ĐẠI HỌC SƯ PHẠM THÀNH PHỐ HỒ CHÍ MINH

**KHOA CÔNG NGHỆ THÔNG TIN**

**🙡★🙣**

****

**BÁO CÁO ĐỀ TÀI**

**NGHIÊN CỨU KHOA HỌC**

**XÂY DỰNG HỆ THỐNG   
NHẬN DẠNG ĐỐI TƯỢNG   
SỬ DỤNG KHẨU TRANG**

Giảng viên hướng dẫn : ThS. LƯƠNG TRẦN NGỌC KHIẾT

Nhóm thực hiện :

+ LÊ TẤN LỘC 45.01.104.135

+ HUỲNH THANH PHONG 45.01.104.172

+ TRẦN THỊ TỨ LINH 45.01.104.127

+ PHẠM DUY MINH 45.01.104.145

**TP HỒ CHÍ MINH – 202**

Mục lục

[**LỜI CẢM ƠN** 1](#_Toc67837032)

[**CHƯƠNG 1. TỔNG QUAN** 2](#_Toc67837033)

[**1.1. Giới thiệu bài toán** 2](#_Toc67837034)

[**1.2. Xử lý hình ảnh** 3](#_Toc67837035)

[**1.3. Nhận dạng đối tượng đeo khẩu trang.** 4](#_Toc67837036)

[**1.4. Phạm vi đề tài** 5](#_Toc67837037)

[**CHƯƠNG 2: CƠ SỞ LÝ THUYẾT** 6](#_Toc67837038)

[**2.1. Nhận diện khuôn mặt** 6](#_Toc67837039)

[**2.2. Nhận diện khẩu trang trên khuôn mặt** 12](#_Toc67837040)

[**2.3. Tình hình nghiên cứu ngoài nước về bài toán Face Detection** 13](#_Toc67837041)

[**CHƯƠNG 3: THỰC NGHIỆM CHƯƠNG TRÌNH** 15](#_Toc67837042)

[**3.1. Môi trường** 15](#_Toc67837043)

[3.1.1 Ngôn ngữ Python 15](#_Toc67837044)

[3.1.2 Thư viện Tensorflow 16](#_Toc67837045)

[3.1.3 Thư viện Numpy 16](#_Toc67837046)

[3.1.4 Thư viện Matplotlib 17](#_Toc67837047)

[3.1.5 Thư viện Keras 17](#_Toc67837048)

[3.1.6 Thư viện OpenCV-Python 18](#_Toc67837049)

[3.1.7 Thư viện Imutils 18](#_Toc67837050)

[3.1.8 Thư viện SciPy 18](#_Toc67837051)

[**3.2. Dữ liệu** 19](#_Toc67837052)

[3.2.1 Yêu cầu dữ liệu 19](#_Toc67837053)

[3.2.2 Phân tích dữ liệu 19](#_Toc67837054)

[**3.3. Lấy dữ liệu không gian thực** 21](#_Toc67837055)

[**CHƯƠNG 4: KẾT LUẬN** 22](#_Toc67837056)

[**4.1. Kết luận** 22](#_Toc67837057)

[**4.2. Hướng phát triển (bổ sung)** 22](#_Toc67837058)

[**TÀI LIỆU THAM KHẢO** 24](#_Toc67837059)

DANH MỤC HÌNH ẢNH

Hình 1: Thống kê các công trình nghiên cứu liên quan đến cụm từ “named entity recognition” trên Springer từ 2002 – tháng 11/2017 2

Hình 2: Ví dụ về cây cú pháp 9

Hình 3: Minh họa về khái niệm trích xuất thông tin (Information Extraction) 11

Hình 4: Minh họa về hoạt động của NER (Named Entity Recoginition) 12

Hình 5: IEEE Spectrum The Top Programming languages 2018 14

Hình 6: Trình tải xuống NLTK để chọn gói nào cần được cài đặt 16

Hình 7: Giao diện trang thu thập các văn bản pháp quy 19

Hình 8: Ví dụ về định dạng của một văn bản pháp quy (.htm) 20

Hình 9: : Thêm cấu trúc link download vào Regex trong Settings   
của Link Klipper 21

Hình 10: Giao diện khi mở file step01\_download\_textfile.py bằng Spyder   
(Python 3.7) 22

Hình 11: File "step01\_download\_textfile.py" đang được thực thi 23

Hình 12: Dữ liệu được tải về thành công được lưu trong folder “download” 24

Hình 13: Giao diện khi mở file "step02\_entity\_nlp.py" bằng Spyder (Python 3.7) 25

Hình 14: Biểu đồ cột thống kê các từ có số lần xuất hiện nhiều nhất  
 trong văn bản pháp quy tháng 1 – năm 2019 26

Hình 15: Biểu đồ tròn phân tích nội dung của văn bản pháp quy   
tháng 1 - năm 2019 26

Hình 16: Biểu đồ cột thống kê các từ có số lần xuất hiện nhiều nhất   
trong văn bản pháp quy tháng 2 – năm 2019 27

Hình 17: Biểu đồ cột phân tích nội dung trong các văn bản pháp quy   
tháng 4 - 2019 27

# **LỜI CẢM ƠN**

Nghiên cứu được thực hiện tại Khoa Công nghệ thông tin – Trường Đại học Sư phạm Thành phố Hồ Chí Minh, dưới sự hướng dẫn khoa học của ThS Lương Trần Ngọc Khiết.

Trước tiên chúng em xin bày tỏ lòng biết ơn sâu sắc tới thầy ThS Lương Trần Ngọc Khiết đã đưa chúng em đến với lĩnh vực nghiên cứu này. Thầy đã tận tình giảng dạy, hướng dẫn chúng em tiếp cận và đạt được những kết quả nhất định trong nghiên cứu của mình. Thầy đã luôn tận tâm động viên, khuyến khích và chỉ dẫn giúp chúng em hoàn thành nghiên cứu này.

Chúng em xin bày tỏ lòng biết ơn tới các Thầy Cô thuộc Khoa Công nghệ thông tin và cán bộ Phòng Khoa học Công nghệ, khoa Công nghệ Thông tin – Trường Đại học Sư Phạm Thành phố Hồ Chí Minh đã tạo mọi điều kiện thuận lợi giúp đỡ chúng em trong quá trình học tập và nghiên cứu.

Sự hướng dẫn của thầy Th. s Lương Trần Ngọc Khiết đã tận tình hướng dẫn, động viên, cổ vũ của gia đình, bạn bè là nguồn động lực quan trọng để chúng em thực hiện đề tài nghiên cứu. Do kiến thức còn hạn chế, nên đề tài nghiên cứu của chúng em không tránh khỏi những thiếu sót, kính mong sự thông cảm, chỉ bảo của quý Thầy Cô.

Chúng em xin chân thành cảm ơn.

**Thay mặt nhóm thực hiện. /**

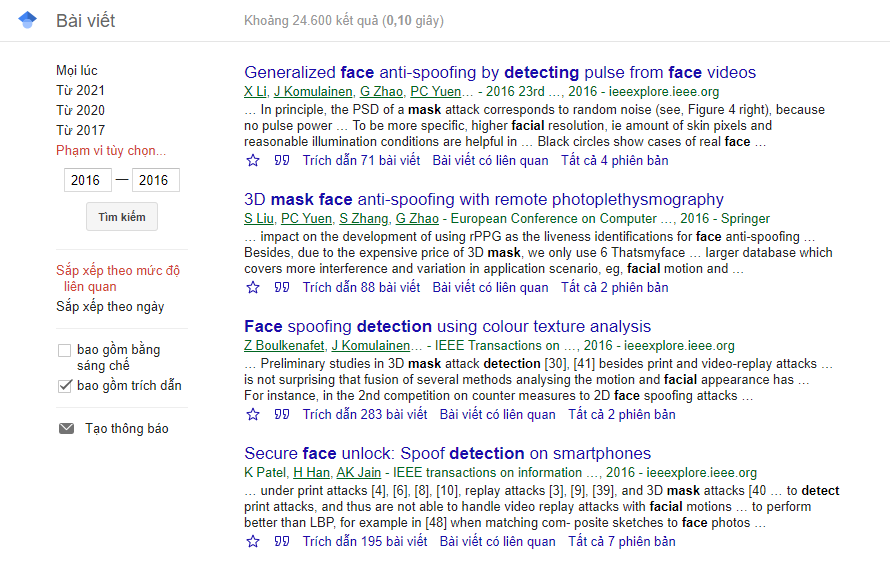
Lê Tấn Lộc

# **CHƯƠNG 1. TỔNG QUAN**

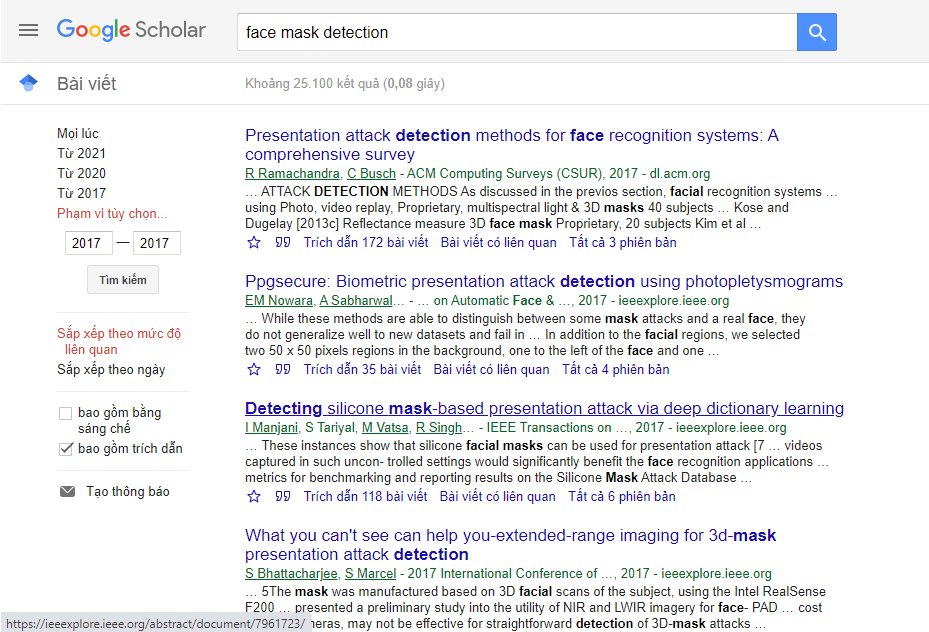
## **Giới thiệu bài toán**

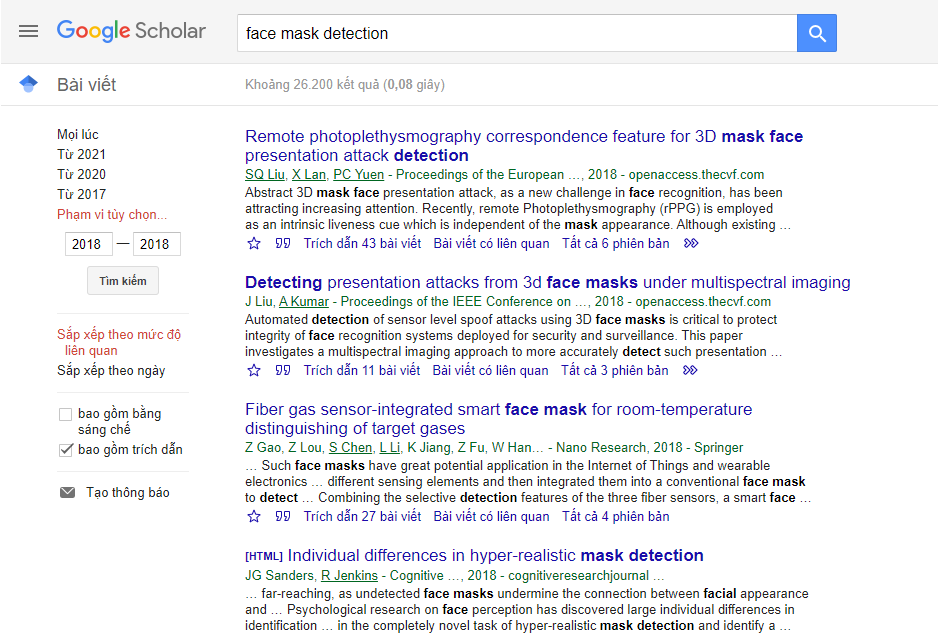
**– Đặt vấn đề**

Nhận dạng khẩu trang có tên (Face mask detection: FMD – còn được gọi là nhận dạng đối tượng khi đeo khẩu trang) là một bài toán chính thuộc lĩnh vực thị giác máy tính - Computer Vision. Đây là một bài toán với một bài toán lớn và một bài nhỏ, từ bài toán nhỏ là nhận dạng đối tượng dẫn đến bài toán lớn là nhận diện đối tượng khi đeo khẩu trang. Ngày nay, cùng với sự phát triển công nghệ kĩ thuật số mà trong vòng 5 năm trở lại đây, bài toàn đã được cộng đồng nghiên cứu quan tâm và cũng đạt nhiều thành tựu nổi bật.

Số bài viết về cụm từ “Face mask dectecion” trên google schoolar (2016-2020)

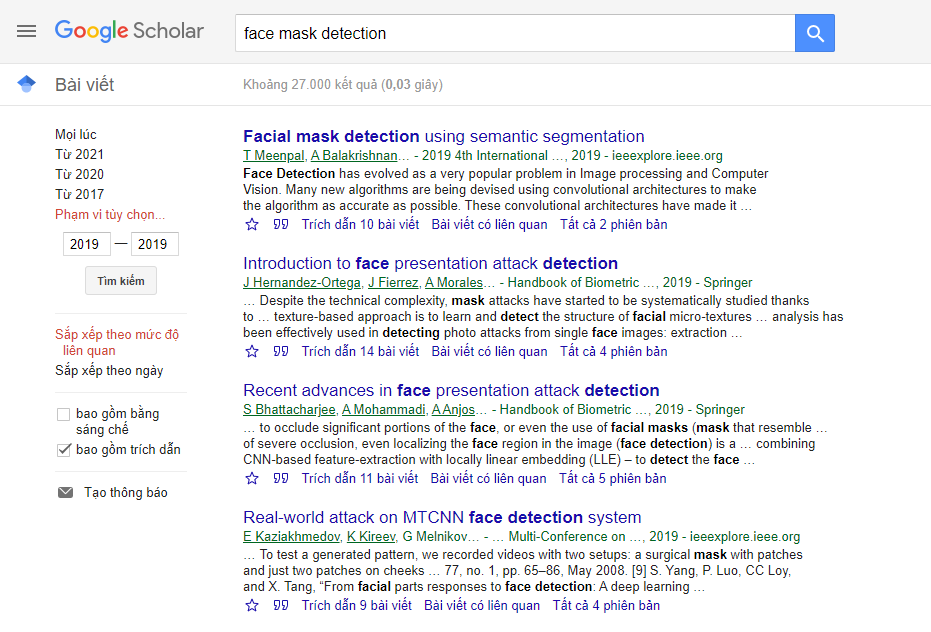
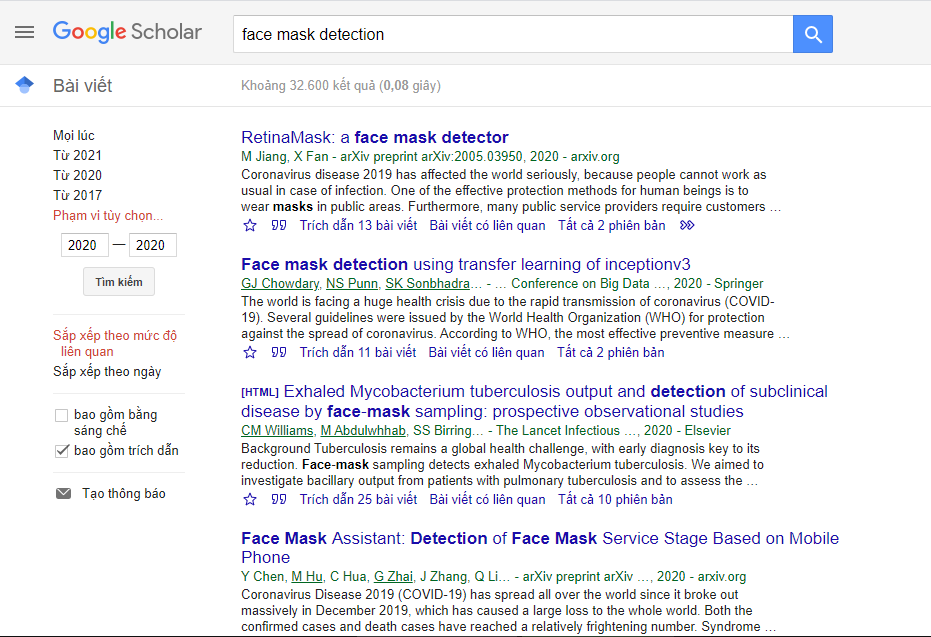
Hình 1: Thống kê các công trình nghiên cứu liên quan đến cụm từ “Face mask detection” 2016





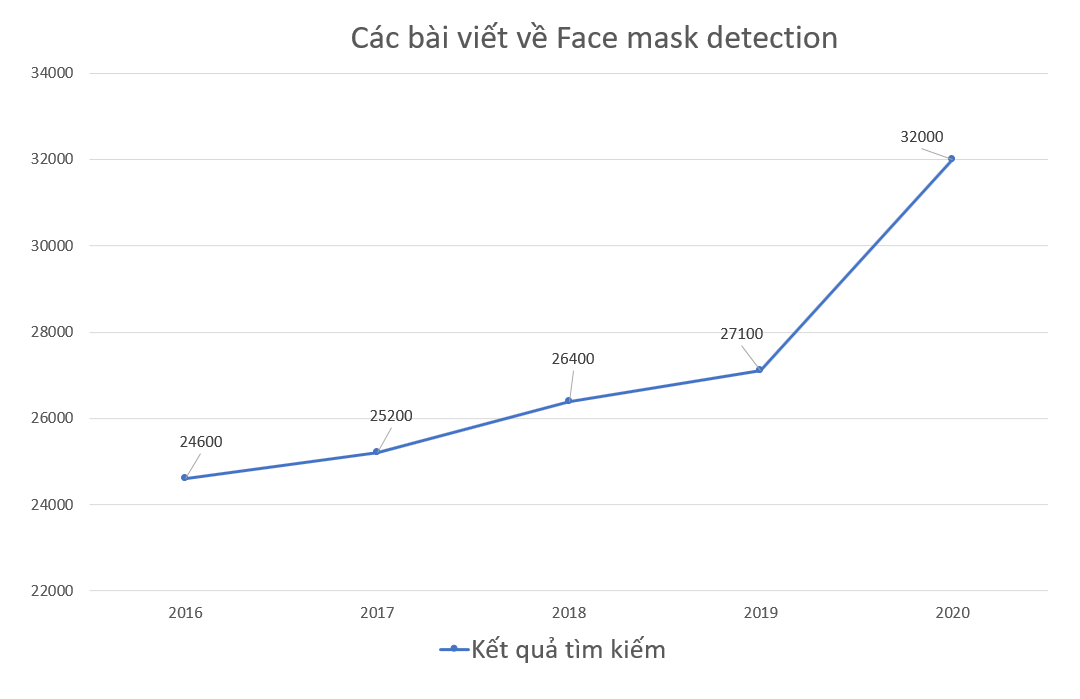
Hình 3: Thống kê các công trình nghiên cứu liên quan đến cụm từ “Face mask detection” 2018

Hình 2: Thống kê các công trình nghiên cứu liên quan đến cụm từ “Face mask detection” 2017



Hình 5: Thống kê các công trình nghiên cứu liên quan đến cụm từ “Face mask detection” 2020

Hình 4: Thống kê các công trình nghiên cứu liên quan đến cụm từ “Face mask detection” 2019



Hình 6: Thống kê các công trình nghiên cứu liên quan đến cụm từ “Face mask detection” 2016 - 2020

Face mask detection dần trở thành một chủ đề nóng cho cộng đồng nghiên cứu khoa học, thường xuyên được đề xuất tuong các hội nghị như: ICASSP, IFIP, CVPR, …và trên các trang tập chí nổi tiếng: Springer

Chỉ trong vòng 5 năm trở lại đây, các thành tựu về nhận điện đối tượng đeo khẩu trang đã nhận được môt số lượng đáng kể, tuy nhiên bài toán này vẫn còn nhiều hạn chế và thách thức liên quan dữ liệu cực lớn, bảo mật thông tin, nhận diện sai khi có các ý tố khác tác động (vật thể khác che mặt mà không phải khẩu trang, ánh sáng….)

Nhờ vào sự phát triển vượt bậc về các công nghệ kỹ thuật số, cùng với các công nghệ khác đã tạo cơ hội cho các lĩnh vực về xử lý hình ảnh dần ăn sâu vào đời sống con người về mọi mặt. Việc ứng dụng có hiệu quả các thành tựu này vào đời sống đã góp phần phát triển kinh tế và nâng cao đời sống xã hội, đồng thời nâng cao chất lượng cuộc sống của mọi người. Gần đây dịch covid đã là mối đe dọa đối với toàn nhân loại về vật chất lẫn tinh thần. Vì vậy bài toán nhận diện đối tượng có đeo khẩu trang rất cần thiết trong công cuộc phòng chống sự lây lan của dịch bệnh covid

* **Mục tiêu cụ thể**

+ Ở nghiên cứu này tập trung vào bài toán nhận diện những người có đeo khẩu trang đảm bảo hạn chế sự lây lan của dịch bệnh

+ Mục tiêu cụ thể của đề tài là phân tích, phát triển vấn đề, đề xuất giải pháp cho việc nhận diện các loại đối tượng có đeo khẩu trang thuộc miền dữ liệu trên

+ Xây dựng một chương trình có hiệu quả, đầy đủ và chính xác với mục đích hỗ trợ việc nhận dạng thực thể song xây dựng một hệ thống huấn luyện, thực nghiệm và ứng dụng.

## **Xử lý hình ảnh**

Xử lý hình ảnh là một lĩnh vực của thị giác máy tính tập trung vào các ứng dụng trên hình ảnh thực. Trong thị giác máy tính thì xử lý hình ảnh là một trong những phần trọng tâm vì nó liên quan đến việc phải xác định, phân tích hình ảnh.

Xử lý hình ảnh là xác định, phân tích và nghiên cứu cấu trúc các điểm ảnh và tạo ra các hệ thống thông minh có khả năng nhận dạng đối tượng từ hình ảnh và video, giúp giải quyết các vấn đề như phân loại từng đối tượng, xác định những vấn đề tiềm ẩn. Phân tích hình ảnh là một nhiệm vụ thiết yếu cho phép chúng ta nhận dạng ra các đối tượng ở trong hình ảnh hoặc video.

Nhờ vào sự tiến bộ của công nghệ kỹ thuật số đã góp phần tạo điều kiện cho chúng ta có thể thu thập được nguồn dữ liệu phong phú về các hình ảnh thử nghiệm, song việc xác định, phân tích và nghiên cứu cũng trở nên chuẩn xác, truyền cảm hứng cho sự đổi mới, phát triển và có thể dẫn đến công nghê nhận dạng khuôn mặt mạnh mẽ hơn.

## **Nhận dạng đối tượng đeo khẩu trang.**

Bài toán nhận dạng đối tượng đeo khẩu trang là bài toán xác định (phát hiện) những đối tượng đeo khẩu trang. Bài toán này sẽ phân tích các điểm ảnh dựa trên các thuộc tính cả việc nhận dạng các điểm ảnh đặc trưng của hình ảnh bằng phương pháp tính chập mạng nơ-ron (CNNs). Ví dụ trong trường hợp cung cấp các bức hình có một hay nhiều người đeo khẩu trang cho chương trình phân tích, sau đó lấy một bức ảnh có người đeo khẩu trang nào đó bất kì đưa cho chương trình kiểm tra, kỳ vọng sau khi kiểm tra là chương trình sẽ phản hồi được là có người đeo khẩu trang trong bước hình đó.

Đối tượng đầu vào cần được nhận diện là một hay nhiều hình ảnh hoặc nhận dạng trực tiếp thông qua các thiết bị ghi hình. Một dữ liệu sau khi được kiểm tra sẽ được xếp vào một trong hai trạng thái là có đeo khẩu trang hoặc không đeo khẩu trang. Như vậy, để có thể nhận diện được một người có đeo khẩu trang hay không, thì hình ảnh là dữ liệu cơ bản nhất, cho dù chương trình nhận diện chạy trên bất kì cấu trúc nào.

Nhận diện người đeo khẩu trang góp phần không nhỏ trong việc đảm bảo an ninh trật tự trong công cộng. Chính vì vậy, bài toán này đã nhận được sự quan tâm sâu sắc của giới khoa học, đặc biệt là trong lĩnh vực thị giác máy tính.

**Các ứng dụng trong thế giới thực của nhận diện đối tượng đeo khẩu trang**

**Trước mùa dịch covid:** Khi chưa bùng phát dịch bệnh, thì bài toàn này được sử dụng để phân tích, phát hiện mặt nạ siêu thật trong công tác an ninh và phòng chống tội phạm [1], giả mạo khuôn mặt [6], các vụ phạm tội liên quan đến máy ATM [5] bằng cách sử dụng xử lý hình ảnh giám sát video. để phát hiện khuôn mặt bị che khuất và bị che, phân tích hành vi bất thường của con người và phát hiện đối tượng bất hợp pháp có thể không hoạt động đối với ATM. Mặc khác, trong thời gian này các bài viết còn đề ra các phương pháp để ngăn chặn sự tấn công giả mạo như phát hiện xung từ video khuôn mặt.[2], chụp ảnh quang tuyến (PPGSecure) [3], học kết cấu động phức hợp sâu [4], ...

**Trong mùa dịch covid:**

Phát hiện mặt nạ bằng MobileNetV2 trong Kỷ nguyên Đại dịch COVID-19 [7]

Đại dịch Corona Virus Disease (COVID-19) đang gây ra môt cuộc khủng hoảng sức khỏe. Một trong những phương pháp hiệu quả để chống lại virus là đeo khẩu trang. Tính năng phát hiện mặt nạ có thể được các cơ quan chức năng sử dụng để lập kế hoạch giảm thiểu, đánh giá, phòng ngừa và hành động chống lại COVID-19. Nhận dạng mặt nạ trong nghiên cứu này được phát triển bằng thuật toán máy học thông qua phương pháp phân loại hình ảnh: MobileNetV2. Các bước để xây dựng mô hình là thu thập dữ liệu, xử lý trước, chia nhỏ dữ liệu, thử nghiệm mô hình và thực hiện mô hình. Mô hình đươc xây dựng có thể phát hiện những người đang đeo mặt nạ và không đeo mặt nạ và không đeo với độ chính xác 96.85%.

Xác định danh tính cá nhân của người đeo mặt nạ [8]

Một công ty Nhật Bản NEC, đã phát triển một hệ thống nhận diện khuôn mặt để danh tính đối tượng đeo khẩu trang. Cảnh sát đã sử dụng hệ thống nhận diện khẩu trang trong thời gian thực của công ty để dò tìm một số đối tượng trong danh sách.

Nhận diện khẩu trang ở thời gian thực trong thời kỳ COVID-19 dựa trên học sâu

Trong thời kỳ COVID-19 đang hoành hành, nó mang đến một mối hiểm họa cho thế giới, song song đó cũng nhắc nhở cho chúng ta cần phải kiểm soát sự hoàng hành của virus. Dẫn đến sự bùng nổ kết nổ không gian mạng giữa các thiết bị giúp cho việc thu nhập dữ liệu. Chẳng hạn như: hệ thống chăm sóc và thông báo cho mọi người chăm sóc sức khỏe. Đặc biệt, trong thời kỳ COVID-19 hiện nay, khẩu trang là một vật dụng không thể thiếu. Vì vậy, chúng ta cần một hệ thống nhận diện khẩu trang trong thời gian thực ngăn chặn sự lây lan. Trong bài viết này, chúng ta sẽ nói tới thư viện nhận dạng khẩu trang dựa trên máy tính (EC Mask) để giúp sức khỏe cộng đồng có thể chạy trên những camera ít năng lượng. Hệ thống EC Mask bao gồm 3 giai đoạn chính: chạy trên camera video, xác định khuôn mặt, nhận dạng danh tính [9].

**Phạm vi đề tài**

Phạm vi nghiên cứu của đề tài là nhận diện được đối tượng có đeo khẩu trang hay không trong video thời gian thực (ghi hình trực tiếp) trên các thiết bị ghi hình.

Phạm vi áp dụng có thể áp dụng trong các siêu thị, cửa hàng tiện lợi, tiệm tạp hóa, trường học, trung tâm thương mai…

Đối tượng được xử lý phải đảm bảo có 2/3 khuôn mặt được xác định qua khung hình camera thì mới có thể nhận diện được.

**Hướng phát triển trước, trong và sau dịch:**

**Trước mùa dịch CoV**: Sản phẩm sẽ tập trung vào những người đeo khẩu trang khi họ đi vào những nơi công cộng vì những người này không muốn người khác nhìn thấy mặt mình nên rất có khả năng họ sẽ làm điều gì đó mờ ám.

**Trong dịch mùa dịch CoV:** Để đảm bảo sức khỏe của mọi người và tránh tình trạng lây lan dịch bệnh thì việc đeo khẩu trang đến các nơi công cộng là điều bắt buộc. Vì vậy những ai không đeo khẩu trang khi đến những nơi công cộng là người có những hành vi bất thường.

**Sau mùa dịch CoV:** Có thể đổi lại cài đặt giống như lúc trước dịch. Vì việc đeo khẩu trang lúc này không còn bắt buộc nữa.

# **CHƯƠNG 2: CƠ SỞ LÝ THUYẾT**

## **2.1. Nhận diện khuôn mặt**

Nhận diện khuôn mặt (Face Detection) là một nhánh của Trí tuệ nhân tạo, tập trung vào việc nghiên cứu các bài toán để máy tính có thể tìm và nhận diện khuôn mặt trên những hình ảnh kỹ thuật số, được phát triển dựa trên những nghiên cứu của Thị giác máy tính (Computer Vision) cùng với sự tiến bộ các lĩnh vực như: máy học (Machine Learning), Mạng nơ-ron nhân tạo (Neural Networks), và nhiều công nghệ khác. Mục tiêu của lĩnh vực này là giúp máy tính tìm và nhận diện có hiệu quả với những đối tượng là khuôn mặt, ngoài việc xử lý để nhận diện khuôn mặt thì lĩnh vực này còn tập trung việc xác định các thông tin thông qua khuôn mặt như: tuổi, giới tính và cả cảm xúc của con người.

Những người tiên phong về nhận dạng khuôn mặt là Woody Bledsoe, Helen Chan Wolf và Charles Bisson. Năm 1964 và 1965, Bledsoe cùng với Wolf và Bisson bắt đầu sử dụng máy tính để nhận dạng khuôn mặt người.

Thừa kế những thành tựu của Bledsoe, bước nhảy tiếp theo của Goldstein, Harmon và Lesk vào những năm 1970 đã được cải tiến thêm việc tự động hóa nhận diện 21 điểm dấu đó bao gồm cả màu tóc, độ dày của miệng.

Cuối những năm 1980, chúng ta mới thấy sự tiến bộ hơn nữa với sự phát triển của nhận dạng khuôn được áp dụng ở các doanh nghiệp. Năm 1988, Sirovich và Kirby bắt đầu áp dụng đại số tuyến tính vào bài toán nhận dạng khuôn mặt.

Năm 1991, Turk và Pentland tiếp bước Sirovich và Kirby bằng cách phát hiện khuôn mặt trong một bức ảnh, dẫn đường cho việc nhận dạng khuôn mặt tự động sớm nhất. Tuy bước đột phá đáng kể này đã bị cản trở bởi các yếu tố công nghệ và môi trường, nhưng nó cũng đã góp phần cho công nghệ Nhận diện khuôn mặt được phát triển hơn trong tương lai.

Vào đầu những năm 1990, để khuyến khích thị trường nhận dạng khuôn mặt thương mại, Cơ quan Chỉ đạo các Dự án Nghiên cứu Quốc phòng Tiên tiến (DARPA) và Viện Tiêu chuẩn và Kỹ thuật quốc gia (NIST) đã triển khai chương trình Công nghệ Nhận dạng Khuôn mặt (FERET).

Vào đầu những năm 2000, Viện Tiêu chuẩn và Kỹ thuật quốc gia (NIST) đã bắt đầu Thử nghiệm nhà cung cấp nhận dạng khuôn mặt (FRVT). Dựa trên FERET và FRVTs được thiết kế để cung cấp các đánh giá độc lập của chính phủ về các hệ thống nhận dạng khuôn mặt có sẵn trên thị trường cũng như các công nghệ nguyên mẫu.

Face Recognition Grand Challenge (FRGC) được ra mắt vào năm 2006 với mục tiêu chính là thúc đẩy và nâng cao công nghệ nhận dạng khuôn mặt, được thiết kế nhằm hỗ trợ các nỗ lực nhận dạng khuôn mặt hiện có của Chính phủ Hoa Kỳ.

Năm 2010, Facebook bắt đầu triên khai tính năng nhận dạng khuôn mặt giúp xác định những người có khuôn mặt có thể xuất hiện trong ảnh mà người dùng Facebook cập nhật hằng ngày. Tuy nhiên, tính năng này ngay lập tức gây tranh cái bởi các phương tiện truyền thông tin tức, làm dấy lên một loạt các bài báo liên quan đến quyền riêng tư. Dù vậy, nhưng người dùng Facebook dường như không bận tâm.

Công nghệ Nhận dạng khuôn mặt đã phát triển nhanh chóng từ năm 2010 trở lại đây và ngày 12 tháng 9 năm 2017 đánh dấu một bước đột phá quan trọng khác để tích hợp tính năng nhận dạng khuôn mặt vào cuộc sống hằng ngày của chúng ta. Song cũng là ngày Apple ra mắt iPhone X, qua đó người dùng iPhone đầu tiên có thể mở khóa điện thoại bằng FaceID – thuật ngữ tiếp thị của Apple về nhận dạng khuôn mặt.

**Một số thuật toán nhận diện khuôn mặt kinh điển:**

Năm 1960, Bled đã phát triển một hệ thống có thể phân loại khuôn mặt bằng tay trên thiết bị là RAND tablet. Là một thiết bị có thể nhập được tọa độ ngang dọc bằng cách xử dụng một cây bút cảm ứng để có thể gửi mã hóa thông tin tọa độ cho máy hiểu. Bằng cách này ông đã có thể ghi lại những đặc điểm quan trọng trên khuôn mặt và những thông tin đó sẽ được lưu trữ trong cơ sở dữ liệu. Và khi đưa một hình ảnh ai đó vào hệ thống, nó sẽ có đủ dữ liệu cũng như khả năng để so sánh giữa thông tin trong dữ liệu và thông tin lấy được từ tấm ảnh mới được đưa vào

Eigenfaces được Sirovich và Kirby nghĩ ra khi áp dụng tuyến tính vào trong vấn đề nhận diện khuôn mặt. Khi mà đưa vào một bức ảnh vào trong máy cần rất nhiều thời gian để có thể nhận dạng được do phải cần thêm thời gian để xác định các đặc điểm nào quan trọng trên khuôn mặt. Và Eigenface đã giúp giảm đi thời gian xác định đó bằng cách làm giảm đi hoặc bớt đi những đặc điểm không quan trọng và giữ những lại những đặc điểm để đi so sánh. Và không những giảm thiểu thời gian cho máy xử lý mà còn làm giảm đi những dự đoán sai của máy đưa ra.

**Các thuật ngữ cơ bản trong nhận diện khuôn mặt:**

Face detection: Phát hiện khuôn mặt trong ảnh

+ Các máy ảnh camera hiện tại điều có chức năng này.

+ Thuật toán được sử dụng phổ biến nhất hiện nay là Viola-Jones(Thư viện OpenCV hỗ trợ nhận diện khuôn mặt theo thuật toán này). Ngoài nó nó còn có thể nhận diện được đồ vật, hình dạng.

+ Cơ chế hoạt động:

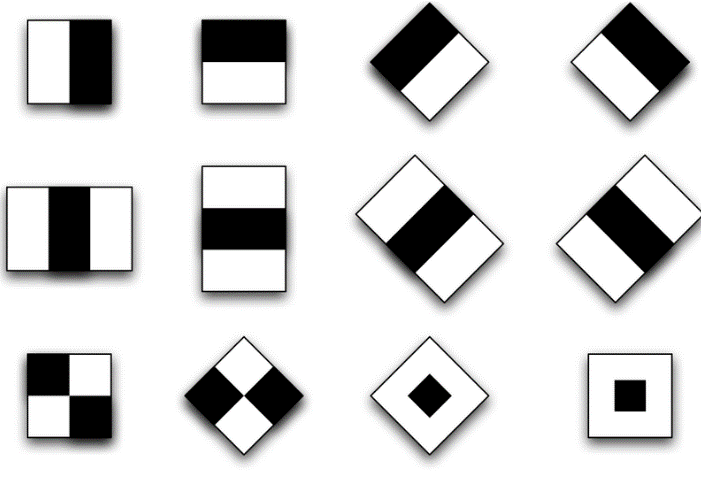
Tạo ra một frame hình vuông, lần lượt di chuyển frame này khắp tấm ảnh gốc

Mổi khi frame di chuyển, check xem khu vực trong frame có phải là khuôn mặt hay đồ vật không (Bằng cách check các vùng sáng tối trong frame)

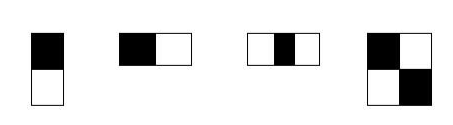
Sau khi di chuyển hết ảnh, tăng kích cở frame lên là scan lại từ đầu

Dừng thuật toán khi frame đủ lớn

Face Recognition: quá trình nhận diện khuôn mặt từ những gương mặt đã quét được trong ảnh. Các khuôn mặt này sẽ được so với các khuôn mặt có trong cơ sở dữ liệu để trả về kết quả

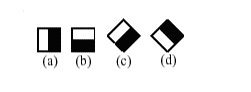
**Haar-Like** là những hình chữ nhật được phân thành các vùng khác nhau như hình:

Đặc trưng do Viola và Jones công bố gồm 4 đặc trưng cơ bản để xác định khuôn mặt người. Mỗi đặc trưng Haar-Like là sự kết hợp của hai hay ba hình chữ nhật trắng hay đen như trong hình sau:

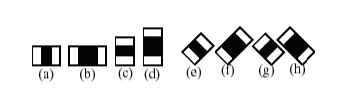


Để sử dụng các đặc trưng này vào việc xác định khuôn mặt người, 4 đặc trưng Haar-Like cơ bản được mở rộng ra và được chia làm 3 tập đặc trưng như sau

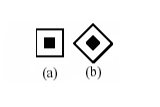
Đặc trưng cạnh (edge feature)



Đặc trưng đường (line feature)



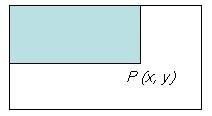
Đặc trưng xung quanh tâm (center-surround features)



Dùng các đặc trưng trên, ta có thể tính được các giá trị của đặc trưng Haar-Like là sự chênh lệch giữa tổng của các pixel của vùng đen và vùng trắng theo công thức:



Viola và Joines đưa ra một khái niệm gọi là Integral Image, là một mảng 2 chiều với kích thước bằng với kích thước của ảnh cần tính đặc trưng Haar-Like, với mỗi phần tử của mảng này được tính bằng cách tính tổng của điểm ảnh phía trên (dòng-1) và bên trái (cột-1) của nó



Công thức tính Intergral Image

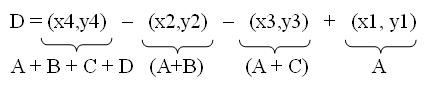


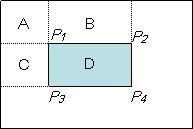
Sau khi tính được Integral Image, việc tính tổng các giá trị mức xám của một vùng bất kỳ nào đó trên ảnh thực hiện rất đơn giản theo cách sau:

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

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:

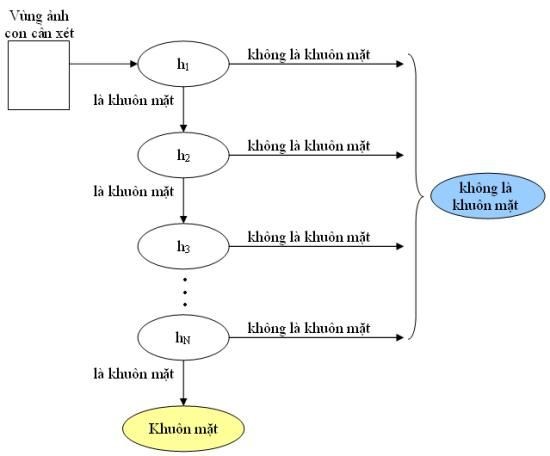




**AdaBoost**

AdaBoost là một bộ phân loại mạnh phi tuyến phức dựa trên hướng tiếp cận boosting được Freund và Schapire đưa ra vào năm 1995. Adaboost cũng hoạt động trên nguyên tắc kết hợp tuyến tính các weak classifiers để hình thành một trong các classifiers

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:

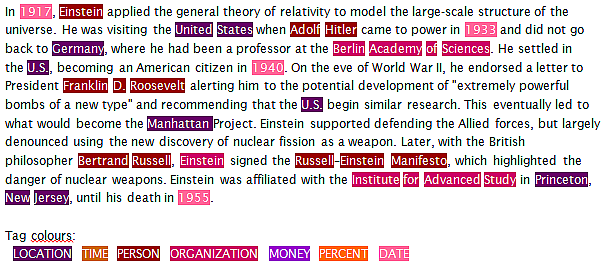


## **2.2. Nhận diện khẩu trang trên khuôn mặt**

**Information Extraction (IE: trích xuất thông tin) là quá trình phân tích, xử lý dữ liệu để trích xuất các thông tin hữu ích, có cấu trúc từ nguồn thông tin phi cấu trúc hoặc bán cấu trúc. Thông thường quá trình này bao gồm ba bước chính là: xác định thực thể (NER: Named Entity Recognition), xác định mối liên hệ (Relation Extraction) và trích xuất sự kiện (Event Extraction).

Hình 3: Minh họa về khái niệm trích xuất thông tin (Information Extraction)

Bước đầu tiên trong quá trình Trích xuất thông tin là việc xác định các thực thể xuất hiện trong đoạn dữ liệu rồi tiến hành phân loại thực thể này (được gọi là Named Entity Recoginition hay NER). Việc xác định thực thể có thể bắt đầu bằng việc xác định các từ trong đoạn văn bản có POS tag là Proper Name (như person, location, organization). Sau đó, việc phân loại thực thể (NER) có thể được thực hiện bằng việc áp dụng các sequence model như MEMM hoặc CRF. Tuy nhiên công việc này thường gặp khó khăn do tính không rõ ràng của các từ. Ví dụ như từ ‘Washington’ có thể được phân loại là Tên người (PERSON) như tổng thống Washington hoặc Địa danh (LOCATION) như thủ đô Washington,…

**Ngoài ra, một khái niệm khác có liên quan đến NER là Coreference Resolution. Coreference Resolution được sử dụng để tìm tất cả các từ đại diện cho cùng một thực thể nhất định trong đoạn văn bản.

Hình 4: Minh họa về hoạt động của NER (Named Entity Recoginition)

Bên cạnh vai trò trong việc xác định mối liên hệ (Relation Extraction) và phát hiện sự kiện (Event Extraction) thì NER còn được áp dụng trong nhiều bài toán liên quan đến Xử lý ngôn ngữ tự nhiên .

## **2.3. Tình hình nghiên cứu ngoài nước về bài toán Face Detection**

**Học có giám sát:** Hiện nay phần lớn các công trình giải quyết bài toán trích rút thực thể NER thường sử dụng kỹ thuật học máy có giám sát như mô hình Markov ẩn, cây quyết định, mô hình Maximum entropy, máy vector hỗ trợ (SVM), mô hình trường ngẫu nhiên có điều kiện CRF. Nhược điểm của học có giám sát là đòi hỏi tập dữ liệu huấn luyện gán nhãn bằng tay cực lớn.

**Học không giám sát:**

Để giải quyết vấn đề gán nhãn bằng tay một số nhà nghiên cứu đã sử dụng phương pháp học không giám sát. Các hệ hống học máy không giám sát không đòi hỏi dữ liệu huấn luyện, nhưng khả năng phát hiện thực thể không cao và độ chính xác thấp.

**Học bán giám sát:** Phương pháp học bán giám sát có thể khắc phục được nhược điểm của hai phương pháp trên. Phương pháp học bán giám sát huấn luyện hệ thống sử dụng cả dữ liệu gán nhãn và không gán nhãn. Có nhiều nghiên cứu về trích rút thực thể sử dụng phương pháp học bán giám, trong đó có Liao:

Liao sử dụng phương pháp học bán giám sát để huấn luyện hệ thống. Hệ thống ban đầu có một tập dữ liệu được gán nhãn thực thể với số lượng nhỏ, sử dụng mô hình CRF để huấn luyện hệ thống. Kết quả của quá trình huấn luyện được sử dụng để phân loại các thực thể trên tập dữ liệu chưa được gán nhãn. Sau đó tập dữ liệu và tập đặc trưng được cải tiến bằng cách tìm các dữ liệu được phân loại có độ tin cậy thấp bởi mô hình CRF bằng tập dữ liệu được gán nhãn bằng tay kích cỡ nhỏ. Tiếp theo, sử dụng mô hình này để phân loại dữ liệu chưa có nhãn và sau đó tập dữ liệu đã được cải thiện và tập đặc trưng được cải thiện bằng cách áp dụng cho các dữ liệu chưa được gán nhãn được phân loại có độ tin cậy thấp nhờ mô hình học phân lớp được huấn luyện theo tập dữ liệu đã được gán nhãn thủ công có kích thước nhỏ. Liao đề xuất các tri thức để tìm các dữ liệu huấn luyện mới, các tri thức bên ngoài được sử dụng là hai giả định ngôn ngữ sau đây:

- Giả định xuất hiện nhiều lần: Những cụm từ (viết hoa) giống hệt nhau cùng xuất hiện trong một văn bản thường có kiểu thực thể giống nhau. Chẳng hạn, trong một văn bản tài chính, *John* có thể xuất hiện nhiều lần với cùng một tên là *“John”*

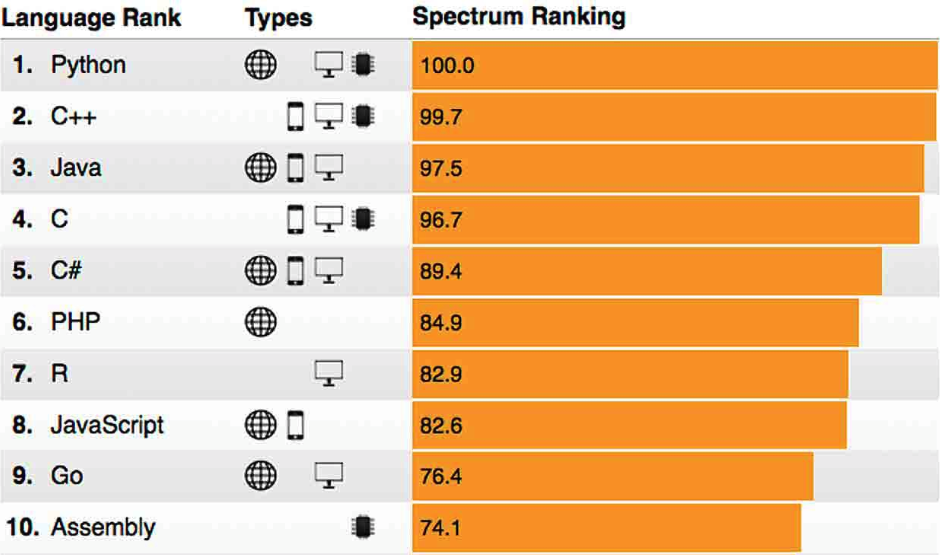
- Giả định ngữ cảnh: Các thực thể như con người, tổ chức, địa điểm thường có các ngữ cảnh xung quanh giúp chúng ta nhận biết được chúng một cách dễ dàng. Chẳng hạn, tên một tổ chức thường được đi cùng bởi các hậu tố như: *Inc., Co.,* …; tên một người thường được đi cùng bởi các tiền tố như: *Mr.*, *CEO*,…

# **CHƯƠNG 3: THỰC NGHIỆM CHƯƠNG TRÌNH**

## **3.1. Môi trường**

### 3.1.1 Ngôn ngữ Python

Python là một ngôn ngữ lập trình thông dịch, hướng đối tượng, ngôn ngữ lập trình cấp cao được giải thích với ngữ nghĩa động. Python với triết lý thiết kế của nó rất thuận tiện cho việc đọc hiểu code, đơn giản và rõ ràng được thiết kế bởi **Guido van Rossum**. Thiết kế bắt đầu vào cuối những năm 1980 và được phát hành lần đầu tiên vào tháng **2 năm 1991**. Đến nay thì cộng đồng người sử dụng ngôn ngữ này rất đông, nếu so sánh từ bảng xếp hạng các ngôn ngữ năm 2018 thì Python đã leo lên vị trí số 1 trên bảng xếp hạng những ngôn ngữ lập trình phổ biến.

Hiện nay, Python cũng là một ngôn ngữ rất phát triển trong lĩnh vực Data Science và Machine Learning. Python cũng cung cấp những hàm và thư viện xử lý hình ảnh. Chính vì vậy, Python là một lựa chọn hợp lý khi thực hiện xử lý hình ảnh và nhận diện khuôn mặt.

Hình 5: IEEE Spectrum The Top Programming languages 2018

### 3.1.2 Thư viện Tensorflow

Tensorflow là một thư viện mã nguồn mở để tính toán và học máy quy mô lớn được sử dụng chủ yếu trong việc đào tạo máy học và suy luận trong các mạng nơ-ron như: thu thập dữ liệu, xây dựng mô hình đào tạo, phục vụ đề án và dự đoán kết quả tương lai.

Tesorflow kết hợp các mô hình và thuật toán Machine Learning và Deep Learning lại với nhau và chạy trên Python.

Cho phép tạo ra một đồ thị tính toán để thực hiện. Mỗi một đỉnh trong biểu đồ đại diện cho một phép toán và mỗi kết nối đại diện cho dữ liệu, giúp có cái nhìn trực quan hơn về lối logic của bài toán.

**-Cài đặt:**

Để cài đặt Tensorflow trên môi trường Python, có thể cài đặt trực tiếp thông qua pip.

pip install tensorflow

### 3.1.3 Thư viện Numpy

Numpy là một package tuyệt vời dành cho tính toán dữ liệu trong Python. Nó là 1 thư viện Python có thể cung cấp 1 mảng đa chiều như mảng đánh dấu hoặc ma trận và cung cấp những phép toán xoay quanh ma trận bao gồm toán logic, định dạng size, sắp xếp, chọn lọc, biến đổi, biểu diển những phép tính cơ bản của tuyến tính, mô phỏng random.

Numpy cho phép làm việc hiệu quả với ma trận và mảng, đặc biệt là dữ liệu ma trận và mảng lớn với tốc độ xử lý nhanh

-**Cài đặt**

Để cài đặt Numpy trên môi trường Python, có thể cài đặt trực tiếp thông qua pip.

pip install numpy

Nếu sử dụng conda:

Sử dụng environment thay vì cài đặt trong base env

conda create -n my-env

conda activate my-env

Muốn cài đặt từ conda-forge

conda config --env --add channels conda-forge

Sau đó nhập lệnh cài đặt trên terminal

conda install numpy

### 3.1.4 Thư viện Matplotlib

Matplotlib là một thư viện viết bằng Python dùng để hổ trợ cho việc vẽ đồ thị trong ngôn ngữ lập trình Python. Nó biểu diển được cả đồ thị 3d và 2d

Hổ trợ rất mạnh mẽ hữu ích trong việc vẽ đồ thị cho những người làm việc với Python và Numpy

**-Cài đặt**

Để cài đặt Numpy trên môi trường Python

python -m pip install -U pip

python -m pip install -U matplotlib

### 3.1.5 Thư viện Keras

Keras là một thư viện mã nguồn mở được viết bằng Python nó cung cấp giao diện cho mạng neural network. Keras hoạt động như một giao diện cho thư viện TensorFlow.

Keras giúp xây dựng dễ dàng các models như deep learning mà không cần có kiến thức sâu về các thuật toán máy học cũng như về mạng neural network

**-Cài đặt**

Bưới 1: Cài thư viện tensorflow

pip install tensorflow

Bưới 2: import thư viện tensorflow trong code

import tensorflow as tf

from tensorflow import keras

### 3.1.6 Thư viện OpenCV-Python

OpenCV-Python là một thư viện của Python, bao gồm các ràng buộc về thiết kế để giải quyết các vấn đề liên quan tới Thị giác máy tính, hỗ trợ rất hiệu quả trong xử hình ảnh, quay video, phân tích và có cả các tính năng như nhận diện khuôn mặt và phát hiện đối tượng.

**-Cài đặt**

Để cài đặt Numpy trên môi trường Python, có thể cài đặt trực tiếp thông qua pip.

pip install opencv-python

### 3.1.7 Thư viện Imutils

Imutils là một thư viện cung cấp một loạt các hàm có sẵn nhằm cung cấp các chức năng sử lý hình ảnh cơ bản như: dịch hình ảnh, xoay ảnh, điều chỉnh kích thước ảnh, xây dựng cấu trúc liên kết của đối tượng trong hình ảnh, hỗ trợ hiện thị hình ảnh Matplotlib, phân loại đường viền, phát hiện các cạnh và còn nhiều tính năng khác, sử dụng dễ dàng với OpenCV và cả Python.

**-Cài đặt**

Để cài đặt Numpy trên môi trường Python, có thể cài đặt trực tiếp thông qua pip.

pip install imutils

### 3.1.8 Thư viện SciPy

Thư viện SciPy là một trong những nền tảng cơ bản để tạo nên ngăn xếp SciPy (SciPy stack). Nó sẽ cung cấp nhiều tính năng nhằm hỗ trợ cho việc phân tích và tính toán các vấn đề liên quan với số học như: tích phân, nội suy, tối ưu hóa, đại số tuyến tính, số liệu thống kê, ...

**-Cài đặt**

Để cài đặt Numpy trên môi trường Python, có thể cài đặt trực tiếp thông qua pip.

python -m pip install --user numpy scipy matplotlib ipython jupyter pandas sympy nose

## **3.2. Dữ liệu**

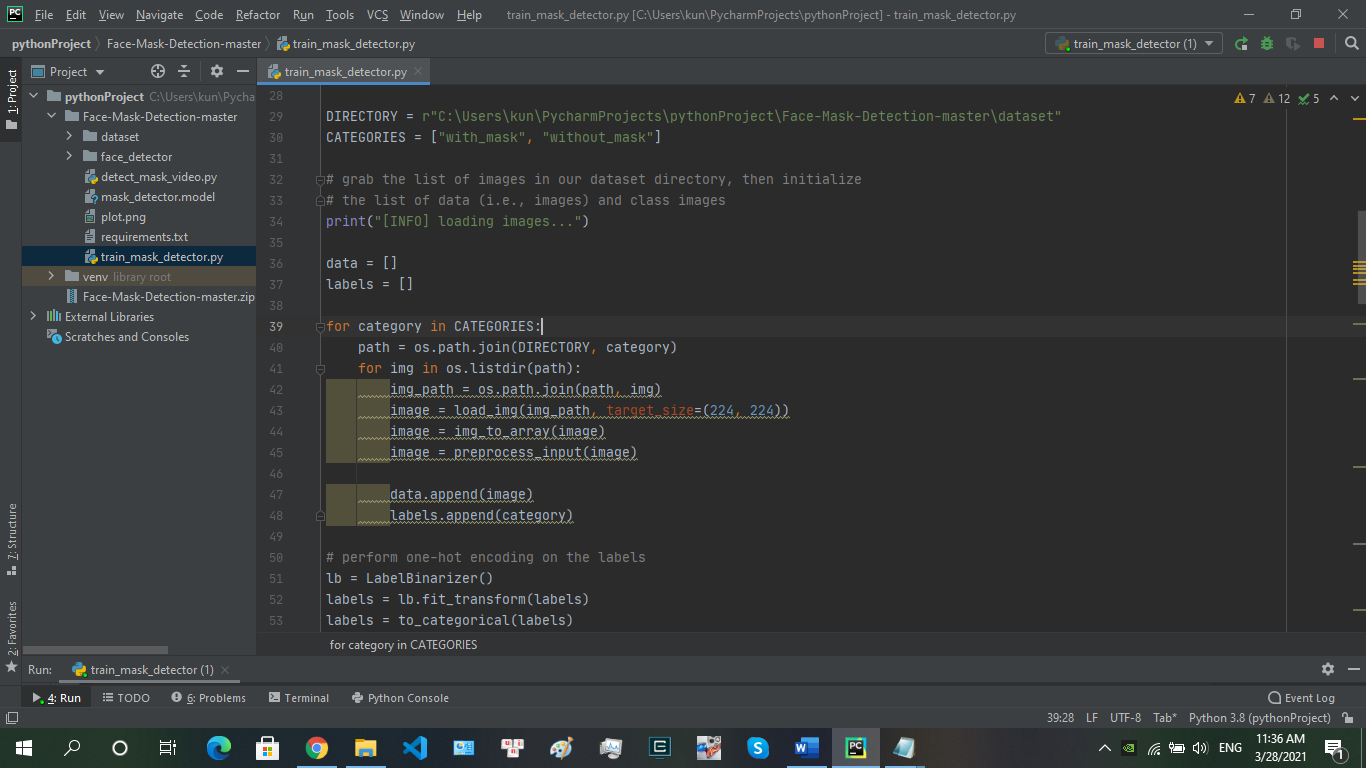
### 3.2.1 Yêu cầu dữ liệu

Dữ liệu để kiểm tra là phải cho máy tính nhìn thấy được trên 2/3 khuôn mặt có đeo hoặc không đeo khẩu trang mà không phải là một đối tượng nào khác khi quay video.

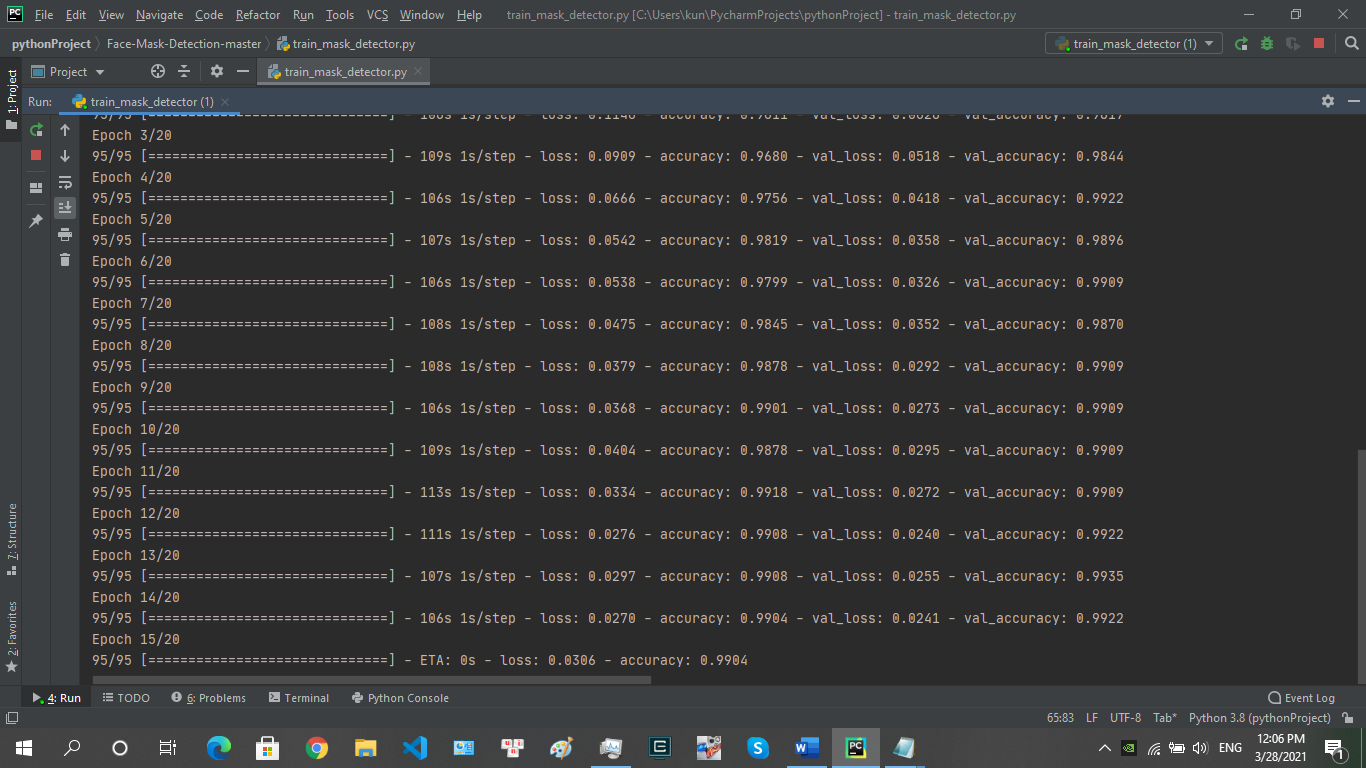
Ví dụ:



### 3.2.2 Phân tích dữ liệu

Mở file *“train\_mask\_detector.py”*:

Hình 10: Giao diện khi mở file train\_mask\_detector.py bằng PyCharm (Python 3.8.6rc1)

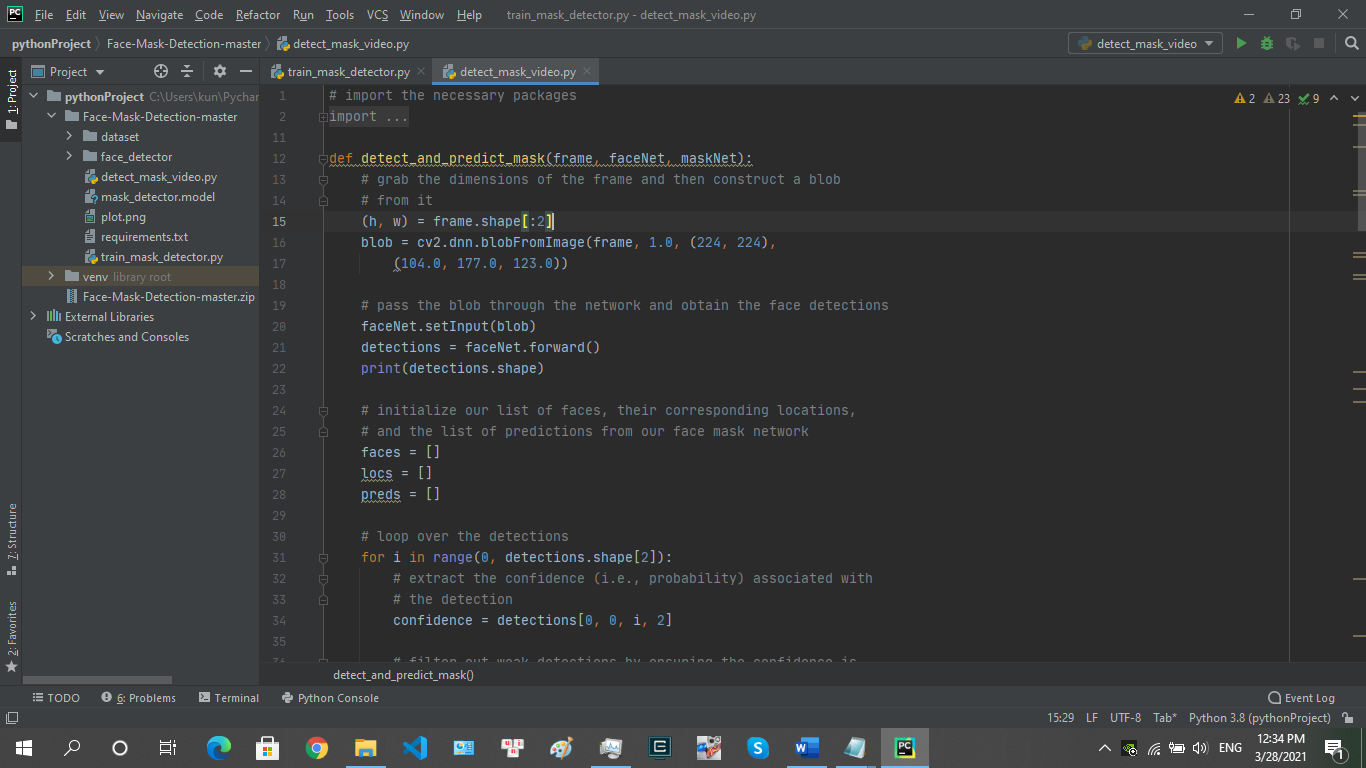
Tiến hành chạy file *“train\_mask\_detector.py”*:

Hình 11: File " train\_mask\_detector.py " đang được thực thi

Sau khi chương trình thực thi xong, chương trình sẽ tạo một file là “mask\_detector\_model”



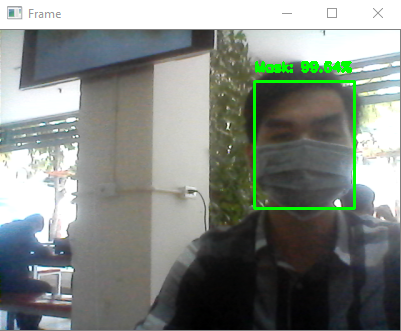
## **3.3. Lấy dữ liệu không gian thực**

Mở file **“***detect\_mask\_video.py”***:**

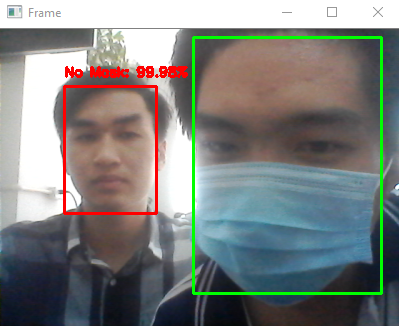
Hình 13: Giao diện khi mở file " detect\_mask\_video.py " bằng PyCharm (Python 3.8.6rc1)

Tiến hành thực thi chương trình trong file *“detect\_mask\_video.py”*:

Khi chạy chương trình thì camera trước của thiết bị chạy chương trình sẽ được bật lên, và nhận dạng khuôn mặt có sử dụng khẩu trang hay không trên không gian thực

*Đối tượng có sử dụng khẩu trang.* *Đối tượng không có sử dụng khẩu trang*.

Một đối tượng sử dụng khẩu trang và một đối tượng không sử dụng khẩu trang



# **CHƯƠNG 4: KẾT LUẬN**

## **4.1. Kết luận**

Nghiên cứu đã trình bày về vấn đề nhận dạng đối tượng sử dụng khẩu trang, một bài toán quan trọng trong lĩnh vực xử lý hình ảnh. Luận án tập trung nghiên cứu, phát triển về lý thuyết và ứng dụng đối với bài toán nhận dạng đối tượng, đề xuất một số mô hình và giải pháp nhằm nâng cao hiệu quả nhận dạng đối tượng sử dụng khẩu trang và đưa ra một số khung làm việc phục vụ cho quá trình nhận dạng đối tượng.

Nâng cao chất lượng nhận dạng đối tượng liên quan tới khuôn mặt sử dụng khẩu trang bằng cách thu nhập số lượng dữ liệu hình ảnh đủ lớn, tiến hành thực nghiệm ngoài thực tế để ghi chép những những sai sót, lập bảng thống kê về các tham số để chương trình được tối ưu hơn. detect\_mask\_video.py detect\_mask\_video.py

## **4.2. Hướng phát triển (bổ sung)**

Hiện tại, do hạn chế về thời gian, nghiên cứu dừng lại thử nghiệm với tập dữ liệu có kích cỡ vừa. Sau này chúng em sẽ mở rộng thử nghiệm trên tập dữ liệu lớn hơn.

Mở rộng và thử nghiệm trên các kiểu thực thể và mối quan hệ thực thể khác.

Cải tiến áp dụng các dạng biểu đồ trực quan khác nhau trong việc phân tích các thực thể.

Tuy nghiên cứu còn những hạn chế và thiếu sót nhất định, nhưng chúng em sẽ tiếp tục cố gắng hoàn thiện trong thời gian tới.

# **TÀI LIỆU THAM KHẢO**

[1] Sanders, J.G., Jenkins, (2018). R Individual differences in hyper-realistic mask detection. Cogn. Research 3, 24

[2] Xiaobai Li, J. Komulainen, G. Zhao, Pong-Chi Yuen and M. Pietikäinen, "Generalized face anti-spoofing by detecting pulse from face videos," 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 2016, pp. 4244-4249

[3] E. M. Nowara, A. Sabharwal and A. Veeraraghavan, "PPGSecure: Biometric Presentation Attack Detection Using Photopletysmograms," 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), Washington, DC, USA, 2017, pp. 56-62

[4] R. Shao, X. Lan and P. C. Yuen, "Deep convolutional dynamic texture learning with adaptive channel-discriminability for 3D mask face anti-spoofing," 2017 IEEE International Joint Conference on Biometrics (IJCB), Denver, CO, USA, 2017, pp. 748-755

[5] Sikandar, T., Ghazali, K.H. & Rabbi, M.F. ATM crime detection using image processing integrated video surveillance: a systematic review. Multimedia Systems 25, 229–251 (2019).

[6] A. Zaliha Abd Aziz, H. Wei and J. Ferryman, "Face anti-spoofing countermeasure: Efficient 2D materials classification using polarization imaging," 2017 5th International Workshop on Biometrics and Forensics (IWBF), Coventry, UK, 2017, pp. 1-6

[7] S. A. Sanjaya and S. Adi Rakhmawan, "Face Mask Detection Using MobileNetV2 in The Era of COVID-19 Pandemic," 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI), Sakheer, Bahrain, 2020, pp. 1-5,

[8] https://www.bbc.com/news/technology-55573802

[9] https://www.bbc.com/news/technology-55573802

[10] pandas.pydata.org

[11] www.govinfo.gov