

TRA CỨU THÔNG TIN SINH VIÊN QUA ẢNH KHUÔN MẶT

Nguyễn Hải Yến¹, Nguyễn Phương Hạc¹, Đinh Thị Mận¹, Nguyễn Văn Thịnh², Trần Thị Vân Anh^{1*}

¹ Khoa Công nghệ Thông tin, Trường Đại học Công thương TP.HCM

² Khoa Công nghệ thông tin, Trường Đại học Sư phạm TP.HCM

yennh@huit.edu.vn, hacnp@huit.edu.vn, mandt@huit.edu.vn, thinhnv@hcmue.edu.vn, anhttv@huit.edu.vn

TÓM TẮT—Để hỗ trợ thuận tiện hơn và tăng tốc độ tra cứu thông tin của sinh viên, bài báo này trình bày phương pháp nhận diện khuôn mặt để thực hiện tra cứu thông tin của sinh viên. Đầu tiên, mạng nơ ron xếp chồng MTCNN (multitask cascaded convolutional networks) được sử dụng để phát hiện khuôn mặt. Sau đó, phương pháp trích xuất đặc trưng HOG được dùng để trích xuất vector đặc trưng của ảnh khuôn mặt và sử dụng thuật toán phân lớp SVM (Support vector machine) huấn luyện mô hình nhận diện khuôn mặt. Kết quả thực nghiệm cho độ chính xác lần lượt 100% trên bộ dữ liệu ảnh YaleFace, 98,44% trên bộ dữ liệu ảnh YaleFaceB và 86,2% trên bộ dữ liệu sinh viên Trường Đại học Công thương TP.HCM (HUIT). Từ đó xây dựng ứng dụng tra cứu điểm rèn luyện của sinh viên qua nhận diện ảnh khuôn mặt. Mô hình nhận diện ảnh khuôn mặt trong bài báo này có thể tích hợp vào các hệ thống khác nhau, qua đó, người dùng có thể thực hiện tra cứu thông tin hoặc thao tác với hệ thống bằng cách xác thực qua ảnh khuôn mặt.

Từ khóa—Nhận diện khuôn mặt, MTCNN, HOG, SVM, tra cứu thông tin sinh viên.

I. GIỚI THIỆU

Trong xã hội hiện đại, công nghệ thông tin đóng vai trò quan trọng trong hầu hết các lĩnh vực đời sống từ quốc phòng, an ninh, kinh doanh, quản lý đến y tế, giáo dục góp phần giúp người dùng tiết kiệm thời gian, công sức và tăng hiệu quả công việc. Trước đây, trong các cơ sở giáo dục, việc thông báo kết quả học tập và rèn luyện của người học được thực hiện thủ công và định kỳ hàng tháng hoặc từng học kỳ. Ngày nay, công việc này đã tiện lợi và kịp thời hơn khi người học muốn xem thông tin kết quả học tập và rèn luyện của mình, chỉ cần đăng nhập vào hệ thống. Nghiên cứu này ứng dụng phương pháp học sâu để phát hiện khuôn mặt từ đó hỗ trợ người dùng đăng nhập hệ thống bằng gương mặt mà không cần phải nhớ mật khẩu. Ứng dụng được xây dựng cho phép định danh sinh viên và trả về kết quả điểm rèn luyện cá nhân từ ảnh đầu vào, góp phần hỗ trợ học sinh, sinh viên biết thông tin điểm rèn luyện của mình theo từng học kỳ, từ đó có kế hoạch tham gia các hoạt động để đảm bảo đạt đủ điểm rèn luyện khi xét tốt nghiệp.

Nhận dạng khuôn mặt (Face Recognition) là một hướng nghiên cứu quan trọng của công nghệ sinh trắc học (Biometrics) và thị giác máy tính (Computer Vision). Hiện nay, có nhiều công trình nghiên cứu về bài toán phát hiện và nhận diện khuôn mặt người từ ảnh đen trắng, ảnh xám, đến ảnh màu [1, 2]. Một trong những phương pháp phát hiện khuôn mặt đầu tiên dựa trên Adaboost kết hợp đặc trưng Haar-Like để phát hiện ảnh khuôn mặt được công bố năm 2001 do Viola-Jones và cộng sự đề xuất. Kỹ thuật phát hiện khuôn mặt này được áp dụng nhanh chóng cho bài toán nhận diện khuôn mặt với tốc độ phát hiện và xử lý khung hình 300px x 300px trong 0,07 giây với độ chính xác 90% [3]. Cho đến những năm gần đây, các mô hình học sâu như mạng nơ ron tích chập CNNs (convolutional neural networks) được sử dụng phổ biến trong các bài toán phân loại ảnh do ưu điểm về tính năng trích chọn đặc trưng và tính năng phân lớp được huấn luyện đồng thời [4-7]. Một so sánh trong nghiên cứu [8] cho thấy hiệu quả của các phương pháp học sâu CNNs trong các bài toán phân loại ảnh. Trong nghiên cứu [9], tác giả so sánh các phương pháp nhận diện ảnh bằng PCA (Principal Component Analysis), K-NN (K-Nearest Neighbour), LBPH (Local Binary Patterns Histograms) và CNN trên bộ dữ liệu ORL. Nghiên cứu trên cho kết quả khẳng định CNN có độ chính xác cao hơn các phương pháp khác. Một trong những kỹ thuật phân lớp ảnh có hiệu quả khác phải kể đến là máy vector hỗ trợ SVM. Trong các công bố [10] sử dụng nhiều thuật toán nhận dạng khuôn mặt như KDA, k-NN, SVM hoặc [11] kết hợp trích xuất đặc trưng HOG với SVM cũng cho hiệu quả nhận diện đáng kể. Công bố [12] sử dụng CNN để trích xuất đặc trưng sau đó dùng SVM để huấn luyện phân lớp nhằm nâng cao độ chính xác khi nhận dạng ảnh.

Các công bố trong nước gần đây cho thấy bài toán nhận diện khuôn mặt đang được nhiều nhà nghiên cứu quan tâm như nghiên cứu [13], sử dụng MTCNN kết hợp với mô hình FaceNet và bộ phân loại SVM để nhận diện khuôn mặt cho độ chính xác cao. Tuy nhiên phương pháp học sâu cần phải có bộ dữ liệu lớn để huấn luyện và tốn nhiều chi phí; công trình [14] đã nghiên cứu ứng dụng lý thuyết Compressive sensing (CS) vào bài toán nhận dạng khuôn mặt nói trên nhằm nâng cao hiệu suất nhận dạng và tối ưu hóa hệ thống. Kết quả cho thấy tỉ lệ nhận dạng khi thay đổi không gian đặc trưng qua các trường hợp đặc trưng khác nhau: với ngưỡng $D = 30, 56, 120, 504$ lần lượt là 98,26%; 98,16%; 99,65%; 100%. Về tốc độ xử lý qua các lần thử nghiệm cho ra các kết quả: 1,09; 1,16; 2,07; 7,07 (giây). Có thể thấy dù có kết quả chính xác cao nhưng đổi lại thời gian xử lý khá chậm làm tiêu tốn tài nguyên và chi phí. Để đáp ứng về mặt chi phí và khắc phục những điểm nêu ở trên, trong nghiên cứu này,

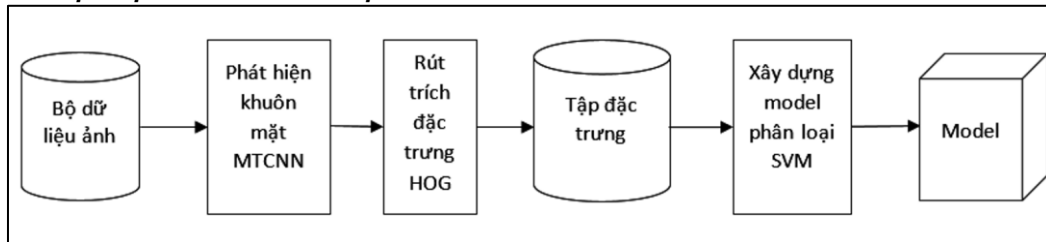
chúng tôi tiếp cận xây dựng mô hình nhận diện ảnh sử dụng MTCCN kết hợp với phương pháp trích xuất đặc trưng HOG và SVM nhằm tăng độ chính xác và tốc độ nhận diện của mô hình.

Đóng góp chính của bài báo gồm: (1) Kết hợp mạng MTCNN và phương pháp HOG để phát hiện và trích xuất đặc trưng khuôn mặt có tính phân loại cao; (2) xây dựng và huấn luyện mô hình nhận dạng ảnh khuôn mặt bằng thuật toán SVM với độ chính xác cao; (3) xây dựng hệ thống tra cứu điểm rèn luyện của sinh viên thông qua ảnh khuôn mặt được nhận diện từ mô hình đã huấn luyện. Phần còn lại của bài báo sẽ trình bày như sau: phần 2 trình bày phương pháp nghiên cứu của mô hình nhận dạng ảnh khuôn mặt; phần 3 kết quả thực nghiệm và cuối cùng là phần kết luận.

II. PHƯƠNG PHÁP NGHIÊN CỨU

Mô hình nhận dạng ảnh khuôn mặt bao gồm các bước: Phát hiện khuôn mặt, rút trích đặc trưng ảnh khuôn mặt sau đó các đặc trưng này được huấn luyện để xây dựng mô hình phân loại ảnh.

A. MÔ HÌNH NHẬN DẠNG ẢNH KHUÔN MẶT



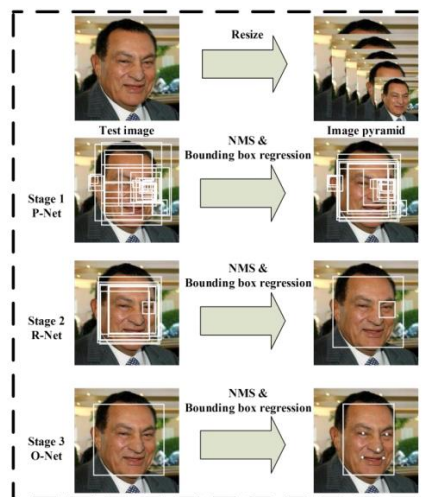
Hình 1. Mô hình phân loại ảnh.

Để xây dựng bộ phân loại gồm có 3 phần chính: Phát hiện khuôn mặt, rút trích đặc trưng, huấn luyện mô hình. Mô hình xây dựng Hình 1 gồm các bước: (1) Phát hiện khuôn mặt bằng mô hình MTCNN; (2) Rút trích đặc trưng HOG; (3) Huấn luyện thuật toán SVM để phân lớp các đặc trưng của dữ liệu ảnh đầu vào.

B. CHI TIẾT PHƯƠNG PHÁP NHẬN DẠNG

1. PHÁT HIỆN KHUÔN MẶT

MTCCN được đề xuất năm 2016 [15] là một kiến trúc hỗ trợ phát hiện khuôn mặt hiệu quả và căn chỉnh khuôn mặt trong thời gian thực, phương pháp này cho hiệu quả hơn so với các kỹ thuật hiện đại dò tìm khuôn mặt trên các dataset FDDB, WIDER FACE và AFLW về phương pháp căn chỉnh khuôn mặt. Kiến trúc mô hình MTCNN bao gồm 3 mạng CNN xếp chồng P-Net, R-Net và O-Net như Hình 2.

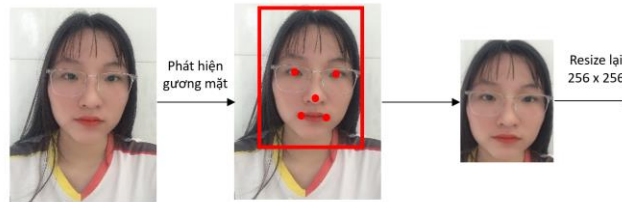


Hình 2. Minh họa kiến trúc MTCNN 3 giai đoạn. P-Net tách các cửa sổ ứng viên. R-Net. Chọn lọc lại các cửa sổ chứa khuôn mặt. O-Net. Tạo bounding box khuôn mặt và đánh dấu vị trí 5 điểm trên khuôn mặt [15]

Ban đầu, ảnh được hiệu chỉnh về các kích thước khác nhau để tạo thành một kim tự tháp hình ảnh làm đầu vào cho mô hình MTCNN. Giai đoạn 1 gọi là mạng đề xuất – P-Net (proposal network) dùng một mạng CNN để tạo ra các cửa sổ ứng viên và các vector hồi qui. Sau đó dùng các vector hồi qui này để hiệu chỉnh các vùng chứa khuôn mặt (bounding box), tiếp theo dùng phương pháp NMS (non-maximum) để hợp nhất các cửa sổ chọn ra các ứng cử viên là ảnh có chứa khuôn mặt. Giai đoạn 2 gọi là mạng tinh chỉnh R-Net (Refine network). Tất cả các ứng cử viên của giai đoạn P-Net được đưa vào một mạng CNN khác R-Net để tiếp tục loại bỏ một số lượng lớn cửa sổ

không chứa khuôn mặt và tiếp tục hiệu chỉnh bounding box bằng vector hồi qui và hợp nhất các ứng viên với NMS. Giai đoạn 3 gọi là mạng đầu ra O-Net (Output network). Thực hiện tương tự giai đoạn 2 để xác định 1 bounding box cuối cùng và đánh dấu 5 vị trí mốc trên khuôn mặt [15].

Hình 3 minh họa cho thuật toán MTCNN phát hiện khuôn mặt trên tập ảnh sinh viên HUIT. Sau khi phát hiện khuôn mặt kết quả trả về là ảnh khuôn mặt đã được resize lại là 256x256 là đầu vào cho bước trích xuất đặc trưng kế tiếp.



Hình 3. Phát hiện khuôn mặt

2. RÚT TRÍCH ĐẶC TRƯNG HOG

HOG là viết tắt của Histogram of Oriented Gradient - một loại bộ mô tả đặc trưng (*feature descriptor*). Mục đích của bộ mô tả đặc trưng là trừu tượng hóa đối tượng bằng cách trích xuất ra những đặc trưng của đối tượng đó và bỏ đi những thông tin không hữu ích. Vì vậy, HOG được sử dụng chủ yếu để mô tả hình dạng và sự xuất hiện của một đối tượng trong ảnh [11]. Có 5 bước cơ bản để xây dựng một vector HOG cho hình ảnh, bao gồm: tiền xử lý, tính gradient; gradient của một trường vô hướng là một vector có chiều hướng về phía mức độ tăng lớn nhất của trường vô hướng, và có độ lớn là mức độ thay đổi lớn nhất, tính vector đặc trưng cho từng ô (cells), chuẩn hóa khối (blocks), tính toán vector đặc trưng HOG.

Đầu vào của thuật toán là ảnh xám (gray scale) nên bước xử lý đầu tiên là chuyển đổi ảnh sang ảnh xám, kích thước ảnh được chuẩn hóa về 256x256. Với mỗi hình ảnh kích thước 256x256, chia thành các block 4x4 chồng nhau, sẽ có 5 block ngang và 5 block dọc, nên sẽ có 5x5 = 25 blocks. Tiếp theo, ta là tính gradient của ảnh bằng công thức 1:

$$|G| = \sqrt{F_x^2 + F_y^2} \text{ và } \theta = \tan^{-1}\left(\frac{F_y}{F_x}\right) \quad (1)$$

Với F_x là đạo hàm theo trục x, F_y là đạo hàm theo trục y được tính bởi công thức 2:

$$F_x = F * k_x; F_y = F * k_y \quad (2)$$

Trong đó $k_x = [-1 \ 0 \ 1]$; $k_y = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

Mỗi block gồm 16 ô. Khi áp dụng biểu đồ 9-bin cho mỗi cell, mỗi block sẽ được đại diện bởi một vector có kích thước 144x1. Vì vậy, khi nối tất cả các vector trong một block lại với nhau, thu được vector đặc trưng HOG của ảnh có kích thước 25x144x1 = 3600x1.

3. HUẤN LUYỆN MÔ HÌNH PHÂN LOẠI SVM

SVM là một phương pháp phân lớp hiệu quả đã được áp dụng trong một số lĩnh vực như phân loại văn bản, nhận dạng chữ viết tay, nhận dạng ảnh và nhận dạng khuôn mặt [12-14]. SVM là thuật toán có thể sử dụng cho việc phân loại và hồi qui. Đối với phân lớp ảnh khuôn mặt, đầu vào của SVM là các điểm dữ liệu và vector đặc trưng của ảnh khuôn mặt. SVM sẽ tìm ra một siêu phẳng tối ưu bằng hàm $f(x)$ và tính toán độ lệch của siêu phẳng với b để tách dữ liệu thành 2 lớp theo công thức (3)[15-16].

$$f(x) = b_0 + \sum (a_i * (x, x_i)) \quad (3)$$

Trong đó:

x : vector đặc trưng ảnh đầu vào

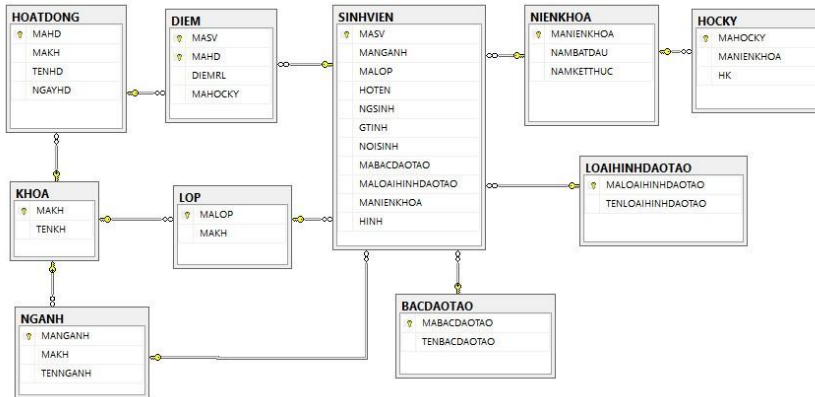
x_i : vector hỗ trợ

b_0, a_i : là các hệ số huấn luyện

Trong nghiên cứu này, thuật toán SVM được mở rộng theo phương pháp one-vs-all để phân loại đa lớp [16]. sử dụng kernel RBF với các tham số: $C = 100$, $\text{Gamma} = 0.001$, số lần lặp tối đa là 1000, epsilon bằng 0.0001. Tập dữ liệu được chia làm 2 phần, 80% dùng cho huấn luyện và 20% dùng để kiểm thử.

4. THIẾT KẾ CƠ SỞ DỮ LIỆU

Để thực hiện tra cứu điểm rèn luyện của sinh viên nhóm tác giả đã phân tích và thiết kế cơ sở dữ liệu quản lý điểm rèn luyện như sau:



Hình 4. Cơ sở dữ liệu quản lý điểm rèn luyện

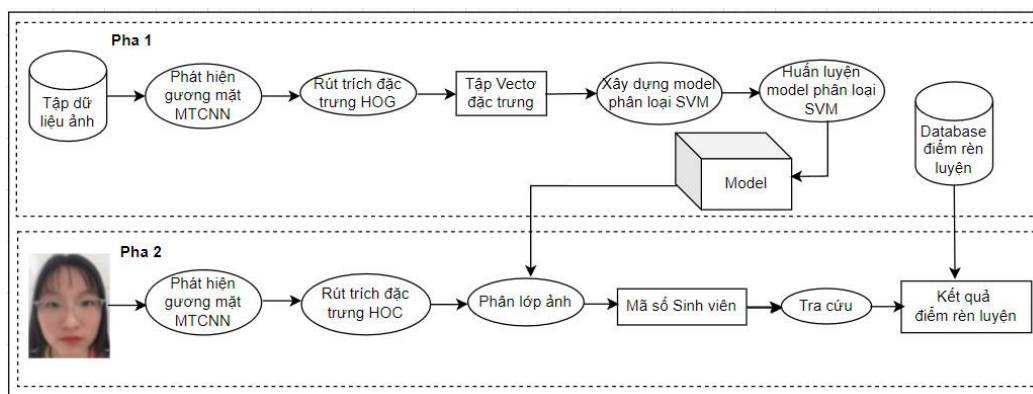
Một số bảng chính: Bảng KHOA lưu thông tin về các khoa trong trường; bảng NGANH lưu thông tin về các ngành đào tạo; bảng LOP lưu thông tin các lớp thuộc khoa quản lý; bảng SINHVIEN lưu thông tin sinh viên, mỗi sinh viên thuộc một khoa, một ngành và một lớp; bảng HOATDONG lưu thông tin về các hoạt động được cộng điểm rèn luyện; bảng DIEM lưu thông tin điểm rèn luyện của sinh viên trên mỗi hoạt động trong từng học kỳ; bảng HOCKY lưu thông tin học kỳ theo niên khóa; bảng NIENKHOA lưu thông tin về năm bắt đầu và năm kết thúc của một niên khóa; bảng BACDAOTAO lưu thông tin về các bậc đào tạo đang đào tạo tại trường (đại học, đại học liên thông, cao đẳng); bảng LOAIHINHDAOTAO lưu thông tin về hình thức đào tạo (chính quy, vừa học vừa làm).

Một số ràng buộc chính: Một sinh viên thuộc về một lớp của một khoa theo một ngành xác định. Một sinh viên có thể tham gia nhiều hoạt động trong một học kỳ, các hoạt động do trường hoặc khoa tổ chức; ứng với mỗi hoạt động tham gia, sinh viên được cộng một số điểm rèn luyện xác định. Mỗi sinh viên sẽ thuộc vào một bậc học và loại hình đào tạo xác định.

III. THỰC NGHIỆM VÀ ĐÁNH GIÁ

Mô hình thực nghiệm **Hình 5** gồm 2 pha: pha 1 xây dựng và huấn luyện mô hình phân loại ảnh, pha 2 nhận diện ảnh khuôn mặt của sinh viên thông qua mô hình và trả về kết quả điểm rèn luyện của sinh viên nhận dạng được từ mô hình.

A. MÔ HÌNH THỰC NGHIỆM



Hình 5. Mô hình hệ thống tra cứu điểm rèn luyện qua hình ảnh

Mô hình hệ thống tra cứu điểm rèn luyện sinh viên qua hình ảnh gồm có 2 pha chính. Pha 1 dùng để huấn luyện mô hình phân loại ảnh khuôn mặt gồm các bước: (1) Phát hiện khuôn mặt cho tập dữ liệu ảnh, (2) Rút trích đặc

trung HOG,(3) Xây dựng và huấn luyện model phân loại SVM để phân loại khuôn mặt của ảnh đầu vào, (4) Thiết kế cơ sở dữ liệu để quản lý điểm rèn luyện của sinh viên. Pha 2 dùng để nhận dạng và tra cứu điểm rèn luyện gồm các bước: (1) Phát hiện khuôn mặt cho ảnh đầu vào, (2) Rút trích đặc trưng HOG, (3) Sử dụng model phân loại SVM đã huấn luyện để nhận dạng ảnh đầu vào, (4) Lấy kết quả nhận dạng được tra cứu với cơ sở dữ liệu để đưa ra kết quả rèn luyện.

Để đo hiệu suất của phương pháp đề xuất, nghiên cứu sử dụng các độ đo gồm: độ chính xác (precision), độ phủ (recall) và độ trung hòa (F1-score) theo các công thức (4):

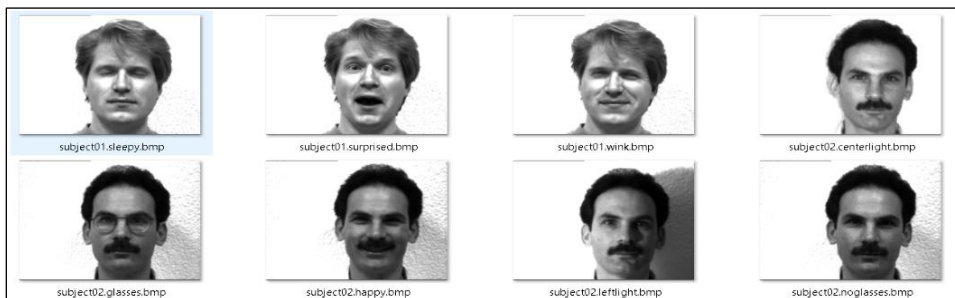
$$precision = \frac{\sum_{C=1}^C TP}{\sum_{C=1}^C (TP+FP)} \quad recall = \frac{\sum_{C=1}^C TP}{\sum_{C=1}^C (TP+FN)} \quad F1 - score = 2 \times \frac{(precision \times recall)}{(precision+recall)} \quad (4)$$

Trong đó, TP (True Positive): số ảnh thuộc lớp positive được phân loại đúng là positive; TN (True Negative): số ảnh thuộc lớp negative được phân loại đúng là negative; FP (False Positive): số ảnh thuộc lớp negative bị phân loại nhầm là positive; FN (False Negative): số ảnh thuộc lớp positive bị phân loại nhầm là negative; C là số lớp của tập dữ liệu ảnh.

B. DỮ LIỆU THỰC NGHIỆM

Nghiên cứu thực nghiệm trên 03 bộ dữ liệu YaleFace, YaleFaceB và bộ dữ liệu ảnh sinh viên HUIT như minh họa ở Hình 6 - Hình 8.

Tập ảnh YaleFace chứa 165 hình ảnh xám ở định dạng GIF của 15 cá nhân. Có 11 hình ảnh cho mỗi chủ thể, một hình ảnh cho mỗi biểu hiện hoặc cấu hình khuôn mặt khác nhau: ánh sáng trung tâm, có đeo kính, hạnh phúc, ánh sáng trái, không đeo kính, bình thường, ánh sáng phải, buồn, buồn ngủ, ngạc nhiên và nháy mắt.



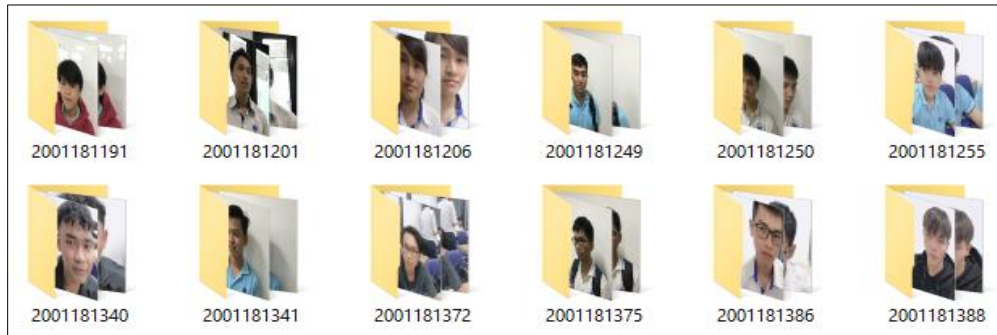
Hình 6. Tập ảnh YaleFace

Cơ sở dữ liệu khuôn mặt YaleFace B chứa 16128 hình ảnh của 28 người dưới 9 tư thế và 64 điều kiện chiếu sáng. Định dạng dữ liệu của cơ sở dữ liệu này giống như bộ ảnh YaleFace. Tất cả dữ liệu hình ảnh thử nghiệm được sử dụng trong các thí nghiệm được căn chỉnh thủ công, cắt xén và sau đó có kích thước lại thành hình ảnh 168x192.



Hình 7. Tập ảnh YaleFaceB

Tập ảnh sinh viên Khoa Công nghệ thông tin, Trường Đại học Công thương TP. Hồ Chí Minh gồm 80 bạn mỗi bạn từ 10-20 ảnh với các biểu cảm khác nhau: vui, buồn, hạnh phúc, ngạc nhiên,



Hình 8. Tập ảnh sinh viên khoa Công nghệ thông tin

C. KẾT QUẢ THỰC NGHIỆM

Thực nghiệm được tiến hành trên máy PC CPU Intel Core i5 - 1035G1 CPU @ 1.00GHz - @1.19 GHz, 4.0GB RAM, hệ điều hành Windows 10 Home 64 bit và google colab sử dụng GPU, sử dụng ngôn ngữ lập trình python, C# và .NET Framework 4.6.

Bảng 1 trình bày hiệu suất của phương pháp đề xuất trên 3 bộ ảnh YaleFace, YaleFaceB và trên bộ ảnh sinh viên HUIT với tổng số ảnh lần lượt là 165, 16128 và 1040 (80% dùng để huấn luyện, 20% dùng để kiểm thử). Thực nghiệm trên 2 bộ ảnh chuẩn cho kết quả chính xác gần tương tự nhau đạt 98,46% trên bộ ảnh YaleFace và 98,44% trên bộ ảnh YaleFaceB. Đối với bộ ảnh sinh viên HUIT độ chính xác đạt 86,2% thấp hơn so với 2 bộ ảnh chuẩn do chất lượng ảnh thu thập được, độ phủ và độ trung hòa đạt ở ngưỡng cao là 82,7% và 81,8%.

Bảng 1. Hiệu suất của phương pháp được đề xuất trên các bộ dữ liệu thực nghiệm.

Tập ảnh	Tổng số ảnh	Số ảnh huấn luyện	Số ảnh kiểm thử	Độ chính xác trung bình (precision)	Độ phủ trung bình (recall)	Độ trung hòa trung bình (f1-score)
YaleFace	165	132	33	100%	98,46%	100%
YaleFaceB	16128	13618	2510	98,44%	98,46%	98,44%
Sinh viên HUIT	1040	832	208	86,2%	82,7%	81,8%

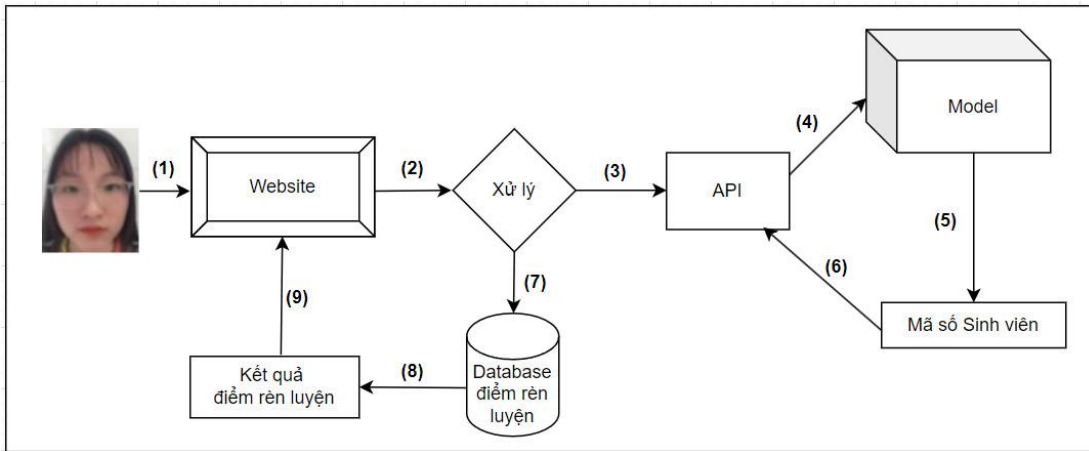
So sánh với các phương pháp đã được công bố trong lĩnh vực nhận diện khuôn mặt trên 2 bộ dữ liệu YaleFace và YaleFaceB thì phương pháp theo đề xuất của nhóm đạt kết quả khả quan hơn với độ chính xác 98,46% trên bộ ảnh YaleFace và 98,44% trên bộ ảnh YaleFaceB thể hiện ở bảng 2.

Bảng 2. So sánh với độ chính xác nhận dạng của các phương pháp khác.

Phương pháp	Tập ảnh YaleFace	Tập ảnh YaleFaceB
Trần Hồng Việt và cộng sự, 2021[13]	95%, 95,1%, 97%	-
Phạm Văn Trường, Trần Thị Thảo [2]	95%	-
Aydin Kayacik, 2008 [17]	78%	-
Nguyễn Thị Nông, 2012 [14]	-	98,26% ;99,65% ; 100%
T. Meenpal, 2018 [18]	-	91.7%
Đề xuất của nhóm	100%	98,44%

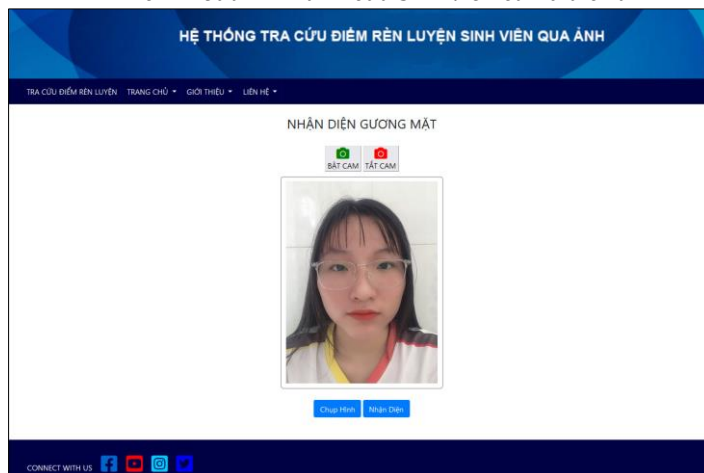
D. ỨNG DỤNG TRA CỨU ĐIỂM

Kiến trúc ứng dụng được thực thi theo các bước trình bày: (1) Lấy dữ liệu ảnh (load ảnh lên website), (2) xử lý ảnh thành chuỗi và mã hoá chuỗi, (3) truyền chuỗi lên web API đã dựng, (4) Web API sẽ gọi hàm nhận dạng khuôn mặt (đưa qua model phân loại đã huấn luyện), (5) trả về kết quả là mã số sinh viên khi nhận dạng thành công, (6) trả lại dữ liệu là kết quả đã nhận diện lên web API, (7) đưa qua lớp xử lý để tiến hành tra cứu kết quả rèn luyện, (8) lấy mã số sinh viên truy vấn vào cơ sở dữ liệu kết quả rèn luyện của sinh viên, (9) lấy được kết quả rèn luyện, trả về giao diện kết quả rèn luyện trên website như hình minh họa 9 -10.

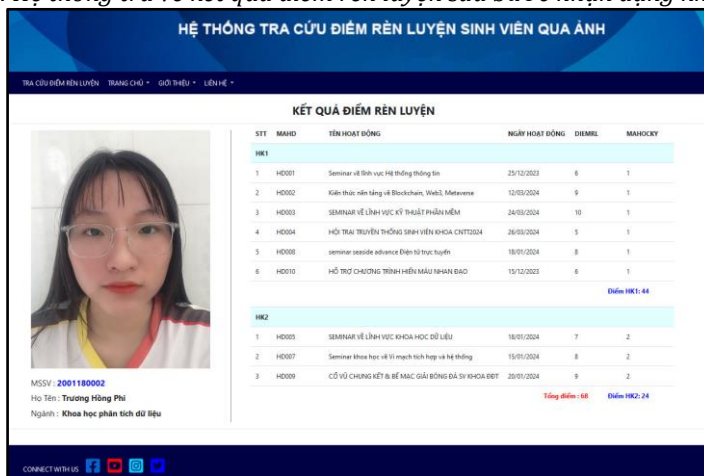


Hình 9. Kiến trúc ứng dụng tra cứu

Bước 1: Load hình ảnh của sinh viên cần tra cứu



Bước 2: Hệ thống trả về kết quả điểm rèn luyện sau bước nhận dạng khuôn mặt



Hình 10. Hình ảnh từ chức năng tra cứu điểm rèn luyện của sinh viên qua hệ thống nhận diện khuôn mặt

IV. KẾT LUẬN

Với mục đích hỗ trợ tra cứu thông tin của sinh viên nhanh chóng và chính xác hơn so với việc phải nhập thông tin khi đăng nhập, nhóm tác giả đã đề xuất giải pháp nhận diện sinh viên qua ảnh khuôn mặt bằng cách dùng MTCNN để phát hiện khuôn mặt sau đó dùng phương pháp trích xuất đặc trưng HOG để tạo vector đặc trưng tối ưu sau đó huấn luyện mô hình nhận diện bằng SVM. Mô hình sử dụng bộ dữ liệu ảnh YaleFace và YaleFaceB để thực nghiệm và đánh giá cho kết quả nhận diện có độ chính xác cao. Từ kết quả thực nghiệm, mô hình huấn luyện được xây dựng dựa trên bộ dữ liệu ảnh khuôn mặt của sinh viên và thông qua mô hình nhận diện tiếp tục

xây dựng hệ thống tra cứu điểm rèn luyện trên nền tảng web để hỗ trợ sinh viên tra cứu điểm nhanh chóng thông qua nhận diện khuôn mặt. Mô hình nhận dạng khuôn mặt trong bài báo này có thể tích hợp vào hệ thống tra cứu thông tin bất kỳ sử dụng ảnh khuôn mặt để đăng nhập.

V. TÀI LIỆU THAM KHẢO

- [1] Nicolas Ortiz, et al (2018), Survey of biometric pattern recognition via machine learning techniques, *Contemporary Engineering Sciences*, Vol. 11, No. 34, pp. 1677-1694.
- [2] P.V. Trường, T. T. Thảo (2020), Nhận dạng khuôn mặt trực tiếp và ứng dụng trong mở cửa tự động, *Tạp chí Điều khiển và Tự động hóa*, Vol.23, No.1-2, pp.52-58.
- [3] Viola P., Jones M.J (2004), Robust real-time face detection, *International Journal of Computer Vision*, Vol.57, No. 2, pp.137-154.
- [4] Khaled Younis, Abdullah A Alkhateeb (2017), A New Implementation of Deep Neural Networks for Optical Character Recognition and Face Recognition. *Proceedings of the New Trends in Information Technology*, 25-27 April 2017, The University of Jordan, Amman, Jordan.
- [5] A. Krizhevsky, I. Sutskever, G. E. Hinton (2012), Imagenet classification with deep convolutional neural networks, *Advances in Neural Information Processing Systems 25 (NIPS 2012)*.
- [6] K. Simonyan, Andrew Zisserman (2015), Very Deep Convolutional Networks for Large-Scale Image Recognition. *In Journal CoRR*, ICLR 2015, May 7-9, 2015, San Diego, CA, USA.
- [7] Lionel Landry S. De o, Elie T. Fute, Emmanuel Tonye (2018), A Convolutional Neural Network System for Face Detection and Recognition. *International Journal of Advanced Computer Science*.
- [8] Huỳnh Phước Hải, Nguyễn Văn Hòa, Đỗ Thanh Nghị (2017), So sánh mô hình học sâu với các phương pháp học tự động khác trong phân lớp dữ liệu biểu hiện gen microarray. *Kỷ yếu Hội nghị Quốc gia lần thứ XII về Nghiên cứu cơ bản và ứng dụng Công nghệ thông tin (FAIR)*.
- [9] Patrik Kamencay, Miroslav Benco, Tomas Mizdos, Roman Radil (2017), A New Method for Face Recognition Using Convolutional Neural Network. *Digital Image Processing and Computer Graphics*, Vol. 15, No. 4, pp.663-672.
- [10] Mustafa Zuhaer Nayef Al-Dabagh, MH Mohammed Alhabibat al.FH Al-Mukhtar(2018), Face recognition system based on kernel discriminant analysis, k-nearest neighbor and support vector machine, *International Journal of Research and Engineering*, Vol.5, No. 3, pp. 335-338.
- [11] J Kulandai Josephine JulinaT Sree Sharmila(2017), Facial recognition using histogram of gradients and support vector machines. *International Conference on Computer, Communication and Signal Processing (ICCCSP)*. IEEE.
- [12] Shanshan Guo, Shiyu Chenat al.Yanjie Li(2016). Face recognition based on convolutional neural network and support vector machine. *IEEE International conference on Information and Automation (ICIA)*.
- [13] Trần Hồng Việt, Đỗ Đình Tiến, Nguyễn Thị Trà ,Trần Lâm Quân (2021), Nhận diện khuôn mặt sử dụng mạng nơ-ron tích chập xếp chồng và mô hình facenet, *Tạp chí Khoa học và Công nghệ*, Vol. 57, No. 3.
- [14] Nguyễn Thị Nông (2012), Ứng dụng kỹ thuật compressive sensing trong nhận dạng khuôn mặt, *Hội thảo nghiên cứu khoa học "CNTT và ứng dụng CNTT trong các lĩnh vực"*.
- [15] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao (2016), Face detection and alignment using multitask cascaded convolutional networks, *IEEE Signal Process. Lett.*, Vol. 23, No. 10, pp. 1499–1503.
- [16] Corinna Cortes, Vladimir Vapnik (1995), Support-vector networks, *Machine learning*, Vol. 20, No. 3, pp. 273-297.
- [17] Aydin Kayacik (2008), Multiview Face Detection Using Gabor Filters and Support Vector Machine, *Bachelor's Thesis in Computer Systems Engineering*, Halmstad University.
- [18] Meenpal, T., Goyal, A., & Meenpal, A. (2018). Face recognition system based on principal components analysis and distance measures. *International Journal of Engineering & Technology*, Vol.7, No.21, pp.15-19.

QUERY STUDENTS' INFORMATION BY FACE IMAGE

Nguyen Hai Yen, Nguyen Phuong Hac, Dinh Thi Man, Nguyen Van Thinh, Tran Thi Van Anh

ABSTRACT— To more conveniently support and speed up student information queries, this paper presents a facial recognition method to perform student information queries. First, multitask cascaded convolutional networks (MTCNN) are used to detect faces. Then, use the HOG feature extraction method to extract the feature vector of the face image and the SVM (Support vector machine) classifier to train the face recognition model. Experimental results and model application give an accuracy of 100% on the YaleFace image data set, 98.44% on the YaleFaceB image data set, and 86.2% on the student data set, respectively. From there, build an application to identify students through facial images, perform searches in the database, and return students' conduct results. The facial image recognition model in this article can be integrated into

various systems, through which users can query information or manipulate the system by authenticating through facial images.

Keywords— Face Recognition, MTCNN, HOG, SVM, Query students' information.



Nguyễn Hải Yến (1979), tốt nghiệp ngành Toán Tin tại Trường Đại học Đà Lạt, năm 2001; tốt nghiệp Thạc sĩ ngành Khoa học máy tính Trường Đại học Công nghệ Thông Tin (UIT) năm 2017. Hiện nay là giảng viên Khoa Công nghệ Thông tin Trường Đại học Công Thương TP. HCM. Lĩnh vực quan tâm hiện nay là sử dụng các mô hình trong học máy và trí tuệ nhân tạo giải quyết các bài toán về xử lý ảnh.



Đinh Thị Mận tốt nghiệp ngành Công nghệ thông tin tại Trường Đại học Mở TP HCM, tốt nghiệp thạc sĩ ngành Công nghệ thông tin tại Trường Đại học Công nghệ TP HCM năm 2015. Hiện nay là giảng viên Khoa Công nghệ thông tin Trường Đại học Công Thương TP HCM. Lĩnh vực nghiên cứu ứng dụng mô hình học sâu trong bài toán chú thích ảnh



Trần Thị Vân Anh (1985), tốt nghiệp ngành Công nghệ thông tin tại Trường Đại học Khoa học tự nhiên TP HCM, tốt nghiệp thạc sĩ ngành Công nghệ thông tin tại Trường Đại học Công nghệ TP HCM năm 2016. Hiện nay là giảng viên Khoa Công nghệ thông tin Trường Đại học Công Thương TP HCM. Lĩnh vực nghiên cứu: ứng dụng mô hình học sâu trong nhận diện khuôn mặt.



Nguyễn Văn Thịnh (1979), tốt nghiệp kỹ sư Công nghệ thông tin tại Đại học Kỹ thuật Công nghệ TP. HCM và thạc sĩ Hệ thống thông tin tại Trường Đại học Khoa học tự nhiên - ĐH Quốc gia TP. HCM năm 2002 và 2011. Ông hiện là giảng viên Khoa Công nghệ thông tin, Trường Đại học Sư phạm TP. HCM và nghiên cứu sinh ngành Khoa học máy tính tại Học viện Khoa học và Công nghệ - Viện Hàn lâm Khoa học và Công nghệ Việt Nam. Lĩnh vực nghiên cứu mà ông quan tâm bao gồm trí tuệ nhân tạo và ứng dụng, truy vấn thông tin, chú thích hình ảnh.



Nguyễn Phương Hạc (1979), tốt nghiệp ngành Toán Tin tại Trường Đại học Khoa học Tự nhiên TP.HCM năm 2002, tốt nghiệp thạc sĩ ngành Công nghệ thông tin tại Trường Đại học Bách Khoa Hà Nội năm 2010. Hiện nay là giảng viên Khoa Công nghệ thông tin Trường Đại học Công Thương TP HCM. Lĩnh vực nghiên cứu: ứng dụng mô hình học sâu trong nhận diện ảnh