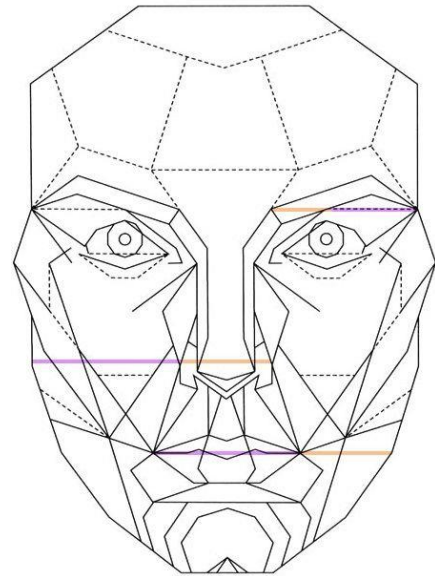


NHẬN DIỆN KHUÔN MẶT SỬ DỤNG CONVOLUTIONAL NEURAL NETWORK (CNN)

Trên thế giới, đa số mọi người có một khuôn mặt hoàn toàn khác nhau. Tuy nhiên để có thể nhận dạng nam hay nữ là một vấn đề. Từ sự bất cập trên, việc áp dụng AI để nhận dạng khuôn mặt là hữu ích. Phương pháp tối ưu nhất đó chính là nhận diện khuôn mặt bằng CNN.

1. Giới thiệu

Nhận diện khuôn mặt được ứng dụng trong nhiều lĩnh vực khác nhau. Cụ thể trong nội dung này chỉ nói về nhận dạng các đặc điểm như là: độ cao của trán, độ to dài của mắt, tỉ lệ mũi, độ rộng của miệng, gò má,... Từ đó, xác định giới tính của mỗi người thông qua camera. Nguồn dữ liệu được thu thập thủ công bằng cách chụp khuôn mặt của mọi người trên thế giới.



2. Phương pháp

Phương pháp được sử dụng là CNN:

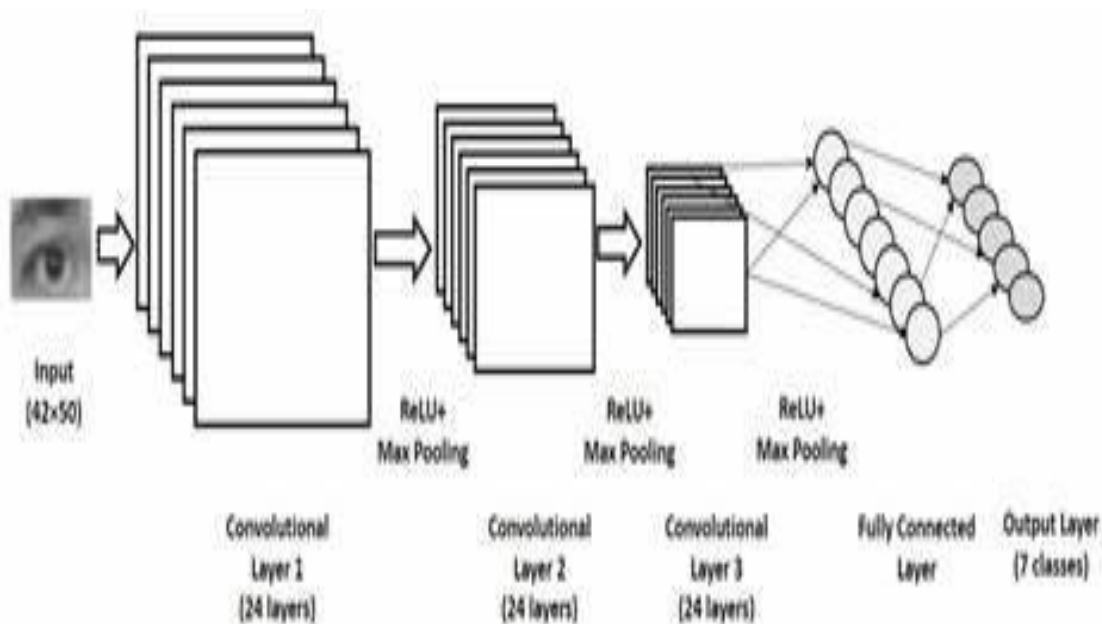
Dựa trên kích thước các bộ phận trên khuôn mặt bao gồm: tỉ lệ khuôn mặt, chiều cao trán, chiều ngang và dài của mũi, độ to của mắt, kích thước của mặt và một số đặc điểm khác như gò má, cằm, tai,...

Convolutional Neural Network là một thuật ngữ với nghĩa tiếng Việt là mạng nơ-ron tích chập, nó thường được viết tắt thành CNN và là một trong những mô hình Deep Learning tiên tiến, hiện đại nhất hiện nay.

Nhờ có CNN mà chúng ta có thể dễ dàng hơn trong việc tạo dựng hệ thống mạng thông minh và có độ chính xác cao. Convolutional là một loại cửa sổ dạng trượt được đặt trên một ma trận.

Những cửa sổ dạng trượt này sẽ có các Parameter để điều chỉnh chúng và nhờ đó có thể lấy được những thông tin một cách chính xác nhất mà không cần phải tiến hành chọn Feature. Convolutional là một trong những phần tử quan trọng trong ma trận. Cửa sổ trượt còn được gọi là Filter, Kernel hoặc Feature Detect hoặc có thể gọi nó là một loại ma trận có kích thước nhỏ.

Từ lớp đầu vào, ảnh sẽ được nhân tích chập với một ma trận để làm nổi bức hình hay còn gọi là phân đoạn ảnh. ảnh đầu ra của lớp đầu sẽ có kích thước nhỏ hơn so với ban đầu. Trải qua lần lượt các lớp chúng ta sẽ có một ma trận ảnh được nổi lên hẳn so với nền. Vì kích thước của ảnh qua nhiều lần nhân tích chập với ma trận sẽ nhỏ dần đi nên thuật toán cần bù thêm vào chỗ bị mất đi các nền. Ở lớp cuối cùng sau khi đã nhân tích chập, ảnh sẽ chuyển sang dạng mảng và đưa vào mạng ANN với các liên kết đầy đủ giữa các lớp (full connected layers).



3. Thực thi

Dựa vào bảng tóm tắt, chương trình tạo ra gồm 2 mạng:

- CNN, có 8 lớp tạo ra:

+ Layer 1 thực thi 32 lần tích chập với đầu vào (input_shape) là 96x96 pixel 3 kênh màu

+ Layer 2 thực thi 64 lần tích chập

+ Layer 3 thực thi 64 lần tích chập

+ Layer 4 thực thi 128 lần tích chập

+ Layer 5 thực thi 128 lần tích chập

+ Layer 6 thực thi 64 lần tích chập

- ANN, được tạo ra 2 lớp:

+ 1024 giá trị đầu vào

+ 2 giá trị đầu ra tương ứng với man và woman

Model: "sequential"

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		

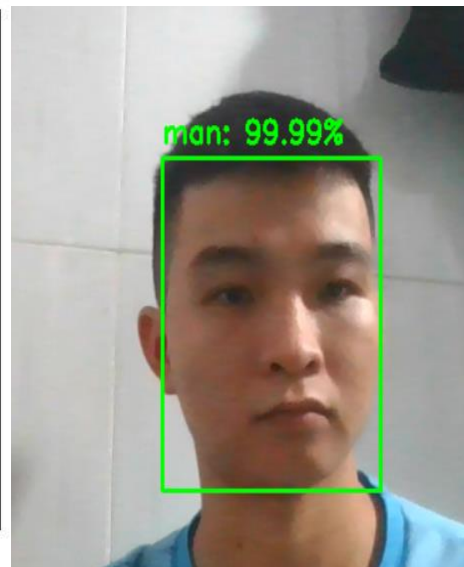
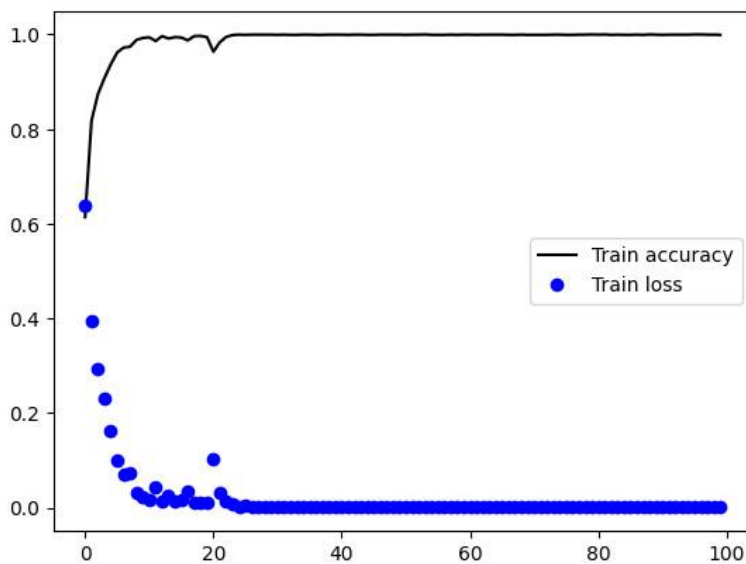
Chương trình cụ thể (mã QR code):



4. Kết quả

Dữ liệu sau khi train sẽ được lưu dưới dạng face_detetion.h5, việc train dữ liệu lớn tốn rất nhiều thời gian. Vì thế điều này sẽ đem lại sự hiệu quả cho những lần nhận diện khuôn mặt tiếp theo mà không cần train lại dữ liệu.

Dưới đây là đồ thị biểu diễn độ chính xác trong lúc train và độ mất mát.



5. Thảo luận

Nhận diện khuôn mặt là chủ đề được thực hiện ở khắp nơi, việc nhận diện đem lại nhiều lợi ích. Phương pháp CNN thường được sử dụng trong mục đích này. Tuy vẫn còn sai số trong nhận diện khuôn mặt nhưng nó đem lại độ chính xác đáng kể với đầu vào là ảnh màu 3 kênh.

Việc nhận diện khuôn mặt có thể được phát triển trong các lĩnh vực khác nhau. Trong đề tài này chỉ đề cập đến việc nhận diện giới tính. Nguồn dữ liệu được thu thập thủ công. Để phát triển thêm có thể cần nhiều dữ liệu đầu vào và các phương pháp kết hợp với nhau.