thần kinh để làm tất cả những điều này tự động và ngay lập tức. Quả thật, con người đã quá giỏi nhưng với máy tính, nó không có khả năng cao cấp như vậy. Do đó chúng ta cần phải dạy cho nó cách làm từng bước một cách riêng biệt.

Trước hết, để có thể phân tích khuôn mặt và nhận diện, cần phải tách khuôn mặt ra khỏi khung cảnh còn lại trước. Tiếp đến, chính là việc phân tích gương mặt. Chúng ta nhận diện gương mặt mọi người qua các đặc điểm, hệ thống nhận diện cũng vậy, nhưng nó thực hiện đánh giá các đặc điểm ở 1 level cao hơn, chính xác hơn là việc cảm nhận bằng mắt như của chúng ta, vì thị giác có thể bị đánh lừa. Mỗi khuôn mặt đều có nhiều điểm mốc, những phần lồi lõm tạo nên các đặc điểm của khuôn mặt. Các hệ thống nhận diện gương mặt định nghĩa những điểm này là những điểm nút. Mỗi mặt người có khoảng 80 điểm nút. Có thể nhận diện một số điểm nút như sau:

- Khoảng cách giữa hai mắt.
- Chiều rộng của mũi.
- Độ sâu của hốc mắt.
- Hình dạng của xương gò má.
- Độ dài của xương hàm.
- ...



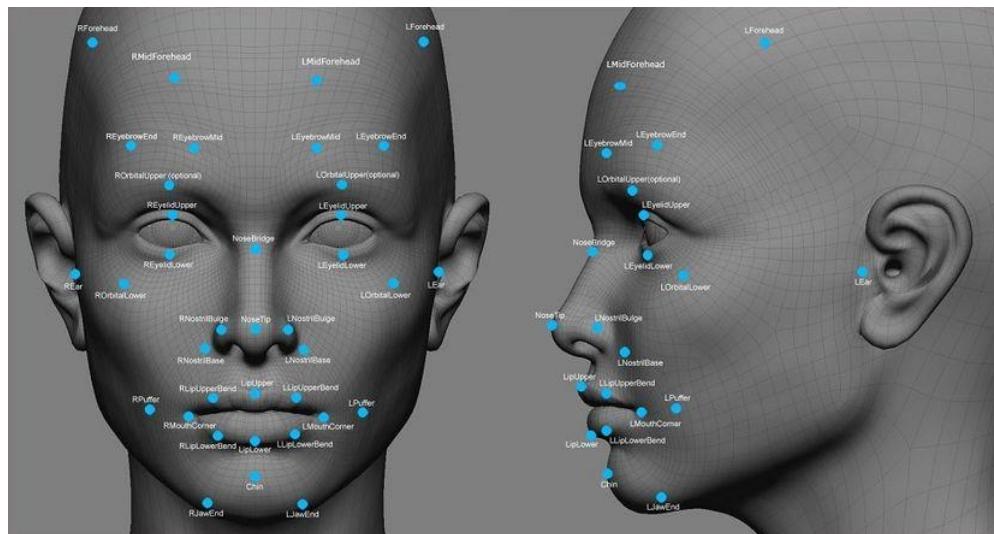
Hình 2.3 Các điểm nút đặc trưng trên một khuôn mặt.

Một số thuật toán nhận dạng khuôn mặt xác định các đặc điểm khuôn mặt bằng cách trích xuất các ranh giới, hoặc đặc điểm, từ một hình ảnh khuôn mặt của đối tượng. Từ đó các thuật toán sẽ trích xuất được các thông tin, và những thông tin này sau đó được sử dụng để tìm kiếm các hình ảnh khác với các thông tin phù hợp. Trong trường hợp sử dụng để nhận diện, cần phải lưu lại thông tin khuôn mặt để ghi nhớ trước. Các thuật toán sẽ đơn giản hóa một tập các hình ảnh khuôn mặt và sau đó nén dữ liệu khuôn mặt, chỉ lưu dữ liệu hình ảnh nào là hữu ích cho việc nhận dạng khuôn mặt.

Khi đó, muốn nhận diện sẽ so sánh hình ảnh mẫu với các dữ liệu khuôn mặt đã lưu. Các thuật toán nhận dạng có thể được chia thành hai hướng chính, là hình học, đó là nhìn vào tính năng phân biệt, hoặc trắc quang (đo sáng), là sử dụng phương pháp thống kê để ‘chưng cất’ một hình ảnh thành những giá trị và so sánh các giá trị với các mẫu để loại bỏ chênh lệch. Tuy nhiên, các trường hợp nhận diện thường không phải lúc nào cũng được đo đạc trong môi trường ổn định, có thể bị

ánh hưởng ngay chỉ bởi sự thiếu sáng, hay góc nghiêng của khuôn mặt, do đó ảnh hưởng đáng kể đến độ chính xác của kết quả.

Ngày nay, với sự phát triển của công nghệ, và với xu hướng 3D đang thịnh hành, thì cũng tương tự vậy, chúng ta có nhận dạng khuôn mặt 3 chiều. Và tất nhiên, nghe là biết kỹ thuật này sẽ cải thiện được độ chính xác nhờ việc phân tích được nhiều thông tin để so sánh hơn rồi. Kỹ thuật này sử dụng các cảm biến 3D để nắm bắt thông tin về hình dạng của khuôn mặt, rồi dùng các điểm nổi bật trên khuôn mặt – nơi những mô cứng và xương nhìn thấy rõ nhất như đường cong của hốc mắt, mũi và cằm — để nhận ra đối tượng. Các đặc điểm này là độc nhất đối với mỗi khuôn mặt và không thay đổi theo thời gian.



Hình 2.4 Công nghệ nhận diện khuôn mặt 3D.

Cách thức sử dụng độ sâu và trực của các phần trên khuôn mặt không bị ảnh hưởng bởi ánh sáng, vì thế việc nhận dạng khuôn mặt 3D có thể được sử dụng cả trong bóng tối và có thể nhận ra khuôn mặt từ nhiều góc độ khác nhau với độ chênh lệch lên tới 90 độ. Và nhờ có việc xác định khuôn mặt dựa trên hình ảnh 3 chiều, nên nhận dạng 3D có khả năng xác định 1 khuôn mặt từ nhiều góc nhìn hơn. Số lượng các điểm dữ liệu 3 chiều khiến cho độ chính xác tăng lên đáng kể bởi sự phát triển của các bộ cảm biến tinh vi giúp nắm bắt hình ảnh chụp khuôn mặt 3D được tốt hơn. Các cảm biến hoạt động bằng cách chiếu ánh sáng có cầu

trúc lên gương mặt. Hàng chục hoặc nhiều hơn nữa các bộ cảm biến hình ảnh này có thể được đặt lên cùng một con chip CMOS-mỗi cảm biến sẽ thu một phần khác nhau của hình ảnh.

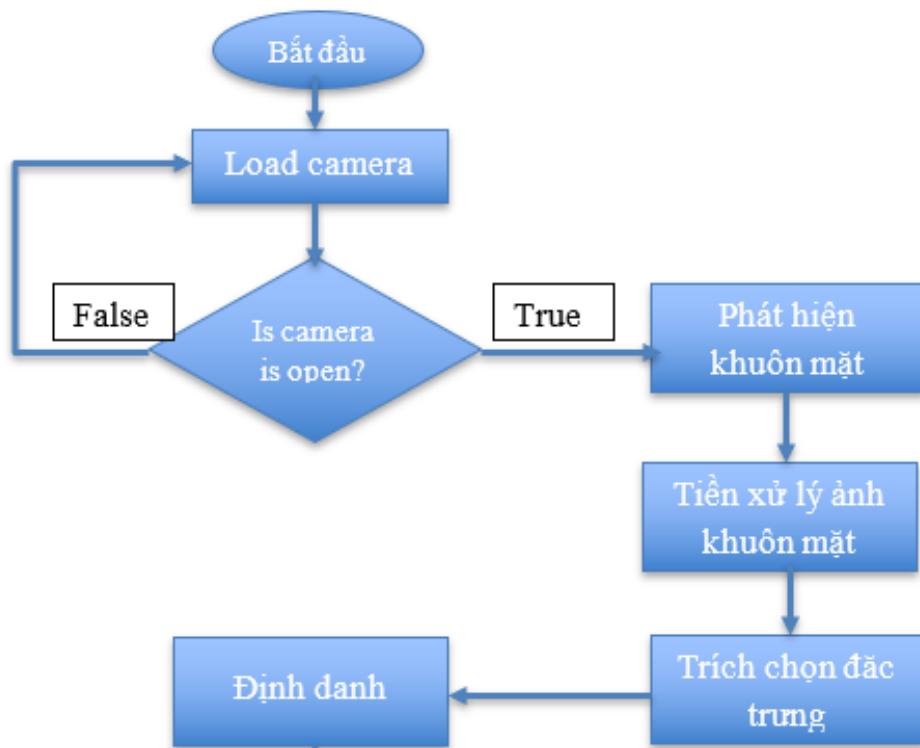
Hiện nay, chúng ta có thể thấy rõ nhất công nghệ nhận diện 3D này ngoài thực tế qua các sản phẩm điện thoại thông minh iPhone do Apple phát triển và sản xuất. Trên điện thoại iPhone có hỗ trợ FaceID, Apple tích hợp hàng loạt cảm biến ở mặt trước máy để có thể quét và tái tạo theo hình khối 3D sau đó ghi lại một cách chính xác nhất khuôn mặt, cuối cùng là mã hóa, lưu trữ vào bộ nhớ máy.

Chúng ta chỉ cần nhìn vào màn hình, FaceID sẽ tự động nhận dạng và mở khóa ngay lập tức.



Hình 2.5 FaceID trên điện thoại iPhone.

Sử dụng công nghệ nhận diện khuôn mặt 3D đòi hỏi kỹ thuật cao, tốn kém chi phí về trang thiết bị và thời gian nghiên cứu. Vì vậy, trong khóa luận này, sinh viên chọn công nghệ nhận diện khuôn mặt 2D, trích xuất từ ảnh, video để nhận diện khuôn mặt của một người



Hình 2.6 Sơ đồ hệ thống nhận diện khuôn mặt.

Các pha trong một hệ thống nhận dạng mặt: để xây dựng một hệ thống nhận dạng mặt, cũng không hề đơn giản, bước đầu tiên cần thực hiện là face detection, tức là phát hiện phần ảnh mặt trong dữ liệu input (ảnh, video) và cắt lấy phần ảnh mặt để thực hiện nhận dạng (face cropping), bước thứ hai là tiền xử lý ảnh (preprocessing) bao gồm các bước căn chỉnh ảnh (face image alignment) và chuẩn hóa ánh sáng (illumination normalization), tiếp đến là bước trích chọn đặc điểm (feature extraction), ở bước này một phương pháp trích chọn đặc điểm nào đó (Local Binary Pattern – LBP, EigenFace, FisherFace) sẽ được sử dụng với ảnh mặt để trích xuất các thông tin đặc trưng cho ảnh, kết quả là mỗi ảnh sẽ được biểu diễn dưới dạng một vector đặc điểm (feature vector), bước tiếp theo là bước nhận dạng (recognition) hay phân lớp (classification), tức là xác định danh tính (identity) hay nhãn (label) của ảnh – đó là ảnh của ai. Ở bước classification, thường thì phương pháp k-láng giềng gần nhất (k-nearest neighbor:kNN) sẽ được sử dụng, thực tế cho thấy việc dùng SVM (Support Vector Machine) không mang lại hiệu quả cao hơn mà còn chậm hơn.

2.3.1 Phát hiện khuôn mặt



Hình 2.7: Phát hiện các khuôn mặt trong ảnh.

Điều cần làm đầu tiên với bức ảnh/ khung hình chúng ta có đó chính là xác định xem trong bức ảnh/ khung hình đó có sự xuất hiện của bao người khuôn mặt (bao nhiêu người) và vị trí của chúng trong bức ảnh. Bài toán trở nên rất giống với bài toán xác định vật thể (Object Detection). Đây là một trong những bài toán rất khó bởi chúng ta sẽ cần nhiều kinh nghiệm cũng như lý thuyết về xử lý ảnh để có thể giải quyết được bước này. Cụ thể hơn, một số kỹ thuật cho bài toán này có thể kể đến như:

- Kỹ thuật xác định dựa vào các kiến thức con người (Knowledge-based):
Kỹ thuật này dựa vào các hiểu biết của con người về khuôn mặt để xác định được một khuôn mặt trong ảnh (Ví dụ như việc một khuôn mặt sẽ phải có mắt, mũi, miệng và khoảng cách giữa chúng thường sẽ phải thỏa mãn các ràng buộc nào đó,...). Điểm khó của kỹ thuật này nằm ở việc chúng ta sẽ phải xây dựng nên 1 bộ quy tắc. Nếu bộ quy tắc quá chung chung hay quá

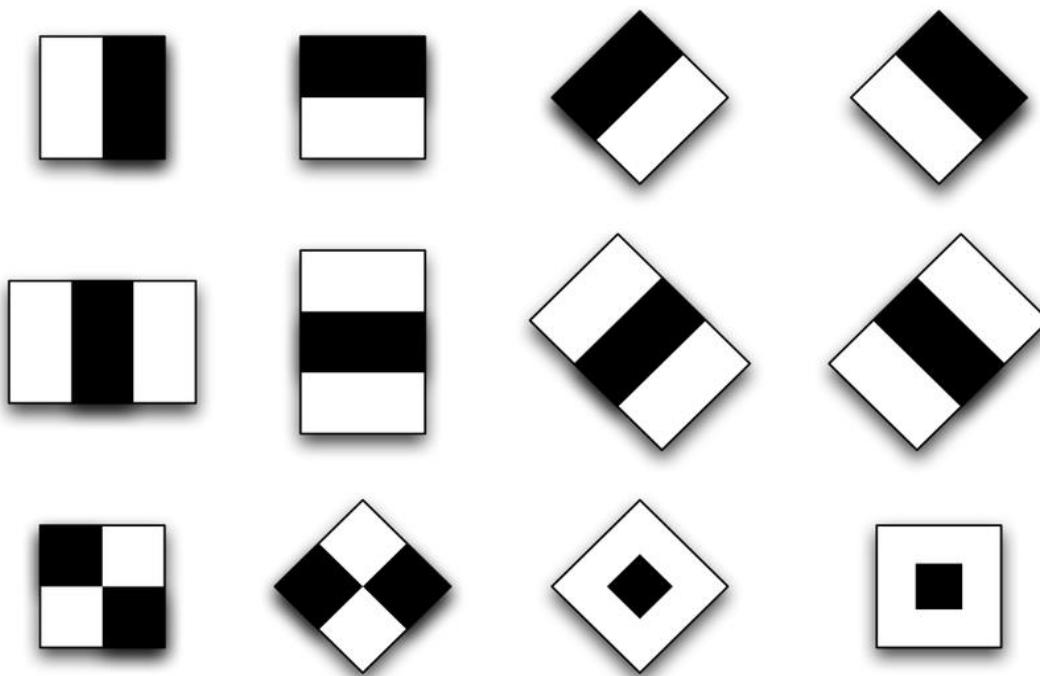
chặt chẽ thì đều không được vì nó sẽ dẫn tới việc nhận dạng nhầm hoặc không nhận dạng được.

- Kỹ thuật xác định dựa vào đặc tính khuôn mặt (Feature-based): Kỹ thuật này tạo ra một mô hình, sau đó chúng ta sẽ huấn luyện mô hình đó như một mô hình để phân loại (classifier) nhằm xác định trong các khung hình cắt ra từ 1 ảnh ban đầu, đâu là các vùng của một khuôn mặt. Điểm yếu lớn nhất của kỹ thuật này là về mặt thời gian. Có vấn đề này là do chúng ta sẽ phải lấy ra rất nhiều vùng trong 1 bức ảnh nhầm đưa qua classifier.
- Kỹ thuật xác định dựa vào mẫu cho trước (Template-Matching): Kỹ thuật này xác định được vị trí của một khuôn mặt trong bức ảnh dựa vào việc so sánh giữa các bức ảnh khuôn mặt chúng ta cho trước (Feature template) và các khung hình được cắt ra. Template-Matching rất dễ dàng để sử dụng tuy nhiên cũng gặp phải vấn đề về thời gian tương tự như kỹ thuật Feature-based ở trên.
- Kỹ thuật xác định dựa vào hình dáng (Appearance-Based): Đây là phương pháp sử dụng các phương pháp hình thái học kết hợp với phân tích từ mô hình machine-learning để xác định trực tiếp về vị trí của các vùng có khuôn mặt trong ảnh.

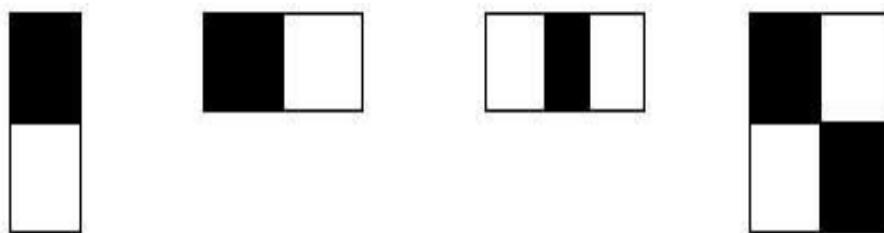
Phát hiện mặt người là bài toán cơ bản được xây dựng từ nhiều năm nay, có nhiều phương pháp được đưa ra như sử dụng khớp mẫu (template matching), mạng nơ-ron (neuron network)... Cho tới nay bài toán này hầu như được giải quyết dựa trên phương pháp sử dụng các đặc trưng Haar-like. Phương pháp này được cho là đơn giản và kết quả phát hiện là tương đối cao, lên tới 98%, các hãng sản xuất máy ảnh như Canon, Samsung... cũng đã tích hợp nó vào trong các sản phẩm của mình.

2.3.1.1 Tìm hiểu về các đặc trưng Haar-Like.

Các đặc trưng Haar-Like là những hình chữ nhật được phân thành các vùng khác nhau như hình:

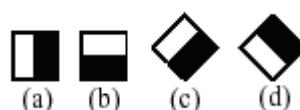
*Hình 2.8 Các đặc hình trung trong Haar-Like.*

Đặc trung do Viola và Jones công bố gồm 4 đặc trung cơ bản để xác định khuôn mặt người. Mỗi đặc trung Haar-Like là sự kết hợp của hai hay ba hình chữ nhật trắng hay đen như trong hình sau:

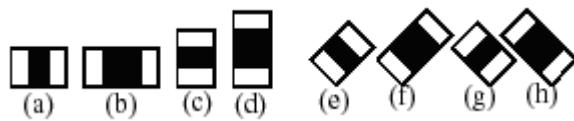
*Hình 2.9 Các đặc trung Haar-Like cho phát hiện khuôn mặt.*

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

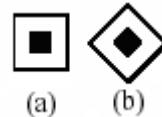
- Đặc trung cạnh(edge feature).



- Đặc trung đường(line feature)



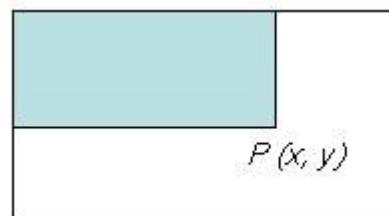
- Đặc trưng xung quanh tâm (center-surround features)



Dùng các đặc trưng trên, ta có thể tính được các giá trị của đặc trưng Haar-Like là sự chênh lệch giữa tổng của các pixel của vùng đen và vùng trắng như trong công thức sau:

$$f(x) = \text{Tổng}_\text{vùng_đen}(\text{các mức xám của pixel}) - \text{Tổng}_\text{vùng_trắng}(\text{các mức xám của pixel})$$

Viola và Jones đưa ra một khái niệm gọi là “Hình ảnh tích hợp” (Integral Image), là một mảng 2 chiều với kích thước bằng với kích thước của ảnh cần tính đặc trưng Haar-Like, với mỗi phần tử của mảng này được tính bằng cách tính tổng của điểm ảnh phía trên (dòng-1) và bên trái (cột-1) của nó.



Hình 2.10 Hình ảnh tích hợp (Integral Image).

Công thức tính Integral Image:

$$P(x, y) = \sum_{x' < x, y' < y} i(x', y')$$

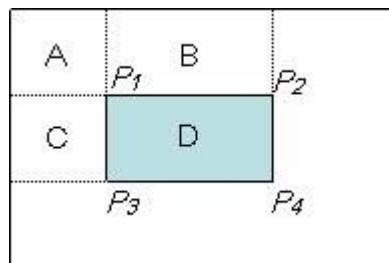
Sau khi tính được Integral Image, việc tính tổng các giá trị mức xám của một vùng bất kỳ nào đó trên ảnh thực hiện rất đơn giản theo cách sau:

Giả sử ta cần tính tổng giá trị mức xám của vùng D như hình dưới, ta có thể tính được như sau:

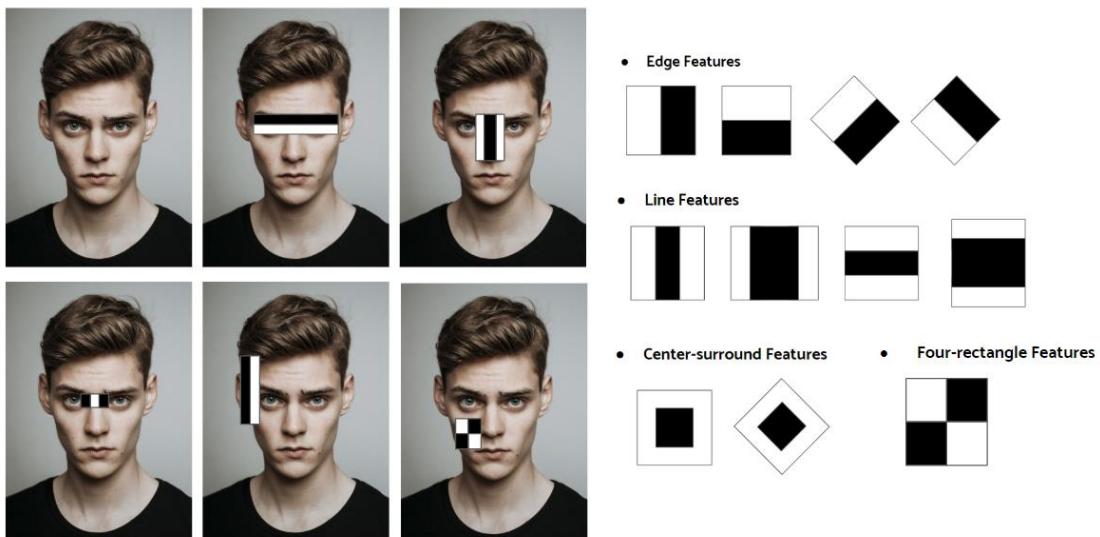
$$D = A + B + C + D - (A + B) - (A + C) + A$$

Với $A + B + C + D$ chính là giá trị tại điểm P4 trên Integral Image, tương tự như vậy $A+B$ là giá trị tại điểm P2, $A+C$ là giá trị tại điểm P3, và A là giá trị tại điểm P1. Vậy ta có thể viết lại biểu thức tính D ở trên như sau:

$$D = \underbrace{(x4,y4)}_{A + B + C + D} - \underbrace{(x2,y2)}_{(A+B)} - \underbrace{(x3,y3)}_{(A + C)} + \underbrace{(x1,y1)}_A$$



Trong đặc trưng Haar-Like, các bộ lọc chiếm một phần cửa sổ trượt. Các cửa sổ trượt này chạy khắp hình ảnh và các bộ lọc Haar-Like này sẽ chạy để tìm các vùng giống với khuôn mặt nhất.



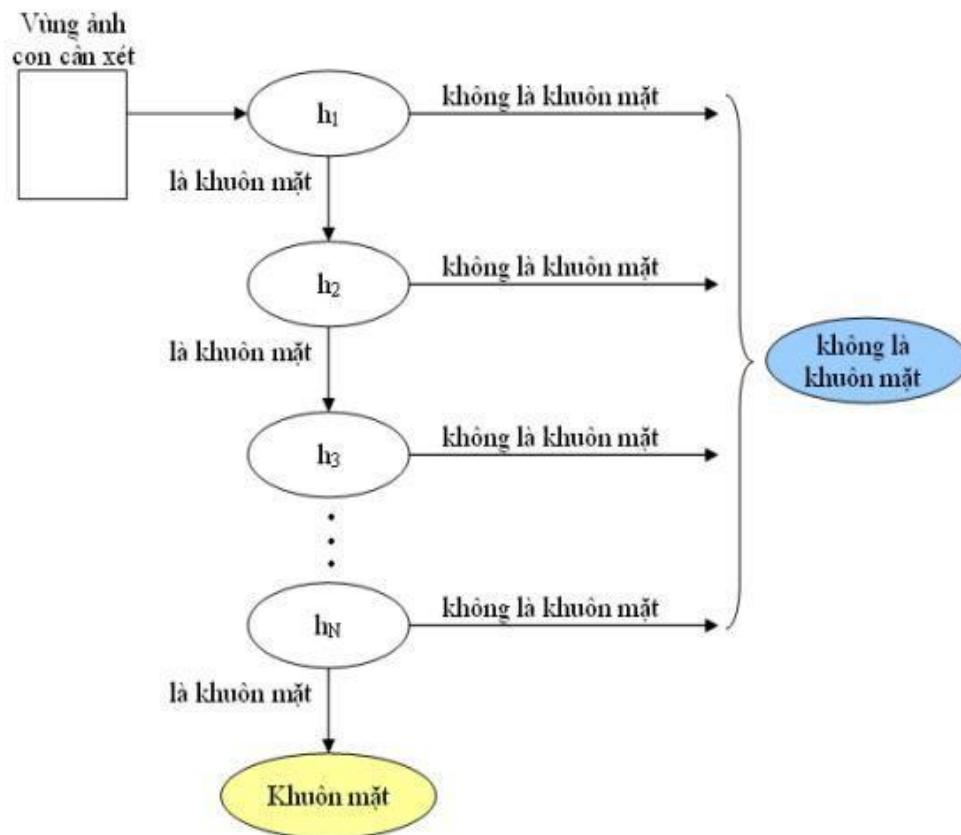
Hình 2.11: Các bộ lọc Haar-Like trượt trên hình ảnh.

Dựa vào tập dữ liệu khuôn mặt được training trước đó. Hệ thống sẽ phân loại để xe vùng cửa sổ được trượt qua có phần nào giống khuôn mặt hay không.

2.3.1.2 AdaBoost.

AdaBoost là một bộ phân loại mạnh phi tuyến phức dựa trên hướng tiếp cận boosting được Freund và Schapire đưa ra vào năm 1995. Adaboost cũng hoạt động trên nguyên tắc kết hợp tuyến tính các weak classifiers để hình thành một trong các classifiers.

Viola và Jones dùng AdaBoost kết hợp các bộ phân loại yếu sử dụng các đặc trưng Haar-like theo mô hình phân tầng (cascade) như sau:



Hình 2.12 Mô hình phân tầng Haar-Cascade.

Trong đó, $h(k)$ là các bộ phân loại yếu, được biểu diễn như sau:

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

Với:

x : cửa sổ con cần xét.

θ_k : ngưỡng.

f_k : giá trị của đặc trưng Haar-Like.

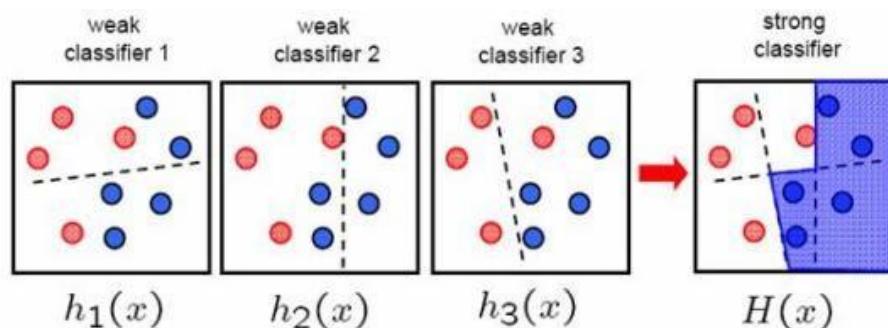
p_k : hệ số quyết định chiều của phương trình.

AdaBoost sẽ kết hợp các bộ phân loại yếu thành bộ phân loại mạnh như sau:

$$H(x) = \sum (\alpha_1 h_1(x) + \alpha_2 h_2(x) + \dots + \alpha_n h_n(x))$$

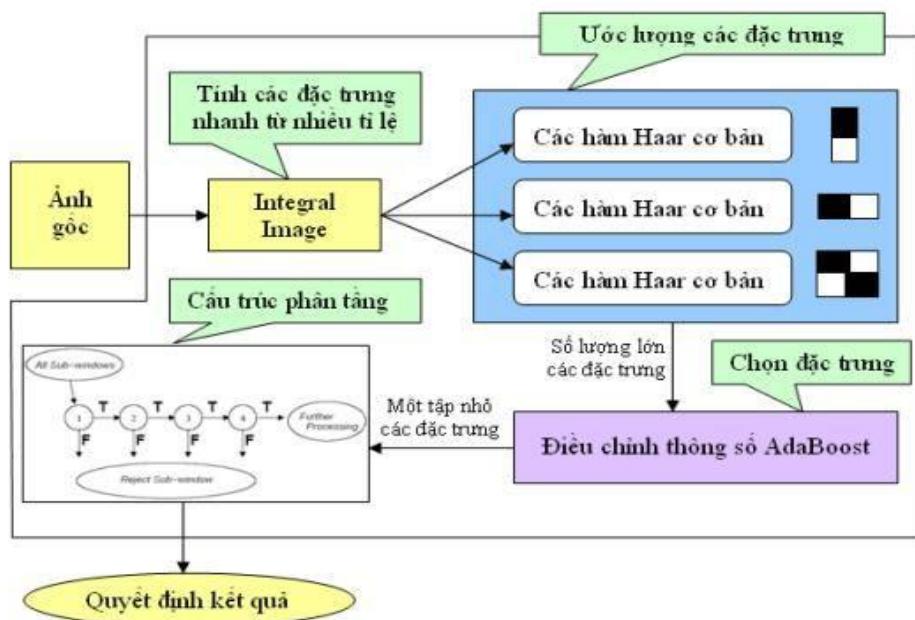
Với: $\alpha_t \geq 0$ là hệ số chuẩn hóa cho các bộ phân loại yếu.

Đây là hình ảnh minh họa việc kết hợp các bộ phân loại yếu thành bộ phân loại mạnh.



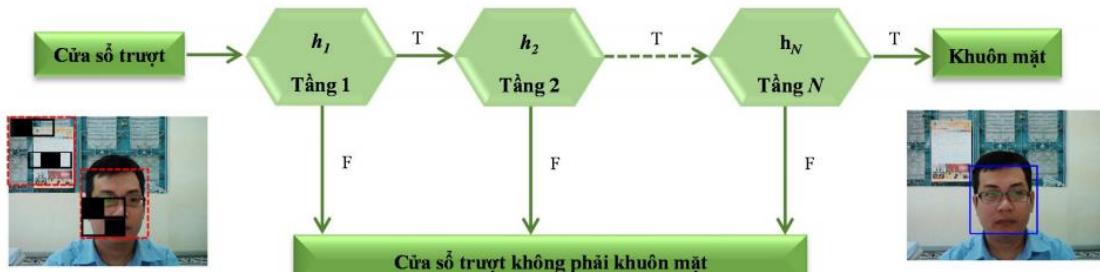
Hình 2.13: kết hợp các bộ phân loại yếu thành bộ phân loại mạnh.

Chúng ta có thể tóm tắt sơ đồ về phát hiện khuôn mặt trong ảnh dưới đây:



Hình 2.14 Sơ đồ hệ thống phát hiện khuôn mặt Viola Jones.

Phương pháp định vị khuôn mặt bằng cách di chuyển cửa sổ trượt trên ảnh (từ trái sang phải, từ trên xuống dưới), rút trích đặc trưng Haar-like của cửa sổ đang xét, đưa vào mô hình phân lớp Adaboost theo thứ tự từ tầng 1 đến tầng thứ N của mô hình phân tầng CBC, nếu ở tầng thứ t mô hình h_t phân lớp vùng ứng viên không phải là khuôn mặt người thì cửa sổ trượt tiếp tục đến vị trí tiếp theo trên ảnh, nếu vùng cửa sổ đang quét được mô hình h_t ở tầng thứ t phân lớp là khuôn mặt người thì vùng của sổ tiếp tục chuyển đến tầng thứ $t+1$ để xét vùng ứng viên có phải mặt người hay không, quá trình tiếp tục cho đến khi tầng N mô hình h_N phân lớp là khuôn mặt người thì vùng cửa sổ đang quét được xác định là khuôn mặt người.



Hình 2.15 Mô hình phân tầng định vị khuôn mặt trong ảnh.

2.3.2 Nhận diện khuôn mặt.

Sau khi phát hiện được các khuôn mặt trong ảnh, chúng ta sẽ thực hiện thuật toán để trích chọn các đặc trưng của khuôn mặt đó để nhận diện, xác định các khuôn mặt này là của ai. Về cơ bản là tính độ giống/ khác nhau giữa các khuôn mặt chúng ta lấy được với các khuôn mặt chúng ta đã training trước đó.

Về phần nhận diện khuôn mặt, chúng ta có thể chia làm hai pha nhỏ nữa là trích chọn đặc trưng của khuôn mặt và phân lớp các khuôn mặt.

Trên thực tế người ta hay chia các phương pháp nhận dạng mặt ra làm 3 loại: phương pháp tiếp cận toàn cục (global, như Eigenfaces-PCA, Fisherfaces-LDA), phương pháp tiếp cận dựa trên các đặc điểm cục bộ (local feature based, như LBP, Gabor wavelets) và phương pháp lai (hybrid, là sự kết hợp của hai

phương pháp toàn cục và cục bộ). Phương pháp dựa trên các đặc điểm cục bộ đã được chứng minh là ưu việt hơn khi làm việc trong các điều kiện không có kiểm soát và có thể nói rằng lịch sử phát triển của nhận dạng mặt là sự phát triển của các phương pháp trích chọn đặc trưng (feature extraction methods) được sử dụng trong các hệ thống dựa trên feature based. Các ứng dụng cụ thể của nhận dạng mặt dựa trên 2 mô hình nhận dạng: identification (xác định danh tính, bài toán 1-N), và verification (xác thực danh tính, bài toán 1-1). Trong bài toán identification, ta cần xác định danh tính của ảnh kiểm tra, còn trong bài toán verification ta cần xác định 2 ảnh có cùng thuộc về một người hay không.

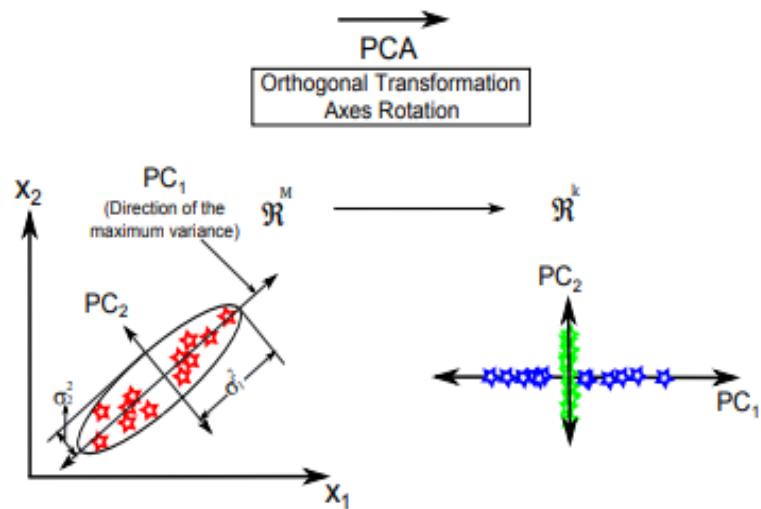
Hiện tại có nhiều phương pháp để trích chọn đặc trưng của khuôn mặt, mỗi thuật toán đều có các ưu điểm và nhược điểm riêng. Chúng ta sẽ làm rõ một số thuật toán phổ biến trong trích chọn đặc trưng ở phần 2.4 dưới đây.

2.4 Các thuật toán phổ biến trong nhận diện khuôn mặt.

2.4.1 Eigen Face Recognizer.

2.4.1.1 Tìm hiểu về PCA.

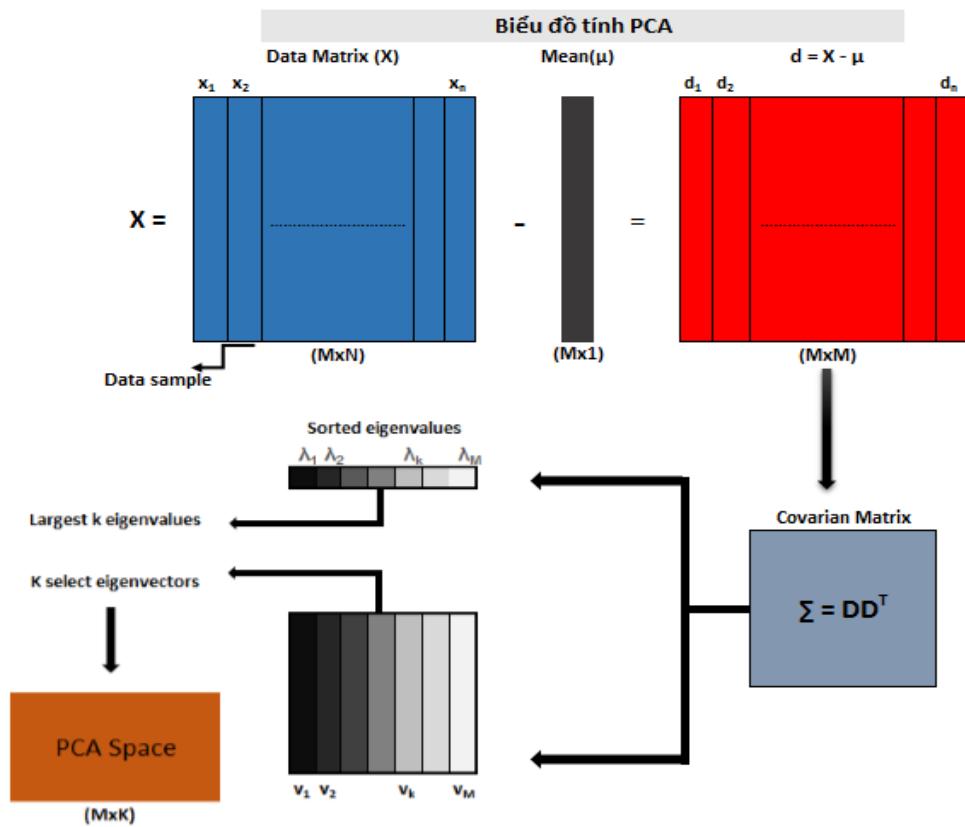
Principal Component Analysis (PCA) là một trong những phương pháp giảm chiều dữ liệu (Dimensionality reduction techniques) phổ biến nhất và được sử dụng trong nhiều lĩnh vực khác nhau. PCA có nhiều ứng dụng như tìm mối tương quan giữa các biến (relationship between observation), trích xuất những thông tin quan trọng từ data, phát hiện và loại bỏ outlier và giảm chiều chiểu dữ liệu. Ý tưởng của phương pháp PCA là tìm ra một không gian mới để chiếu (project) data sao cho variation giữ lại là nhiều nhất. Ta có thể hình dung qua hình vẽ dưới đây.



Hình 2.16 Mô hình giảm chiều dữ liệu với PCA.

Có 2 phương pháp tiếp cận PCA là covarian matrix và SVD chúng ta chỉ tìm hiểu về covarian matrix trong bài này.

Phương pháp Covarian matrix : Các bước thực hiện thuật toán như sau :



Hình 2.17: Mô hình biểu đồ tính PCA.

- X data có chiều $M \times N$ (với N là số sample, M là số feature).
- Tính mean của X : $\mu = \frac{1}{N} \cdot \sum_{i=1}^N x_i$.
- Trừ X với mean của X : $D = \{d_1, d_2, \dots, d_N\} = \sum_{i=1}^N x_i - \mu$.
- Tính toán covarian: $\sum_{N-1}^1 D \cdot D^T$
- Tính toán EigenVector V và EigenValue λ của Covarian \sum .
- Sắp xếp EigenValue tương ứng với EigenVector theo thứ tự λ giảm dần.
- Chọn những EigenVector tương ứng với EigenValue lớn nhất $W = v_1, v_2, \dots, v_k$. EigenVector W sẽ làm đại diện để chiểu X vào không gian PCA.
- Tất cả mẫu X sẽ được chiểu vào không gian nhỏ hơn theo công thức $Y = W^T \cdot D$.

Variable	X	μ	D	$D \cdot D^T$	λ_k	v_k	V	Y
Dimension	$(M \times N)$	$(M, 1)$	(M, N)	$(M \times M)$	scale	$(M \times 1)$	$(M \times M)$	(k, N)
M	số feature							
N	Số sample							
μ	Mean theo feature							
D	$X - \mu$							
$D \cdot D^T$	Covarian							
λ_k	Eigen Value thứ k							
v_k	Eigen Vector thứ k							
V	Matrix toàn bộ Eigen Vector							

Bảng 2.1 Bảng chú thích các biến và kích thước của biến trong PCA.

Để xây dựng không gian nhỏ hơn (từ M thành k), trong đó k là số eigen value mà ta chọn. Khi đó không gian PCA được định nghĩa là $W = v_1, v_2, \dots, v_k$. Ta viết lại biến Y là project của X qua W như sau: $Y = W^T \cdot D = \sum_{i=1}^N x_i - \mu$.

2.4.1.2 Tìm hiểu về EigenFace.

Được giới thiệu vào năm 1987, thuật toán Eigenfaces tuy đã cũ nhưng vẫn rất đáng để tìm hiểu. Cần biết rằng thuật toán Eigenfaces không hề bị giới hạn trong bài toán Face Recognition, nó có thể được áp dụng cho bất kỳ tập hợp đối tượng nào.

Dataset chúng ta cần có là một tập hợp các bức ảnh chứa khuôn mặt của những người cần nhận diện, càng nhiều ảnh cho mỗi người càng tốt. Mỗi bức ảnh cần có kích thước giống nhau và vị trí của các đặc trưng trên khuôn mặt phải tương ứng trên từng bức ảnh. Các bức ảnh phải được resize về một kích thước đồng nhất.

EigenFace hiểu một cách đơn giản là nó dùng PCA để trích chọn đặc trưng sau đó mới đưa vào model để training. Mỗi image có chiều chǎng hạn $28 \times 28 = 784$ pixel nếu đưa hết vào model thì có một số nhược điểm sau :

- Thời gian training lâu vì chiều dữ liệu lớn.
- Không phải tất cả các vị trí trên image đều quan trọng.

Vì vậy PCA giúp ta khắc phục các nhược điểm này , nó giúp ta giảm chiều dữ liệu mà vẫn giữ lại được những thông tin quan trọng trên image.

Để thực hiện Eigenfaces, mỗi bức ảnh có kích thước $W \times H$ trong dataset sẽ được chuyển về dạng vector $[W \times H]$ -dim. Nếu dataset chứa N bức ảnh thì chúng ta sẽ tạo được một ma trận M có kích thước $N \times [W \times H]$.

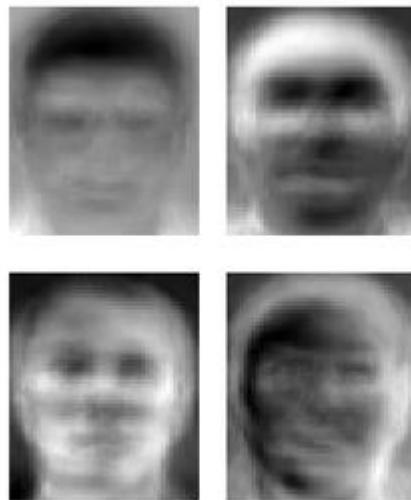
12	43	7
5	250	45
10	58	76
98	89	55

12	43	7	5	250	45	10	58	76	98	89	55
----	----	---	---	-----	----	----	----	----	----	----	----

Khi đã có được ma trận M chứa tất cả các vector biểu diễn khuôn mặt trong dataset, chúng ta sẽ áp dụng Principal Component Analysis (PCA) để giảm chiều dữ liệu xuống còn K -dim thay vì $[W \times H]$ -dim (với K nhỏ hơn nhiều so với $W \times H$). Như vậy ma trận M sẽ được biến đổi thành ma trận Z mới có kích thước $N \times K$. Lưu ý rằng: số chiều dữ liệu K được giữ lại phải thỏa mãn điều kiện: $K < \min(N, W \times H)$ để đảm bảo giữ được tối thiểu các đặc trưng của khuôn mặt.

Ví dụ: Chúng ta có một tập hợp gồm 200 bức ảnh khuôn mặt có kích thước 360×240 , từng bức ảnh sẽ được chuyển về dạng vector 86400-d. Như vậy, bức ảnh này được biểu diễn bởi 86400 giá trị trong hệ cơ sở đơn vị, với kích thước vector như vậy thì thời gian tính toán sẽ rất lớn. Để khắc phục nhược này, chúng

ta sử dụng PCA để giảm chiều dữ liệu, ví dụ ở đây chúng ta chọn giữ lại 100 chiều ($K=100$), lúc này các bức ảnh sẽ được biểu diễn lại bởi 100 giá trị ở một cơ sở khác được tạo bởi 100 Eigenvectors (còn gọi là Eigenfaces).



Hình 2.18 Khuôn mặt được giảm chiều dữ liệu với PCA.

Khi kích thước chỉ còn K -dim, các vector có thể dễ dàng được phân loại bởi k-NN (sử dụng khoảng cách Euclidean) hoặc SVM classifier. Thông thường, chúng ta sử dụng khoảng cách Euclidean giữa hai vectơ để tìm vectơ láng giềng gần nhất tương ứng với khuôn mặt trong tập dữ liệu training phù hợp nhất với khuôn mặt cần nhận diện. Cách tính khoảng cách Euclidean sẽ được trình bày trong phần dưới thuật toán Mẫu nhị phân cục bộ phần 2.4.2

Các bước thực hiện thuật toán EigenFace có thể tóm tắt lại như sau:

- Chuẩn bị dữ liệu : Face nên được căn chỉnh và có cùng kích thước $N \times N$ sau đó chuẩn hóa bằng cách chia 255.
- Image sau đó được làm phẳng (Flatten) thành $1 \times N^2$ pixel, chúng ta có M image nên data sẽ có chiều $M \times N^2$.
- Sau đó chúng ta tính mean và tính toán covariance như ở thuật toán PCA ở trên.
- Điểm khác biệt ở đây là covarian có chiều $N^2 \times N^2$ quá lớn để tính trực tiếp eigen vector và eigen value nên người ta sẽ tính eigen

vector của $M \times M$ (vì $M \times M$ có kích thước nhỏ hơn nhiều so với $N^2 \times N^2$). Sau đó tính ngược lại cho $N^2 \times N^2$.

- Cuối cùng chọn số eigen vector để chiếu data sang không gian mới lấy nó là đặc trưng để training.

Điểm trừ lớn nhất của Eigenfaces là để nhận được kết quả tốt nhất, các bức ảnh khuôn mặt phải được căn chỉnh sao cho vị trí của các đặc trưng trên khuôn mặt phải tương ứng trên mọi bức ảnh trong dataset, và tất nhiên điều này là không thể trong thực tế.

2.4.2 Local Binary Pattern Histogram.

2.4.2.1 Local Binary Pattern (LBP) là gì?

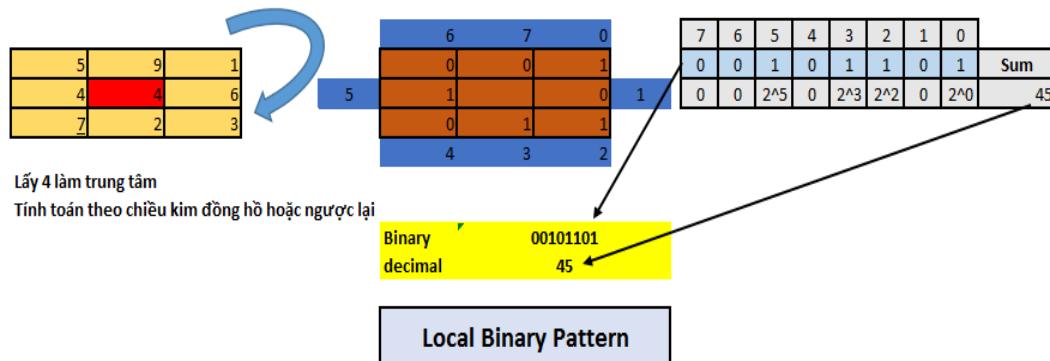
Với các thuật toán tiếp cận toàn cục như EigenFace, cường độ ánh sáng, độ phân giải của ảnh hay cảm xúc khuôn mặt ảnh hưởng rất lớn đến độ chính xác của thuật toán. Vì vậy, trong nhiều trường hợp, hai thuật toán trên tỏ ra kém hiệu quả. Với tính chất đẽ tài mà sinh viên chọn, thuật toán Local Binary Pattern Histogram (LBPH) tỏ ra hiệu quả hơn so với thuật toán EigenFace.

LBP dịch tiếng Việt là “mẫu nhị phân cục bộ” được Ojala trình bày vào năm 1996 như một đơn vị đo độ tương phản cục bộ của ảnh. Nhiều nghiên cứu đã xác định thêm rằng, khi LBP kết hợp với HOG, nó sẽ cải thiện hiệu suất đáng kể. Việc sử dụng LBP với histogram, chúng ta có thể biểu diễn hình ảnh khuôn mặt bằng một vectơ dữ liệu đơn giản.

Ý tưởng cơ bản của nó là mô phỏng lại cấu trúc cục bộ (local texture) của image bằng cách so sánh mỗi pixel với các pixel lân cận nó (neighbors). Ta sẽ đặt một pixel là trung tâm(center) và so sánh với các pixel lân cận với nó, nếu pixel trung tâm lớn hơn hoặc bằng pixel lân cận thì nó sẽ trả về giá trị 1, ngược lại 0. Ví dụ chúng ta lấy bán kính 8 pixel lân cận thì lbp sẽ có dạng 00101101, là một chuỗi nhị phân đẽ đơn giản và dễ đọc hơn ta sẽ chuyển về dạng decimal là bằng 45.

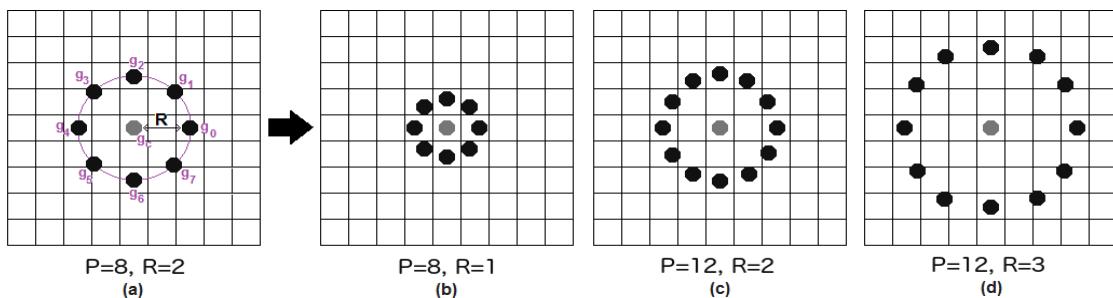
Lưu ý: Chúng ta có thể quy định việc tính toán theo chiều ngược chiều kim đồng hồ và giá trị pixel lân cận là 1 hoặc 0 so với giá trị pixel trung tâm. Ở đây

sinh viên chọn tính theo chiều kim đồng hồ và nếu pixel trung tâm lớn hơn hoặc bằng pixel lân cận thì nó sẽ trả về giá trị 1, ngược lại 0.



Hình 2.19: Tính toán giá trị LBP.

Cách tính này có hạn chế đó là chỉ giới hạn 3×3 pixel không đủ để mô tả các cấu trúc quy mô lớn (large scale) nên người ta mở rộng khái niệm LBP bằng cách định nghĩa thêm 2 tham số là (P, R) trong đó P là số pixel lân cận xem xét và R là bán kính ta quét từ pixel trung tâm. Như hình bên dưới.



Hình 2.20: Giá trị bán kính R và số pixel lân cận P .

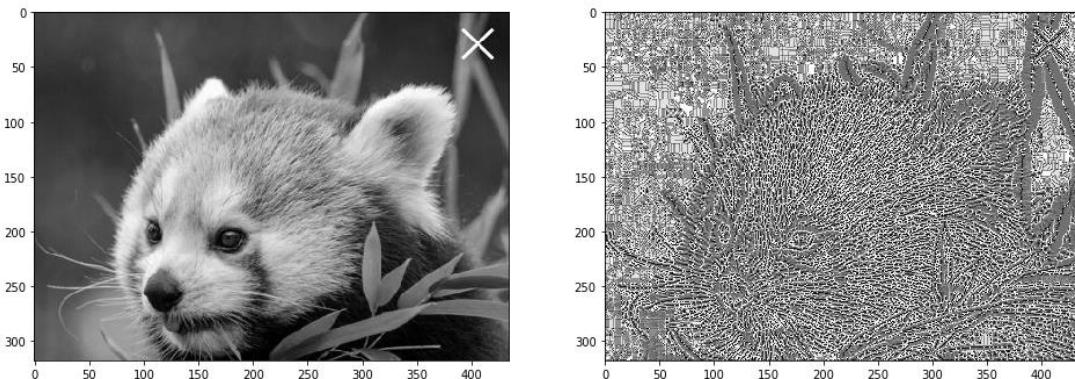
Công thức tính như sau:

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

Trong đó:

$$S(x) = \begin{cases} 1, & \text{if } x \geq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Với mỗi điểm ảnh ta xét sau khi tính toán thì sẽ ra một giá trị nhị phân, chuyển giá trị nhị phân này thành giá trị thập phân và các giá trị thập phân này sẽ tạo thành một bức ảnh xám.

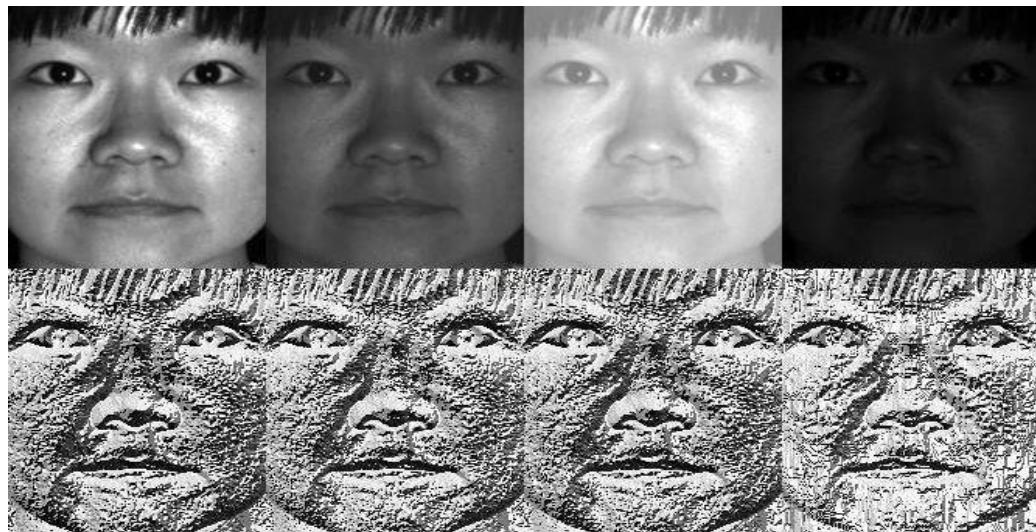


Hình 2.21: Hình ảnh gốc (trái) và sau khi được biểu diễn bằng LBP.

2.4.2.2 LBPH trong nhận diện khuôn mặt.

Sau khi có được các khuôn mặt được cắt ra từ hình ảnh gốc. Bước tính toán đầu tiên là chúng ta cần phải tính toán LBPH của các ảnh này để tạo ra một hình ảnh trung gian mô tả hình ảnh gốc. Để làm được điều này, thuật toán sử dụng khái niệm cửa sổ trượt dựa trên bán kính R và các pixel lân cận P.

Trong phương pháp LBP để nhận diện khuôn mặt, sự xuất hiện của mã LBP trong ảnh được thu thập thành biểu đồ. Việc phân loại sau đó được thực hiện bằng cách tính các điểm tương đồng biểu đồ đơn giản. Tuy nhiên, xem xét một cách tiếp cận tương tự cho biểu diễn hình ảnh khuôn mặt dẫn đến mất thông tin không gian và do đó, người ta nên mã hóa thông tin kết cấu trong khi vẫn giữ lại vị trí của chúng. Một cách để đạt được mục tiêu này là sử dụng các mô tả kết cấu LBP để xây dựng một số mô tả cục bộ của khuôn mặt và kết hợp chúng thành một mô tả toàn cầu. Những mô tả địa phương như vậy đã được quan tâm gần đây là điều dễ hiểu với những hạn chế của các đại diện tổng thể. Các phương pháp dựa trên tính năng cục bộ này mạnh mẽ hơn đối với các biến thể trong tư thế hoặc chiếu sáng so với các phương pháp tổng thể.

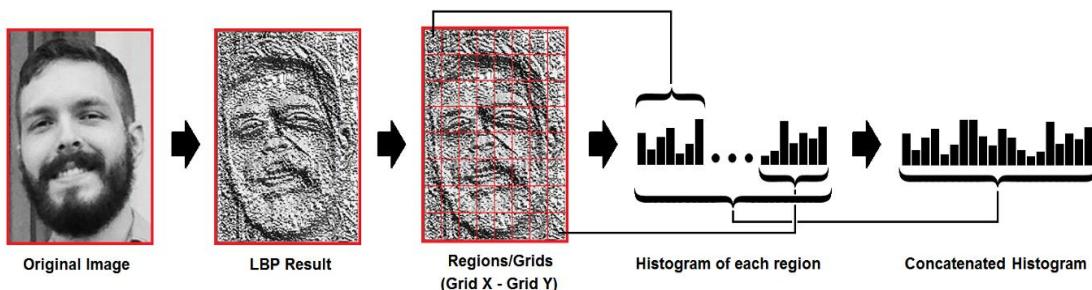


Hình 2.22: Hình ảnh khuôn mặt được mô tả với LBPH trong các điều kiện sáng khác nhau.

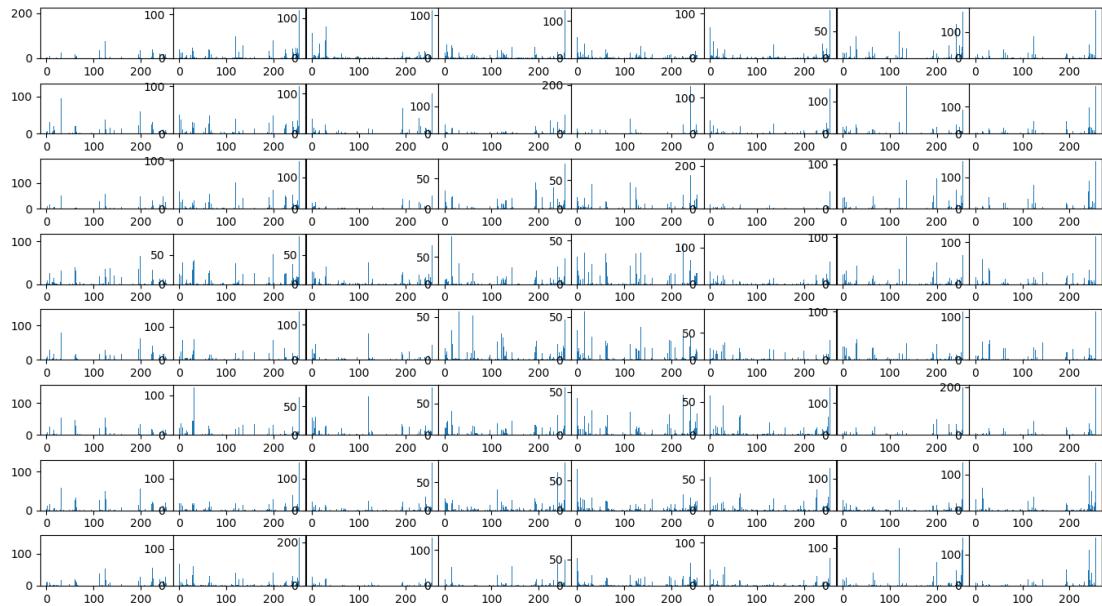
Qua hình trên cho ta thấy, ảnh với các mức sáng khác nhau khi được biểu diễn với LBPH cho ra cùng một kết quả vì giá trị tại một pixel phụ thuộc vào giá trị các pixel lân cận nó. Vì vậy, thuật toán LBPH trong nhận diện khuôn mặt tỏ ra hiệu quả trong thực tế vì cường độ ánh sáng trong thực tế luôn thay đổi.

Phương pháp cơ bản cho mô tả khuôn mặt dựa trên LBP được đề xuất bởi Ahonen (2006) như sau: Hình ảnh khuôn mặt được chia thành các vùng cục bộ và mô tả kết cấu LBP được trích xuất từ mỗi vùng một cách độc lập. Các mô tả sau đó được ghép nối để tạo thành một mô tả toàn bộ về khuôn mặt.

Sau khi tạo được ảnh trung gian biểu diễn ảnh gốc khuôn mặt, chúng ta đến bước trích xuất biểu đồ.



Hình 2.23: Chia khuôn mặt thành nhiều block cục bộ.



Hình 2.24 Các biểu đồ của từng khối (block)

Dựa vào hình ảnh trên, chúng ta có thể trích xuất biểu đồ của từng vùng như sau:

- Khi chúng ta có một hình ảnh ở thang độ xám, mỗi biểu đồ (từ mỗi block) sẽ chỉ chứa 256 vị trí ($0 \sim 255$) đại diện cho các lần xuất hiện của từng cường độ pixel.
- Sau đó, chúng ta cần nối từng biểu đồ để tạo ra biểu đồ mới và lớn hơn. Giả sử chúng ta có các lưới 8×8 , chúng ta sẽ có các vị trí $8 \times 8 \times 256 = 16.384$ trong biểu đồ cuối cùng. Biểu đồ cuối cùng đại diện cho các đặc điểm của hình ảnh khuôn mặt gốc.

Sau khi có được biểu đồ đại diện cho các đặc điểm của hình ảnh khuôn mặt gốc, bước tiếp theo chúng ta sẽ tiến hành nhận dạng khuôn mặt này.

Trong bước này, thuật toán đã được đào tạo. Mỗi biểu đồ được tạo ra được sử dụng để thể hiện mỗi hình ảnh từ tập dữ liệu training. Vì vậy, khi đầu vào cung cấp một hình ảnh mới, chúng ta sẽ thực hiện thuật toán LBPH cho hình ảnh này và tạo ra một biểu đồ mới đại diện cho hình ảnh đó.

Để tìm hình ảnh phù hợp với hình ảnh đầu vào, chúng ta chỉ cần so sánh hai biểu đồ và trả lại hình ảnh với biểu đồ gần nhất. Chúng ta có thể sử dụng các

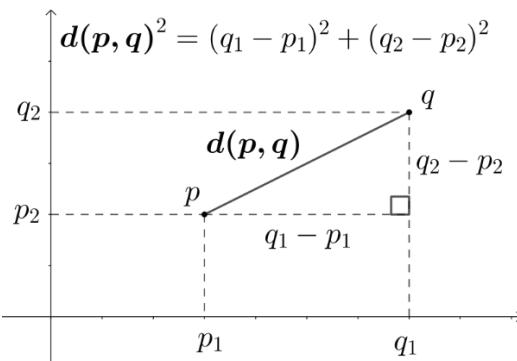
cách tiếp cận khác nhau để so sánh biểu đồ (tính khoảng cách giữa hai biểu đồ). Ví dụ: khoảng cách euclidean, chi bình phương, giá trị tuyệt đối, và nhiều phương pháp khác. Trong bài luận này, sinh viên sử dụng khoảng cách Euclidean (khá phổ biến).

Khoảng cách Euclidean là khoảng cách giữa hai điểm trong không gian.

Ví dụ:

- Với không gian hai chiều thì khoảng cách giữa hai điểm p, q là độ dài đoạn thẳng \overline{pq} được tính bằng công thức Pytago.

$$D(p, q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}$$

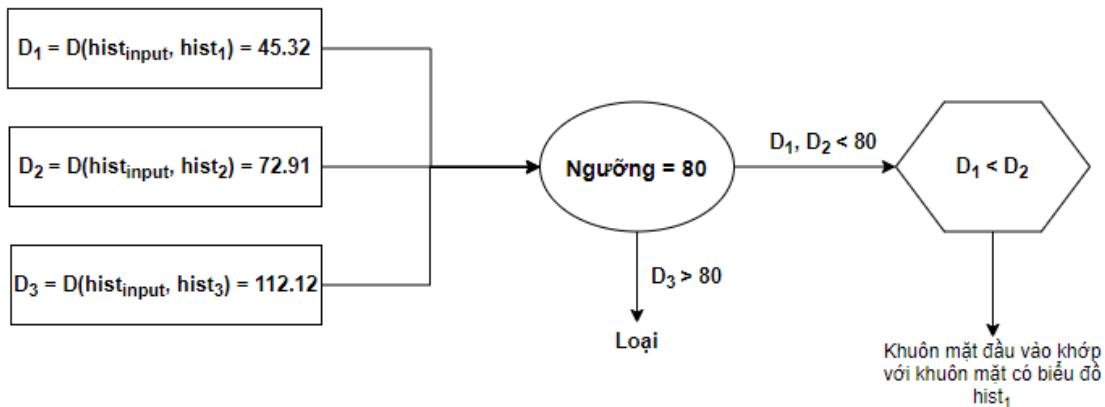


- Với không gian n chiều, hay khoảng cách giữa hai histogram thì công thức tính khoảng cách Euclidean được tổng quát như sau:

$$D = \sqrt{\sum_{i=1}^n (hist1_i - hist2_i)^2}$$

Sau khi tính được khoảng cách giữa hai biểu đồ, chúng ta có thể coi khoảng cách này là sự tương đồng giữa hai khuôn mặt. Khoảng cách D càng nhỏ thì hai khuôn mặt càng tương đồng nhau, có nghĩa là tỷ lệ hai khuôn mặt này là của cùng một người càng cao. Sau đó, chúng ta có thể quy định một “ngưỡng” khoảng cách để tự động ước lượng nhận diện khuôn mặt. Nếu khoảng cách D nhỏ hơn “ngưỡng” thì được coi như khuôn mặt đầu vào phù hợp với khuôn mặt được so sánh.

Ví dụ: ảnh khuôn mặt cần nhận diện có histogram là $hist_{input}$, các ảnh training có histogram là $hist_1, hist_2, hist_3$. D_1, D_2, D_3 lần lượt là khoảng cách euclidean giữa các histogram của ảnh cần nhận diện với từng ảnh training.



Hình 2.25 Chọn khuôn mặt phù hợp bằng khoảng cách euclidean.

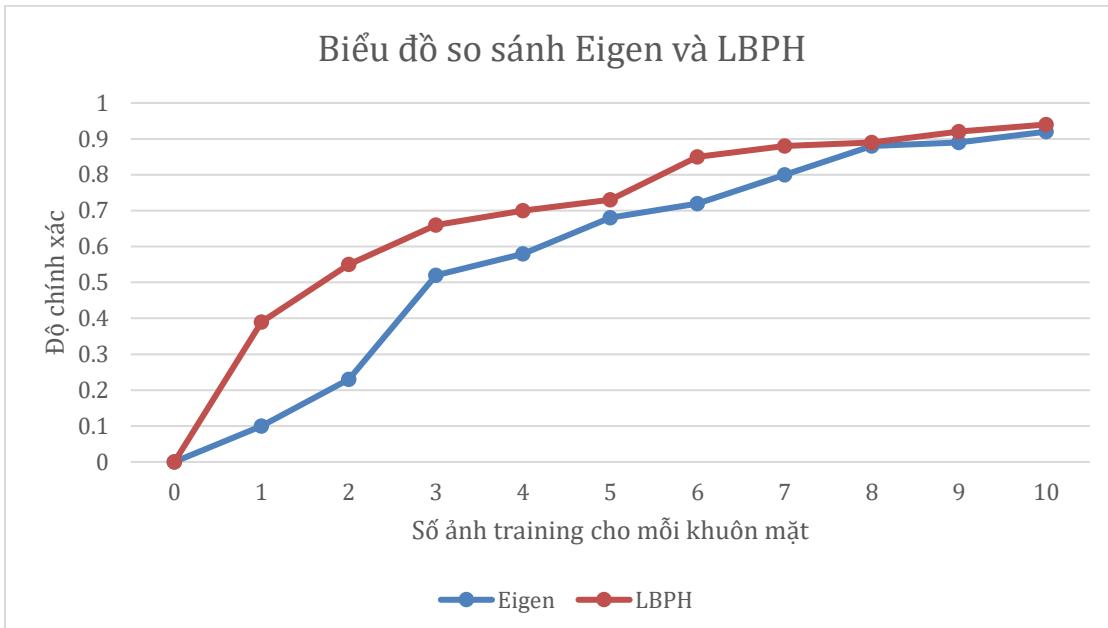
Các ưu điểm của thuật toán LBPH:

- LBPH là một trong những thuật toán trích chọn đặc trưng tốt nhất.
- Là một phương pháp trích chọn đặc trưng cục bộ nên nó không ảnh hưởng bởi cường độ ánh sáng như các phương pháp trích chọn đặc trưng toàn cục.
- Không cần quá nhiều ảnh training, do vậy thời gian xử lý phân loại, nhận dạng nhanh hơn.
- Thuật toán LBPH phân tích độc lập từng ảnh training trong khi các phương pháp như Eigen hay Fisher phải phân tích toàn bộ dữ liệu.

2.5 So sánh độ chính xác giữa các thuật toán và bài toán tối ưu hóa.

Độ chính xác của các thuật toán nhận diện khuôn mặt phụ thuộc vào rất nhiều yếu tố. Trong đó, có những yếu tố quan trọng ảnh hưởng rất lớn đến độ chính xác là mức sáng của ảnh, số lượng ảnh training của từng khuôn mặt và hướng của khuôn mặt. Biểu cảm của khuôn mặt cũng là một trong những yếu tố ảnh hưởng đến độ chính xác.

Hai thuật toán có độ chính xác phụ thuộc rất lớn vào số lượng ảnh training. Qua thực nghiệm sinh viên thấy, trong cùng một môi trường với điều kiện ánh sáng tốt, khuôn mặt chính diện, độ chính xác của các thuật toán được cải thiện dần theo số lượng ảnh training.



Biểu đồ 2.1 Độ chính xác của các thuật toán với số lượng ảnh training.

Với các khuôn mặt khi bị khuất nửa mặt thì hai thuật toán gần như không thể nhận diện được. Ảnh bị nghiêng so với trực chính từ 30 độ trở lên thì thuật toán Eigen Face tỏ ra kém chính xác hơn LBPH

So với Eigen thì thuật toán LBPH tỏ ra chính xác hơn trong trường hợp ảnh đầu vào có điều kiện sáng khác so với tập ảnh được training. thậm chí với ảnh bị lật gương (flip) thì thuật toán LBPH cũng nhận diện chính xác. Trong ứng dụng mà sinh viên xây dựng, với mỗi ảnh chụp được training thì luôn lưu thành 2 bản là bản gốc và bản lật gương (flip) để đảm bảo trong trường hợp các camera chụp ảnh bị lật gương thì thuật toán vẫn có thể nhận diện chính xác.



Hình 2.26 Ảnh training gốc (trái) và ảnh training được lật gương (phải).

Các thuật toán nhận diện khuôn mặt có điểm chung là độ chính xác sẽ giảm khi số lượng khuôn mặt khác nhau trong tập ảnh training tăng. Có nghĩa là số lượng người được training càng nhiều thì ảnh cần được nhận diện càng khó tìm được khuôn mặt chính xác. Vì vậy trong ứng dụng mà sinh viên xây dựng, thay vì lấy toàn bộ ảnh của tập dữ liệu training làm ảnh mồi cho thuật toán thì sinh viên chỉ truy vấn và chọn những ảnh training của các sinh viên trong lớp học đó. Điều này làm giảm số lượng khuôn mặt khác nhau (vì mỗi lớp thực tế luôn có số sinh viên nhỏ hơn 100) mà vẫn giữ được số ảnh training của từng khuôn mặt. Điều này cải thiện đáng kể độ chính xác.

CHƯƠNG 3 TỔNG QUAN VỀ C#, MICROSOFT SQL, REACTJS, THU'

VIỆN EMGUCV VÀ ỨNG DỤNG WEB

3.1 Ngôn ngữ lập trình C# và .NET Framework.

3.1.1 Giới thiệu ngôn ngữ lập trình C#.

C# hay C-sharp là một ngôn ngữ được các kỹ sư của Microsoft phát triển và ra mắt vào năm 2001. C# là ngôn ngữ đơn giản, hiện đại và hướng đối tượng, cung cấp tính linh hoạt và tính năng cho các lập trình viên xây dựng phần mềm nhanh và chính xác nhất.

Các ưu điểm chính của ngôn ngữ C# bao gồm:

- *Cú pháp hiện đại, dễ học:* ngôn ngữ C# được xây dựng C++ và Java nên ngôn ngữ C# khá đơn giản. Diện mạo, cú pháp, biểu thức, toán tử và những chức năng khác được kế thừa từ C và C++ và được cải tiến để làm cho ngôn ngữ thân thiện, đơn giản hơn.
- *Mã nguồn mở:* C# là dự án mã nguồn mở dưới .NET Foundation, được quản lý và điều hành độc lập với Microsoft. Các tính năng được dẫn dắt và thiết kế bởi các kỹ sư của Microsoft, trong khi các thông số kỹ thuật, trình biên dịch và các công cụ liên quan đều là các dự án mã nguồn mở trên Github - một cộng đồng mã nguồn mở năng động, luôn cải tiến ngôn ngữ.
- *An toàn và hiệu quả:* C# là một ngôn ngữ an toàn, không cho phép chuyển đổi loại có thể dẫn đến mất dữ liệu hoặc các vấn đề khác. C# cho phép các nhà phát triển viết mã an toàn. C# cũng tập trung vào việc viết mã hiệu quả.
- *Đa nền tảng:* các ứng dụng được viết bởi C# có thể chạy được trên các hệ điều hành khác nhau như Window, Linux, Mac platform.
- *Đa năng:* Trong khi các ngôn ngữ khác hầu hết được xây dựng cho một mục đích nhất định. Thì với C# chúng ta có thể xử lý hầu hết các yêu cầu như xây dựng ứng dụng máy tính, ứng dụng điện thoại, website, game, trí tuệ nhân tạo, máy học, điện toán đám mây.

- *Phát triển và nâng cấp liên tục:* Các tính năng mới liên tục được xây dựng và phát triển giúp lập trình viên viết ứng dụng nhanh hơn, chính xác hơn.

Hiện nay có nhiều ngôn ngữ mới hơn được ra đời, nhưng C# vẫn luôn là một ngôn ngữ được giới lập trình viên ưa chuộng vì ưu điểm của nó cùng sự hỗ trợ mạnh mẽ đến từ Microsoft.

Về mặt cú pháp, C# được xây dựng trên C++ và Java nên có cú pháp khá đơn giản và thân thiện. Không quá khó để làm quen với một ngôn ngữ như C#. Dưới đây là một chương trình C# đơn giản với ví dụ “Hello World” điển hình:

```
using System;
class ExampleClass
{
    static void Main(string[] args)
    {
        Console.WriteLine("Hello World!");
    }
}
```

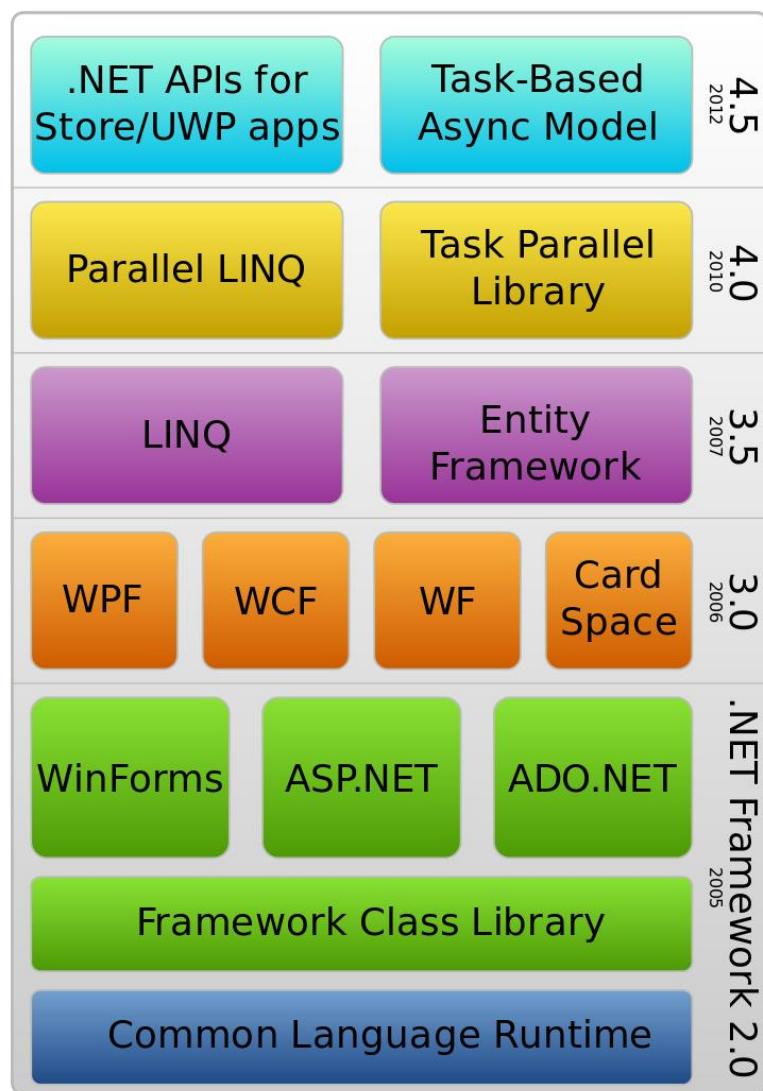
3.1.2 .NET Framework.

.NET Framework là một nền tảng lập trình và cũng là một nền tảng thực thi ứng dụng chủ yếu trên hệ điều hành Microsoft Windows được phát triển bởi Microsoft. Các chương trình được viết trên nền .NET Framework sẽ được triển khai trong môi trường phần mềm (ngược lại với *môi trường phần cứng*) được biết đến với tên Common Language Runtime (CLR). Môi trường phần mềm này là một máy ảo trong đó cung cấp các dịch vụ như an ninh phần mềm (*security*), quản lý bộ nhớ (*memory management*), và các xử lý lỗi ngoại lệ (*exception handling*).

.NET framework bao gồm tập các thư viện lập trình lớn, và những thư viện này hỗ trợ việc xây dựng các chương trình phần mềm như lập trình giao diện; truy cập, kết nối cơ sở dữ liệu; ứng dụng web; các giải thuật, cấu trúc dữ liệu; giao tiếp mạng... CLR cùng với bộ thư viện này là 2 thành phần chính của.NET framework.

.NET framework đơn giản hóa việc viết ứng dụng bằng cách cung cấp nhiều thành phần được thiết kế sẵn, người lập trình chỉ cần học cách sử dụng và tùy theo sự sáng tạo mà gắn kết các thành phần đó lại với nhau. Nhiều công cụ được tạo ra để hỗ trợ xây dựng ứng dụng.NET, và IDE (*Integrated Development Environment*) được phát triển và hỗ trợ bởi chính Microsoft là Visual Studio.

Dưới đây là sơ đồ các thành phần trong .NET Framework:



Hình 3.1 Sơ đồ các thành phần trong .NET Framework.

3.2 Hệ quản trị dữ liệu Microsoft SQL Server.

Microsoft SQL Server (MSSQL) là một hệ quản trị cơ sở dữ liệu được phát triển bởi Microsoft, nó cũng là phần mềm dùng để lưu trữ dữ liệu dựa trên chuẩn Hệ quản trị cơ sở dữ liệu quan hệ (Relational Database Management System – RDBMS).

MS SQL thuộc hệ quản trị cơ sở dữ liệu quan hệ đối tượng. Các ngôn ngữ dùng để truy vấn trong MS SQL là: T-SQL, ANSI SQL

MS SQL cung cấp đầy đủ công cụ để quản lý, từ giao diện cho đến việc sử dụng ngôn ngữ truy vấn SQL. Ngoài ra điểm mạnh của nó là Microsoft có khá nhiều nền tảng kết hợp hoàn hảo với SQL Server như ASP.NET, C# xây dựng Winform, bởi vì nó hoạt động hoàn toàn độc lập.

Các ưu điểm của MS SQL:

- Cho phép tạo nhiều cơ sở dữ liệu
- Duy trì lưu trữ bền vững
- Bảo mật cao
- Phân tích dữ liệu bằng SSAS - SQL Server Analysis Services
- Tạo được báo cáo bằng SSRS - SQL Server Reporting Services
- Thực hiện quá trình ETL (Extract-Transform-Load) bằng SSIS - SQL Server Integration Services.

3.3 Ngôn ngữ lập trình JavaScript và ReactJs Library.

3.3.1 JavaScript là gì?

JavaScript là một ngôn ngữ lập trình kịch bản được phát triển bởi Brendan Eich tại hãng truyền thông Netscape với cái tên ban đầu là Mocha, sau đó đổi tên thành LiveScript, cuối cùng là JavaScript như hiện tại. Ngôn ngữ này được sử dụng cho các trang web, nhưng cũng dùng để tạo khả năng viết Script sử dụng các đối tượng nằm sẵn trong ứng dụng. Cú pháp của JavaScript giống ngôn ngữ C. .js là phần mở rộng của các tập tin mã nguồn của JavaScript.

JavaScript trước đây được sử dụng nhiều cho việc lập trình front-end của web. Nhưng dạo gần đây với sự hỗ trợ mạnh mẽ đến từ các thư viện, framework thì nó đã đa dụng hơn như xây dựng ứng dụng di động, ứng dụng cho máy chủ server.

3.3.2 ReactJs Library.

ReactJS một thư viện của ngôn ngữ JavaScript được xây dựng bởi các lập trình viên của Facebook. ReactJS giúp tạo ra những ứng dụng website hấp dẫn với tốc độ nhanh và hiệu quả cao với những mã hóa tối thiểu.

Với ReactJs bạn có thể lồng ghép các đoạn mã của JavaScript và HTML vào với nhau nhờ vào JSX (JavaScript XML - là một loại cú pháp mở rộng dành cho ngôn ngữ JavaScript viết theo kiểu XML) giúp việc xây dựng front-end của ứng dụng website dễ dàng hơn, nhanh hơn.

Cơ chế hoạt động của ReactJs đó là nó tạo một Virtual DOM (Virtual Document Object Model) chứa đầy đủ các dữ liệu để tạo ra một DOM thật trên màn hình nên khi có sự thay đổi dữ liệu thì nó sẽ tính toán sự sai khác giữa virtual DOM và DOM thật nên việc re-render lại sẽ nhanh hơn, hiệu năng tốt hơn.

3.4 Tổng quan về thư viện EmguCv.

OpenCV (Open Computer Vision) là một thư viện mã nguồn mở chuyên dùng để xử lý các vấn đề liên quan đến thị giác máy tính. Nhờ một hệ thống các giải thuật chuyên biệt, tối ưu cho việc xử lý thị giác máy tính, vì vậy tính ứng dụng của OpenCV là rất lớn.

EmguCv là:

- Là một cross platform .NET, một thư viện xử lý hình ảnh mạnh dành riêng cho ngôn ngữ C#. Cho phép gọi được chức năng của OpenCV là từ .NET.
- Tương thích ngôn ngữ như: C#, VB, VC ++, Iron Python...

- Wrapper có thể được biên dịch bởi Visual Studio, Xamarin Studio và Unity.
- Nó có thể chạy trên Windows, Linux, Mac OS X, iOS, Android và Windows Phone.

Những lợi thế của EmguCv

- EmguCV được viết hoàn toàn bằng C#. Có thể chạy trên bất kỳ nền tảng hỗ trợ bao gồm iOS, Android, Windows Phone, Hệ điều hành Mac OS X và Linux.
- EmguCV có thể được sử dụng từ nhiều ngôn ngữ khác nhau, bao gồm C#, VB.NET, C ++ và Iron Python.
- Nhận dạng ảnh: nhận dạng khuôn mặt, các vật thể ...
- Xử lý ảnh: khử nhiễu, điều chỉnh độ sáng ...
- Nhận dạng cử chỉ.
- Hỗ trợ tài liệu XML và intellisense.
- Sự lựa chọn để sử dụng hình ảnh lớp hoặc trực tiếp gọi chức năng từ OpenCV.

Với những lợi ích to lớn và mạnh mẽ của thư viện EmguCv trong xử lý ảnh, trong khóa luận này, sinh viên ứng dụng thư viện EmguCv để hỗ trợ các thuật toán xử lý ảnh phức tạp.

3.5 Ứng dụng Web.

3.5.1 Ứng dụng web là gì và những tiện ích của ứng dụng web.

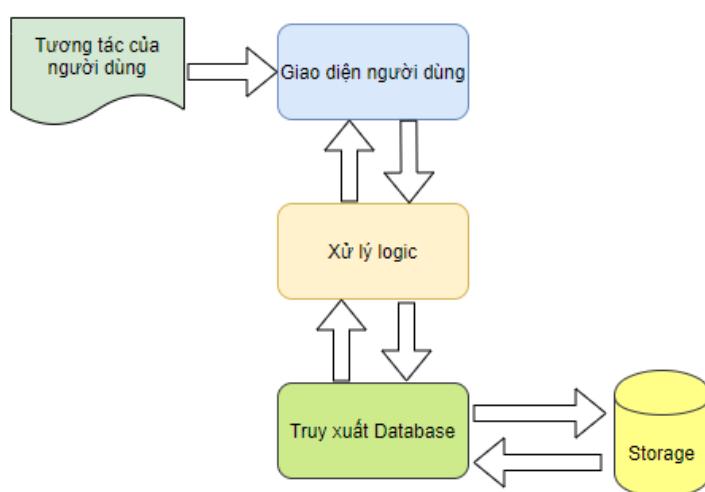
Trong kỹ thuật phần mềm, một Ứng dụng web hay WebApp là một trình ứng dụng mà có thể tiếp cận qua web thông qua mạng như Internet hay Intranet. Ứng dụng web phổ biến nhờ vào sự có mặt vào bất cứ nơi đâu của một chương

trình. Khả năng cập nhật và bảo trì ứng dụng Web mà không phải phân phối và cài đặt phần mềm trên hàng ngàn máy tính là lý do chính cho sự phổ biến của nó.

Trong dạng tính toán chủ-khách trước đây, mỗi ứng dụng có chương trình khách riêng của nó sẽ phục vụ như giao diện người dùng và phải được cài đặt riêng rẽ trên mỗi máy tính cá nhân của người dùng. Sự nâng cấp phần mềm chủ của ứng dụng sẽ cần nâng cấp tất cả máy khách đã được cài trên mỗi máy trạm người dùng, thêm vào đó là chi phí hỗ trợ và giảm năng suất.

Ngược lại, ứng dụng web linh hoạt tạo ra một loạt các tài liệu Web ở định dạng chuẩn được hỗ trợ bởi những trình duyệt phổ biến như HTML / XHTML. Ngôn ngữ kịch bản phía người dùng ở dạng ngôn ngữ chuẩn như JavaScript thường được thêm vào để có thêm những yếu tố động trong giao diện người dùng. Nói chung, mỗi trang Web đơn lẻ được gửi tới người dùng như một tài liệu ổn định, nhưng thứ tự các trang có thể cung cấp cảm giác trực quan, khi những gì người dùng nhập vào sẽ được trả về thông qua thành phần mẫu Web được nhúng vào trong đánh dấu trang. Trong quá trình giao dịch đó, trình duyệt Web sẽ thông dịch và hiển thị trang, và hoạt động như một người dùng chung cho bất kỳ ứng dụng Web nào.

3.5.2 Cấu trúc của một ứng dụng web.



Hình 3.2 Cấu trúc một ứng dụng Web

Về mặt cấu trúc của một ứng dụng web, dù có nhiều biến thể, một ứng dụng Web thông thường được cấu trúc như một ứng dụng ba lớp. Ở dạng phổ biến nhất, một trình duyệt Web là lớp thứ nhất, một bộ máy sử dụng công nghệ nội dung Web động (như ASP, ASP.NET, CGI, ColdFusion, JSP/Java, PHP, Python, hoặc Ruby On Rails) là lớp giữa, và một cơ sở dữ liệu là lớp thứ ba. Trình duyệt sẽ gửi yêu cầu (HttpRequest) đến lớp giữa, lớp giữa sẽ phục vụ bằng cách tạo ra truy vấn và cập nhật cơ sở dữ liệu và tạo ra giao diện người dùng.

3.5.3 Dịch vụ Web và Web API.

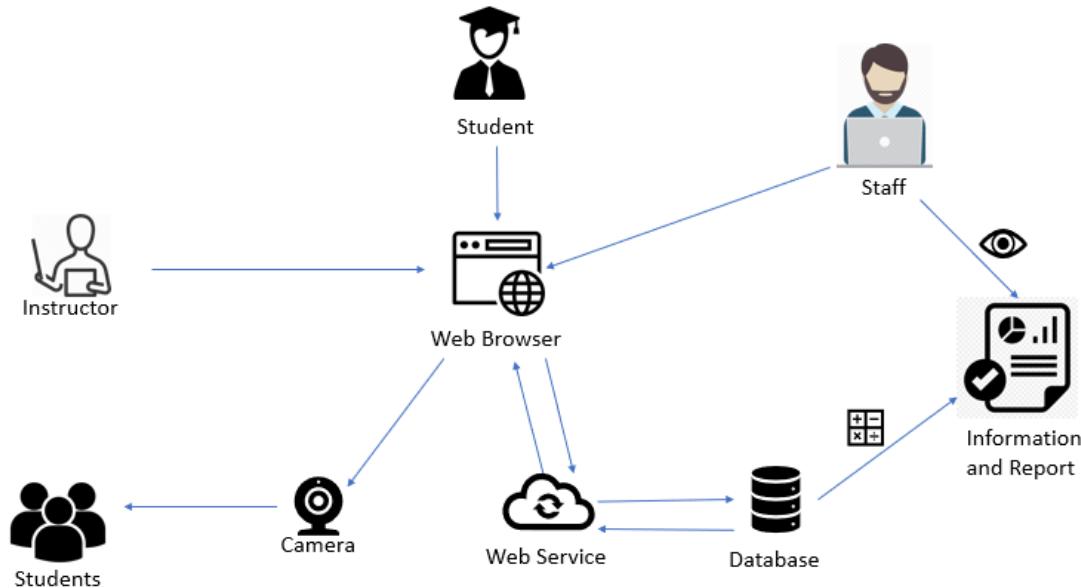
Dịch vụ web (*Web service*) là sự kết hợp các máy tính cá nhân với các thiết bị khác, các cơ sở dữ liệu và các mạng máy tính để tạo thành một cơ cấu tính toán ảo mà người sử dụng có thể làm việc thông qua các trình duyệt mạng.

Web API là giao diện lập trình ứng dụng cho máy chủ web hoặc trình duyệt web. Đây là khái niệm phát triển web, thường giới hạn ở phía máy khách (client-side) của ứng dụng web (bao gồm mọi khung web đang được sử dụng) và do đó thường không bao gồm máy chủ Web (Web server) hoặc chi tiết triển khai trình duyệt như SAPI hoặc API trừ khi có thể truy cập công khai từ xa ứng dụng web. Đây là một trong những framework mới sẽ giúp ích cho chúng ta trong việc xây dựng các HTTP service một cách rất đơn giản và nhanh chóng. Mã nguồn mở nên bạn có thể được sử dụng bởi bất kỳ một client nào hỗ trợ XML, JSON. Nó cũng có khả năng hỗ trợ đầy đủ các thành phần HTTP: URI, request/response headers, caching, versioning, content forma. Bạn có thể sử dụng các host nằm trong phần ứng dụng hoặc trên IIS. Một kiểu kiến trúc vô cùng phù hợp dành cho các thiết bị trang bị băng thông giới hạn như smartphone, tablet. Thường nó có định dạng dữ liệu là JSON, XML hoặc một kiểu dữ liệu bất kỳ.

CHƯƠNG 4 HỆ THỐNG ĐIỂM DANH TỰ ĐỘNG BẰNG NHẬN DIỆN KHUÔN MẶT

4.1 Thiết kế hệ thống.

4.1.1 Phân tích nghiệp vụ.



Hình 4.1 Sơ đồ hệ thống điểm danh sinh viên bằng nhận diện khuôn mặt.

Trong khóa luận này, sinh viên phát triển Hệ thống điểm danh sinh viên bằng nhận diện khuôn mặt nhằm đưa ra giải pháp để khắc phục nhược điểm của các phương pháp điểm danh đã và đang được triển khai. Việc áp dụng công nghệ vào điểm danh giúp đơn giản hóa quy trình điểm danh cho giảng viên và sinh viên cũng như quản lý sinh viên cho phía nhà trường.

Hệ thống sẽ phân quyền thành 2 user. User cho phần giảng viên và user cho phần sinh viên.

Khi tham gia lớp học, giảng viên sẽ đăng nhập tài khoản của mình và chọn lớp học tương ứng. Danh sách sinh viên trong lớp học đó sẽ được hiển thị ra. Giảng viên có thể bật webcam trên laptop hoặc kết nối với camera trong lớp học để chụp hình phục vụ cho việc điểm danh.

Hiện nay, hầu hết mỗi phòng học đều có lắp camera phục vụ cho công tác quản lý sinh viên từ phòng Đào Tạo. Vì vậy, hệ thống trở nên lý tưởng và hoạt động hiệu quả nhất là khi camera trong phòng có thể kết nối với máy tính cá nhân của giảng viên. Giảng viên sẽ thực hiện việc chụp hình, sau đó, hình ảnh được gửi về Web Service và xử lý. Hệ thống sẽ tiến hành nhận diện các sinh viên có trong ảnh và đánh dấu điểm danh các sinh viên nhận diện thành công đồng thời trả hình ảnh được gán nhãn các mã số sinh viên tương ứng với mỗi khuôn mặt được nhận diện về cho Web Browser để giảng viên có thể xem một cách trực quan.

Việc training các khuôn mặt được thực hiện ở phân quyền user sinh viên. Các sinh viên sẽ có tài khoản để đăng nhập vào hệ thống và thực hiện tải ảnh khuôn mặt của mình lên để training một cách chủ động cho hệ thống. Các sinh viên luôn được yêu cầu training tối thiểu 10 tấm ảnh khuôn mặt của mình để đảm bảo hệ thống có thể hoạt động chính xác trong việc nhận diện để điểm danh.

Hệ thống cũng cung cấp thêm các chức năng phục vụ cho việc quản lý các sinh viên. Giảng viên có thể xem các khuôn mặt được sinh viên tải lên training cho hệ thống và có quyền xóa các ảnh không rõ ràng hay khác với khuôn mặt còn lại.

4.1.2 Thiết kế cơ sở dữ liệu.

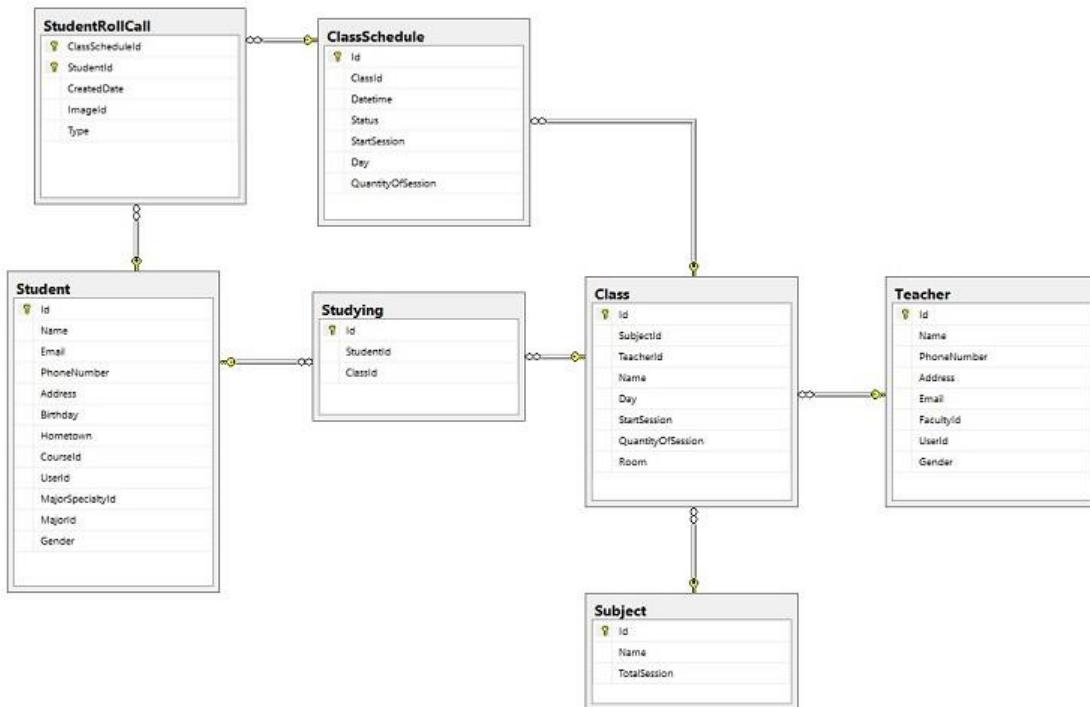
Với các hệ thống phần mềm, cơ sở dữ liệu (database) là không thể thiếu để hệ thống có thể hoạt động một cách đồng nhất. Cơ sở dữ liệu giúp lưu trữ dữ liệu theo một cấu trúc nhất định, có tính nhất quán cao. Bên cạnh đó, việc truy xuất dữ liệu cũng cực kỳ dễ dàng và nhanh chóng. Với đặc điểm này, database giúp người dùng thuận tiện trong việc tạo lập, lưu trữ, tìm kiếm và sử dụng dữ liệu một cách chính xác và nhanh chóng.

Với Hệ thống điểm danh bằng nhận diện khuôn mặt mà sinh viên thực hiện trong khóa luận này, cơ sở dữ liệu của hệ thống được thiết kế có 13 bảng:

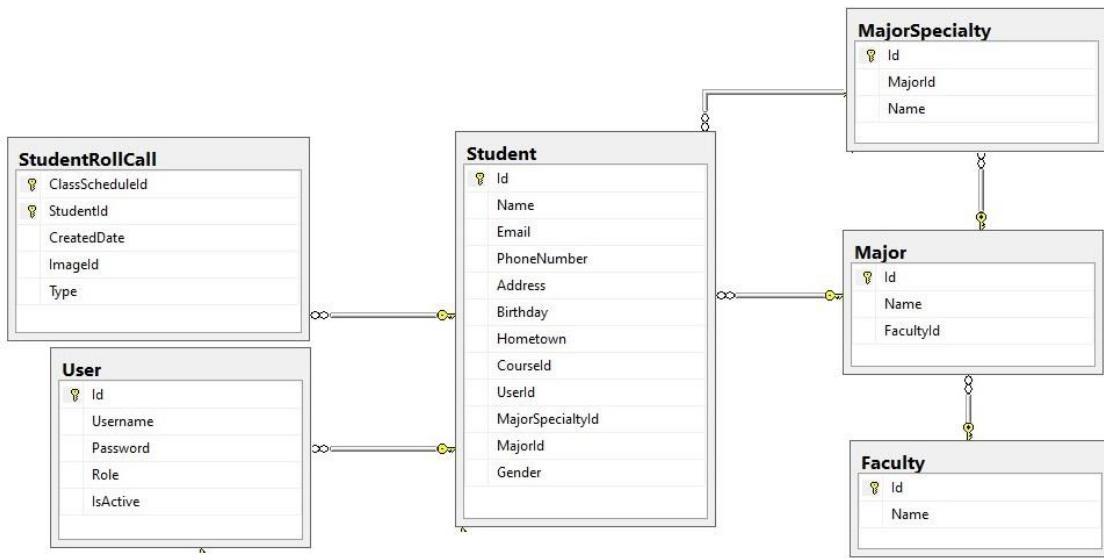
- Class: Các lớp học được mở.
- ClassSchedule: Kế hoạch giảng dạy của lớp đó.

- Faculty: bảng thông tin tên và mã khoa.
- FileAttachment: Các file có trong hệ thống.
- FileData: Dữ liệu của file.
- Major: Ngành học của một khoa.
- MajorSpecialty: Chuyên ngành của một ngành.
- Student: Thông tin sinh viên.
- StudentRollCall: Các lần điểm danh đi học của học sinh.
- Studying: Các lớp mà học sinh đăng ký học.
- Subject: Các môn học.
- Teacher: Thông tin giáo viên.
- TrainImage: Dữ liệu nhận diện khuôn mặt của sinh viên.

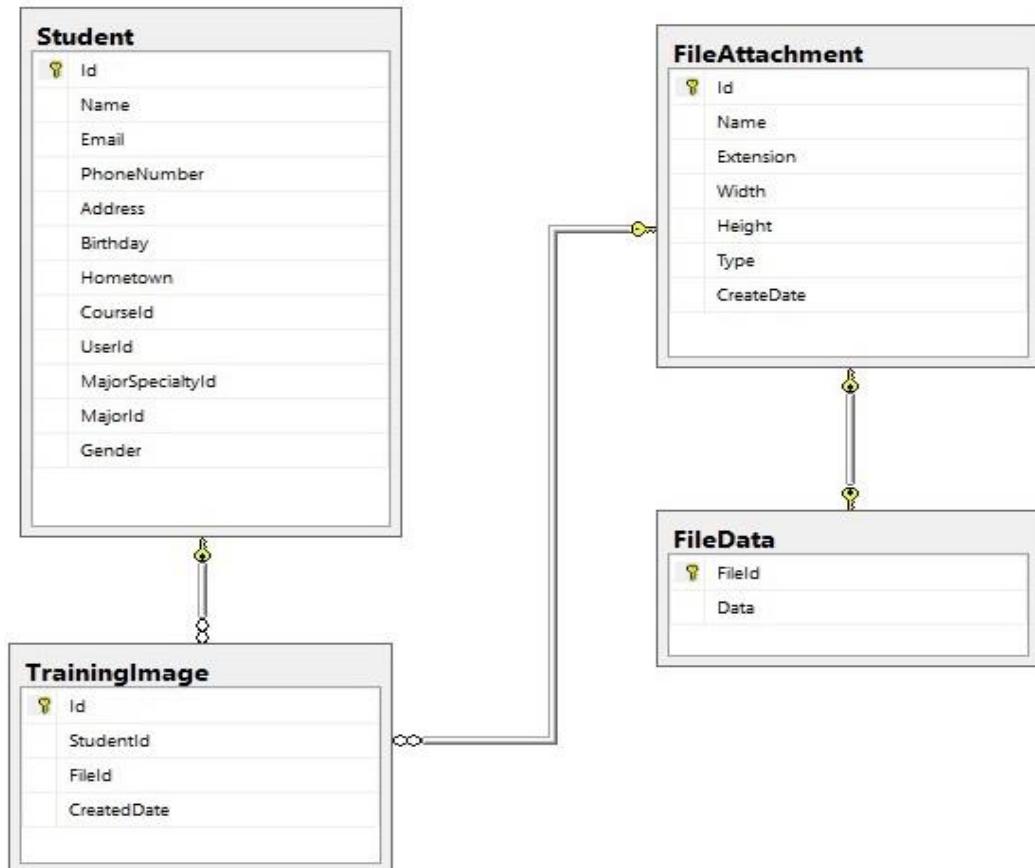
Đây là cơ sở dữ liệu quan hệ, các bảng có thông tin liên quan đến nhau được ràng buộc dữ liệu một cách chặt chẽ để trong quá trình vận hành không xảy ra sai sót về dữ liệu. Việc ràng buộc dữ liệu giữa các bảng được thể hiện trong hình dưới đây:



Hình 4.2 Cơ sở dữ liệu và ràng buộc quan hệ giữa các bảng.



Hình 4.2.1 Cơ sở dữ liệu và ràng buộc quan hệ giữa các bảng.



Hình 4.2.2 Cơ sở dữ liệu và ràng buộc quan hệ giữa các bảng.

4.1.3 Nền tảng công nghệ.

Với nền công nghiệp 4.0, Internet dần trở thành một phần không thể thiếu trong cuộc sống của con người. Cùng với sự bùng nổ của Internet là sự phát triển của các ứng dụng trên nền tảng này. Ứng dụng web phổ biến nhờ vào sự có mặt vào bất cứ nơi đâu của một chương trình. Khả năng cập nhật và bảo trì ứng dụng Web mà không phải phân phối và cài đặt phần mềm trên hàng ngàn máy tính là lý do chính cho sự phổ biến của nó. Ứng dụng web được dùng để hiện thực Webmail, bán hàng trực tuyến, đấu giá trực tuyến, wiki, diễn đàn thảo luận, Weblog, MMORPG, Hệ quản trị nội dung, Phần mềm quản lý nguồn nhân lực và nhiều chức năng khác.

Trong dạng tính toán chủ-khách trước đây, mỗi ứng dụng có chương trình khách riêng của nó sẽ phục vụ như giao diện người dùng và phải được cài đặt riêng rẽ trên mỗi máy tính cá nhân của người dùng. Sự nâng cấp phần máy chủ của ứng dụng sẽ cần nâng cấp tất cả máy khách đã được cài trên mỗi máy trạm người dùng, thêm vào đó là chi phí hỗ trợ và giảm năng suất.

Ngược lại, ứng dụng web linh hoạt tạo ra một loạt các tài liệu Web ở định dạng chuẩn được hỗ trợ bởi những trình duyệt phổ biến như HTML/XHTML. Ngôn ngữ kịch bản phía người dùng ở dạng ngôn ngữ chuẩn như JavaScript thường được thêm vào để có thêm những yếu tố động trong giao diện người dùng. Nói chung, mỗi trang Web đơn lẻ được gửi tới người dùng như một tài liệu ổn định, nhưng thứ tự các trang có thể cung cấp cảm giác trực quan, khi những gì người dùng nhập vào sẽ được trả về thông qua thành phần mẫu Web được nhúng vào trong đánh dấu trang. Trong quá trình giao dịch đó, trình duyệt Web sẽ thông dịch và hiển thị trang, và hoạt động như một người dùng chung cho bất kỳ ứng dụng Web nào.

Chính vì sự tiện lợi này mà sinh viên chọn ứng dụng Web để triển khai cho hệ thống của mình.

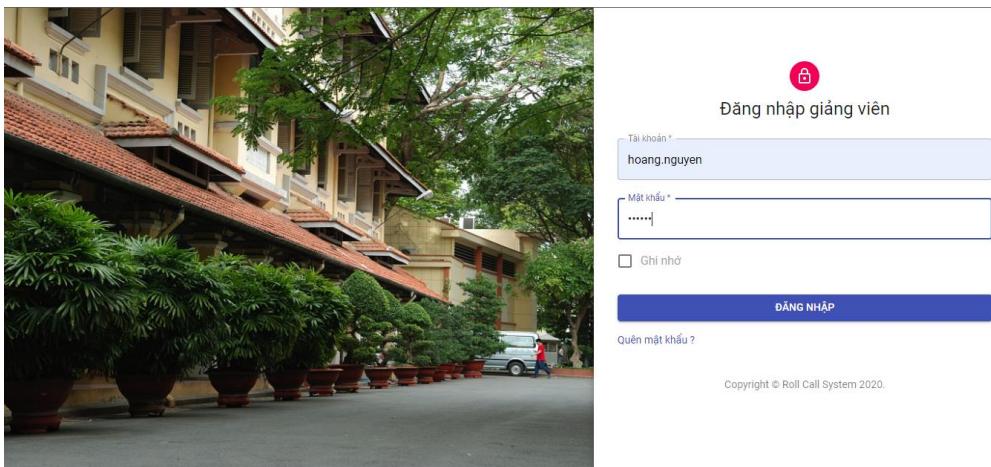
Một ứng dụng web được chia làm 3 phần chính là Front-end, Back-end và Database. Front-end là sử dụng HTML, CSS và JavaScript để thiết kế và xây dựng

giao diện của web cho người dùng có thể xem và tương tác trực tiếp trên đó. Back-end là nơi xử lý mọi logic, nghiệp vụ phức tạp phía sau Front-end. Đây là nơi xử lý dữ liệu từ các hành động tương tác của người dùng phía Front-end. Database là nơi lưu trữ các dữ liệu cần để cho hệ thống web hoạt động.

Trong đề tài này, sinh viên sử dụng Library ReactJs để phát triển phần Front-end. .NET Framework cho phát triển Back-end và database sử dụng Microsoft SQL Server.

4.2 Một số chức năng chính của hệ thống.

4.2.1 Giao diện đăng nhập giảng viên.



Hình 4.3 Giao diện đăng nhập của giảng viên.

Với các hệ thống có phân quyền người dùng (user) thì phần đăng nhập là không thể thiếu. Các giảng viên sẽ được Phòng Đào Tạo cấp tài khoản riêng để đăng nhập vào hệ thống. Giao diện đăng nhập của hệ thống được thiết kế tối giản. Khi giảng viên nhập tài khoản và mật khẩu, dưới back-end, hệ thống xác thực tài khoản và mật khẩu có khớp với cơ sở dữ liệu hay không, đồng thời xác thực phân quyền của tài khoản đăng nhập xem tài khoản thuộc loại user nào (giảng viên hoặc sinh viên) từ đó trả về các token tương ứng để truy cập vào các giao diện và chức năng tương ứng với phân quyền.

Hệ thống sử dụng hàm băm (hash) một chiều để mã hóa mật khẩu nhằm nâng cao tính bảo mật. Mật khẩu sau khi được mã hóa ra một chuỗi ký tự, và từ chuỗi ký tự này không thể giải mã thành mật khẩu ban đầu được.

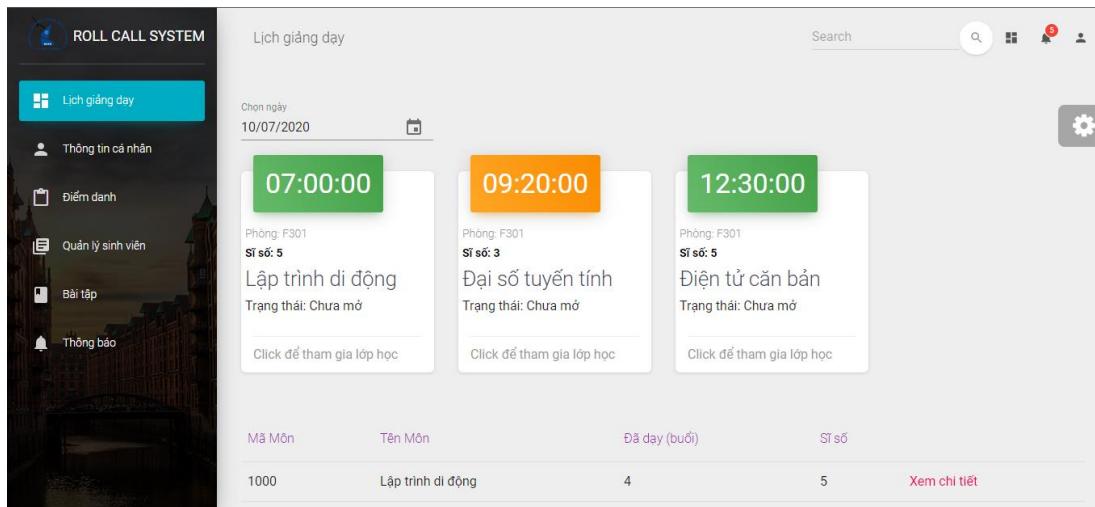
Ví dụ: Tài khoản có mật khẩu là :

TkGiangVien123

Sau khi băm mật khẩu trên bằng hàm băm MD5, mật khẩu trên được mã hóa thành chuỗi sau:

89ea8b275d491dce85d024d4695cd477

4.2.2 Quản lý lịch giảng dạy của giảng viên.



Hình 4.4 Giao diện lịch giảng dạy của giảng viên.

Sau khi đăng nhập thành công vào hệ thống, trên giao diện màn hình chính, giảng viên sẽ thấy các lớp học mà giảng viên phải tham gia trong ngày hiện tại. Các lớp học có 3 trạng thái là: Chưa mở, Đang mở và Đã đóng kèm theo đó là các thông tin chính về lớp học như : giờ bắt đầu, tên môn học, phòng học, sĩ số của lớp. Ngoài ra, giảng viên cũng có thể chọn xem các lớp học theo ngày.

Khi click vào một lớp học:

- Nếu lớp học đó có trạng thái chưa mở: nó sẽ chuyển trạng thái thành đang mở và chuyển sang trang “Điểm danh sinh viên” (phần 4.2.3).

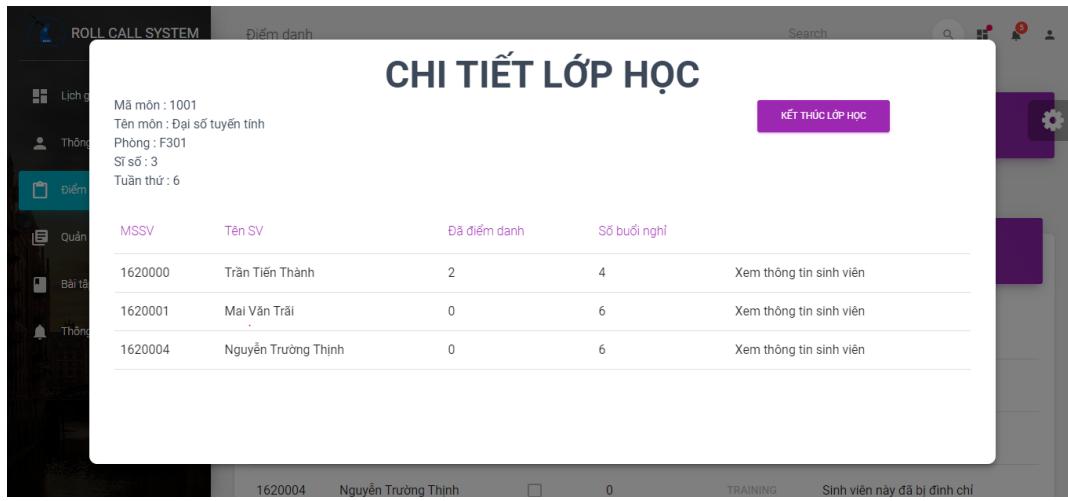
- Nếu lớp học ở trạng thái đang mở: sẽ hiện thông báo lớp học đang được mở và xem thông tin chi tiết của lớp học bằng cách click vào icon “Lớp” trên thanh Navigation.
- Nếu lớp học có trạng thái đã đóng: hiển thị thông báo lớp học đã đóng.

Bên dưới màn hình này là bảng hiển thị danh sách tất cả các môn học mà giảng viên đó đang quản lý giảng dạy. Bảng này chứa một số các trường thông tin cơ bản về môn học như “Số buổi đã dạy” hay “Số sinh viên đăng ký”.

Mã Môn	Tên Môn	Đã dạy (buổi)	Sĩ số	
1000	Lập trình di động	5	5	Xem chi tiết
1001	Đại số tuyến tính	6	3	Xem chi tiết
1002	Điện tử căn bản	3	5	Xem chi tiết
1003	Toán rời rạc	4	3	Xem chi tiết
1004	Java	6	5	Xem chi tiết

Hình 4.5 Danh sách các môn học mà giảng viên quản lý giảng dạy.

Khi click chọn “Xem chi tiết” của một môn sẽ chuyển tới trang “Quản lý sinh viên” (Phần 4.2.4).



Hình 4.6 Giao diện chi tiết lớp học đang mở.

Khi mở modal “Chi tiết lớp học”, modal hiển thị các thông tin của môn học đó như: “Mã môn”, “Tên môn”, “Phòng học”, “Sĩ số của lớp”, “Tuần dạy” và danh

sách sinh viên đăng ký học môn học đó. Nút “Kết thúc lớp học” là để đóng lớp học. Khi kết thúc lớp học đó, hệ thống sẽ thông kê xem tại tuần dạy được đóng có bao nhiêu sinh viên mà tổng số buổi nghỉ học từ 3 trở lên thì sẽ đánh đinh chỉ sinh viên đó.

Khi giảng viên quên “Kết thúc môn học”, khi qua ngày hôm sau thì hệ thống sẽ tự động đóng các lớp có trạng thái đang mở. Việc này để tránh giảng viên quên đóng lớp học và không thể thông kê được các buổi đi học, nghỉ học của sinh viên.

4.2.3 Điểm danh sinh viên.



Hình 4.7 Màn hình điểm danh sinh viên.

Sau khi click chọn mở một lớp từ màn hình Lịch giảng dạy, màn hình Điểm danh sẽ được mở ra và lấy tất cả các sinh viên đã đăng ký môn học này. Giảng viên sẽ nhấn Button Điểm danh tự động để mở webcam. Sau khi nhấn Button chụp, hệ thống sẽ tiến hành nhận diện các sinh viên có trong ảnh. Những sinh viên nhận diện thành công sẽ được tự động đánh điểm danh ở bảng bên dưới.

Trong trường hợp không nhận diện được hết các khuôn mặt sinh viên có trong ảnh (ở xa, khuất góc). Ngoài điểm danh tự động bằng nhận diện khuôn mặt, giảng viên cũng có thể điểm danh thủ công bằng cách tích chọn vào các nút Checkbox tương ứng với từng sinh viên.

4.2.4 Thống kê, quản lý các môn học, sinh viên.

The screenshot shows a user interface for managing students in a class. At the top left, there is a dropdown menu labeled 'Chọn lớp học của bạn:' with 'Java' selected. To the right of the dropdown is a search bar with the placeholder 'Nhập lớp học –'. Below this, a purple header bar displays the text 'Danh sách sinh viên của lớp học'. On the far right of the header is a gear icon. The main content area is a table with the following columns: MSSV, Tên SV, Khoa, Số ảnh training, and 'XEM CHI TIẾT' (View Detail) button. The table contains three rows of data:

MSSV	Tên SV	Khoa	Số ảnh training	XEM CHI TIẾT
1620000	Trần Tiến Thành	Điện tử viễn thông	6	XEM CHI TIẾT
1620001	Mai Văn Trãi	Vật lí	0	XEM CHI TIẾT
1620002	Huỳnh Quốc Khoa	Vật lí	0	XEM CHI TIẾT

Hình 4.8 Màn hình quản lý sinh viên của một lớp.

Giảng viên nhập tên môn học, hệ thống sẽ lấy tất cả các sinh viên đã đăng ký học môn học đó. “Xem chi tiết” từng sinh viên sẽ hiển thị modal thông tin chi tiết sinh viên về môn học này như: danh sách các buổi đã điểm danh, biểu đồ tương quan buổi học và buổi nghỉ, các ảnh sinh viên đã tải lên để training.

This screenshot shows a detailed view of a student's information. On the left, there is a sidebar with the student's details: Mã số SV : 1620000, Tên sinh viên : Trần Tiến Thành, Email : thanhtran@gmail.com, and Khóa : 2016. Below this is a circular pie chart divided into two segments: blue (labeled 'Các buổi đã điểm danh : 5') and purple (labeled 'Các buổi chưa điểm danh : 6'). To the right of the sidebar is a grid of seven small portrait photos labeled 'Ảnh trained'. Below the sidebar is a table titled 'Danh sách các buổi điểm danh của sinh viên' with columns: Ngày, Phòng học, and Tuần thứ. The table lists six attendance entries:

Ngày	Phòng học	Tuần thứ
01/06/2020	C32	1
17/06/2020	F301	2
19/06/2020	F301	4
20/06/2020	F301	5
21/06/2020	F301	6

Hình 4.9 Thông tin chi tiết sinh viên.

Giảng viên có quyền xóa các ảnh training không phù hợp (không rõ khuôn mặt, khuôn mặt khác so với ảnh gốc).

4.2.5 Giao diện đăng nhập sinh viên.

Với phân quyền user Sinh viên cũng có màn hình đăng nhập riêng. Dưới hệ thống backend sẽ dựa vào phân quyền user đăng nhập để trả về token và giao diện chức năng tương ứng để phân biệt quyền Giảng viên và Sinh viên.

Về phần giao diện thì màn hình đăng nhập của Sinh viên không khác so với giao diện đăng nhập của Giảng viên

4.2.6 Quản lý lịch học sinh viên.

Sinh viên có thể xem danh sách các môn học mình đã đăng ký, đồng thời click chọn từng môn sẽ ra thông tin chi tiết của môn học cũng như các buổi điểm danh / không điểm danh của môn học này.

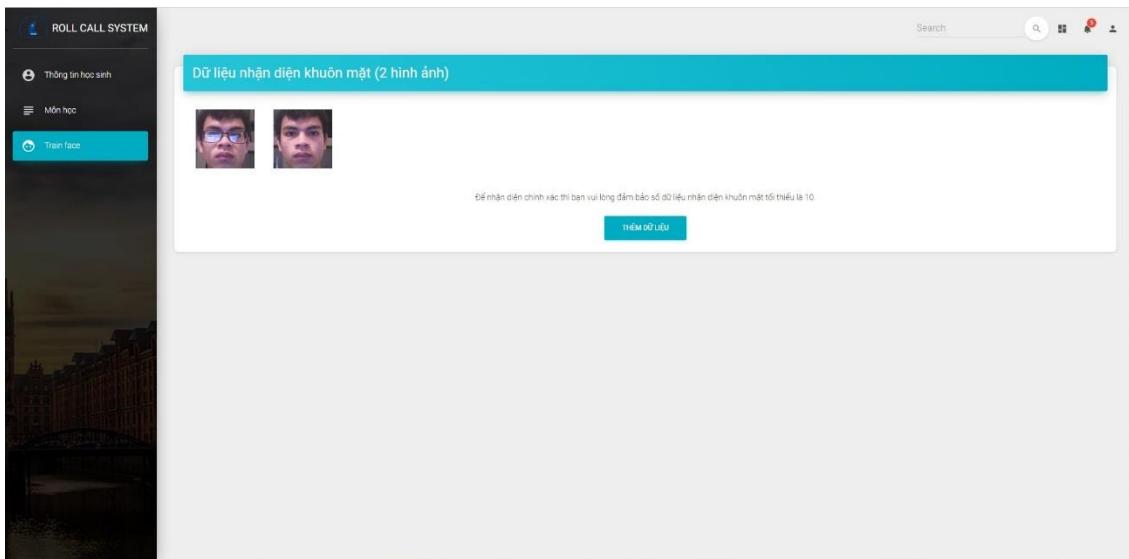
LỊCH HỌC		
Ngày học	Trạng thái	Điểm danh
22/06/2020	Đã kết thúc	<input checked="" type="checkbox"/>
29/06/2020	Đã kết thúc	<input type="checkbox"/>
06/07/2020	Đã kết thúc	<input checked="" type="checkbox"/>
13/07/2020	Đã kết thúc	<input checked="" type="checkbox"/>
20/07/2020	Đã kết thúc	<input checked="" type="checkbox"/>
27/07/2020	Chưa mở	<input type="checkbox"/>

Hình 4.10 Giao diện quản lý lịch học của sinh viên.

4.2.7 Upload ảnh training.

Với một hệ thống nhận diện khuôn mặt thì luôn phải có một tập ảnh để training cho hệ thống gọi là ảnh mồi. Trong ứng dụng điểm danh mà sinh viên xây dựng, các ảnh mồi được tải lên và training hoàn toàn chủ động từ phía sinh viên. Sinh viên sẽ chụp ảnh khuôn mặt của mình và upload lên hệ thống để lưu vào database như một ảnh mồi để phục vụ cho việc nhận diện và điểm danh trên lớp. Các ảnh sẽ được yêu cầu chụp ở nhiều góc khác nhau và không được khuất quá

nhiều. Đồng thời mỗi sinh viên luôn được yêu cầu chụp ít nhất 10 ảnh để đảm bảo hệ thống nhận diện hoạt động tốt.



Hình 4.11 Giao diện upload ảnh training cho sinh viên.

CHƯƠNG 5 KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN.

5.1 Kết luận.

Nhận diện khuôn mặt là một bài toán thiết thực, được áp dụng rộng rãi trong thực tế trong những năm gần đây. Thông qua đề tài với những kiến thức đã được học về nhận diện khuôn mặt và ứng dụng Web, sinh viên đã áp dụng vào bài toán điểm danh giúp giải quyết các vấn đề mà những phương pháp điểm danh cũ đang gặp phải. “Hệ thống điểm danh bằng nhận diện khuôn mặt” được sinh viên xây dựng đã đáp ứng được các nhu cầu căn bản của một hệ thống điểm danh cũng như quản trị. Tuy nhiên vẫn còn những hạn chế nhất định cần được phát triển thêm để hoàn thiện hơn.

Hệ thống được xây dựng trên nền tảng web đem lại sự linh hoạt vì có thể truy cập trên đa thiết bị, đa nền tảng, không cần phải cài ứng dụng trên từng thiết bị. Tuy nhiên, ứng dụng luôn đòi hỏi phải có kết nối Internet để có thể truy cập vào hệ thống. Trong trường hợp không có Internet thì giảng viên không thể quản lý được lớp học hay điểm danh.

Điểm danh bằng nhận diện khuôn mặt tỏ ra hiệu quả trong bài toán điểm danh vì nhanh chóng và tự động nhưng đôi khi có sự chính xác chưa cao với các khuôn mặt mà có ít ảnh training hoặc các khuôn mặt ở xa với ảnh được chụp với chất lượng kém.

Thuật toán “Mẫu nhị phân cục bộ” cho ra kết quả khá cao với các ảnh có các mức sáng tối khác nhau. Tuy nhiên thuật toán xử lý lâu hơn do phải so sánh ảnh cần nhận diện với từng ảnh trong tập ảnh training. Một trong những vấn đề chung và lớn nhất của nhận diện khuôn mặt là các khuôn mặt bị khuất góc hay nửa mặt thì hệ thống chưa nhận diện được. Do đó cần phải cải thiện thêm về thuật toán để có thể phát hiện các khuôn mặt bị khuất và nhận diện các khuôn mặt này. Hiện tại, với sự phát triển của mô hình Học Sâu (Deep Learning) cùng các ngôn ngữ xử lý toán mạnh mẽ như Python đã cho ra kết quả nhận diện khá chính xác

với các khuôn mặt khi khuất nửa và gần như tuyệt đối với các khuôn mặt chính diện, đủ sáng.

5.2 Hướng phát triển.

Hệ thống cần được phát triển thêm để khắc phục các nhược điểm hiện có. Trong tương lai, sinh viên có thể xây dựng các ứng dụng di động để cải thiện sự tiện lợi và linh hoạt. Các giảng viên có thể cầm điện thoại của mình và trực tiếp chụp ảnh và điểm danh. Ngay cả khi không có internet, giảng viên vẫn có thể chụp ảnh và khi có Internet, có thể nhập (import) ảnh đã chụp để nhận diện và điểm danh sau. Ngoài ra, ứng dụng di động có thể đồng bộ lịch giảng dạy với lịch trên máy giúp giảng viên tiện theo dõi những lớp nào cần dạy trong ngày và các lớp có bao nhiêu sinh viên.

Ứng dụng có thể cập nhật thêm các chức năng như giao bài tập về nhà, chấm điểm, tạo các thông báo về lớp học như đổi lịch học, đổi phòng học và gửi thông báo này tới các sinh viên trong lớp.

Điểm danh bằng nhận diện khuôn mặt có thể phát triển thành hệ thống nhận diện real-time, qua đó có thể bỏ các bước chụp ảnh và sinh viên chỉ cần đi qua cửa lớp có gắn camera thì đã điểm danh xong.

Hệ thống không chỉ áp dụng trong lĩnh vực điểm danh sinh viên ở trường học, mà còn có thể phát triển để áp dụng cho các cơ quan làm việc để chấm công nhân viên, check-in ở sân bay, lực lượng an ninh theo dõi tội phạm và nhiều hệ thống khác có quy trình tương tự.

DANH MỤC TÀI LIỆU THAM KHẢO

Trang web

[1] Face Recognition: Understanding LBPH Algorithm.

<https://medium.com/dev-genius/face-recognition-based-on-lbph-algorithm-17acd65ca5f7>

[2] Face Recognition Based On LBPH Algorithm.

<https://towardsdatascience.com/face-recognition-how-lbph-works-90ec258c3d6b>

[3] Tìm hiểu về phương pháp nhận diện khuôn mặt của Violas & John.

<https://viblo.asia/p/tim-hieu-ve-phuong-phap-nhan-dien-khuon-mat-cua-violas-john-ByEZkNVyKQ0>

[4] Face Recognition Using Eigenfaces (PCA Algorithm).

<https://www.geeksforgeeks.org/ml-face-recognition-using-eigenfaces-pca-algorithm/>

[5] Haar Cascade là gì? Luận về một kỹ thuật chuyên dùng để nhận biết các khuôn mặt trong ảnh.

<https://viblo.asia/p/haar-cascade-la-gi-luan-ve-mot-ky-thuat-chuyen-dung-de-nhan-bietcac-khuon-mat-trong-anh-E375zamdlGW>

[6] Face detection with OpenCV and deep learning.

<https://www.pyimagesearch.com/2018/02/26/face-detection-with-opencv-and-deep-learning/>

[7] .NET Framework

https://vi.wikipedia.org/wiki/.NET_Framework

[8] Getting Started with ReactJS

<https://reactjs.org/docs/getting-started.html>

[9] Camera Face Detection in C# using Emgu CV

<https://www.codeproject.com/articles/462527/camera-face-detection-in-csharp-using-emgu-cv-and>

[10] Easy application to detect and recognize faces with EmugCV

<https://www.codeproject.com/Articles/1120266/Easy-application-to-detect-and-recognize-faces-wit>

[11] Quickstart: Use Visual Studio to create your first ASP.NET Core web app

<https://docs.microsoft.com/en-us/visualstudio/ide/quickstart-aspnet-core?view=vs-2019>

[12] Web application

[https://en.wikipedia.org/wiki/Web_application#:~:text=A%20web%20application%20\(or%20web,with%20an%20active%20internet%20connection.](https://en.wikipedia.org/wiki/Web_application#:~:text=A%20web%20application%20(or%20web,with%20an%20active%20internet%20connection.)

[13] Quyết định 25/2006-QĐ-BGDDT

<https://thuvienphapluat.vn/van-ban/giao-duc/Quyet-dinh-25-2006-QD-BGDDT-quy-che-dao-tao-dai-hoc-cao-dang-he-chinh-quy-12885.aspx>

Sách

Madhura Datta, Pradipta Kumar Banerjee, Asit Kumar Datta, “Face Detection and Recognition: Theory and Practice”, 2015, pp.64-142