BỘ GIÁO DỤC VÀ ĐÀO TẠO



**TRƯỜNG ĐẠI HỌC CỬU LONG**

**BÁO CÁO TỔNG KẾT**

**ĐỀ TÀI NGHIÊN CỨU KHOA HỌC CỦA SINH VIÊN GIAI ĐOẠN 2023-2024**

**NGHIÊN CỨU KỸ THUẬT NHẬN DẠNG ĐỐI TƯỢNG TRONG ẢNH VÀ ỨNG DỤNG ĐẾM SỐ LƯỢNG NGƯỜI HOẶC XÁC ĐỊNH DANH TÍNH CỦA NGƯỜI CÓ TRONG ẢNH**

### Chủ nhiệm đề tài: Phạm Quốc Lịnh

Vĩnh Long - 2024

BỘ GIÁO DỤC VÀ ĐÀO TẠO

**TRƯỜNG ĐẠI HỌC CỬU LONG**



**BÁO CÁO TỔNG KẾT**

**ĐỀ TÀI NGHIÊN CỨU KHOA HỌC CỦA SINH VIÊN GIAI ĐOẠN 2023-2024**

**NGHIÊN CỨU KỸ THUẬT NHẬN DẠNG ĐỐI TƯỢNG TRONG ẢNH VÀ ỨNG DỤNG ĐẾM SỐ LƯỢNG NGƯỜI HOẶC XÁC ĐỊNH DANH TÍNH CỦA NGƯỜI CÓ TRONG ẢNH**

**Chủ nhiệm đề tài:Phạm Quốc Lịnh**

MSSV:2111020025 Ngành: Công Nghệ Thông Tin

Khoa: Công Nghệ Thông Tin-Truyền Thông Năm học:2023-2024

**Thành viên cùng tham gia thực hiện: Châu Hồng Bảo Ngọc**

MSSV:2111020010 Ngành: Công Nghệ Thông Tin

Khoa: Công Nghệ Thông Tin-Truyền Thông Năm học:2023-2024

**Người hướng dẫn**: **ThS.Nguyễn Hữu Thể**

**Vĩnh Long,2024**

**LỜI CẢM ƠN**

Để được hoàn thành đồ án này, em đã nhận được sự hướng dẫn, giúp đỡ, hỗ trợ của quý thầy cô trường Đại Học Cửu Long.

Trước tiên em xin gửi lời cảm ơn chân thành, sâu sắc đến các quý thầy cô trong Khoa **Công Nghệ Thông Tin – Truyền Thông trường Đại Học Cửu Long**. Đã tận tình giảng dạy, truyền đạt cho em những kinh nghiệm quý báu trong suốt thời gian học tập tại trường.

Đặc biệt, em xin gửi lời cảm ơn đến thầy **Nguyễn Hữu Thể** - giảng viên hướng đã tận tình giúp đỡ, hướng dẫn em trong suốt quá trình làm báo cáo thực tập tốt nghiệp. Trong thời gian thực tập và làm việc với thầy, em không ngừng tiếp thu thêm nhiều kiến thức bổ ích mà còn học hỏi được tinh thần làm việc, thái độ học tập và làm việc nghiêm túc, … Đây là những điều rất cần thiết cho em trong suốt thời gian làm việc sau này.

Sau cùng, em xin gửi lời cảm ơn chân thành đến gia đình và bạn bè đã không ngừng động viên, đóng góp ý kiến và giúp đỡ em trong quá trình hoàn thành báo cáo tổng kết này.

Mặt dù đã cố gắng hoàn thành, song không tránh khỏi những sai sót, rất mong quý thầy cô cho em những nhận xét và ý kiến quý báu, để em ngày một hoàn thiện và phát triển hơn.

Sinh viên thực hiện

**Phạm Quốc Lịnh**

**Châu Hồng Bảo Ngọc**

**LỜI MỞ ĐẦU**

Trong thế giới ngày nay, sự phát triển không ngừng của công nghệ đã mở ra một lĩnh vực rộng lớn cho nghiên cứu và ứng dụng các phương pháp nhận dạng đối tượng trong ảnh. Điều này không chỉ làm thay đổi cách chúng ta tương tác với thế giới xung quanh mà còn mở ra những cơ hội mới và tiềm năng không giới hạn trong nhiều lĩnh vực, từ công nghiệp đến y tế, an ninh, và nhiều lĩnh vực khác.

Đề tài nghiên cứu kỹ thuật về nhận dạng đối tượng trong ảnh và ứng dụng đếm số lượng người hoặc xác định danh tính của họ là một phần trong cơn sóng đổi mới này. Trong báo cáo này, chúng em sẽ đào sâu vào quy trình nghiên cứu, các phương pháp và công nghệ tiên tiến được sử dụng, cùng với những ứng dụng tiềm năng của đề tài này trong thực tế.

Chúng em tin rằng thông qua báo cáo này, người đọc sẽ được truyền đạt một cái nhìn toàn diện về tầm quan trọng và tiềm năng của việc áp dụng các kỹ thuật nhận dạng đối tượng trong ảnh vào cuộc sống hàng ngày, cụ thể là trong việc đếm số lượng người hoặc xác định danh tính của họ. Đồng thời, chúng em cũng mong muốn nhận được sự đóng góp ý kiến và phản hồi từ quý thầy cô, nhằm tạo ra một môi trường nghiên cứu và ứng dụng tiếp tục phát triển và tiên tiến hơn nữa.

Chân thành cảm ơn sự quan tâm và hỗ trợ của quý thầy cô đối với đề tài này.

**CHƯƠNG 1: TỔNG QUAN**

* 1. **Giới thiệu đề tài**

Trong thời đại số hóa hiện nay, nghiên cứu về kỹ thuật nhận dạng đối tượng trong ảnh và ứng dụng đếm số lượng người hoặc xác định danh tính của họ đang trở thành một lĩnh vực nghiên cứu vô cùng hứa hẹn và đầy thách thức. Đây là một trong những đề tài mang tính đột phá, đem lại những ứng dụng thực tiễn rộng lớn và tiềm năng trong nhiều lĩnh vực khác nhau.

Trong đề tài nghiên cứu này, chúng em tập trung vào việc phát triển và áp dụng các phương pháp, thuật toán và công nghệ nhận dạng đối tượng trong ảnh, nhằm mục đích đếm số lượng người có trong ảnh hoặc xác định danh tính của họ. Điều này có thể bao gồm việc sử dụng các thuật toán học máy, học sâu (deep learning), xử lý ảnh số, và các kỹ thuật phân loại đối tượng.

Mục tiêu của đề tài là nghiên cứu và phát triển các phương pháp chính xác và hiệu quả để nhận dạng và đếm số lượng người trong ảnh, cũng như xác định danh tính của họ dựa trên các đặc điểm nhận dạng như khuôn mặt, dáng đi, hoặc các đặc điểm khác. Ứng dụng của đề tài có thể được thấy rõ trong các lĩnh vực như an ninh, giám sát giao thông, quản lý sự kiện, và nhiều lĩnh vực khác yêu cầu giám sát và quản lý nguồn nhân lực hoặc phân tích hành vi con người.

Chúng em tin rằng việc nghiên cứu và phát triển trong lĩnh vực này sẽ đem lại nhiều lợi ích cho cả xã hội và doanh nghiệp, từ việc nâng cao hiệu suất làm việc đến cải thiện an ninh và quản lý. Qua đó, đề tài cũng mong muốn đóng góp vào sự tiến bộ của công nghệ và phát triển bền vững của xã hội.

* 1. **Mục tiêu đề tài**

Thông qua máy học nghiên cứu các thuật toán nhận dạng khuôn mặt mạng nơron tích chập mang lại kết quả có độ chính xác cao. Từ đó áp dụng vào thực tế xây dựng ứng dụng ứng dụng đếm số lượng người hoặc xác định danh tính của họ:

Phía người quản lý:

- Xác định yêu cầu và mục tiêu

- Thu thập dữ liệu

- Tùy chỉnh và huấn luyện mô hình

- Thiết kế và triển khai ứng dụng

- Kiểm tra và đánh giá

- Quản lý và cập nhật

- Thống kê

- Xuất danh sách sang tệp excel.

* 1. **Nội dung đề tài**

Đề tài sẽ tập trung vào nghiên cứu và phát triển một ứng dụng toàn diện nhằm nhận dạng danh tính từ camera và đếm số người từ ảnh, cũng như từ camera. Với việc sử dụng cơ sở dữ liệu MySQL, sẽ lưu trữ thông tin liên quan đến danh tính và số lượng người. Giao diện người dùng sẽ được xây dựng bằng thư viện Tkinter của Python, cung cấp các chức năng nhập liệu, hiển thị kết quả và điều khiển camera một cách dễ dàng và trực quan.

Thuật toán nhận dạng và đếm số người sẽ được phát triển dựa trên các công nghệ tiên tiến nhất từ thư viện Dlib và OpenCV. Điều này sẽ đảm bảo tính chính xác và hiệu suất của hệ thống trong việc nhận dạng khuôn mặt và đếm số lượng người một cách nhanh chóng và chính xác.

Sau khi hoàn thành, đề tài sẽ trải qua các bước kiểm thử và đánh giá kỹ lưỡng để đảm bảo tính đáng tin cậy và tính ứng dụng trong các tình huống thực tế. Điều này sẽ đảm bảo rằng ứng dụng của chúng em có thể được triển khai và sử dụng một cách hiệu quả và hiệu quả trong nhiều lĩnh vực, từ giám sát an ninh đến quản lý sự kiện và nhiều ứng dụng khác.

* 1. **Giới hạn đề tài**

Đề tài sẽ tập trung vào việc phát triển một ứng dụng nhận dạng và đếm số người từ camera và ảnh, sử dụng thư viện Dlib và OpenCV. Tuy nhiên, đề tài sẽ không xem xét các trường hợp đặc biệt như ánh sáng yếu, góc chụp không lý tưởng, hoặc nền ảnh phức tạp. Ứng dụng sẽ được giới hạn trong phạm vi nhận dạng và đếm số người trong một phạm vi cụ thể, chẳng hạn như trong một phòng họp, sự kiện, hoặc khu vực công cộng. Giao diện người dùng sẽ được xây dựng bằng Tkinter để đáp ứng các yêu cầu cơ bản như nhập liệu, hiển thị kết quả và điều khiển camera. Tuy nhiên, việc tối ưu hóa giao diện để tương tác một cách dễ dàng và hiệu quả có thể là một điều cần xem xét trong các phiên bản cập nhật sau này. Cuối cùng, mặc dù sẽ cố gắng tối ưu hóa hiệu suất và tính chính xác của thuật toán, nhưng với các điều kiện về phần cứng và môi trường thử nghiệm cụ thể, việc đạt được một mức độ hiệu suất và tính chính xác nhất định có thể gặp hạn chế.

* 1. **Cấu trúc đồ án**

Cấu trúc đồ án gồm 5 chương:

* + - Chương 1: Tổng quan
    - Chương 2: Cơ sở lý thuyết
    - Chương 3: Phân tích thiết kế hệ thống
    - Chương 4: Hiện thực hệ thống
    - Chương 5: Kết luận

## CHƯƠNG 2: CƠ SỞ LÝ THUYẾT VÀ CÁC NGHIÊN CỨU LIÊN QUAN

### Cơ sở lý thuyết

### Nhận diện

### Giới thiệu

Công nghệ nhận diện khuôn mặt đã xuất hiện từ những năm 1960 [1]. Cho đến những năm gần đây, với sự phát triển không ngừng của khoa học – công nghệ đã dẫn đến sự phát triển mạnh mẽ của dạng công nghệ này.

Nhận dạng khuôn mặt là một loại phần mềm sinh trắc học ánh xạ các đặc điểm khuôn mặt của một cá nhân về mặt toán học và lưu trữ dữ liệu dưới dạng faceprint (dấu khuôn mặt). Công nghệ sử dụng các thuật toán Deep Learning để so sánh ảnh chụp trực tiếp hoặc hình ảnh kỹ thuật số với faceprint được lưu trữ để xác minh danh tính của một cá nhân.

### Ứng dụng

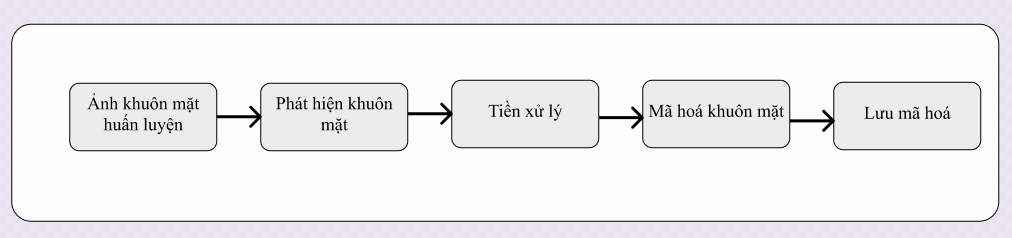
Nhận diện gương mặt có thể được sử dụng trong nhiều lĩnh vực, nhiều việc từ bảo mật đến quảng cáo như: mở khoá trên thiết bị di động, mạng xã hội, bảo mật doanh nghiệp (vì doanh nghiệp có thể sử dụng gương mặt của mình để vào toà nhà, nơi làm việc), điểm danh sinh viên trong giáo dục, chấm công cho nhân viên, tìm người mất tích, thất lạc, …

### Lợi ích khi ứng dụng nhận dạng gương mặt

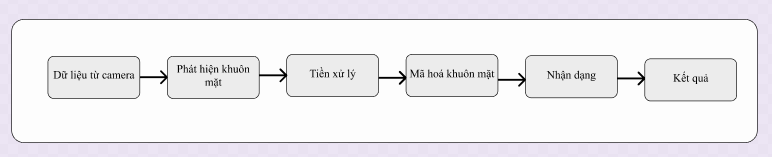
Khi chúng ta ứng dụng nhận dạng gương mặt vào thực tế. Nó sẽ đem lại nhiều lợi ích như: không cần thực hiện tiếp xúc trực tiếp với cá nhân hay thiết bị vẫn có thể xác thực, không đòi hỏi người dùng phải chờ đợi một thời gian dài hay làm bất cứ điều gì khác ngoài việc nhìn vào camera, cải thiện mức độ bảo mật một cách tối ưu nhất, yêu cầu xử lý ít hơn so với các kỹ thuật khác, dễ dàng kết hợp với các tính năng kết hợp sẵn có.

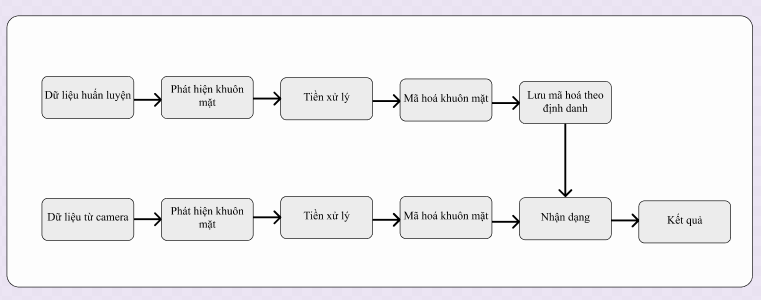
### Hệ thống nhận diện khuôn mặt

Đầu tiên phải đào tạo khuôn mặt để có dữ liệu:



Hình 1. Sơ đồ huấn luyện khuôn mặt Sau khi đã có khuôn mặt, chúng ta có thể nhận diện:



Hình 2. Sơ đồ nhận diện khuôn mặt Một hệ thống nhận diện khuôn mặt:

Hình 3. Hệ thống nhận diện khuôn mặt

### Phát hiện khuôn mặt

Phát hiện khuôn mặt là quá trình xác định vị trí khuôn mặt (vị trí của mắt, mũi, miệng, …) trong hình ảnh và sẽ lấy ra tất cả khuôn mặt trong một hình ảnh đó. Ở bước này có thể sử dụng thuật toán Viola Jones hay HOG and SVM để phát hiện khuôn mặt.

Phát hiện mặt người với các đặc trưng HOG

* + - * + Để phát hiện mặt người với các đặc trưng HOG ta tiến hành các bước sau:
        + Chuẩn bị P mẫu là ảnh mặt người và trích xuất các vector đặc trưng HOG từ ảnh này
        + Chuẩn bị N mẫu không phải là ảnh mặt người (N lớn hơn P) và trích xuất các vector đặc trưng HOG từ các ảnh này.
        + Sử dụng bộ phân loại SVM tuyến tính để học các vector của các mẫu tích cực (các vector của mẫu ảnh có xuất hiện mặt người) và tiêu cực (các vector của mẫu ảnh không có xuất hiện mặt người).
        + Đối với mỗi bức ảnh trong bộ tiêu cực, sử dụng một cửa sổ trượt di chuyển tất cả các vị trí có thể của ảnh input. Tại mỗi vị trí của cửa sổ trượt, tính vector HOG của cửa sổ và đưa vào bộ phân lớp. Nếu bộ phân lớp, phân lớp sai một

cửa sổ là ảnh có xuất hiện mặt người thì ghi lại vector tương ứng cùng với xác suất phân lớp.

* + - * + Lấy các mẫu nhận diện sai ở bước 4 và sắp xếp chúng theo mức sắp xếp của nhận diện sai và cho bộ phân lớp học lại các mẫu sai này.
        + Áp dụng bộ phân lớp học lại các ảnh cần phát hiện khuôn mặt người.

### Đánh giá phương pháp

Nhận dạng khuôn mặt theo tầng Haar: Dò tìm khuôn mặt dựa trên tầng Haar là công nghệ tiên tiến nhất trong nhận diện khuôn mặt trong nhiều năm kể từ năm 2001 khi nó được Viola và Jones giới thiệu trong bài báo của họ, Đã có nhiều cải tiến trong những

năm gần đây. Phương pháp này có kiến trúc đơn giản, hoạt động gần như thời gian thực trên CPU. Ngoài ra, nó có thể phát hiện hình ảnh ở các tỉ lệ khác nhau với tốc độ phát hiện khá nhanh. Nhưng nhược điểm chính là nó cho kết quả sai lệch cũng như không hoạt động trên ảnh không chính diện.

Nhận dạng HOG. Đây là mô hình dựa trên các tính năng thường được sử dụng để trích xuất các tính năng từ dữ liệu hình ảnh. Nó không có khả năng phát hiện hình ảnh nhỏ và xử lý khớp cắn. Ngoài ra, nó thường loại trừ một số phần của cằm và trán trong khi phát hiện. Tuy nhiên, đây là phương pháp nhanh nhất trên CPU có thể hoạt động trên các hình ảnh trực diện và hơi không trực diện.

### Các bộ dữ liệu:

Cơ sở dữ liệu LFW: Bao gồm các khuôn mặt được gắn nhãn trong tự nhiên. Bộ dữ liệu gồm 13233 hình ảnh khuôn mặt của 5749 người được thu thập từ web. Mỗi khuôn mặt được gắn nhãn với tên của người đó.

Face95: một chuỗi 20 hình ảnh cho mỗi cá nhân được chụp. Trong trình tự, các đối tượng tiến một bước về phía máy ảnh. Chuyển động này được sử dụng để giới thiệu các biến thể đầu (tỉ lệ) đáng kể giữa các hình ảnh của cùng một cá nhân. Có khoảng 0,5 giây giữa các khung hình liên tiếp trong chuỗi. Số lượng cá thể là 72, độ phân giải hình ảnh (180 x 200 pixel định dạng dọc), chứa hình ảnh các đối tượng nam và nữ.

Bộ dữ liệu khuôn mặt quy mô lớn UTKFace: Tập dữ liệu UTKFace là tập dữ liệu về khuôn mặt quy mô lớn với độ tuổi lâu dài, nằm trong khoảng từ 0 đến 116 tuổi. Các hình ảnh bao gồm sự thay đổi lớn về tư thế, nét mặt, ánh sáng, khớp cắn, độ phân giải và những thứ khác. Với kích thước bao gồm hơn 24.000 ảnh với chú thích về độ tuổi, giới tính và dân tộc.

### Đánh giá phương pháp:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tập dữ liệu | Tổng số khuôn mặt | Phương pháp  HOG | | Phương pháp Haar-Like AdaBoost | |
| Số khuôn mặt phát hiện | Tỉ lệ (%) | Số khuôn mặt phát hiện | Tỉ lệ (%) |
| LFW | 13233 | 13232 | 99,98 | 12624 | 95,4 |
| Face95 | 1440 | 1439 | 99,93 | 1372 | 95.27 |
| UTKFace | 24108 | 24106 | 99.98 | 19912 | 82,59 |

Vậy là ta đã thấy rõ phương pháp HOG and SVM có tỉ lệ phát hiện khuôn mặt cao và chính xác. Với phương pháp Haar-Like AdaBoost có tỉ thấp hơn, phát hiện sai hoặc không phát hiện được khuôn mặt trong ảnh.

Cho nên đối với việc phát hiện khuôn mặt trong đồ án này, tôi đã sử dụng phương pháp HOG và SVM nhầm phát hiện khuôn mặt tốt hơn và chính xác hơn.

Trong đề tài này toàn bộ phần phất hiện khuôn mặt sử dụng các dữ liệu phân lớp đã được học trước được cung cấp bởi thư viện dlib.

### Tiền xử lý

Bước này nhằm mục đích lọc nhiễu, nâng cao chất lượng ảnh, trong bước này bao gồm các bước: Căn chỉnh ảnh, chuẩn hóa ánh sáng, màu sắc của ảnh.

### Mã hoá khuôn mặt

Mã hoá các vị trí mắt, mũi, miệng, mài, … của khuôn mặt. Lưu giá trị của các khái niệm này, thành ma trận của khuôn mặt.

### Nhận diện

Giai đoạn cuối cùng là đưa ra quyết định xem các đặc điểm của khuôn mặt từ camera hay webcam có khớp với khuôn mặt từ cơ sở dữ liệu hay không. Cũng như xác định danh tính của khuôn mặt đang nhận diện. Và tôi đã sử dụng phương pháp Euclide. Đây là một phương pháp phân loại đối tượng dựa trên cách tính toán khoảng cách giữa các nút trên khuôn mặt và khuôn mặt có sự khác biệt nhỏ nhất giữa các giá trị khoảng cách này được coi là khớp.

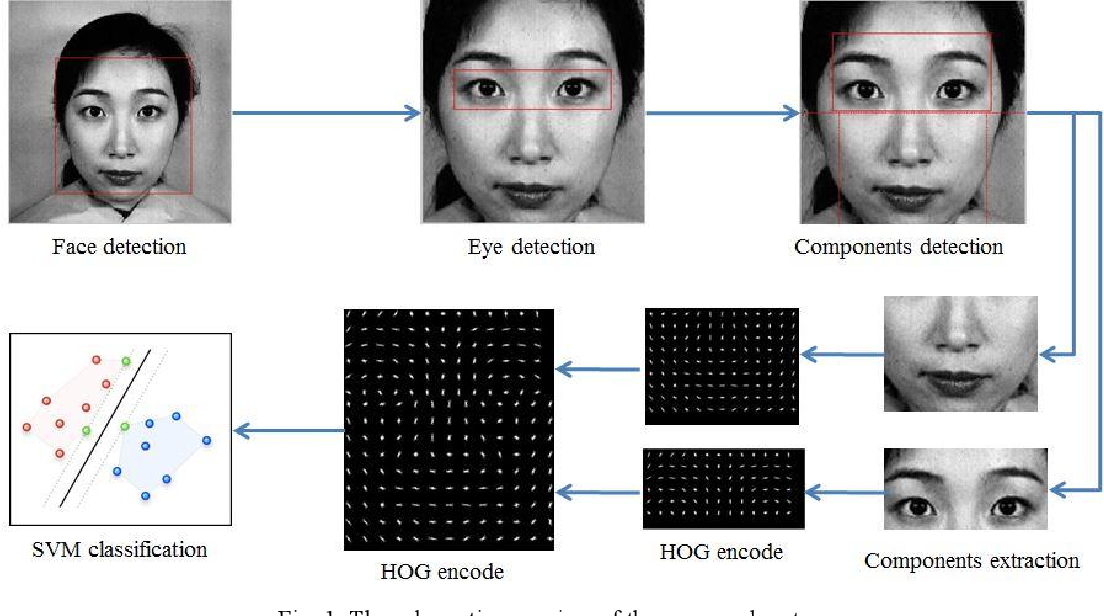
### Các nghiên cứu liên quan

#### *Phương pháp HOG*

Theo thông tin từ trang web [11], các dặc trưng HOG (histogram of oriented gradients) là một tính năng được sử dụng trong xử lý hình ảnh, dùng để phát hiện một đối tượng. Các khái niệm về HOG được nêu ra từ năm 1986 tuy nhiên cho đến năm 2005 HOG mới được sử dụng rộng rãi sau khi Navneet Dalal và Bill Triggs công bố những bổ sung về HOG.

Phương pháp HOG dựa trên việc đếm số lần xuất hiện của các hướng đạo hàm trong các vùng cục bộ của ảnh. Bản chất của HOG là các thông tin về hình dáng và vẻ bề ngoài của các đối tượng cục bộ trong ảnh có thể được mô tả bằng cách sử dụng thông tin về sự phân bố của các cường độ gradient và các hướng biên. Các toán tử HOG được cài đặt bằng cách chia nhỏ một bức ảnh thành một vùng con gọi là cells. Với mỗi cell, một histogram về các hướng của gradient sẽ được tính cho các điểm nằm trong cell. Ghép các histogram này lại, ta sẽ có một biểu diễn cho bức ảnh ban đầu.

Ví dụ như hình dưới đây:

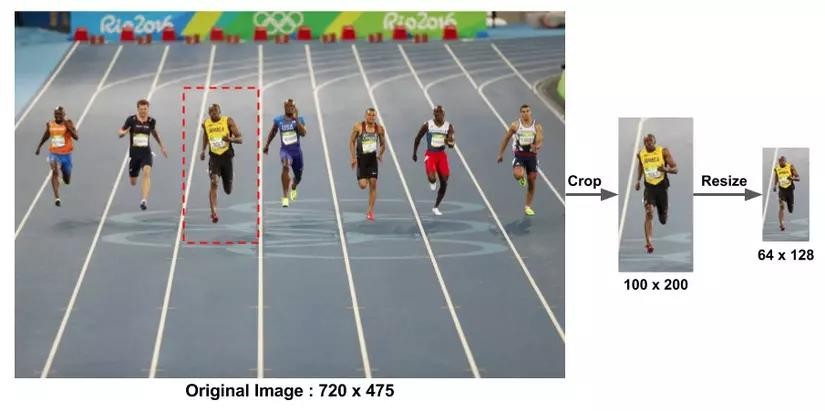


Hình 4. Ví dụ về phương pháp HOG

Các bước cơ bản của việc tính một vector HOG cho một ảnh gồm có: chuẩn hoá hình ảnh trước khi xử lý, tính gradient, gán hướng vào các giá trị khác nhau, tính histogram của các khối và cuối cùng là chuẩn hoá các khối.

### Chuẩn hoá hình ảnh trước khi xử lý

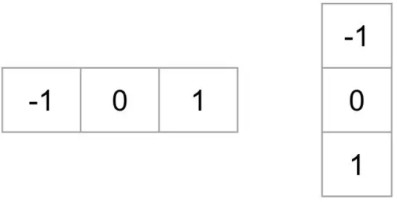
Để thuận tiện cho việc chia đều hình ảnh thành các khối, ô và tính toán đặc trưng ở các bước tiếp theo, chúng ta cần resize kích thước tất cả các hình ảnh trong tập dữ liệu về một kích thước chung.



Hình 5. Minh hoạ resize hình ảnh Ở đây, kích thước chung sẽ là 64 x 128.

### Tính gradient

Ở bước này, được thực hiện bằng hai phép nhân chập ảnh gốc với 2 chiều, tương ứng với các toán tử lấy đạo hàm theo hai hướng Ox và Oy. Trong đó, 2 hướng tương ứng đó là:



Hình 6. Hai chiều theo hai hướng Ox, Oy

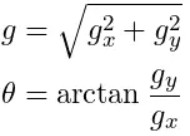


Trong đó, T là phép toán chuyển vị ma trận.

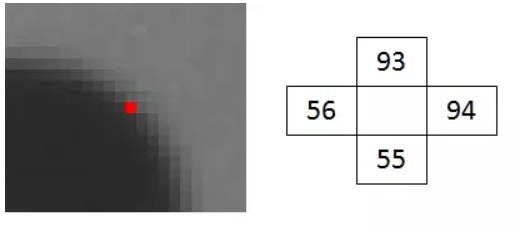
Giả sử, ta có một ảnh I, thì ta sẽ có 2 ảnh đạo hàm riêng theo 2 hướng đó, tương ứng với công thức:



Ta có thể tính được gradient bao gồm hai thành phần cường độ (gradient magnitude) và hướng (gradient direction) theo công thức sau:

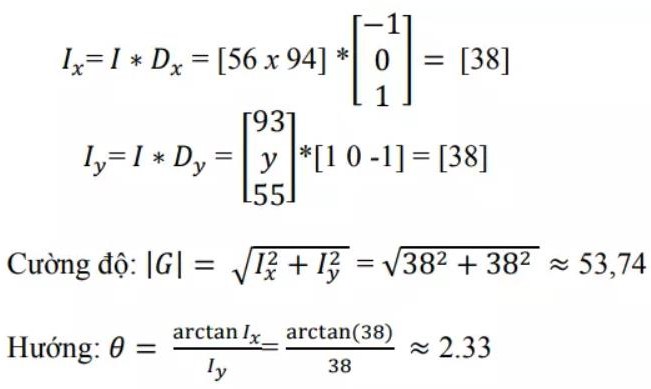


Ví dụ: Ta có một điểm ảnh:

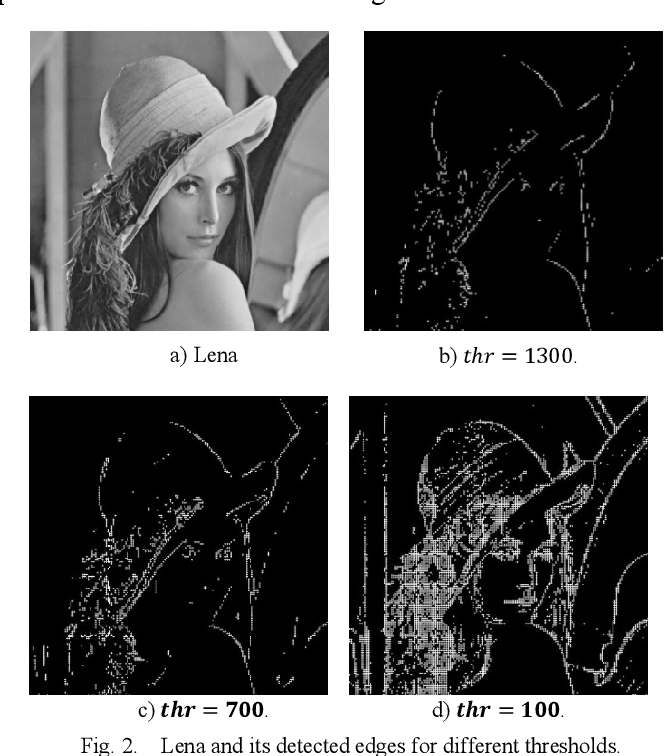


Hình 7. Ví dụ về một điểm ảnh

Chúng ta sẽ áp dụng các công thức trên để tính được gradient của điểm ảnh này.

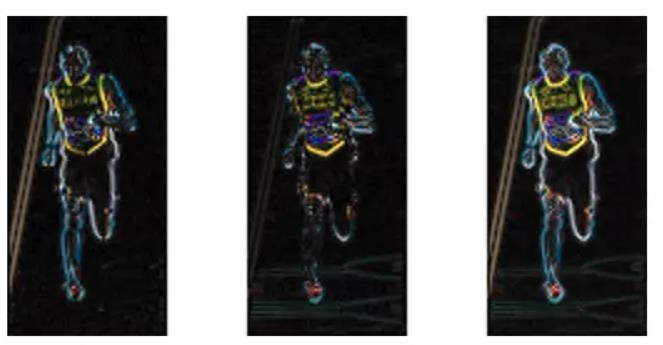


Kết quả của bước này, đối với ảnh trắng đen:



Hình 8. Kết quả tính gradient với hình trắng đen

Đối với ảnh màu gradient (red, green và blue) được đánh giá. Độ lớn của gradient tại một điểm ảnh là giá trị lớn nhất của cường độ gradient, và góc là góc tương ứng với gradient tối đa.



Hình 9. Kết quả tính gradient với hình màu

### Tính vector đặc tưng cho từng ô

Để tính toán vector đặc trưng cho từng ô, chúng ta cần chia hình ảnh thành các khối, mỗi khối lại chia đều thành các ô. Để xác định được số khối, chúng ta sẽ sử dụng công thức sau:

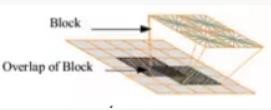


Trong đó:

W image là chiều rộng của ảnh W block là chiều rộng khối

W cell là chiều rộng của ô H image là chiều dài của ảnh H block là chiều dài của khối H cell là chiều dài của ô

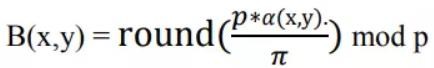
Các khối có thể xếp chồng lên nhau như hình:



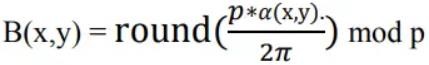
Hình 10. Hình ảnh minh hoạ các khối xếp chồng lên nhau

Sau khi xác định số block và kích thước mỗi khối, ô, để tính toán vector đặc trưng cho từng cell, chúng ta cần:

* + Chia không gian hướng thành p bin (số chiều vector đặc trưng của ô).
  + Rời rạc hóa góc hướng nghiêng tại mỗi điểm ảnh vào trong các bin. Giả sử góc hướng nghiêng tại pixel ở vị trí (x,y) có độ lớn là alpha(x,y): Trường hợp rời rạc hóa unsigned-HOG với p=9, ứng với công thức:



Trường hợp rời rạc hóa signed-HOG với p=18:



Giá trị bin được định lượng bởi tổng cường độ biến thiên của các pixels thuộc về bin đó. Sau khi tính toán đặc trưng ô, ta sẽ nối các vector đặc trưng ô để thu được vector đặc trưng khối. Số chiều vector đặc trưng khối tính theo công thức:



Trong đó:

Size block là số chiều vector đặc trưng khối n là số ô trong khối

size cell là số chiều của vector dặc trưng ô (sizecell = 9 nếu sử dụng unsigned-HOG, size cell = 18 nếu sử dụng signed-HOG)

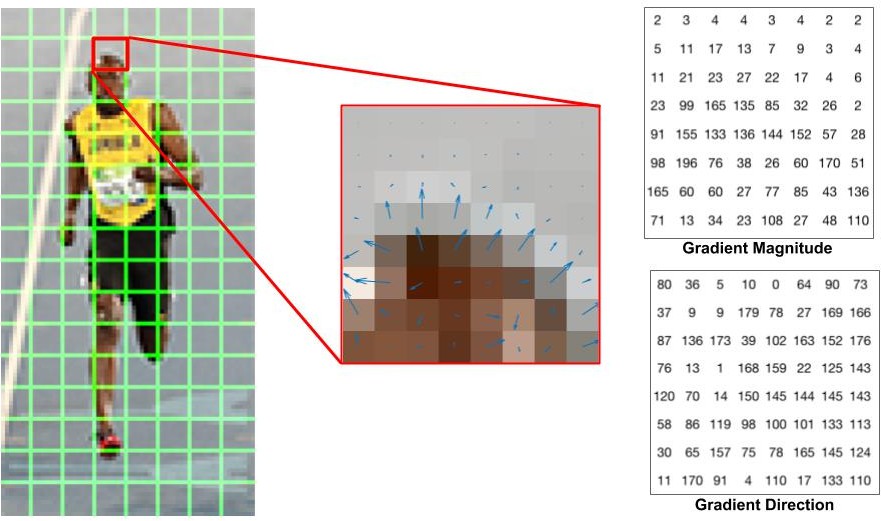
Như ví dụ trên, hình ảnh của chúng ta có kích thước 64x128, ta sẽ chia mỗi hình ảnh thành các khối có kích thước 16x16. Mỗi khối sẽ bao gồm 4 ô, mỗi ô có kích thước là 8x8.



Hình 11. Minh hoạ về xác định số khối

Tiếp theo, tiến hành tính toán đặc trưng HOG tại mỗi cell *sử dụng không gian hướng 9 bin*, trường hợp “unsigned-HOG”. Hướng gradient sẽ chạy trong khoảng 0 độ đến 180 độ.

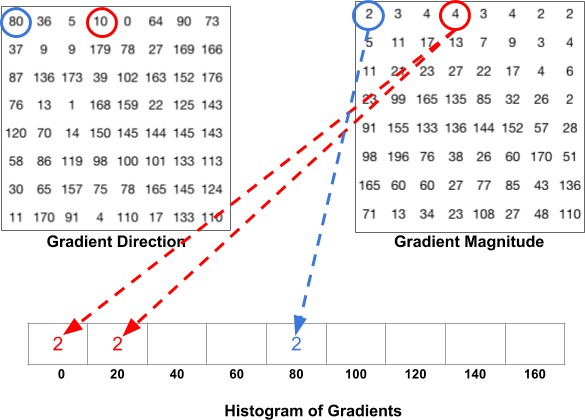
Ví dụ:



Hình 12. Vector đặc trưng của khối

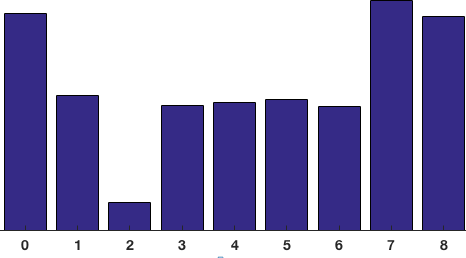
### Tính histogram của các khối

Tại mỗi ô, xây dựng một biểu đồ cường độ gradient bằng cách vote các pixel vào biểu đồ. Trọng số vote của mỗi pixel phụ thuộc hướng và cường độ gradient (được tính toán từ bước 2) của pixel đó.



Hình 13. Vote các pixel

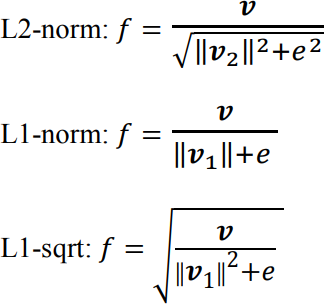
Như trong hình ảnh trên, đầu tiên là pixel có bao quanh màu xanh lam. Nó có hướng 80 độ và cường độ là 2, vì vậy ta thêm 2 vào bin thứ 5 (hướng 80 độ). Tiếp theo là pixel có bao quanh màu đỏ. Nó có hướng 10 độ và cường độ 4. Vì không có bin 10 độ, nên ta vote cho bin 0 độ và 20 độ, mỗi bin thêm 2 đơn vị. Sau khi vote hết các pixel trong một cell kích thước 8x8 vào 9 bin, ta có thể thu được kết quả như sau:



Hình 14. Histogram of gradients

### Chuẩn hoá các khối

Có nhiều phương pháp có thể được dùng để chuẩn hóa khối. Gọi v là vector cần chuẩn hóa chứa tất cả các histogram của một khối. ‖v(k)‖ là giá trị chuẩn hóa của v theo các chuẩn k=1, 3 và e là một hằng số nhỏ. Khi đó, các giá trị chuẩn hóa có thể tính bằng một trong những công thức sau:



Ghép các vector đặc trưng khối sẽ thu được vector đặc trưng R-HOG cho ảnh. Số chiều vector đặc trưng ảnh tính theo công thức:

size image = n \* size block Trong đó:

n là số khối của hình ảnh

size block là số chiều của vector đặc trưng khối

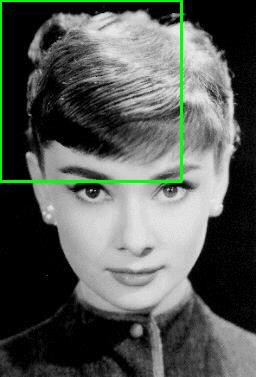
#### *Thuật toán Viola-Jones*

* + - 1. *Nội dung thuật toán*

Năm 2001, Paul Viola kết hợp Michael Jones đã xuất bản bài báo về thuật toán Viola-Jones lần đầu tiên của họ [5], bài báo này đã trở thành một trong những bài báo được trích dẫn nhiều nhất trong tài liệu về thị giác máy tính. Trong bài báo này, Viola và Jones đề xuất một thuật toán có khả năng phát hiện các vật thể trong hình ảnh, bất kể vị trí và tỷ lệ của chúng trong một hình ảnh. Hơn nữa, thuật toán này có thể chạy trong thời gian thực, giúp phát hiện các đối tượng trong video stream.

Cụ thể, Viola và Jones tập trung vào việc phát hiện khuôn mặt trong ảnh.

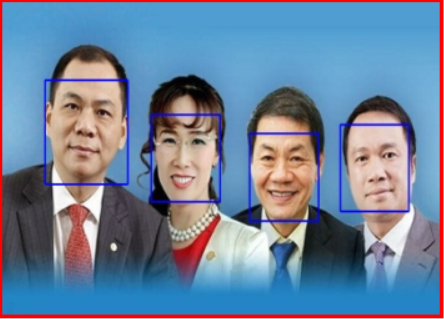
Hình ảnh tích hợp (Kernel Image) và cách trượt một ma trận nhỏ trên hình ảnh từ trái sang phải và từ trên xuống dưới, tính toán một giá trị đầu ra cho mỗi pixel trung tâm của kernel? Cách tiếp cận cửa sổ trượt này cũng cực kỳ hữu ích trong bối cảnh phát hiện các đối tượng trong một hình ảnh.



Hình 15. Trượt ma trận nhỏ trên hình ảnh

Trong hình trên, có thể thấy rằng đang trượt một cửa sổ kích thước cố định trên hình ảnh ở nhiều tỉ lệ. Tại mỗi giai đoạn này, cửa sổ dừng lại, tính toán một số tính năng và sau đó phân loại khu vực là có khuôn mặt hay không.

Rõ ràng, điều này đòi hỏi một chút về machine learning. Chúng ta cần một bộ phân loại được đào tạo sử dụng các mẫu positive và negative của khuôn mặt. Điểm dữ liệu positive sẽ là các ví dụ khu vực có chứa một khuôn mặt. Các điểm dữ liệu negative sẽ là ví dụ về các vùng không chứa khuôn mặt. Dựa vào những điểm dữ liệu này, chúng ta có thể đào tạo một bộ phân loại để nhận biết liệu một vùng nhất định của hình ảnh có chứa một khuôn mặt hay không.



Hình 16. Ví dụ về phát hiện gương mặt

### Lựa chọn tính năng Haar

Theo như Viola và Jones, đã công bố gồm có 4 đặc trưng cơ bản để xác định khuôn mặt [6]. Mỗi đặc trưng gồm hai hay ba hình chữ nhật trắng hoặc đen như trong hình:



Hình 17. Các đặc trưng cơ bản của Haar Like

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

Đặc trưng cạnh (edge features):

A picture containing text  Description automatically generated

Hình 18. Minh hoạ đặc trưng cạnh Đặc trưng đường (line features):

A close up of a logo  Description automatically generated with low confidence

Hình 19. Minh hoạ đặc trưng đường Đặc trưng xung quanh tâm (center surround):

Qr code  Description automatically generated

Hình 20. Minh hoạ đặc trưng xung quanh tâm

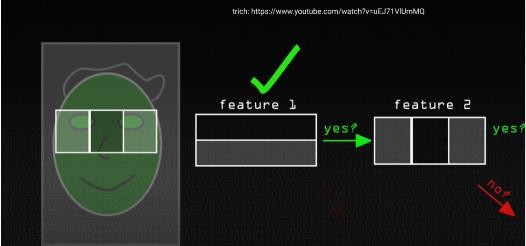
Dùng các đặc trưng trên ta có thể tính được các giá trị của đặc trưng haar-like là sự chênh lệnh giữa tổng các pixel của các vùng đen và các vùng trắng với công thức:

F(x) = X – Y

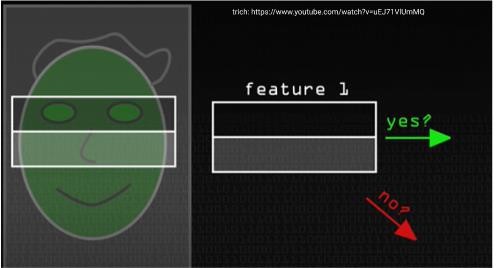
Trong đó: F(x) là giá trị đặc trưng, X là tổng các mức xám của pixel ở vùng đen, Y là tổng các mức xám của pixel ở vùng trắng.

Sử dụng giá trị này so sánh với các giá trị của pixel thô để đánh giá các đặc trưng của ảnh. Nếu xem xét tất cả các tham số của các đặc trưng, ta tính được khoảng 160.000+ đặc trưng cho mỗi cửa sổ.

Ví dụ: Vùng mũi sáng hơn vùng mắt.



Hình 21. Minh hoạ vùng mũi sáng hơn vùng mắt [10] Vùng mắt có xu hướng tối hơn vùng má.

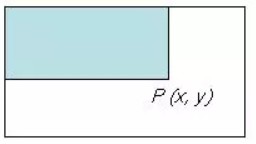


Hình 22. Minh hoạ vùng mắt tối hơn vùng má [10]

### Hình ảnh tích hợp

Bởi vì sử dụng tính năng haar-like, ở tất cả các kích thước và vị trí, dẫn đến kết quả cuối cùng là khoảng hai trăm nghìn tính năng để tính toán đó là một con số quá lớn Cho nên Viola-Jones đã đưa ra một khái niệm gọi là Integral Image (Hình ảnh tích hợp). Là một mảng hai chiều với kích thước bằng với kích thước của ảnh cần tính đặc trưng Mỗi phần tử của mảng hay một pixel nhất định trong hình ảnh tích phân (x, y) là tổng

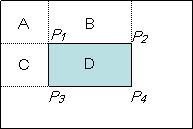
của tất cả pixel ở bên trái và tất cả các pixel ở trên nó. Cho phép tính tổng của các pixel trong bất kỳ hình chữ nhật chỉ với bốn giá trị ở bốn góc.



Hình 23. Biểu diễn hình ảnh tích hợp 1 Công thức tính Integral Image: A picture containing text, watch  Description automatically generated

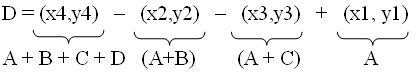
Sau khi tính được Integral Image, việc tính tổng các giá trị mức xám của một vùng bất kỳ nào đó trên ảnh thực hiện rất đơn giản theo cách sau:

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



Hình 24. Biểu diễn hình ảnh tích hợp 2

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:

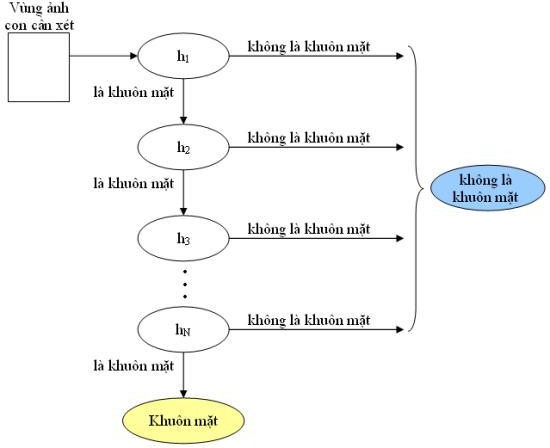


### Đào tạo AdaBoost

Có rất nhiều đặc trưng được lấy ra từ 1 cửa sổ nhưng chỉ có 1 số ít là hữu dụng trong việc nhận diện khuôn mặt. Tiếp theo để chọn các đặc trưng Haar-Like dùng cho việc thiết lập ngưỡng. Viola và Jones sử dụng một phương pháp máy học được gọi là AdaBoost. AdaBoost sẽ kết hợp các đặc trưng yếu tạo thành một bộ phân loại các đặc trưng mạnh. Bộ phân loại yếu chỉ cho ra câu trả lời ngẫu nhiên còn bộ phân loại mạnh sẽ cho ra câu trả lời chính xác hơn 60%.

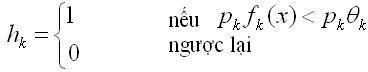
Sử dụng thuật toán AdaBoost để tìm những đặc trưng tốt nhất. Sau đó các đặc trưng này được gán cho các trọng số để tạo nên hàm đánh giá quyết định xem một cửa sổ có là khuôn mặt hay không. Mỗi đặc trưng chọn nếu chúng ít nhất thể hiện tốt hơn đoán ngẫu nhiên (phát hiện nhiều hơn một nửa).

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:



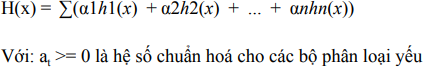
Hình 25. Mô hình phân tầng

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



với: x là cửa sổ con cần xét, 𝜃k là ngưỡng, fk là giá trị của đặc trưng Haar-Like, pk là hệ số quyết định chiều của phương trình.

Các đặc trưng được gọi là các bộ phân lớp yếu. Chúng được tổ hợp tuyến tính để tạo ra một bộ phân lớp mạnh.



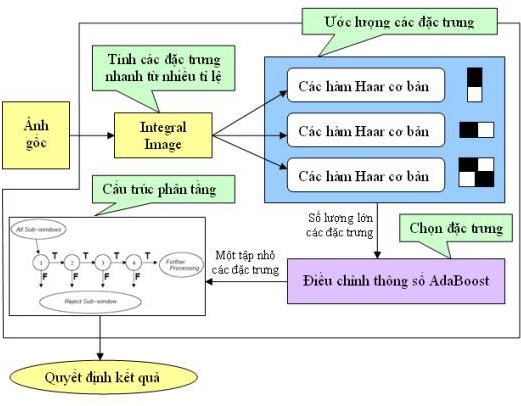
Mặc dù một ảnh có thể chứa một hoặc nhiều khuôn mặt nhưng số lượng vật không phải khuôn mặt vẫn lớn hơn rất nhiều. Vì vậy, thuật toán tập trung vào việc bỏ những vật không phải khuôn mặt một cách nhanh chóng.

Một bộ phân lớp cascade (cascade classifier) được sử dụng tất cả các đặc trưng được đưa vào các giai đoạn. Mỗi giai đoạn gồm một số các đặc trưng.

Mỗi giai đoạn được sử dụng để xác định một cửa sổ có phải là khuôn mặt hay

không.

Hệ thống nhận diện gương mặt:



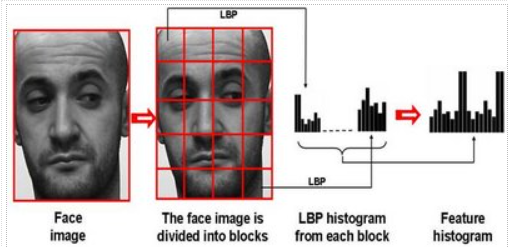
Hình 26. Mô tả hệ thống nhận diện gương mặt bằng Viola-Jones

* + - 1. Ưu điểm
         * Tỷ lệ phát hiện khuôn mặt cao.
         * Có thể phân biệt khuôn mặt với khuôn mặt hay không phải khuôn mặt từ các hình ảnh tùy ý.
         * Tỷ lệ nhận diện sai thấp, tỷ lệ nhận diện chính xác cao hơn.
         * Áp dụng trong thời gian thực.
      2. Nhược điểm
         * Phụ thuộc vào ánh sáng, khoảng cách.
    1. *Thuật toán Local Binary Pattern*
       1. Nội dung

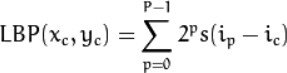
Thuật toán Local Binary Pattern được đề xuất Ojala và các cộng sự vào năm 2006 [7]. Nó dựa trên toán tử nhị phân cục bộ. Nó được sử dụng rộng rãi trong xử lý ảnh đặc

biệt trong nhận dạng khuôn mặt do tính đơn giản và sức mạnh phân biệt của nó. Thuật toán LBP là một phần của opencv.

Ý tưởng cơ bản của LPB là tóm tắt cấu trúc cục bộ trong một hình ảnh bằng cách so sánh từng pixcel với khu vực lân cận của nó. Lấy một pixcel làm trung tâm và tính toán các giá trị lân cận so với nó. Nếu cường độ của một pixcel lớn hơn hoặc bằng pixcel trung tâm thì chỉ định là 1 và nếu cường độ của một pixcel nhỏ hơn pixcel trung tâm thì chỉ định là 0. Sau đó ta sẽ thu được một dãy nhị phân. Đổi giá trị nhị phân sang thập phân để lưu trữ. Lặp hết toàn ảnh ta có kết quả đầu ra bằng kích thước với ảnh đầu vào. Mỗi giá trị trên ảnh đầu ra là đặc trưng LBP và phạm vi từ 0-255.



Hình 27. Ví dụ về thuật toán LBP Công thức của LBP được mô tả như sau:

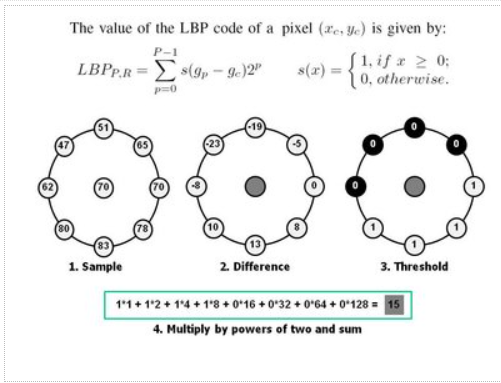


Trong đó '(Xc, Yc)' là pixel trung tâm với cường độ 'Ic'. Và 'Ip' là cường độ của pixel lân cận, s được xác định bởi:

op

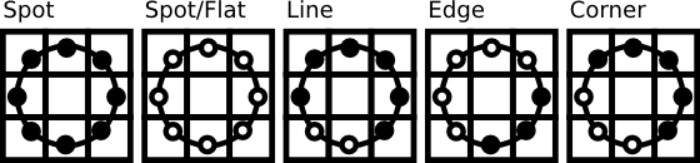
Do đó, chúng tôi nhận được tổng cộng 8 giá trị nhị phân từ 8 giá trị lân cận.

Sau đó, cần lưu ý rằng một vùng lân cận cố định không thể mã hóa các chi tiết khác nhau về tỷ lệ. Thuật toán đã được Ojala đã đưa ra giải pháp cải tiến thuật toán này là sử dụng số lượng bán kính và vùng lân cận khác nhau, bây giờ nó được gọi là LBP tròn.



Hình 28. LBP tròn

Ý tưởng ở đây là sắp xếp một số lượng lân cận tùy ý trên một vòng tròn có bán kính thay đổi. Bằng cách này, các vùng lân cận sau đây được chụp:



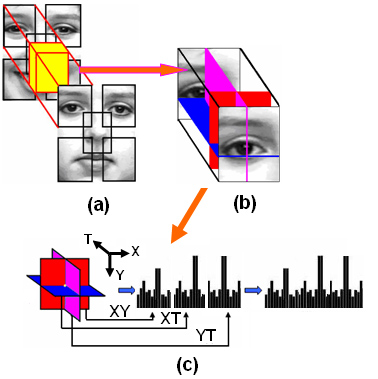
Hình 29. Các vùng lân cận được chụp



Hình 30. Hình ảnh khi áp dụng thuật toán

Sau khi tạo biểu đồ giá trị LBP của vùng được tạo bằng cách đếm số lượng các giá trị LBP tương tự trong vùng.

Sau khi tạo biểu đồ cho mỗi vùng, tất cả các biểu đồ được hợp nhất để tạo thành một biểu đồ duy nhất và đây được gọi là vectơ đặc trưng của hình ảnh.



Hình 31. Biểu đồ thể hiện vector đặc trưng của hình ảnh

Bây giờ chúng ta so sánh biểu đồ của hình ảnh thử nghiệm và các hình ảnh trong cơ sở dữ liệu và sau đó chúng tôi trả về hình ảnh với biểu đồ gần nhất. (Điều này có thể được thực hiện bằng nhiều kỹ thuật như khoảng cách euclid, chi-square, giá trị tuyệt đối, v.v.)

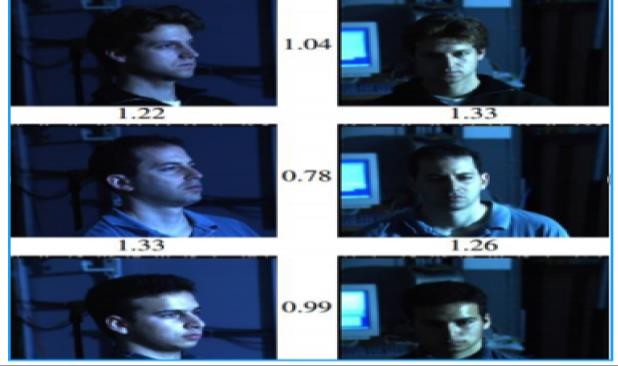
Khoảng cách Euclid được tính bằng cách so sánh các đặc điểm của hình ảnh thử nghiệm với các đặc điểm được lưu trữ trong bộ dữ liệu. Khoảng cách tối thiểu giữa hình ảnh thử nghiệm và hình ảnh gốc cho tỷ lệ khớp. Như một đầu ra, ta nhận được một id của hình ảnh từ cơ sở dữ liệu nếu hình ảnh thử nghiệm được nhận dạng.

LBPH có thể nhận dạng cả mặt bên và mặt trước và nó không bị ảnh hưởng bởi các biến thể chiếu sáng, có nghĩa là nó linh hoạt hơn.

* + - 1. Ưu điểm
         * Phương pháp LBPH là một trong những công cụ mô tả kết cấu hoạt động tốt.
         * Có thể nhận dạng cả mặt bên và mặt trước và nó không bị ảnh hưởng bởi các biến thể chiếu sáng.
         * Phương pháp LBPH có thể sẽ hoạt động tốt hơn so với trong các môi trường và điều kiện ánh sáng khác nhau, …
      2. Nhược điểm
         * Cần lưu trữ dữ liệu khuôn mặt lớn.
         * Tốn nhiều thời gian xử lý để nhận dạng.
         * Nó cũng phụ thuộc vào tập dữ liệu đào tạo
    1. *Nhận dạng khuôn mặt sử dụng thuật toán Face Net*
       1. Nội dung

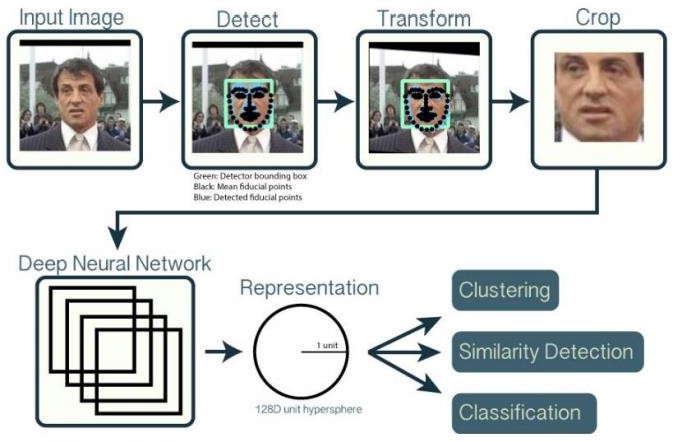
Google đã phát triển một thuật toán có tên là FaceNet [8]. Thuật toán này sẽ học cách ánh xạ từ ảnh khuôn mặt vào không gian Euclide compact với khoảng cách đo được tương ứng với độ tương đồng của khuôn mặt. Thuật toán này có thể tạo ra vector đặc trưng và nhúng vào bài toán nhận dạng khuôn mặt, kiểm tra khuôn mặt và phân cụm khuôn mặt. Nhóm tác giả sử dụng Mạng Tích Chập Sâu (Deep Convolution Network -

DNN) được huấn luyện để tự tối ưu hóa bài toán. Mạng được huấn luyện sao cho khoảng cách bình phương trong không gian nhúng tương ứng với mức độ tương đồng của khuôn mặt: Mặt cùng người sẽ có khoảng cách nhỏ, mặt khác người sẽ có khoảng cách lớn.



Hình 32. Hình minh hoạt output khoảng cách khi sử dụng FaceNet

Sau khi thực hiện phép nhúng, thu được vector đặc trưng thì ta có thể thực hiện được 3 bài toán: Kiểm tra khuôn mặt, ta chỉ cần phân ngưỡng khoảng cách giữa 2 vector đặc trưng của 2 khuôn mặt, sử dụng thuật toán K-Mean để phân cụm khuôn mặt.



Hình 33. Tóm tắt quy trình nhận dạng gương mặt sử dụng FaceNet

Ban đầu ta có ảnh đầu vào (input image). Xác định những điểm chính trên khuôn mặt (detect). Canh chỉnh lại khuôn mặt (Transform). Cắt khuôn mặt ra khỏi ảnh (crop). Đưa vào mạng Neuron Sâu (Deep Neural Network). Thu được vector đặc trưng 128

chiều dùng để biểu diễn khuôn mặt (Representation). Từ vector đặc trưng này có thể dùng để phân cụm khuôn mặt (Clustering), xác định tính tương đồng (Similarity Detection) và phân loại (Classification).

FaceNet sử dụng DNN, giả sử cấu trúc mô hình là một khối lớp, sau khi sử dụng cấu trúc DNN, vấn đề quan trọng nằm ở kết quả sau khi huấn luyện. Do đó, sử dụng đến bộ ba sai số có thể giúp kiểm tra, nhận dạng và phân cụm khuôn mặt.

FaceNet sử dụng CNN bằng cách sử dụng hàm f(x) và nhúng hình ảnh x vào không gian Euclid d chiều sao cho khoảng cách giữa các hình ảnh của một người không phụ thuộc vào điều kiện bên ngoài [9], khoảng cách của các khuôn mặt giống nhau của cùng một người là nhỏ. Trong khi, khoảng cách của các khuôn mặt khác nhau sẽ có khoảng cách lớn.

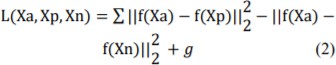
Ta biểu diễn phép nhúng là hàm 𝑓(𝑥) ∈ ℝ𝑑 có chức năng nhúng ảnh 𝑥 vào không gian Euclide 𝑑 chiều. Ví dụ, tại đây có 3 bức ảnh: Anchor (ảnh gốc), Positive (ảnh gần giống với ảnh gốc) và Negative (ảnh khác với ảnh gốc). Sau khi biểu diễn vào không

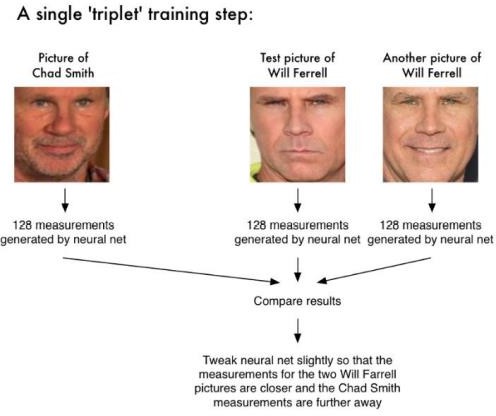
gian Euclide thì tương ứng với 3 bức ảnh trên là Xa, Xp, Xn. Để nhận dạng tốt thì khoảng cách từ Xa đến Xb sẽ phải nhỏ hơn Xa đến Xn: d (Xa, Xp) < d (Xa, Xn)

Do đó, dẫn đến biểu thức (1) với g là giá trị biên (1):



Lúc đó hàm Triplet loss sẽ có dạng như sau (2):



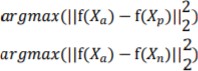


Hình 34. Ví dụ về bộ ba sai số

Khi huấn luyện mô hình cần xác định trước cặp ảnh (Xa, Xp), thuộc về cùng một người. Ảnh Xn là ảnh khác với ảnh gốc của người đó thường sẽ được lựa chọn ngẫu nhiên từ các bức ảnh thuộc các lớp còn lại. Do đó, tập hợp ảnh Xn thường được thu thập nhiều hơn một ảnh trên một người để có thể chuẩn bị được tập dữ liệu huấn luyện. Nếu một người chỉ có một ảnh thì có thể đưa những tập dữ liệu như vậy làm bộ ảnh Xn, khi huấn luyện.

Khi sử dụng triplet loss vào mô hình CNN có thể tạo ra các vector đặc trưng tốt nhất cho mỗi một bức ảnh. Các vector đặc trưng này sẽ cho phép phân biệt ảnh Negative (Ảnh khác với ảnh gốc). Hơn nữa, khoảng cách giữa các bức ảnh thuộc cùng một lớp sẽ trở nên gần nhau hơn trong không gian chiếu Euclidean.

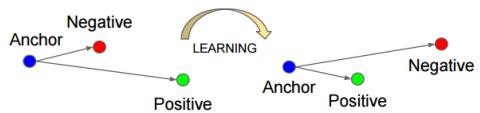
Tuy vậy, việc sử dụng bộ ba như trên sẽ khiến cho quá trình hội tụ chậm. Do đó, cần chọn bộ ba thích hợp trong quá trình huấn luyện để cải thiện được hiệu suất và độ chính xác của mô hình. Để khắc phục được việc hội tụ chậm, thường sẽ chọn bộ ba sai số sao cho khoảng cách giữa ảnh gốc và ảnh gần với ảnh gốc (ảnh của cùng 1 người) là lớn nhất và khoảng cách giữa ảnh gốc và ảnh của người khác là gần nhất:



Việc chọn hình ảnh như trên có thể xảy ra trường hợp:



Lúc này ta sẽ huấn luyện làm sao cho biểu thức (3) trở về biểu thức (2). Việc huấn luyện sẽ giúp khoảng cách giữa hai ảnh của cùng 1 người là nhỏ nhất và ngược lại ảnh của 2 người sẽ có khoảng cách là lớn nhất.



Hình 35. Bộ ba sai số tối thiểu hoá khoảng cách

Việc lựa chọn bộ ba sai số sẽ ảnh hưởng đến hiệu quả của mô hình, nếu giá trị bộ ba sai số được xác định tốt thì quá trình hội tụ khi huấn luyện sẽ nhanh hơn và kết quả sẽ cho độ chính xác cao hơn.

* + - 1. Ưu điểm

Thuật toán này đã lập nên kỷ lục mới trong nhận dạng khuôn mặt dưới nhiều điều kiện ảnh khác nhau.

* + - 1. Nhược điểm

FaceNet huấn luyện với một số lượng lớn hình ảnh (hơn 200 triệu ảnh của 8 triệu đối tượng), lớn gấp 3 lần so với các bộ dữ liệu hiện có. Để xây dựng bộ dữ liệu lớn như vậy rất khó thực hiện trong các phòng thiết bị, học thuật do đòi hỏi kiến trúc máy lớn.

### Ngôn ngữ lập trình python

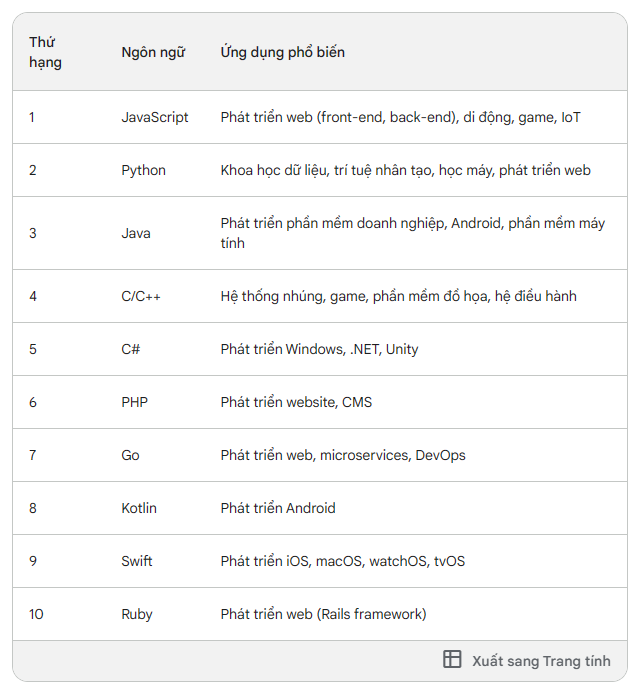
Python là ngôn ngữ lập trình đa năng, dễ học và dễ sử dụng, được sử dụng rộng rãi trong nhiều lĩnh vực khác nhau. Python được Guido van Rossum tạo ra vào năm 1991 và được phát triển bởi cộng đồng Python Software Foundation.Trong những năm gần đây, với sự bùng nổ của các lĩnh vực như Trí Tuệ Nhân Tạo (Artificial Intelligence) và Dữ Liệu Lớn (Big Data) đã góp phần gia tăng lớn nhu cầu sử dụng python.

Theo nguồn [2], cho biết ban đầu ngôn ngữ lập trình Python được thiết kế bởi [Guido van Rossum](https://twitter.com/gvanrossum?ref_src=twsrc%5Egoogle%7Ctwcamp%5Eserp%7Ctwgr%5Eauthor) vào năm 1991 và được duy trì, phát triển bởi tổ chức [Python](https://www.python.org/psf/) [Software Foundation](https://www.python.org/psf/). Ông đã nghiên cứu các sản phẩm dựa trên ứng dụng của mình tháng 12 năm 1989. Trong thời gian này, ông đã được tiếp xúc với ngôn ngữ lập trình ABC. Bằng sự sáng tạo và trợ giúp của đồng đội, ông đã kết hợp cú pháp của ngôn ngữ ABC và một số tính năng của nó để tạo ra ngôn ngữ lập trình mới và đặt tên là Python.

Cái tên này được lấy cảm hứng từ chương trình truyền hình Monty Python, Flying Circus

Circus mà ông rất hâm mộ.

Python được sử dụng phổ biến trong các ngôn ngữ lập trình. Theo khảo sát từ nguồn [2] cho biết, vào năm 2018 trên trang web [Stack Overflow](https://insights.stackoverflow.com/survey/2018), Python được xếp hạng thứ 7 trong những ngôn ngữ lập trình phổ biến nhất và xếp hạng thứ nhất trong những công nghệ được ưa thích nhất năm. Không những vậy, [những công ty phần mềm hàng](https://realpython.com/world-class-companies-using-python/) [đầu thế giới sử dụng Python trong các dự án của họ mỗi ngày.](https://realpython.com/world-class-companies-using-python/) Theo một báo cáo của [Dice](https://insights.dice.com/2016/02/01/whats-hot-and-not-in-tech-skills/), Python tính thời điểm hiện tại được xem như là một ngôn ngữ mà lập trình viên nào cũng nên biết và là ngôn ngữ phổ biến trong bảng xếp hạng của [Popularity](https://pypl.github.io/PYPL.html) [of Programming Language Index](https://pypl.github.io/PYPL.html).



Hình 36.Bảng xếp hạng mức độ phổ biến của các ngôn ngữ lập.

Python thì dễ học so với các ngôn như lập trình cao cấp như C++, C#, Javascript, Java... Trẻ nhỏ cũng có thể học Python nhờ vào cú pháp đơn giản dễ hiểu và gần với

ngôn ngữ giao tiếp thường ngày. Chính vì thế, ở nhiều quốc gia đã đưa ngôn ngữ lập trình này vào trong giáo dục. Nhiều trường học giảng dạy từ tiểu học cho đến trung học để cho các học sinh có thể dễ dàng tiếp thu và khơi dậy niềm đam mê lập trình.

Trình thông dịch của Python được phát triển dưới giấy phép mã nguồn mở (OSI- approved open-source license), chính vì vậy ta có thể tải và cài đặt nó hoàn toàn miễn phí, kể cả dùng cho các mục đích thương mại khác.

Python có để được viết trong một ngôn ngữ khác và có thể dùng trình biên dịch/thông dịch để thực thi chương trình. Ví dụ, chúng ta có thể viết mã nguồn Python và biên dịch nó bên trong ngôn ngữ C/C++.

Python có một đặc trưng khác biệt so với các ngôn ngữ lập trình C, C++, java… Các biến trong trong chương trình python khi khai báo không cần chỉ rõ kiểu dữ liệu cho nó và một biến có thể nắm giữ đối tượng có kiểu dữ liệu khác nhau

Python có đến hàng ngàn thư viện hỗ trợ ngôn ngữ lập trình với mục tiêu nhất định. Python được ứng dụng trong các lĩnh vực như: Phát triển web, internet, game, đồ hoạ 3D, phục vụ cho khoa học và tính toán đặc biệt sử dụng trong lĩnh vực trí tuệ nhân tạo.

### Giao diện tkinter

Theo tham khảo từ nguồn [3], tkinter là một module của python, được người dùng sử dụng để tạo ra giao diện. Đây là module được sử dụng phổ biến nhất cho các ứng dụng GUI trong python trong máy tính và các thiết bị đa phương tiện.

Tkinter là thư viện GUI tiêu chuẩn cho Python. Khi kết hợp với Tkinter, Python sẽ được cung cấp các công cụ một cách nhanh chóng và dễ dàng để tạo các ứng dụng GUI. Tkinter cung cấp giao diện hướng đối tượng mạnh mẽ đến các bộ công cụ Tk GUI.

### MySQL

**MySQL là một hệ thống quản trị cơ sở dữ liệu mã nguồn mở (gọi tắt là RDBMS) hoạt động theo mô hình client-server**. Với **RDBMS** là viết tắt của **Relational Database Management System**. **MySQL** được **tích hợp apache**, **PHP**. **MySQL** quản lý dữ liệu thông qua các cơ sở dữ liệu. Mỗi cơ sở dữ liệu có thể có nhiều bảng quan hệ chứa dữ liệu. **MySQL** cũng có cùng một cách truy xuất và mã lệnh tương tự với ngôn ngữ SQL.

* + 1. *Lịch sử phát triển*

Quá trình hình thành và phát triển của MySQL được tóm tắt như sau:

- Công ty Thuy Điển MySQL AB **phát triển MySQL** vào năm 1994.

**- Phiên bản đầu tiên của MySQL** phát hành năm 1995

- Công ty **Sun Microsystems** mua lại **MySQL AB** trong năm 2008

- Năm 2010 tập đoàn **Oracle** thâu tóm Sun Microsystems. Ngay lúc đó, đội ngũ phát triển của MySQL tách MySQL ra thành 1 nhánh riêng gọi là **MariaDB**. Oracle tiếp tục phát triển MySQL lên phiên bản 5.5.

- 2013 **MySQL** phát hành phiên bản 5.6

- 2015 **MySQL** phát hành phiên bản 5.7

**- MySQL** đang được phát triển lên phiên bản 8.0

## CHƯƠNG 3: PHÂN TÍCH THIẾT KẾ

### Yêu cầu hệ thống