



TRƯỜNG ĐẠI HỌC  
**SƯ PHẠM KỸ THUẬT TP. HỒ CHÍ MINH**  
HCMC University of Technology and Education

**KHOA CƠ KHÍ CHẾ TẠO MÁY**  
**BỘ MÔN CƠ ĐIỆN TỬ**

□□×□□

**BÁO CÁO CUỐI KÌ**

# **FACE DETECTION**

**GVHD: PGS.TS NGUYỄN TRƯỜNG THỊNH**

**SVTH: NGUYỄN HOÀNG THANH**

**MSSV: 20146133**

**LỚP: Chiều thứ 2**

**Tp.Hồ Chí Minh, Tháng 05 Năm 2023**

## LỜI MỞ ĐẦU

- Lời đầu tiên em xin chân thành cảm ơn đến thầy Nguyễn Trường Thịnh đã giúp em rất nhiều trong quá trình thực hiện đồ án này. Trong quá trình thực hiện đồ án, được sự giúp đỡ tận tình của thầy em đã thu được nhiều kiến thức quý báu giúp em rất nhiều trong quá trình học và làm việc trong tương lai.
- Trong quá trình thực hiện đồ án không tránh khỏi một số sai sót. Mong nhận được sự góp ý của các thầy để hoàn thiện hơn.
- Một lần nữa em xin chân thành cảm ơn sự giúp đỡ của thầy trong quá trình thực hiện đồ án để em hoàn thành đồ án này

**NHẬN XÉT**  
**(Của giảng viên hướng dẫn)**

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

*Chữ ký của giáo viên hướng dẫn*

## MỤC LỤC

Chương 1 GIỚI THIỆU .....	1
1.1 Tổng Quan.....	1
1.1.1 Mở đầu .....	1
1.1.2 Mục tiêu của đồ án .....	1
1.1.3 Đối tượng nghiên cứu.....	1
1.1.4 Phạm vi nghiên cứu.....	2
1.2 Phương pháp – kết quả.....	2
1.2.1 Phương pháp.....	2
1.2.2 Kết quả .....	2
Chương 2 CƠ SỞ LÝ THUYẾT .....	3
2.1 Tổng quan về Deep Learning và bài toán nhận diện khuôn mặt trong thị giác máy tính .....	3
2.1.1 Tổng quan về Deep Learning .....	3
2.2 Giới thiệu Convolutional Neural Network .....	5
Chương 3 XÂY DỰNG ỨNG DỤNG .....	7
3.1. Môi trường thực hiện.....	7
3.2 Các bước thực hiện .....	7
3.2.1 Chuẩn bị dữ liệu .....	7
3.2.3 Cài đặt một số thư viện khác. ....	7
3.2.4 Phân chia bộ train và bộ test.....	8
3.2.5 Model summary .....	8
3.2.6 Biểu đồ train .....	9
3.2.7 Test kết quả.....	9
Chương 4 KẾT LUẬN .....	10
4.1 Kết quả.....	10
4.2 Ưu điểm - nhược điểm.....	10
4.3 Hướng phát triển.....	11

# Chương 1 GIỚI THIỆU

## 1.1 Tổng Quan

### 1.1.1 Mở đầu

Hiện nay, cùng với sự phát triển của xã hội, vấn đề an ninh bảo mật đang được yêu cầu khắt khe tại mọi quốc gia trên thế giới. Các hệ thống nhận dạng con người được ra đời với độ tin cậy ngày càng cao. Một trong các bài toán nhận dạng con người rất được quan tâm hiện nay là nhận dạng khuôn mặt. Vì nhận dạng khuôn mặt là cách mà con người sử dụng để phân biệt nhau. Bên cạnh đó, ngày nay việc thu thập, xử lý thông tin qua ảnh để nhận biết đối tượng đang được quan tâm và ứng dụng rộng rãi. Với phương pháp này, chúng ta có thể thu nhận được nhiều thông tin từ đối tượng mà không cần tác động nhiều đến đối tượng nghiên cứu. Sự phát triển của khoa học máy tính tạo môi trường thuận lợi cho bài toán nhận dạng mặt người từ ảnh số. Các hệ thống nhận dạng offline đã ra đời và có độ tin cậy cao, tuy nhiên các hệ thống nhận dạng online lại chưa đáp ứng được nhiều.

Trong khuôn khổ đề án này, chúng em sẽ tiếp tục giải quyết bài toán nhận dạng offline. Trong đó đối tượng được thu thập thành các file dữ liệu và được chuyển về trung tâm. Tại đó, các số liệu sẽ được phân tích xử lý. Trong phần đầu, chúng em sẽ giải quyết bài toán nhận dạng thông thường, phần hai sẽ là phần nhận dạng giới tính.

### 1.1.2 Mục tiêu của đề án

Sử dụng phương thức Convolutional Neural Network để xác định khuôn mặt từ một bức ảnh đầu vào ở camera cho trước sau đó sử dụng model đã được train bên collab và từ đó xác định giới tính của người đó.

- Input: Ảnh chụp mặt người.
- Output: Giới tính của người đó.

### 1.1.3 Đối tượng nghiên cứu

- Ngôn ngữ Python
- Tìm hiểu về Machine Learning và Deep Learning

- Các thuật toán nhận diện

#### **1.1.4 Phạm vi nghiên cứu**

- Xây dựng chương trình nhận dạng giới tính con người sử dụng Machine Learning.
- Ứng dụng đề tài phục vụ cho việc nghiên cứu về nhiều lĩnh vực.

### **1.2 Phương pháp – kết quả**

#### **1.2.1 Phương pháp**

##### ***\*Phương pháp chủ đạo***

- Phương pháp nghiên cứu lý thuyết:
  - + Tìm hiểu tổng quan về Machine Learning và bài toán nhận diện khuôn mặt trong thị giác máy tính.
- + Tìm hiểu phương pháp CNN.
  - + Tìm hiểu thư viện Keras và một số thư viện khác.
- Phương pháp nghiên cứu thực nghiệm:
  - + Tiến hành phân tích và cài đặt trên Python
  - + Tiến hành training data cho máy học.

#### **1.2.2 Kết quả**

- Tạo ra chương trình nhận dạng giới tính thông qua ảnh có sẵn thực trên hệ điều hành Window.

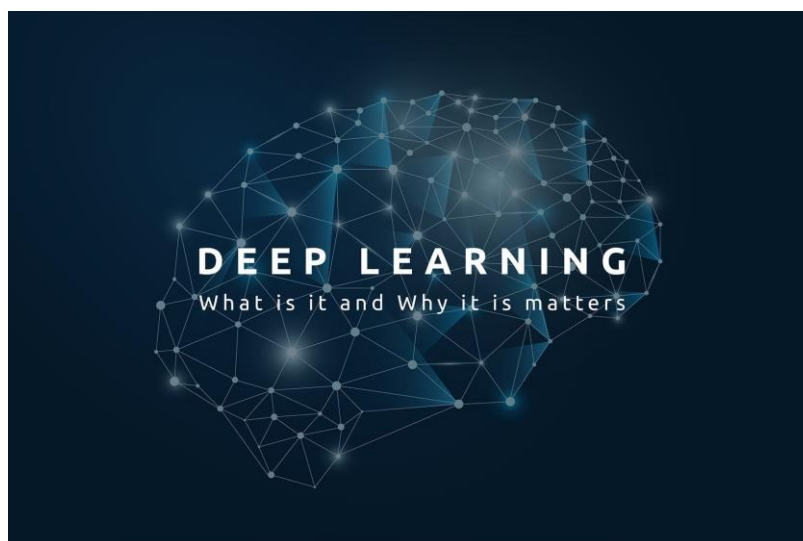
## Chương 2 CƠ SỞ LÝ THUYẾT

### 2.1 Tổng quan về Deep Learning và bài toán nhận diện khuôn mặt trong thị giác máy tính

#### 2.1.1 Tổng quan về Deep Learning

Deep Learning là một tập hợp con của Trí tuệ nhân tạo – một kỹ thuật học máy dạy cho máy tính và các thiết bị hoạt động một cách logic. Tại sao lại đặt tên là Deep Learning? Bởi thực tế là nó liên quan đến việc đi sâu vào một số lớp của mạng, bao gồm cả một lớp ẩn. Bạn càng học sâu, bạn càng trích xuất ra những thông tin phức tạp.

Phương pháp Deep learning dựa vào các chương trình phức tạp khác nhau để bắt chước trí thông minh của con người. Phương pháp đặc biệt này dạy cho máy móc nhận biết các họa tiết để có thể phân loại chúng thành các loại khác nhau. Nhận dạng mẫu là một phần thiết yếu của Deep learning và nhờ Machine learning, máy tính thậm chí không cần phụ thuộc vào lập trình mở rộng. Thông qua Deep learning, máy móc có thể sử dụng các tệp hình ảnh, văn bản hoặc âm thanh để xác định và thực hiện bất kỳ tác vụ nào theo cách giống như con người.



Nhận dạng khuôn mặt (Face Recognition) là một phương pháp sinh trắc để xác định hoặc xác minh một cá nhân nào đó bằng cách so sánh dữ liệu hình ảnh chụp trực tiếp hoặc hình ảnh kỹ thuật số với bản ghi được lưu trữ cho người đó. Nó được xem là một lĩnh

vực nghiên cứu của ngành Biometrics (tương tự như nhận dạng vân tay – Fingerprint Recognition, hay nhận dạng mống mắt – Iris Recognition).

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

Các hệ thống nhận dạng khuôn mặt thường được sử dụng cho các mục đích an ninh như kiểm soát an ninh tại tòa nhà, sân bay, máy ATM, tra cứu thông tin của tội phạm, phát hiện tội phạm ở nơi công cộng, ...và ngày càng được ứng dụng rộng rãi trong cuộc sống.

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

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



Phát hiện khuôn mặt (Face Detection): Chức năng chính của bước này là phát hiện ra khuôn mặt xem nó có xuất hiện ở trong một bức hình hay một đoạn video hay không?



Tỉ lệ phát hiện ra khuôn mặt phụ thuộc nhiều vào điều kiện về độ sáng, hướng khuôn mặt, biểu hiện cảm xúc trên khuôn mặt hay các yếu tố môi trường khác. Để hệ thống nhận dạng hoạt động đạt hiệu quả cao thì hình ảnh khuôn mặt sau khi được phát hiện cần chuẩn hóa về kích thước, ánh sáng.

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

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

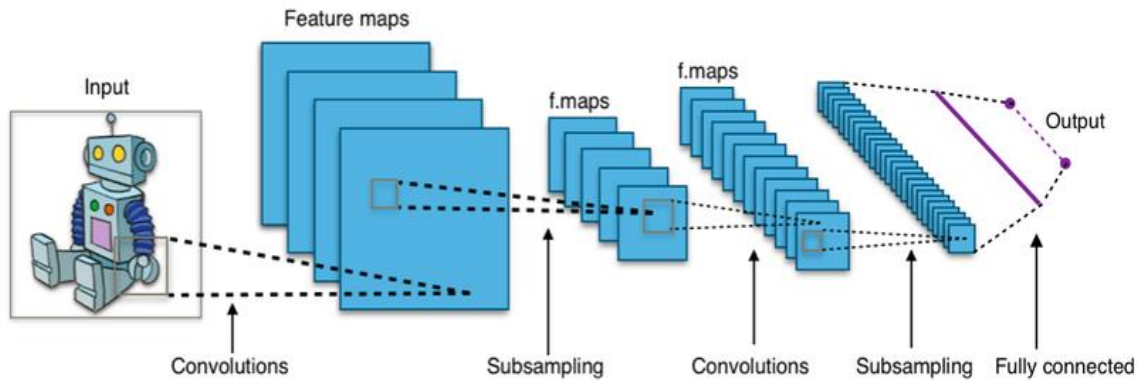
Bài toán nhận dạng khuôn mặt được ứng dụng nhiều trong các lĩnh vực đời sống đặc biệt ở những lĩnh vực công nghệ cao, yêu cầu về an ninh, bảo mật. Do đó để hệ thống nhận dạng khuôn mặt hoạt động mạnh mẽ với tốc độ và độ tin cậy thì có rất nhiều các phương pháp về nhận dạng khuôn mặt được đưa ra. Các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

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

## **2.2 Giới thiệu Convolutional Neural Network**

Convolutional Neural Network (CNN) bao gồm một hoặc nhiều lớp chập (thường với một bước lấy mẫu con) và sau đó theo sau bởi một hoặc nhiều hơn các lớp kết nối như trong một mạng nơ-ron đa lớp chuẩn. Kiến trúc của một CNN được thiết kế để tận dụng lợi thế của cấu trúc 2 chiều của một hình ảnh đầu vào (hoặc đầu vào 2 chiều khác như một tín hiệu tiếng nói). Điều này đạt được với các kết nối cục bộ và trọng số ràng buộc theo một số hình thức tổng hợp mà kết quả là các đặc trưng không thay đổi. Một lợi ích khác

của CNN là dễ dàng huấn luyện hơn và có ít thông số so với các mạng kết nối đầy đủ với cùng một số đơn vị ẩn.



Chập hình ảnh đầu vào với các bộ lọc huấn luyện khác nhau và các bias bổ sung, nhiều bản đồ đặc trưng được tạo trong lớp C1. Mỗi bản đồ đặc trưng trong S2 thu được bởi thao tác tổng hợp các bản đồ đặc trưng tương ứng trong lớp C1. Chập và tổng hợp cực đại trong lớp C3 và S4 thì giống trong lớp C1 và S2. Trong bước nhận diện cuối cùng, các đặc trưng thu được sau khi tổng hợp cực đại trong lớp S4 thì được mã hóa thành một vector 1 chiều.

## Chương 3 XÂY DỰNG ỨNG DỤNG

### 3.1. Môi trường thực hiện

- Ngôn ngữ Python
- Hệ điều hành: Window 10

### 3.2 Các bước thực hiện

#### 3.2.1 Chuẩn bị dữ liệu

Dữ liệu khoảng 2200 bức ảnh gồm nam và nữ.

Phân chia bộ dữ liệu: Trainingset: 80% và Testset: 20%

#### 3.2.3 Cài đặt một số thư viện khác.

```
from keras.preprocessing.image import ImageDataGenerator
from keras.optimizers import Adam
from keras.utils import img_to_array, load_img
from keras.utils import to_categorical, plot_model
from keras.models import Sequential, Model
from keras.layers import BatchNormalization, Conv2D, MaxPooling2D, LeakyReLU
from keras.layers import Activation, Flatten, Dropout, Dense
from keras.callbacks import EarlyStopping, ModelCheckpoint
from keras.utils import np_utils

from sklearn.model_selection import train_test_split
from sklearn.utils import validation

import matplotlib.pyplot as plt
import numpy as np
from os import listdir
from numpy import asarray, save
```

### 3.2.4 Phân chia bộ train và bộ test

```
folder = 'drive/MyDrive/face/'
datas, labels = list(), list()
for file in listdir(folder):
    if file.startswith('Man_'):
        output = 0
    if file.startswith('Woman_'):
        output = 1
    data = load_img(folder+file, target_size=(96,96))
    data = img_to_array(data)
    datas.append(data)
    labels.append(output)

data_train = asarray(datas)
label_train = asarray(labels)
print(data_train.shape, label_train.shape)
```

(2307, 96, 96, 3) (2307,)

```
folder = 'drive/MyDrive/test/'
datas, labels = list(), list()
for file in listdir(folder):
    if file.startswith('Man_'):
        output = 0
    if file.startswith('Woman_'):
        output = 1
    data = load_img(folder+file, target_size=(96,96))
    data = img_to_array(data)
    datas.append(data)
    labels.append(output)

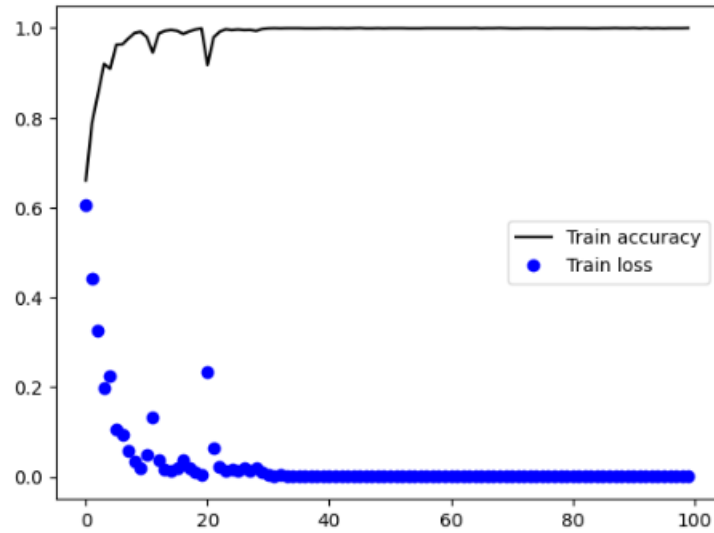
data_test = asarray(datas)
label_test = asarray(labels)
print(data_test.shape, label_test.shape)
```

(500, 96, 96, 3) (500,)

### 3.2.5 Model summary

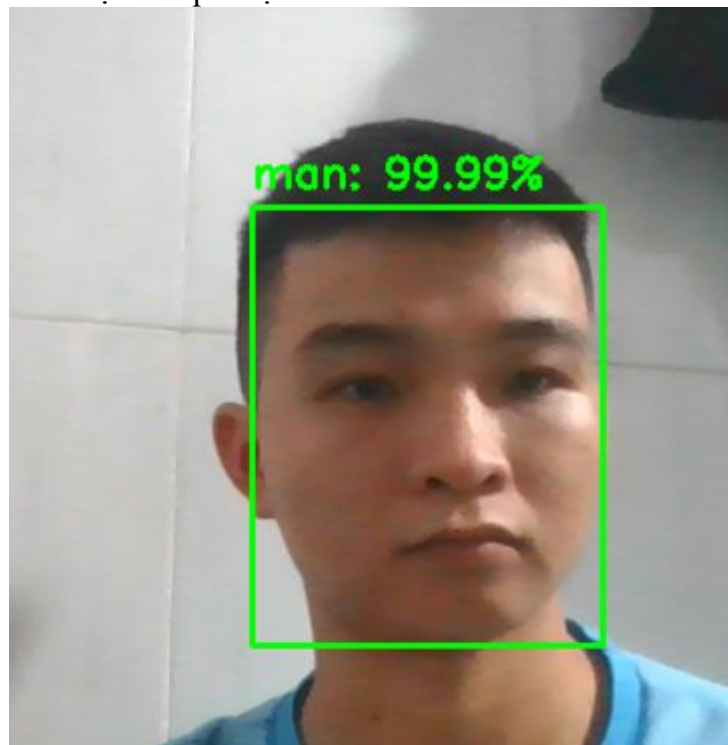
Layer (type)	Output Shape	Param #
=====		
conv2d (Conv2D)	(None, 96, 96, 32)	896
leaky_re_lu (LeakyReLU)	(None, 96, 96, 32)	0
max_pooling2d (MaxPooling2D)	(None, 32, 32, 32)	0
conv2d_1 (Conv2D)	(None, 32, 32, 64)	18496
leaky_re_lu_1 (LeakyReLU)	(None, 32, 32, 64)	0
max_pooling2d_1 (MaxPooling2D)	(None, 11, 11, 64)	0
conv2d_2 (Conv2D)	(None, 11, 11, 64)	36928
leaky_re_lu_2 (LeakyReLU)	(None, 11, 11, 64)	0
max_pooling2d_2 (MaxPooling2D)	(None, 6, 6, 64)	0
conv2d_3 (Conv2D)	(None, 6, 6, 128)	73856
leaky_re_lu_3 (LeakyReLU)	(None, 6, 6, 128)	0
max_pooling2d_3 (MaxPooling2D)	(None, 3, 3, 128)	0
conv2d_4 (Conv2D)	(None, 3, 3, 128)	147584
leaky_re_lu_4 (LeakyReLU)	(None, 3, 3, 128)	0
max_pooling2d_4 (MaxPooling2D)	(None, 2, 2, 128)	0
flatten (Flatten)	(None, 512)	0
dense (Dense)	(None, 1024)	525312
activation (Activation)	(None, 1024)	0
dropout (Dropout)	(None, 1024)	0
dense_1 (Dense)	(None, 2)	2050
activation_1 (Activation)	(None, 2)	0
=====		
Total params: 805,122		
Trainable params: 805,122		
Non-trainable params: 0		

### 3.2.6 Biểu đồ train



### 3.2.7 Test kết quả

Dưới đây là kết một vài quả dự đoán chính xác của model.



## **Chương 4 KẾT LUẬN**

### **4.1 Kết quả.**

Với kết quả thu được là chương trình nhận giới tính là nam hay nữ khi được cung cấp ảnh, có thể được sử dụng để ứng dụng vào nhiều vấn đề thực tế như nhận diện khuôn mặt ở các cửa hàng tiện lợi, mở khóa điện thoại, nhận diện học sinh, quản lý nhân sự, xu hướng khách hàng...

### **4.2 Ưu điểm - nhược điểm**

Ưu điểm:

- Chương trình nhận diện khá chính xác khi được cung cấp ảnh chất lượng tốt như rõ nét, chân dung,...
- Chương trình có thể nhận diện được nam giả nữ hoặc ngược lại khi chất lượng ảnh tốt.
- Nhận diện được đa số mọi lứa tuổi.
- Nhanh chóng, thuận tiện, chỉ cần ảnh có sẵn.

Nhược điểm:

- Vẫn còn có trường hợp dự đoán giới tính sai như khi chất lượng ảnh thấp ảnh bị mờ nhiều chỗ khuất.
- Chương trình có khả năng dự đoán sai ở 1 vài trường hợp như: độ tuổi của người được cung cấp quá nhỏ để phân biệt, người trong ảnh trang điểm hoặc ăn mặc quá đậm quá khác dẫn đến kết quả dự đoán sẽ bị sai.
- Tuy nhiên sự sai sót này có thể bỏ qua vì bài toán nhận dạng giới tính thường dùng để áp dụng cho các vấn đề về quản lý nhân sự, xu hướng khách hàng,... và độ tuổi trẻ em thì thực sự chưa có tác động nhiều trên những vấn đề này.
- Dữ liệu nhận diện còn ít, thiếu sự đa dạng.

### **4.3 Hướng phát triển.**

Có thể phát triển chương trình theo hướng nhận dạng khuôn mặt, xây dựng một hệ thống để học các đặc trưng của những người cần nhận dạng. Khi thực hiện, đầu tiên ta đưa qua bức ảnh qua chương trình phát hiện mặt người để phát hiện nhanh các khuôn mặt có trong ảnh, sau đấy so sánh các khuôn mặt đó với các khuôn mặt mà chương trình đã được “học” từ trước, so sánh các đặc trưng của hai khuôn mặt, nếu trùng thì đưa ra thông tin về khuôn mặt được nhận dạng.