

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



NGUYỄN QUANG HUY

HỆ THỐNG NHẬN DIỆN KHUÔN MẶT QUA CAMERA

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT

HÀ NỘI - 2020

Luận văn được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: TS. NGUYỄN ĐÌNH HÓA

Phản biện 1: PGS.TS.Lê Hữu Lập

Phản biện 2: TS.Tạ Quang Hùng.....

Luận văn sẽ được bảo vệ trước Hội đồng chấm luận văn thạc sĩ tại Học viện
Công nghệ Bưu chính Viễn thông

Vào lúc: 8 giờ ngày 09 tháng 01 năm 2021

Có thể tìm hiểu luận văn tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông

MỤC LỤC

DANH MỤC CÁC BẢNG	ii
DANH MỤC CÁC HÌNH	iii
MỞ ĐẦU	1
CHƯƠNG 1. TỔNG QUAN VỀ NHẬN DIỆN KHUÔN MẶT	3
1.1. Tổng quan về nhận diện khuôn mặt cùng với các ứng dụng thực tế của các kỹ thuật nhận dạng khuôn mặt.	3
1.1.1 Tổng quan	3
1.1.2 Kiến trúc tổng quát hệ thống nhận diện	3
1.1.3 Ứng dụng	3
1.2. Một số phương pháp trong nhận diện khuôn mặt thường được áp dụng trong thực tế và nghiên cứu	3
1.2.1 Phân tích thành phần chính(PCA)	3
1.2.2 Phân tích sự khác biệt tuyến tính(LDA)	3
1.2.3 Cây quyết định (Decision Tree)	4
1.2.4 Mạng nơ-ron nhân tạo	4
1.2.5 Mạng nơ-ron tích chập	8
1.3 Phương pháp xác định vị trí khuôn mặt với mạng tích chập MTCNN	8
1.4 Kết luận	9
CHƯƠNG 2. HỆ THỐNG NHẬN DIỆN KHUÔN MẶT DỰA TRÊN MẠNG NƠ RON TÍCH CHẬP	9
2.1 Sơ đồ thiết kế hệ thống nhận diện khuôn mặt	9
2.2 Mạng Inception-ResNet sử dụng cho việc trích chọn đặc trưng khuôn mặt	10
2.3 Rừng ngẫu nhiên	11
2.3.1 Giới thiệu	11
2.3.2 Kiến trúc	12
2.4 Kết luận	12
CHƯƠNG 3. THỬ NGHIỆM VÀ ĐÁNH GIÁ	13
3.1 Bộ dữ liệu đầu vào	13
3.2 Quá trình huấn luyện	13
3.3 Thử nghiệm chạy hệ thống nhận diện khuôn mặt nhận diện khách hàng VIP của khách sạn	13
3.4 Đánh giá	15
3.5 Kết luận	17
KẾT LUẬN	18
DANH MỤC CÁC TÀI LIỆU THAM KHẢO	19

DANH MỤC CÁC BẢNG

Bảng 1.1. Các hàm kích hoạt.....	5
Bảng 3.1. Bảng đánh giá độ chính xác giữa các mô hình	16

DANH MỤC CÁC HÌNH

Hình 1.1. Kiến trúc tổng quát về hệ thống nhận diện	3
Hình 1.2. Cấu tạo của Perceptrons	4
Hình 1.3. Cấu trúc của nơ-ron nhân tạo	5
Hình 1.4. Cấu tạo của mạng truyền thẳng	7
Hình 1.5. Cấu tạo của mạng MLP	7
Hình 1.6. Kiến trúc mạng CNN.....	8
Hình 2.1. Sơ đồ hoạt động của hệ thống nhận diện khuôn mặt	10
Hình 2.2. Kiến trúc mạng Inception-ResNet.....	11
Hình 2.3. Kiến trúc của rừng ngẫu nhiên	12
Hình 3.1. Hệ thống nhận diện khuôn bình thường	14
Hình 3.2. Hệ thống nhận diện khuôn mặt có đeo kính.....	15

MỞ ĐẦU

Công nghệ thông tin ngày càng phát triển và đã là một thành phần không thể thiếu trong hầu hết mọi lĩnh vực trên thế giới. Những người máy thông minh được con người tạo ra đã có khả năng phân tích và xử lý được các công việc của con người một cách tự động và đem lại lợi ích kinh tế rất lớn. Trong thời gian gần đây, một trong những bài toán được nghiên cứu, ứng dụng nhiều nhất vào trong cuộc sống đó là bài toán nhận diện. Tuy mới xuất hiện chưa lâu nhưng nó đã rất được quan tâm vì tính ứng dụng thực tế của bài toán như nhận dạng chữ viết, nhận dạng giọng nói, nhận dạng hình dáng, nhận diện khuôn mặt. Trong đó, bài toán nhận diện khuôn mặt là một chủ đề đang được khá nhiều nhà đầu tư, doanh nghiệp quan tâm đến. Dù đã được nghiên cứu từ rất lâu nhưng bài toán nhận diện khuôn mặt vẫn đang gặp phải nhiều thách thức và vẫn chưa có phương pháp cụ thể nào có thể giải quyết hết các vấn đề trong bài toán này.

Bài toán nhận diện khuôn mặt là một trong những chủ đề đang được quan tâm nhiều nhất. Ứng dụng từ bài toán này được áp dụng trong rất nhiều lĩnh vực khác nhau. Các ứng dụng liên quan đến nhận diện khuôn mặt có thể kể như: tra cứu thông tin tội phạm, phát hiện tội phạm tại các nơi công cộng, tìm người lạc, điểm danh học sinh ...

Từ những phân tích và khảo sát ở trên, em nhận thấy hệ thống nhận diện khuôn mặt rất có ý nghĩa trong thực tiễn cuộc sống và em xin chọn đề tài nghiên cứu “Hệ thống nhận diện khuôn mặt qua camera”. Kết quả của luận văn hướng tới việc xây một hệ thống nhận diện khuôn mặt có khả năng mở khả năng mở rộng cao, dễ dàng tích hợp.

Nội dung luận văn được trình bày trong ba chương với các nội dung chính như sau:

Chương 1. Tổng quan về nhận diện khuôn

Chương này sẽ trình bày một số nội dung nền tảng về bài toán nhận diện khuôn mặt, các ứng dụng tương tác người máy liên quan đến nhận diện khuôn mặt, và một số kỹ thuật hay được sử dụng trong bài toán nhận diện khuôn mặt. Nội dung của

chương bao gồm ba phần chính. Phần đầu tiên giới thiệu tổng quan về bài toán nhận diện khuôn mặt cùng với các ứng dụng thực tế. Phần thứ hai giới thiệu một số phương pháp trong nhận diện khuôn mặt thường được áp dụng trong thực tế và nghiên cứu. Phần cuối cùng giới thiệu một số mạng tích chập thường được sử dụng trong bài toán nhận diện khuôn mặt.

Chương 2. Hệ thống nhận diện khuôn mặt dựa trên mạng nơ ron tích chập

Các kỹ thuật cơ bản được sử dụng để xây dựng hệ thống nhận diện khuôn mặt của luận văn được trình bày trong chương này. Nội dung của chương trình bày về các phương pháp trích chọn đặc trưng phục vụ quá trình nhận diện khuôn mặt, phương pháp định danh khuôn mặt và mô hình học máy được sử dụng để phân loại dữ liệu nhận diện khuôn mặt. Chương này cũng bao gồm các thông tin về mô hình, kiến trúc mạng nơ ron tích chập Inception-ResNet sử dụng cho việc trích chọn đặc trưng khuôn mặt của luận văn.

Chương 3. Thử nghiệm và đánh giá

Chương này mô tả chi tiết về bộ dữ liệu được sử dụng, cùng các kịch bản và kết quả các quá trình huấn luyện mô hình. Các kết quả thực nghiệm kèm theo đánh giá mô hình sau khi huấn luyện cũng được trình bày trong chương này.

Nội dung của luận văn được kết thúc bằng phần Kết luận, trong đó trình bày tóm lược các nội dung và kết quả đã đạt được trong luận văn, từ đó đề xuất các hướng phát triển trong tương lai.

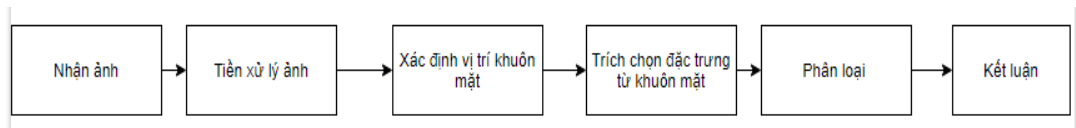
CHƯƠNG 1. TỔNG QUAN VỀ NHẬN DIỆN KHUÔN MẶT

1.1. Tổng quan về nhận diện khuôn mặt cùng với các ứng dụng thực tế của các kỹ thuật nhận dạng khuôn mặt.

1.1.1 Tổng quan

Nhận diện khuôn mặt là một bài toán tổng hợp. Trong đó ta cần các mô đun quan trọng như như xác định vị trí khuôn mặt, trích chọn đặc trưng rồi phân loại. Từ đó ta có thể xác định danh tính người trong ảnh.

1.1.2 Kiến trúc tổng quát hệ thống nhận diện



Hình 1.1. Kiến trúc tổng quát về hệ thống nhận diện

1.1.3 Ứng dụng

Bài toán nhận diện khuôn mặt có rất nhiều ứng dụng trong cuộc sống. Trong đó, một số ứng dụng tiêu biểu không thể không kể đến của bài toán này là hệ thống phát hiện, truy vết tội phạm, hệ thống tìm trẻ lạc, hệ thống điểm danh, chấm công hay ứng dụng nhận diện đối tác, khách hàng VIP. Các bài toán trên hiện đang được sử dụng rất nhiều và thành một phần không thể thiếu trong cuộc sống của mỗi người.

1.2. Một số phương pháp trong nhận diện khuôn mặt thường được áp dụng trong thực tế và nghiên cứu

1.2.1 Phân tích thành phần chính(PCA)

PCA (Principle Components Analysis) là một thuật toán được sử dụng để tạo ra một ảnh mới từ ảnh ban đầu. Ảnh mới này có kích thước nhỏ hơn nhiều so với ảnh ban đầu nhưng vẫn mang những đặc trưng cơ bản nhất của ảnh cần nhận dạng

1.2.2 Phân tích sự khác biệt tuyến tính(LDA)

LDA có thể được coi là một phương pháp giảm chiều dữ liệu (dimensionality reduction), và cũng có thể được coi là một phương pháp phân lớp (classification), và cũng có thể được áp dụng đồng thời cho cả hai, tức giảm chiều dữ liệu sao cho việc phân lớp hiệu quả nhất.

1.2.3 Cây quyết định (Decision Tree)

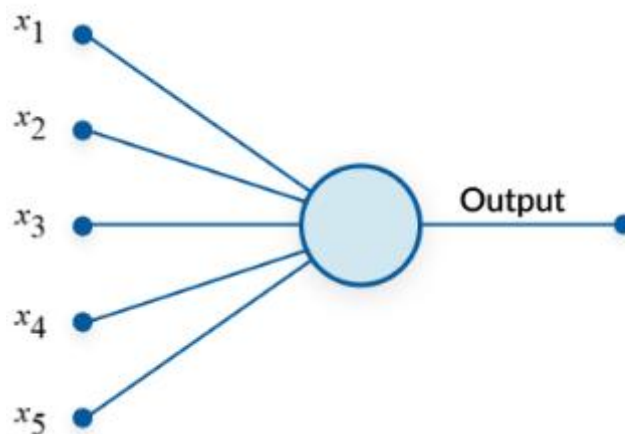
Việc quan sát, suy nghĩ và đưa ra các quyết định của con người thường được bắt đầu từ các câu hỏi. Trong học máy cũng có mô hình đưa ra quyết định dựa vào các câu hỏi như cây quyết định. Cây quyết định (Decision Tree) là một trong những thuật toán phổ biến của học máy thuộc nhánh học có giám sát.

1.2.4 Mạng nơ-ron nhân tạo

Mạng nơ-ron nhân tạo (Neural Network - NN) là một mô hình lập trình rất đẹp lấy cảm hứng từ mạng nơ-ron thần kinh và nó đang trở thành một công cụ rất mạnh mẽ đem lại hiệu quả rất tốt trong các bài toán khó có thể kể đến như xử lý ngôn ngữ, phân loại ảnh, giọng nói.

a. Perceptrons

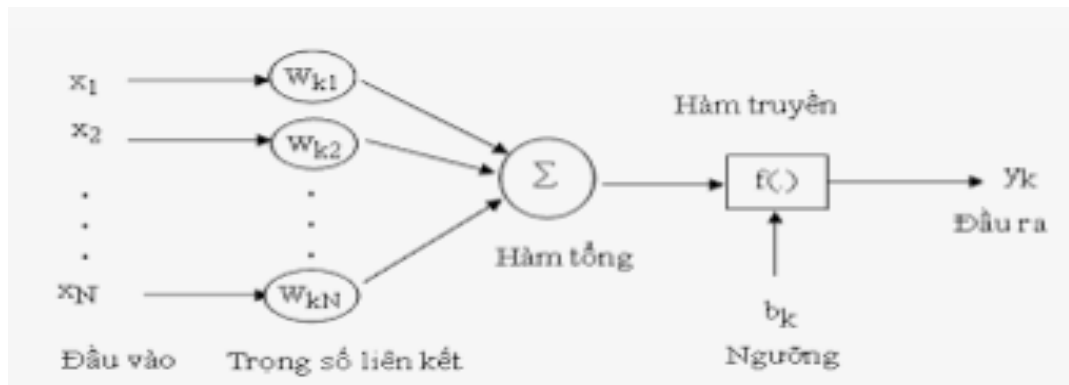
Perceptron là một mạng nơ-ron được cấu tạo dựa trên các nơ-ron đơn lẻ. Mô hình của perceptron cũng giống như một nơ-ron. Chúng đều nhận nhiều đầu vào và cho ra một kết quả duy nhất:



Hình 1.2. Cấu tạo của Perceptrons

Một perceptron sẽ nhận một hoặc nhiều đầu vào dạng nhị phân và cho ra một kết quả dưới dạng nhị phân duy nhất. Dữ liệu đầu vào đều chịu ảnh hưởng bởi các trọng số tương ứng của nó. Kết quả đầu ra quyết định dựa vào một ngưỡng quyết định.

b. Cấu trúc của nơ-ron nhân tạo



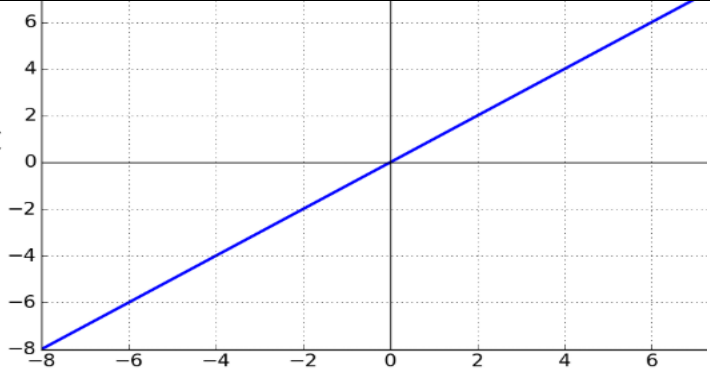
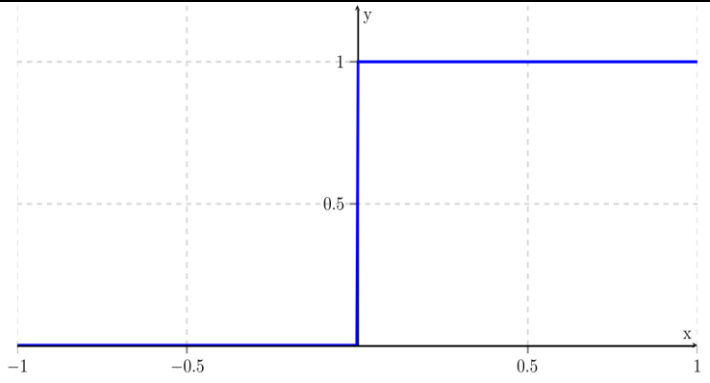
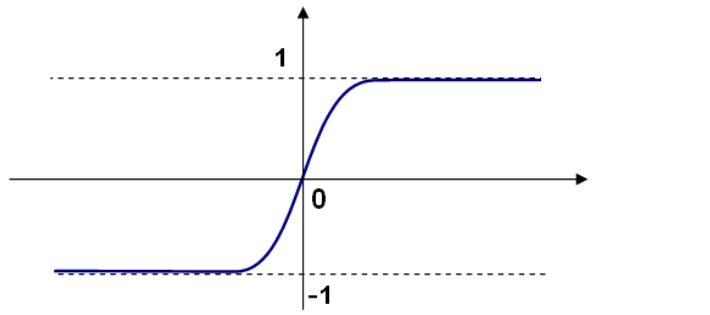
Hình 1. 3. Cấu trúc của nơ-ron nhân tạo

Trong đó, dữ liệu đầu vào là các tín hiệu dưới dạng một vectơ N chiều được truyền vào. Trọng số liên kết là thường được gọi là weight. Chúng được khởi tạo một cách ngẫu nhiên. Hàm tổng là sử dụng để tính tổng các tích của dữ liệu đầu vào với trọng số liên kết của nó. Ngưỡng là được đưa vào để sử dụng trong hàm kích hoạt. Hàm kích hoạt là đầu vào và là kết quả của hàm tổng và ngưỡng. Hàm kích hoạt được sử dụng để giới hạn đầu ra của các nơ-ron. Các hàm kích hoạt là các hàm tuyến tính hay phi tuyến và chúng rất đa dạng. Để chọn lựa được hàm kích hoạt tốt sẽ tùy thuộc vào kinh nghiệm của người thiết kế mạng. Đầu ra là tín hiệu đầu ra duy nhất của mỗi nơ-ron.

c. Các hàm kích hoạt phổ biến

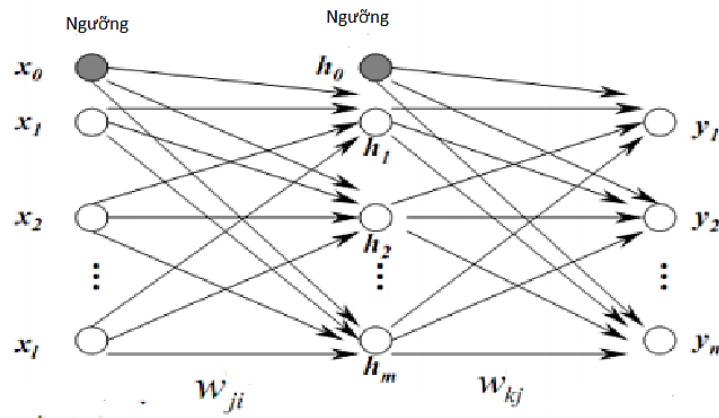
Bảng 1.1. Các hàm kích hoạt

Tên hàm	Mô tả hàm	Đồ thị
Sigmoid	$f = \frac{1}{1 + \exp(-x)}$	

Tên hàm	Mô tả hàm	Đồ thị
Linear	$f = x$	
Hard Limit	$f = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$	
Bipolar sigmoid function	$f = \frac{1 - e^{-x}}{1 + e^{-x}}$	

d. Mạng truyền thẳng

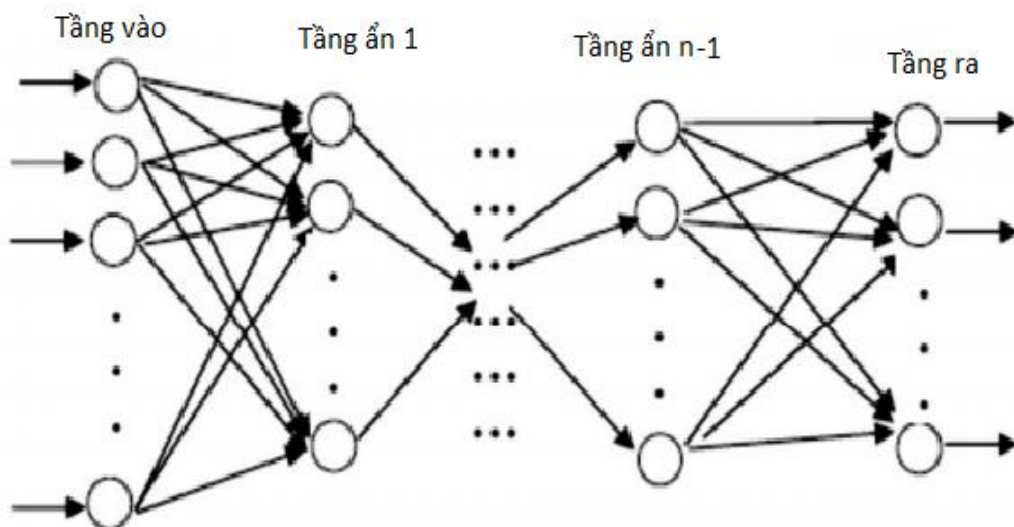
Dữ liệu đầu vào sẽ được truyền đến đơn vị đầu ra bằng cách truyền thẳng. Việc xử lý dữ liệu có thể mở rộng ra nhiều lớp nhưng ở đây chúng không có các liên kết để phản hồi lại.



Hình 1.4. Cấu tạo của mạng truyền thẳng

e. Mạng Multi Layer Perceptron

Mô hình mạng nơ-ron được sử dụng rộng rãi nhất là mô hình mạng nhiều tầng truyền thẳng. Một mạng MLP tổng quát là mạng có n ($n \geq 2$) tầng (thông thường tầng đầu vào không được tính đến): trong đó gồm một tầng đầu ra (tầng thứ n) và $(n-1)$ tầng ẩn.



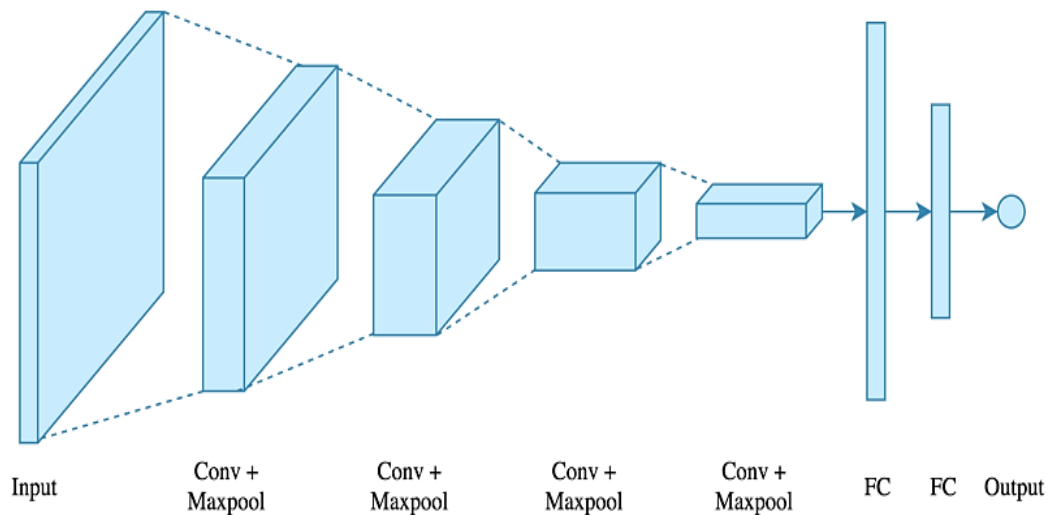
Hình 1.5. Cấu tạo của mạng MLP

1.2.5 Mạng nơ-ron tích chập

a. Khái niệm

Mạng Nơ-ron Tích chập (Convolutional Neural Network - CNN) là một họ các mạng nơ-ron ưu việt. Các kiến trúc dựa trên CNN hiện nay xuất hiện trong mọi góc ngách của lĩnh vực thị giác máy tính, và đã trở thành kiến trúc chủ đạo mà hiếm ai ngày nay phát triển các ứng dụng thương mại hay tham gia một cuộc thi nào đó liên quan tới nhận dạng ảnh, phát hiện đối tượng, hay phân vùng theo ngữ cảnh mà không xây nền móng dựa trên phương pháp này.

b. Kiến trúc mạng CNN



Hình 1.6. Kiến trúc mạng CNN

Kiến trúc mạng CNN bao gồm các thành phần chính là: lớp tích chập, lớp pooling, lớp kết nối đầy đủ, hàm kích hoạt.

1.3 Phương pháp xác định vị trí khuôn mặt với mạng tích chập MTCNN

Trong đó nhiều nghiên cứu chỉ ra MTCNN là một trong những mạng hiệu quả nhất. Dù còn nhiều hạn chế có thể nói đến như tốc độ xử lý, mạng phức tạp nhưng MTCNN ngày càng được sử dụng rộng rãi trong thực tiễn. MTCNN một mạng lớn gồm nhiều mạng con. Mỗi mạng con đều xử lý một bài toán nhỏ trong việc xác định vị trí của khuôn mặt. Trong kiến trúc, MTCNN sử dụng lần lượt ba mạng P-Net, R-Net, O-Net để xác định khuôn mặt.

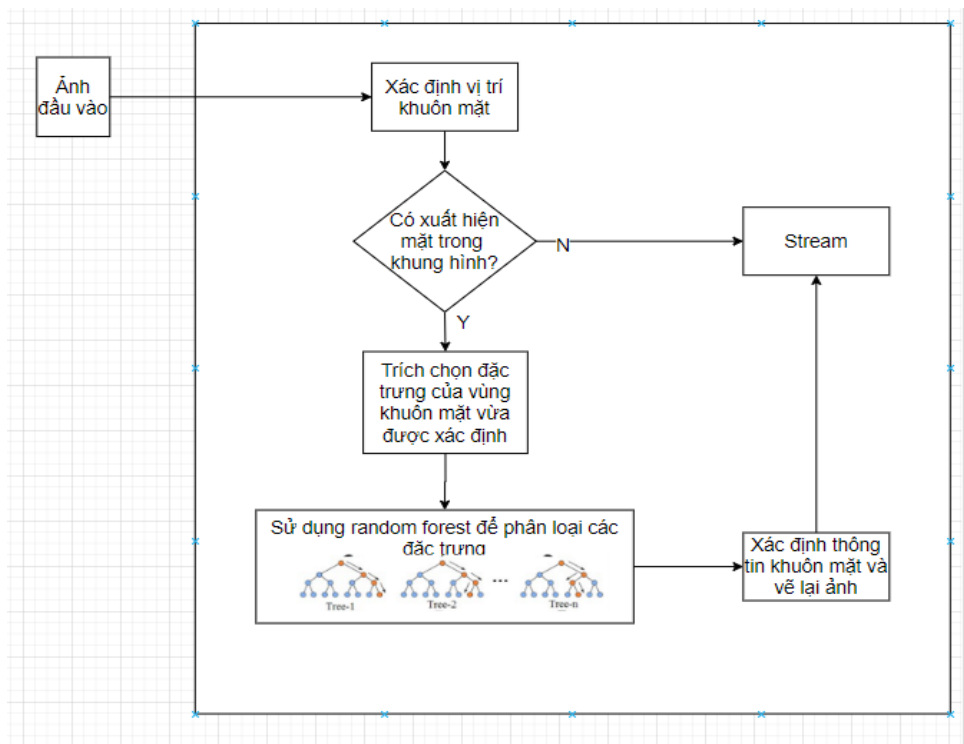
1.4 Kết luận

Chương này đã giới thiệu về tổng quan về nhận diện khuôn mặt, kiến trúc hệ thống nhận diện khuôn mặt cũng như ứng dụng của hệ thống trong thực tế. Bên cạnh đó nội dung của chương cũng trình bày về một số phương pháp hay được sử dụng trong bài toán nhận diện khuôn mặt như PCA, LDA, cây quyết định, mạng nơ ron nhân tạo, nơ ron tích chập. Từ đó có thể đánh giá được ưu nhược điểm của từng phương pháp trong bài toán nhận diện khuôn mặt.

CHƯƠNG 2. HỆ THỐNG NHẬN DIỆN KHUÔN MẶT DỰA TRÊN MẠNG NƠ RON TÍCH CHẬP

2.1 Sơ đồ thiết kế hệ thống nhận diện khuôn mặt

Trong hệ thống có ba bước xử lý chính để có thể nhận diện khuôn mặt. Đó là xác định vị trí khuôn mặt, trích chọn đặc trưng của khuôn mặt và phân loại xác định khuôn mặt.



Hình 2.1. Sơ đồ hoạt động của hệ thống nhận diện khuôn mặt

Trong hệ thống có ba bước xử lý chính để có thể nhận diện khuôn mặt. Đó là xác định vị trí khuôn mặt, trích chọn đặc trưng của khuôn mặt và phân loại xác định khuôn mặt

Về nghiệp vụ xử lý thì sau khi xác định khuôn mặt và định danh thành công khách hàng. Hệ thống sẽ lưu thông tin khuôn mặt và ghi lên hình để phát trực tiếp. Thời gian lưu sẽ là 10s. Sau 10s nếu khách hàng đó tiếp tục xuất hiện trên video thì hệ thống sẽ tiếp tục hiển thị khuôn mặt, thông tin khách hàng đó để lễ tân có thể nắm được.

Các bước để huấn luyện mô hình phân loại:

Bước 1: Chuẩn bị dữ liệu ảnh khuôn mặt để huấn luyện

Bước 2: Xử lý xác định vị trí và trích chọn đặc trưng khuôn mặt từ tập huấn luyện

Bước 3: Huấn luyện rừng ngẫu nhiên để phân loại đặc trưng

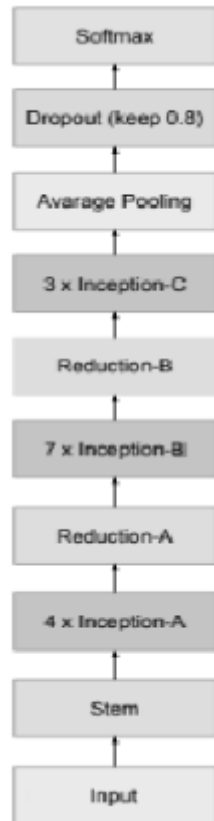
Trong hệ thống có ba bước xử lý chính để có thể nhận diện khuôn mặt. Đó là xác định vị trí khuôn mặt, trích chọn đặc trưng của khuôn mặt và phân loại xác định khuôn mặt

Về nghiệp vụ xử lý thì sau khi xác định khuôn mặt và định danh thành công khách hàng. Hệ thống sẽ lưu thông tin khuôn mặt và ghi lên hình để phát trực tiếp. Thời gian lưu sẽ là 10s. Sau 10s nếu khách hàng đó tiếp tục xuất hiện trên video thì hệ thống sẽ tiếp tục hiển thị khuôn mặt, thông tin khách hàng đó để lễ tân có thể nắm được.

2.2 Mạng Inception-ResNet sử dụng cho việc trích chọn đặc trưng khuôn mặt

Inception-ResNet là một mạng học rất sâu. Cấu tạo của mạng cũng rất đặc biệt. Kiến trúc mạng inception-ResNet sử dụng nhiều khối mạng con như Inception-A, Inception-B, Inception-C, Reduction-A, Reduction-B, Stem. Chính nhờ đặc điểm trên cho một mạng inception phân dư với hiệu năng rất tốt.

Sau đây là kiến trúc mạng chi tiết:



Hình 2.2. Kiến trúc mạng Inception-ResNet

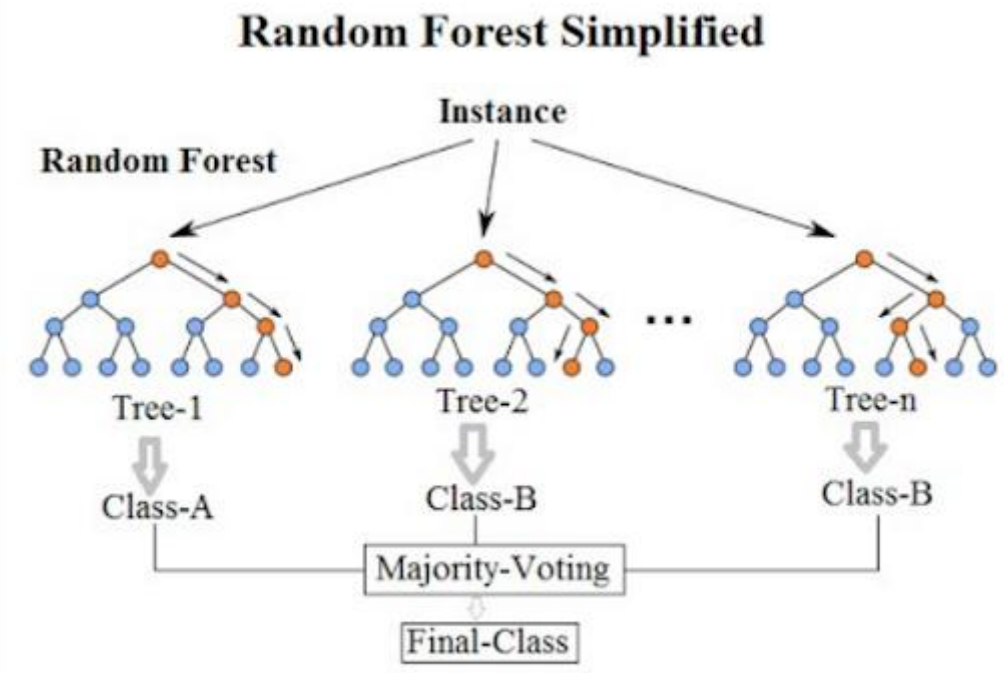
2.3 Rừng ngẫu nhiên

2.3.1 Giới thiệu

Rừng ngẫu nhiên (Random Forest) hay còn được gọi là rừng quyết định ngẫu nhiên là một kỹ thuật học tập được sử dụng để giải quyết các nhiệm vụ học tập có giám sát như phân loại và hồi quy. Một tính năng ưu việt của rừng ngẫu nhiên là nó có thể khắc phục được sự quá khớp (overfitting) trên tập dữ liệu đào tạo của mình.

Kiến trúc mô hình của rừng ngẫu nhiên là một tập hợp của nhiều cây quyết định. Mỗi một cây quyết định sẽ trả ra một kết quả dự báo. Ngoài ra mô hình cũng được chạy trên rất nhiều các sub-sample.

2.3.2 Kiến trúc



Hình 2.3. Kiến trúc của rừng ngẫu nhiên

2.4 Kết luận

Chương này trình bày về sơ đồ thiết kế hệ thống nhận diện khuôn mặt, mạng trích chọn đặc trưng và kỹ thuật phân loại rừng ngẫu nhiên. Trong đó, nội dung chương có đi sâu về mạng học sâu googlenet, resnet từ đó đưa ra ưu nhược điểm khi kết hợp hai mạng thành mạng inception resnet. Chương 3 sẽ trình bày về kết quả thu được khi sử dụng mạng inception resnet kết hợp với rừng ngẫu nhiên cho bài toán nhận diện khuôn mặt.

CHƯƠNG 3. THỬ NGHIỆM VÀ ĐÁNH GIÁ

3.1 Bộ dữ liệu đầu vào

Bộ dữ liệu được sử dụng trong luận văn là hai bộ dữ liệu faces94 [10] và CASIA-WebFace [16]. Trong đó bộ dữ liệu faces94 được thu thập từ trung tâm nghiên cứu Center for Machine Perception thuộc đại học kỹ thuật Séc và bộ dữ liệu CASIA-WebFace được sưu tầm từ viện tự động hóa tại học viện khoa học Trung Quốc(CASIA). Luận văn sẽ sử dụng hai bộ dữ liệu vào các mục đích khác nhau. Bộ dữ liệu faces94 chứa khoảng 1300 ảnh của 153 người. Đây là bộ dữ liệu nhỏ được dùng để thử nghiệm và đánh giá hiệu quả của kiến trúc mạng trước khi sử dụng bộ dữ liệu chính để huấn luyện. Bộ dữ liệu CASIA-WebFace chứa khoảng 500.000 ảnh được thu thập từ 10.000 người. Trước khi đưa vào huấn luyện cho mô hình trích chọn đặc trưng thì những bức ảnh này sẽ được duyệt qua mô hình mtcnn để xác định vị trí tọa độ khuôn mặt của mỗi người trong tập dữ liệu.

3.2 Quá trình huấn luyện

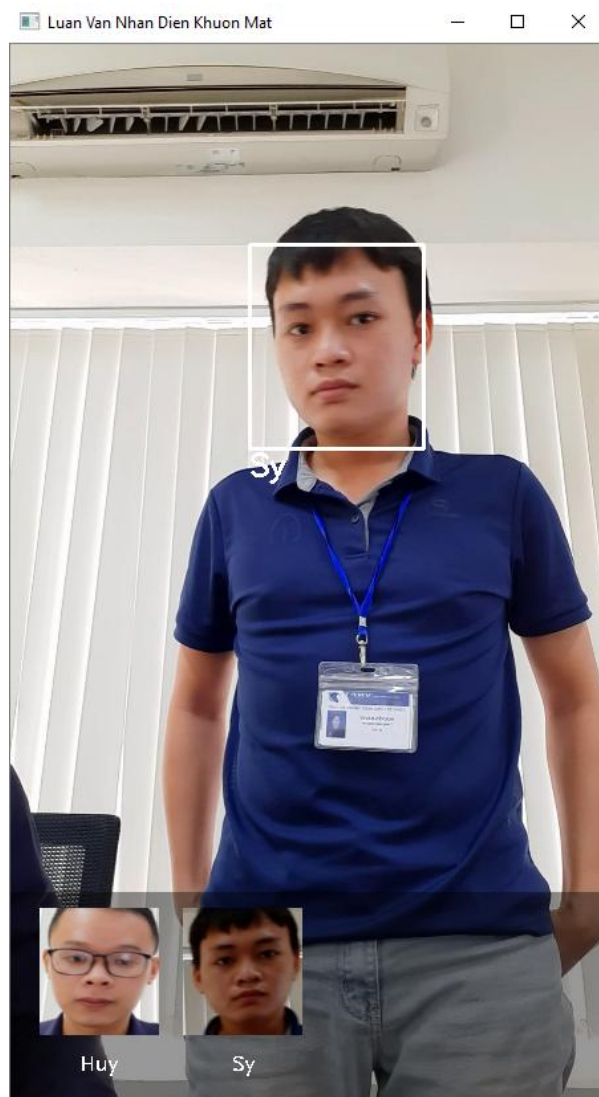
Dữ liệu được sử dụng cho việc huấn luyện này là CASIA-WebFace. Đây là một tập dữ liệu về khuôn mặt rất lớn được thu thập từ học viện khoa học Trung Quốc. Để có một mô hình trích chọn đặc trưng thật tốt thì số lượng dữ liệu về khuôn mặt phải đủ lớn, đủ độ đa dạng về hình dáng, kích thước cũng như màu sắc. Thời gian diễn ra khoảng 350 giờ với 90 lần duyệt. Sau khi đã có được mô hình trích chọn đặc trưng tùy vào từng bài toán ta sẽ tiến hành trích chọn đặc trưng của các nhãn đầu vào để mô hình phân loại bằng thuật toán random forest học các đặc tính đó.

3.3 Thử nghiệm chạy hệ thống nhận diện khuôn mặt nhận diện khách hàng VIP của khách sạn

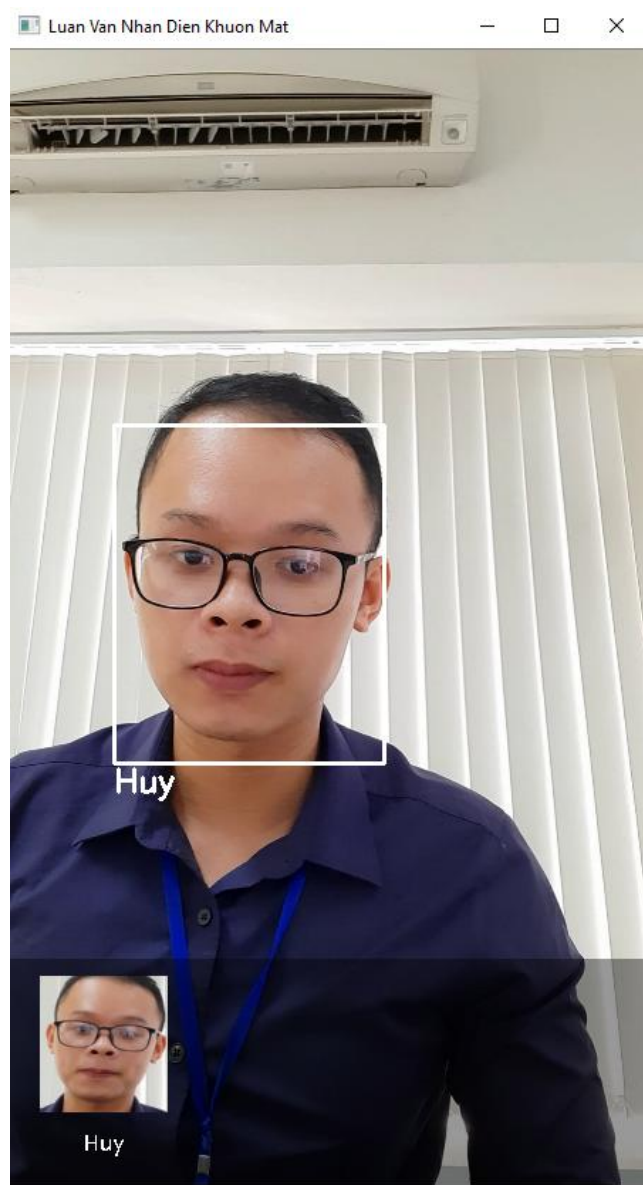
Máy chủ được sử dụng cho việc thực nghiệm sử dụng vi xử lý Intel Core i5-8400 với xung nhịp cơ bản là 2.8GHz. Dung lượng bộ nhớ trong là 16 GB và bộ xử lý đồ họa được sử dụng là GTX 1070 Ti 8G GDDR5. Hệ thống nhận diện khách hàng VIP sẽ được phát triển trên nền tảng Ubuntu và ngôn ngữ lập trình được sử dụng là

python version 3.6 cùng với các bộ thư viện như opencv, keras, tensorflow để phục vụ cho bài toán xử lý ảnh.

Mô hình thử nghiệm được huấn luyện với tập dữ liệu khoảng 200 khách hàng, mỗi vị khách có khoảng 10-30 ảnh khuôn mặt trong cơ sở dữ liệu. Trong quá trình phân loại em đã chọn ra được ngưỡng là 40% để xác định khuôn mặt. Ngưỡng này sử dụng được sử dụng để tránh việc xác suất khuôn mặt trả ra quá thấp gây ra việc nhận nhầm thông tin. Sau khi thiết lập địa chỉ luồng phát video trực tiếp thì ta có thể khởi động hệ thống. Sau đây là một số hình ảnh hệ thống hoạt động :



Hình 3.1. Hệ thống nhận diện khuôn bình thường



Hình 3.2. Hệ thống nhận diện khuôn mặt có đeo kính

Với mỗi khách hàng có trong danh sách sau khi xuất hiện sẽ được lưu lại 10 giây để dễ dàng quan sát. Hệ thống có thể nhận diện khuôn mặt một cách bình thường cả khi người đó đang đeo kính. Từ đó có thể thấy mô hình trích rút đặc trưng và mô hình phân loại đang hoạt động rất tốt trong thực tế. Hệ thống nhận diện khách hàng mỗi giây có thể xử lý 5 khung hình với độ phân giải fullhd.

3.4 Đánh giá

Để đánh phân loại ta sẽ thực hiện lấy vector đặc trưng từ mô hình trích chọn đặc trưng và đưa vào randomforest để huấn luyện. Thiết lập trong random forest của

em hiện tại đang là 100 cây. Các tập dữ liệu được sử dụng để đánh giá trong luận văn này là Faces94 [10], Faces95[11], Faces96[12], Grimace[13]. Tập dữ liệu sẽ được dùng 75% cho việc huấn luyện và 25% còn lại cho việc kiểm thử mô hình phân loại. Sau đó ta sẽ so sánh kết quả của mô hình vừa được huấn luyện với kết quả của một số phương pháp được sử dụng trong nghiên cứu [14] [15] như là LDA, PCA, LBP, SVM based on LDA và MLP kết hợp với PCA và DCT.

Bảng 3.1. Bảng đánh giá độ chính xác giữa các mô hình

	Faces94 dataset	Faces95 dataset	Faces96 dataset	Grimace dataset
Inception Resnet V1+ Random forest	99.1%	99.5%	98.2%	99.95
Multilayer perceptron + PCA + DCT	100%	-	-	100%
SVM based on LDA (RBF kernel)	97.4%	-	95.1%	100 %
LBP	85.93%	80.47%	84.145	86.45%
PCA	72.1%	69.87%	70.95%	74.79%
LDA	79.39%	76.61%	78.34%	81.93%

Sự kết hợp giữa hai mô hình Inception Resnet V1 và Random forest đã cho lại kết quả khá cao với các tập dữ liệu trên. Sự khác biệt lớn nhất giữa các phương pháp là việc trích chọn đặc trưng. Mạng Inception Resnet V1 có thể học được các đặc trưng trên khuôn mặt tốt hơn hẳn so với các phương pháp học máy truyền thống. Kiến trúc mạng tích chập này có khả năng học đặc trưng cực kì tốt vì nó vẫn có thể trích rút các đặc trưng chính của khuôn mặt trong các trường hợp thiếu sáng, khuôn mặt không đầy đủ hay như người đang đeo kính. Inception Resnet V1 kết hợp với Random forest cho ta độ chính xác thấp hơn khoảng 1% so với phương pháp sử dụng multilayer perceptron kết hợp với PCA và DCT nhưng thời gian để huấn luyện và nhận diện

khuôn mặt nhỏ hơn rất nhiều. Từ đó ta thấy phương pháp sử dụng mạng Inception Resnet V1 kết hợp với Random forest dễ dàng triển khai thực tế hơn dù độ chính xác.

3.5 Kết luận

Chương này trình bày về quá trình huấn luyện, kiểm thử, đánh giá chất lượng của mô hình trích chọn đặc trưng inception resnet và thử nghiệm chạy hệ thống nhận diện khách hàng VIP tại khách sạn. Kết quả của hệ thống khá tốt với thời gian xử lý khoảng 0.2 giây một khung hình cùng với khả năng nhận diện được khuôn mặt từ nhiều góc độ, sắc thái, điều kiện khác nhau, điều mà các mô hình học máy truyền thống chưa xử lý tốt.

KẾT LUẬN

Bài toán nhận diện khuôn mặt không còn là một vấn đề mới nhưng nhận diện khuôn mặt dựa trên các mạng học sâu đang rất được quan tâm.

Trên cơ sở tìm hiểu và nghiên cứu các phương pháp nhận diện khuôn mặt áp dụng vào hệ thống nhận diện khuôn mặt qua camera, luận văn đã đạt được những kết quả sau. Đó là tìm hiểu, thực nghiệm các mô hình mạng học sâu, học máy như nơ-ron tích chập và rừng ngẫu nhiên để xử lý bài toán nhận diện. Phân tích kết quả thu được và tìm ra mô hình mạng học sâu thích hợp cho bài toán nhận diện khuôn mặt. Xây dựng thành công hệ thống nhận diện khuôn mặt qua camera có chức năng phát hiện khuôn mặt trực tiếp qua video.

Hệ thống sau khi được phát triển cho thấy rằng việc mô hình mạng trích chọn đặc trưng đang đem tới kết quả rất là tốt tuy nhiên cũng xuất hiện một số mặt hạn chế. Đó là mô hình phân loại này không phù hợp với tập dữ liệu quá lớn. Điều này không phải là do kết quả phân loại kém mà do thời gian huấn luyện lại mô hình phân loại khá lâu.

Qua những kết quả và hạn chế của ứng dụng đã cho thấy việc xây dựng hệ thống nhận diện khuôn mặt còn đòi hỏi phải thực hiện, nghiên cứu và áp dụng thêm nhiều phương pháp mới. Về hướng phát triển tương lai, luận văn sẽ được đề xuất nghiên cứu thêm một số hướng mới. Hướng thứ nhất đó là tiến hành nghiên cứu, áp dụng và đánh giá các mô hình phân loại mới khi áp dụng vào bài toán thay rừng ngẫu nhiên như là Sparse Representation-based classification, Support Vector Machine, Linear Classifier. Hướng tiếp theo là nghiên cứu một số phương pháp sử dụng khoảng cách khi huấn luyện với triplet loss. Phương pháp này sẽ cần dung lượng bộ nhớ lớn nhưng không cần huấn luyện và sử dụng mô hình phân loại.

Do thời gian thực hiện luận văn không nhiều nên chắc chắn luận văn không thể tránh khỏi những hạn chế và thiếu sót. Em rất mong nhận được các ý kiến đóng góp.

DANH MỤC CÁC TÀI LIỆU THAM KHẢO

- [1] Abdulrahman Alkandari, Soha Jaber Aljaber (2015), “*Principle Component Analysis algorithm (PCA) for image recognition*”, ICCTIM, pp. 76-80.
- [2] Liton Chandra Paul, Abdulla Al Sumam (2012), “*Face Recognition Using Principal Component Analysis Method*” International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) 1, pp. 135-139.
- [3] LDA Juwei Lu, Kostantinos N. Plataniotis, and Anastasios N. Venetsanopoulos (2003), "Face Recognition Using LDA-Based Algorithms", IEEE TRANSACTIONS ON NEURAL NETWORKS, pp. 195-200.
- [4] Alaa Eleyan, Hasan Demirel (2006), “*PCA and LDA Based Face Recognition Using Feedforward Neural Network Classifier*” Conference: Multimedia Content Representation, pp. 200-206.
- [5] Philip H. Swain, Hans Hauska (1977), “*The decision tree classifier: Design and potential*” IEEE Transactions on Geoscience Electronics 15, pp. 142-147.
- [6] Georgios Karalis (2020), “*Decision Trees and Applications*” Advances in Experimental Medicine and Biology 1194, pp. 239-242.
- [7] Manish Mishra, Monika Srivastava (2014), “*A view of Artificial Neural Network*”, IEEE ICAETR - 2014, pp. 1-3.
- [8] Saad Albawi, Tareq Abed Mohammed (2017), “*Understanding of a Convolutional Neural Network*”, International Conference on Engineering and Technology (ICET).
- [9] Yushi Chen, Hanlu Jiang, Chunyang Li, Xiuping Jia, Pedram Ghamisi (2016), “*Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks*” IEEE Transactions on Geoscience and Remote Sensing 54, pp. 6232 – 6251.
- [10] Musab Coşkun, Ayşegül Uçar, Özal Yildirim, Yakup Demir (2017), “*Face recognition based on convolutional neural network*” 2017 International Conference on Modern Electrical and Energy Systems (MEES), pp. 376-379.

[11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun (2016), “*Deep Residual Learning for Image Recognition*” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778.

[12] Jehad Ali, Rehanullah Khan, Nasir Ahmad (2012), “Random Forests and Decision Trees” JCSI International Journal of Computer Science Issues 9, pp. 272-276.

[13] Haiyan Guan, Jonathan Li (2012), “*RANDOM FORESTS-BASED FEATURE SELECTION FOR LAND-USE CLASSIFICATION USING LIDAR DATA AND ORTHOIMAGERY*” ISPRS - International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences, pp. 203-208.

[14] Faces94, <http://cmp.felk.cvut.cz/~spacelib/faces/faces94.html>. Truy cập ngày 01/11/2020.

[15] Faces95, <http://cmp.felk.cvut.cz/~spacelib/faces/faces95.html>. Truy cập ngày 01/11/2020.

[16] Faces96, <http://cmp.felk.cvut.cz/~spacelib/faces/faces96.html>. Truy cập ngày 01/11/2020.

[17] Grimace, <http://cmp.felk.cvut.cz/~spacelib/faces/grimace.html>. Truy cập ngày 01/11/2020.

[18] A Vinay, Abhijay Gupta, Aprameya Bharadwaj, Arvind Srinivasan, K N Balasubramanya Murthy, S Natarajan (2018), “*Deep Learning on Binary Patterns for Face Recognition*”, International Conference on Computational Intelligence and Data Science, pp. 77-83.

[19] Nawaf Hazim (2016), “*Face Recognition using PCA-BPNN with DCT Implemented on Face94 and Grimace Databases*”, International Journal of Computer Applications, pp. 8-13.

[20] CASIA-WebFace, <https://pgram.com/dataset/casia-webface/>. Truy cập ngày 01/11/2020.