

# NGHIÊN CỨU GIẢI THUẬT PHÁT HIỆN KHẨU TRANG CHO ROBOT TRONG PHÒNG CHỐNG BỆNH DỊCH

Trần Hoàn\*, Hoàng Đắc Huy

Trường Đại học Công Thương Thành phố Hồ Chí Minh

\*Email: hoant@huit.edu.vn

Ngày nhận bài: 17/01/2023; Ngày chấp nhận đăng: 27/02/2023

## TÓM TẮT

Tính năng phát hiện khẩu trang đã đạt được những tiến bộ đáng kể trong lĩnh vực xử lý ảnh và thị giác máy tính, góp phần quan trọng vào công tác phòng chống dịch Covid-19. Bài báo này trình bày phương pháp sử dụng mô hình Caffe kết hợp với kiến trúc MobileNetV2 để áp dụng cho các thiết bị nhúng có khả năng tính toán hạn chế như Raspberry Pi, NVIDIA Jetson Nano... làm bộ phát hiện khẩu trang theo thời gian thực ngay cả khi đối tượng đội nón hoặc đeo kính. Kết quả bài toán đã được thực nghiệm trên cơ sở dữ liệu tự xây dựng từ sinh viên Khoa Công nghệ Điện - Điện tử trường Đại học Công Thương Thành phố Hồ Chí Minh với độ chính xác lên đến 93%.

*Từ khóa:* Khẩu trang, phát hiện, Caffe, MobileNetV2, Raspberry Pi, Covid-19.

## 1. GIỚI THIỆU

Ở Việt Nam, hàng loạt Robot hỗ trợ chống dịch Covid-19 đã ra đời như Robot CD1.0 khử khuẩn bằng tia UV của trường Đại học Tôn Đức Thắng, Robot phun xịt khử khuẩn phòng cách ly điều khiển từ xa của Trung tâm công nghệ thông tin thuộc Bệnh viện Quân dân y miền Đông, Robot Vibot hỗ trợ vận chuyển thức ăn, thuốc men của Học viện Kỹ thuật quân sự Bộ Quốc phòng, Robot tự hành RaRoVid của Viện Ứng dụng Công nghệ [1].

Trong số đó, chỉ có Robot của trường Đại học Công nghệ (thuộc khối Đại học Quốc gia) chế tạo là có chức năng nhắc nhở khi phát hiện người không đeo khẩu trang [2]. Nguyên lý hoạt động của Robot dựa trên việc phát hiện mũi và miệng của đối tượng đang bị che, có nghĩa là người đó đang đeo khẩu trang. Còn trong trường hợp miệng bị che nhưng mũi hở, hoặc cả mũi và miệng đều hở, nghĩa là người đó đeo chưa đúng cách hoặc không đeo. Giải thuật này có khuyết điểm lớn là không phân biệt được trường hợp đối tượng cố tình dùng tay hoặc vật gì đó che lên mũi, miệng để đánh lừa Robot rằng đang đeo khẩu trang đúng cách. Ngoài ra, nếu đối tượng đội nón hoặc đeo kính sẽ dẫn đến khó phát hiện chính xác khuôn mặt hoặc các chi tiết chính trên khuôn mặt như mắt, mũi, miệng...

Kể từ khi đại dịch Covid-19 bùng phát, hầu hết các thuật toán phát hiện khẩu trang được đề xuất chủ yếu là các thuật toán dựa trên học sâu, là một tập hợp con của các thuật toán máy học [3-6]. Mạng dựa trên học sâu bao gồm mạng nơ-ron sâu (DNN), mạng nơ-ron hồi quy (RNN) và mạng bộ nhớ ngắn - dài hạn (LSTM). Mạng neuron tích chập (CNN) là một loại mạng DNN, đã được sử dụng rộng rãi trong các kỹ thuật phát hiện khẩu trang. Một CNN điển hình có năm đến bảy lớp. Tuy nhiên còn nhiều loại CNN khác như VGGNet (16 lớp) [7], GoogLeNet (19 lớp) [8], ResNet-50 (50 lớp) [9], MobileNetV1 (28 lớp) [10], MobileNetV2 (53 lớp) [11].

Trong đó, MobileNetV1 và MobileNetV2 do nhẹ nên được sử dụng rất nhiều trong việc phát hiện khẩu trang để đạt hiệu suất tối ưu [6, 12, 13]. Ngoài ra, thuật toán phát hiện đối tượng

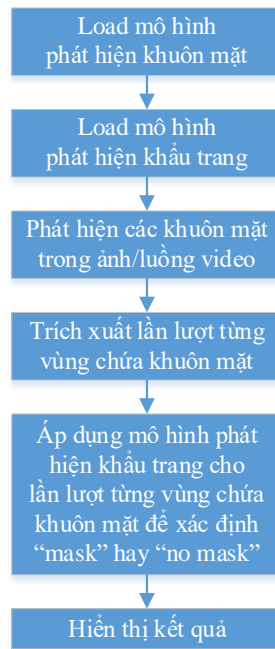
khác nhau như You Only Look Once (YOLO) cũng được sử dụng dù tiêu tốn nhiều tài nguyên, đòi hỏi bộ nhớ lớn và khả năng tính toán cao của thiết bị [14]. Bên cạnh đó, kiến trúc Feature Pyramid Network (FPN) cũng được đề xuất trong [15, 16].

Việc phát hiện khẩu trang trong công tác phòng chống dịch bệnh truyền nhiễm lây qua đường hô hấp, đặc biệt là Covid-19 thường do các Robot được lắp đặt tại các cổng kiểm soát thực hiện. Tuy nhiên trong thực tế, đối tượng có thể đeo kính hoặc đeo khẩu trang sai cách dẫn đến việc phát hiện khẩu trang bị sai sót. Do đó, đây là một thách thức lớn và mang tính cấp thiết, nhất là trong bối cảnh dịch Covid-19 đang có nguy cơ bùng phát trở lại do sự xuất hiện của các biến thể mới, sự chú quan trong việc tuân thủ các biện pháp phòng dịch, cũng như tốc độ tiêm chủng mũi bổ sung còn hạn chế ở một số khu vực. Ngoài ra, việc tổ chức các hoạt động cộng đồng và các sự kiện lớn mà không có các biện pháp giám sát đeo khẩu trang hiệu quả cũng tạo điều kiện cho sự lây lan của virus. Thêm vào đó, với sự gia tăng của các hoạt động giao thương, du lịch quốc tế, việc bảo đảm an toàn y tế và duy trì các biện pháp phòng ngừa hiệu quả càng trở nên quan trọng hơn bao giờ hết. Việc tổ chức và thực hiện các biện pháp ứng phó kịp thời, toàn diện không chỉ giúp kiểm soát dịch bệnh truyền nhiễm lây qua đường hô hấp mà còn giảm thiểu những ảnh hưởng tiêu cực đến đời sống kinh tế và xã hội.

Bài báo này đề xuất một phương pháp phát hiện khẩu trang bằng cách kết hợp mô hình Caffe và kiến trúc MobileNetV2. Cách tiếp cận này giúp phát hiện khẩu trang theo thời gian thực, ngay cả khi đối tượng đeo kính hoặc nón với độ chính xác cao và không quá nặng về tài nguyên để sử dụng trên các thiết bị nhúng có khả năng tính toán hạn chế như Raspberry Pi, NVIDIA Jetson Nano...

## 2. GIẢI THUẬT ĐỀ XUẤT

Để dự đoán một người có đeo khẩu trang hay không, giai đoạn đầu tiên sẽ là huấn luyện mô hình bằng cách sử dụng bộ dữ liệu phù hợp. Sau đó, cần có một mô hình nhận diện khuôn mặt chính xác để phát hiện khuôn mặt rồi mới phân loại xem người đó có đeo khẩu trang hay không như trong Hình 1.



Hình 1. Lưu đồ giải thuật đề xuất

## 2.1. Mô hình phát hiện khuôn mặt

Nhiệm vụ đặt ra trong bài báo này là nâng cao độ chính xác của việc phát hiện đeo khẩu trang đúng cách mà không đòi hỏi nhiều về tài nguyên phần cứng, đáp ứng được yêu cầu hoạt động theo thời gian thực. Do đó, nhóm tác giả đề xuất sử dụng Caffè [17], là một khung học sâu được phát triển như một giải pháp thay thế nhanh hơn, mạnh mẽ và hiệu quả hơn so với các phương pháp phát hiện đối tượng khác.

Nhóm tác giả đã sử dụng mô hình Caffè để triển khai mô hình phát hiện khuôn mặt bằng cách sử dụng cv2.dnn.readNet (“path/to/prototxtfile”, “path/to/caffemodelweights”). Sử dụng phương pháp này cho phép phát hiện các khuôn mặt theo thời gian thực ở các hướng khác nhau như trái, phải, trên, dưới với độ chính xác cao mà không cần sử dụng nhiều tài nguyên.

Kết quả nhận được là số lượng khuôn mặt được phát hiện, vị trí của các hộp giới hạn và điểm tin cậy trong các dự đoán đó. Những đầu ra này sau đó được sử dụng làm đầu vào cho bộ mô hình phát hiện khẩu trang.

## 2.2. Xây dựng bộ dữ liệu huấn luyện

Nhóm tác giả tiến hành thu thập dữ liệu với sự hỗ trợ của 1500 sinh viên Khoa Công nghệ Điện – Điện tử trường Đại học Công Thương Thành phố Hồ Chí Minh. Mỗi sinh viên chụp 20 tấm hình không đeo khẩu trang lần lượt ở các góc chính diện, lệch trái, lệch phải khuôn mặt như trong Hình 2. Tương tự, chụp 20 tấm hình đeo khẩu trang không đúng cách (đeo hở mũi hoặc hở cả mũi và miệng) như trong Hình 3. Cuối cùng chụp 40 tấm hình đeo khẩu trang đúng cách như trong Hình 4. Toàn bộ hình chụp có thể đeo kính hoặc không đeo kính, sử dụng các loại khẩu trang y tế hoặc khẩu trang vải với đủ loại màu sắc tùy ý.



Hình 2. Thu thập ảnh không đeo khẩu trang

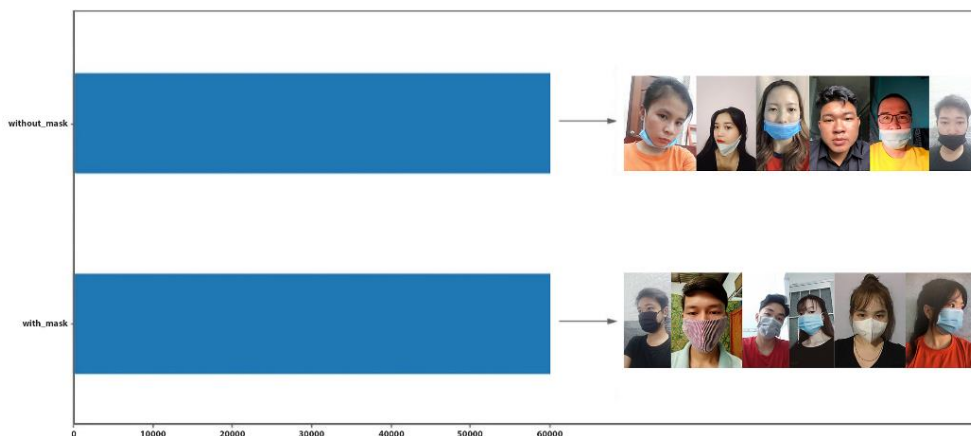


Hình 3. Thu thập ảnh đeo khẩu trang không đúng cách



Hình 4. Thu thập ảnh đeo khẩu trang đúng cách

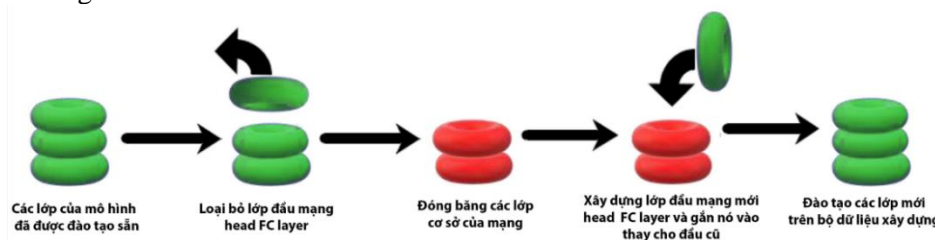
Do yêu cầu bài toán đặt ra là xây dựng bộ phát hiện khẩu trang áp dụng trong phòng chống dịch Covid-19 nên việc đeo khẩu trang sai cách cũng được tính thành không đeo khẩu trang. Cuối cùng, một bộ dữ liệu gồm 60000 ảnh đeo khẩu trang được gán nhãn “with\_mask” và nhãn “without\_mask” cũng được gán cho 60000 ảnh không đeo khẩu trang hoặc đeo khẩu trang sai cách để tạo thành một bộ dữ liệu cân bằng như trong Hình 5.



Hình 5. Bộ dữ liệu huấn luyện

### 2.3. Mô hình phân loại phát hiện khẩu trang

MobileNetV2 là mạng nơ-ron sâu đã được triển khai để giải quyết các vấn đề phân loại [9]. Ở đây, nhóm tác giả sử dụng kỹ thuật tinh chỉnh mô hình có sẵn để tránh các chi phí tính toán không cần thiết và tận dụng các trọng số sai lệch mà không làm mất đi các tính năng đã học như trong Hình 6.



Hình 6. Tinh chỉnh mô hình MobileNetV2

Đầu tiên, mạng MobileNetV2 với các trọng số do ImageNet đào tạo trước được tải về từ TensorFlow. Sau đó, lớp đầu mạng Fully-Connected được loại bỏ và các lớp cơ sở được đồng bằng để tránh làm suy yếu các tính năng đã học. Tiếp đó, nhóm tác giả xây dựng lớp đầu mạng Fully-Connected mới và gắn nó vào thay thế cho đầu cũ. Trọng số của các lớp cơ sở đã đồng bằng sẽ không được cập nhật trong quá trình lan truyền ngược, trong khi trọng số của lớp đầu sẽ được điều chỉnh. Cuối cùng, mô hình đã tinh chỉnh này được huấn luyện trên bộ dữ liệu tự xây dựng theo giải thuật sau:

Input: Các hình ảnh huấn luyện.

Output: mô hình phân loại phát hiện khẩu trang.

Bước 1: đưa ảnh vào, đọc giá trị kích thước pixel của ảnh.

Bước 2: tiền xử lý ảnh như thay đổi kích thước ảnh, chuẩn hóa ảnh và chuyển sang định dạng mảng.

Bước 3: nối ảnh được tiền xử lý trước đó và nhãn được liên kết vào danh sách dữ liệu và nhãn tương ứng.

Bước 4: áp dụng kỹ thuật tăng cường dữ liệu trên các ảnh để có độ chính xác cao hơn.

Bước 5: chia dữ liệu thành 2 tập train và test theo tỷ lệ 80% và 20% rồi đem huấn luyện với mô hình MobileNetV2 đã tinh chỉnh. Trình tối ưu hóa Adam được sử dụng để biên dịch toàn bộ mô hình.

Bước 6: lưu lại mô hình để sử dụng.

### 3. KẾT QUẢ VÀ THẢO LUẬN

#### 3.1. Đánh giá mô hình phân loại phát hiện khẩu trang sau khi huấn luyện

Ngôn ngữ lập trình Python 3.7.3 đã được chọn trong nghiên cứu này. Nhóm tác giả triển khai huấn luyện mô hình phân loại phát hiện khẩu trang đề xuất trên Laptop trang bị bộ xử lý Intel i7 - 8850H (4.1GHz), RAM 16 GB, sử dụng hệ điều hành Windows 10 Pro bản 64 bit. Kết quả huấn luyện mô hình phân loại khẩu trang đề xuất với bộ dữ liệu tự xây dựng được thể hiện trong Bảng 1, đạt độ chính xác 93%.

Bảng 1. Kết quả huấn luyện mô hình đề xuất với bộ dữ liệu tự xây dựng

	Precision	recall	F1 Score
with_mask	1.00	0.86	0.93
without_mask	0.87	1.00	0.93
accuracy			0.93
macro average	0.94	0.93	0.93
weighted average	0.94	0.93	0.93

Tiếp đó, nhóm tác giả thử huấn luyện mô hình đề xuất với bộ dữ liệu nhân tạo của Prajna Bhandary được cung cấp trên PyImageSearch. Bộ dữ liệu này gồm 1376 ảnh được chia thành 2 lớp là có đeo khẩu trang gồm 690 ảnh và không đeo khẩu trang là 686 ảnh. Kết quả huấn luyện được thể hiện trong Bảng 2, với độ chính xác đạt tới 99%.

Bảng 2. Kết quả huấn luyện mô hình đề xuất với bộ dữ liệu của Prajna Bhandary

	Precision	recall	F1 Score
with_mask	0.99	1.00	0.99
without_mask	1.00	0.99	0.99
accuracy			0.99
macro average	0.99	0.99	0.99
weighted average	0.99	0.99	0.99

Như vậy, có thể thấy mô hình đề xuất hoạt động rất tốt với bộ dữ liệu của Prajna Bhandary do bộ dữ liệu chỉ bao gồm 2 trường hợp đeo hoặc không đeo khẩu trang. Đối với bộ dữ liệu tự xây dựng của nhóm tác giả, do trường hợp đeo khẩu trang không đúng cách được gộp tính chung thành không đeo khẩu trang để đảm bảo an toàn phòng chống dịch nên gây ra khó khăn, dẫn đến nhiều dự đoán sai khác nhau trong quá trình huấn luyện. Tuy nhiên, với độ chính xác đạt được 93% thì mô hình đề xuất vẫn đáp ứng được yêu cầu đặt ra.

#### 3.2. Kết quả thực nghiệm

Mô hình được triển khai thực nghiệm trên Raspberry Pi4 4GB RAM. Các khuôn mặt tìm thấy trong khung hình sẽ được đóng khung bằng ô hình chữ nhật màu xanh lá cây với điểm số dự đoán nếu đeo khẩu trang đúng cách. Ngược lại, nếu không đeo khẩu trang hoặc đeo khẩu

trang sai cách thì khuôn mặt sẽ được đóng khung bằng ô hình chữ nhật màu đỏ. Video demo thực nghiệm giải thuật đề xuất được đăng tải tại <https://youtu.be/Ip4p60ir4qY>.



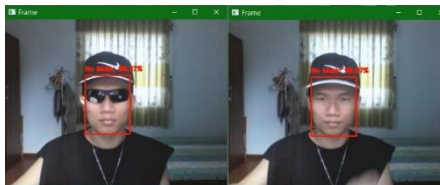
Hình 7. Đối tượng đeo kính, không đội nón và không mang khẩu trang

Hình 7 thể hiện kết quả thực nghiệm phương pháp đề xuất trong trường hợp đối tượng đeo kính, không đội nón và không mang khẩu trang với các góc mặt lần lượt từ trái qua phải, từ trên xuống dưới.



Hình 8. Đối tượng đeo kính và mang khẩu trang y tế

Hình 8 thể hiện kết quả thực nghiệm trong trường hợp đối tượng không đội nón, đeo kính và mang khẩu trang y tế với các góc mặt lần lượt từ trái qua phải, từ trên xuống dưới.



Hình 9. Đối tượng không mang khẩu trang, có đội nón và đeo kính

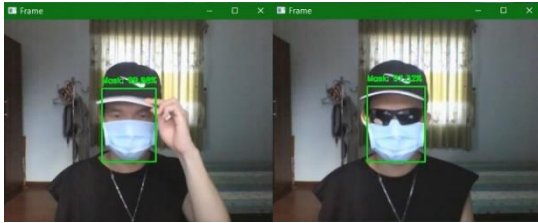


Hình 10. Đối tượng mang khẩu trang, có đội nón và đeo kính

Hình 9 và Hình 10 thể hiện thực nghiệm trong trường hợp đối tượng đội nón và đeo kính trong hoặc tối màu. Kết quả cho thấy phương pháp đề xuất vẫn phân loại phát hiện được chính xác đối tượng có mang khẩu trang hay không.

Hình 11 thể hiện kết quả thực nghiệm trong trường hợp đối tượng có đội nón và đeo kính, mang khẩu trang y tế màu xanh. Tương tự kết quả thực nghiệm đối với trường hợp đối tượng

thay khẩu trang y tế bằng khẩu trang vải đơn sắc hoặc khẩu trang vải có họa tiết phức tạp được thể hiện trong Hình 12.



Hình 11. Đối tượng mang khẩu trang y tế màu xanh, có đội nón và đeo kính



Hình 12. Đối tượng mang khẩu trang vải, có đội nón và đeo kính

Trường hợp đối tượng quên đeo khẩu trang nhưng dùng tay che miệng hoặc che cả mũi và miệng để đánh lừa hệ thống là có đeo khẩu trang cũng được thực nghiệm với kết quả phát hiện tốt như trong Hình 13.



Hình 13. Đối tượng dùng tay che mũi hoặc che mũi và miệng

Cuối cùng, nhóm tác giả tiến hành thực nghiệm tại xưởng thực hành Khoa Công nghệ Điện - Điện tử với sự hỗ trợ của 100 sinh viên. Kết quả thực nghiệm được thể hiện như trong Bảng 3.

Bảng 3. Kết quả thực nghiệm

Trường hợp	Số lần thực hiện	Số lần phát hiện đúng	Số lần phát hiện sai	Độ chính xác (%)
Không mang khẩu trang	100	100	0	100
Mang khẩu trang y tế đúng cách	100	99	1	99
Mang khẩu trang vải đúng cách	100	98	2	98
Đội nón, mang khẩu trang y tế đúng cách	100	96	4	96
Đội nón, mang khẩu trang vải đúng cách	100	93	7	93
Mang khẩu trang sai cách	100	79	21	79

Trong thực tế, do lo sợ dịch bệnh nên nếu có mang khẩu trang thì đối tượng sẽ đeo khẩu trang đúng cách để tự bảo vệ bản thân chứ ít khi nào đeo khẩu trang sai cách. Do đó, có thể nói phương pháp đề xuất đáp ứng được các yêu cầu đề ra, phát hiện được khẩu trang trong các trường hợp ra vào cổng kiểm soát như có đội nón hoặc không đội nón, có đeo kính hoặc không đeo kính mà không cần phải yêu cầu đối tượng tháo nón, cất kiếng gây phiền phức và tốn thời gian.

#### 4. KẾT LUẬN

Bài báo này đề xuất một phương pháp phân loại phát hiện khẩu trang theo thời gian thực và không đòi hỏi nhiều tài nguyên, dựa trên cơ sở kết hợp mô hình Caffee với kiến trúc MobileNetV2. Tính khả thi của phương pháp được thực nghiệm trên thiết bị nhúng Raspberry Pi với cơ sở dữ liệu tự xây dựng từ sinh viên Khoa Công nghệ Điện - Điện tử trường Đại học Công Thương Thành phố Hồ Chí Minh, đạt độ chính xác lên đến 93%.

Hướng phát triển của bài toán là xây dựng Robot phòng chống dịch bệnh truyền nhiễm lây qua đường hô hấp, ngoài chức năng phát hiện đeo khẩu trang sẽ phát triển thêm tính năng đo nhiệt độ cơ thể đối tượng, quét mã QR để khai báo y tế, phun cồn để khử khuẩn tay và phát khẩu trang trong trường hợp đối tượng không đeo khẩu trang. Ngoài ra, đối với trường hợp đối tượng không đeo khẩu trang hoặc có dấu hiệu cảm sốt sẽ chụp hình lại lưu về server để làm cơ sở truy vết khi cần thiết.

Cuối cùng, nếu có thể thu thập dữ liệu khuôn mặt không đeo và khuôn mặt đeo khẩu trang của sinh viên toàn trường, nhóm tác giả hướng tới phát triển hệ thống định danh sinh viên ngay cả khi đeo khẩu trang, có thể áp dụng điểm danh sinh viên toàn trường khi qua cổng kiểm soát hoặc điểm danh ngay trong phòng học.

**Lời cảm ơn:** Nghiên cứu này do Trường Đại học Công nghiệp Thực phẩm TP. Hồ Chí Minh (nay là Trường Đại học Công Thương Thành phố Hồ Chí Minh) bảo trợ và cấp kinh phí theo Hợp đồng số 28/HĐ-DCT ngày 01 tháng 03 năm 2022.

#### TÀI LIỆU THAM KHẢO

1. Trọng Nhân - Bùng nổ Robot chống dịch COVID-19, 2020. Truy cập ngày 01/02/2023 từ <https://tuoitre.vn/bung-no-robot-chong-dich-covid-19-20200420090125689.htm>.
2. Lưu Quý - Dùng robot để phát hiện người không đeo khẩu trang, 2020. Truy cập ngày 01/02/2023 từ <https://vnexpress.net/dung-robot-de-phat-hien-nguoi-khong-deo-khau-trang-4099618.html>.
3. Ge S., Li J., Ye Q., Luo Z. - Detecting masked faces in the wild with LLE-CNNs, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 21-26 July 2017, 2682-2690. <https://doi.org/10.1109/CVPR.2017.53>
4. Chavda A., Dsouza J., Badgujar S., Damani A. - Multi-stage CNN architecture for face mask detection, In 2021 6th International Conference for Convergence in Technology (2021) 1-8. <https://doi.org/10.48550/arXiv.2009.07627>
5. Yadav S. - Deep learning based safe social distancing and face mask detection in public areas for COVID-19 safety guidelines adherence, Int. J. Res. Appl. Sci. Eng. Technol, **8** (7) 1368-1375. <https://doi.org/10.22214/ijraset.2020.30560>
6. Qin B., Li D. - Identifying Facemask-Wearing Condition Using Image Super-Resolution with Classification Network to Prevent COVID-19, Sensors **20** (18) (2020) pp. 5236. <https://doi.org/10.3390/s20185236>
7. Simonyan K., Zisserman A. - Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv: 1409.1556 (2014). <https://doi.org/10.48550/arXiv.1409.1556>
8. Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. - Going deeper with convolutions, In Proceedings of the IEEE conference on computer vision and pattern recognition (2015) 1-9. <https://doi.org/10.48550/arXiv.1409.4842>
9. He K., Zhang X., Ren S., Sun J. - Deep residual learning for image recognition, In Proceedings of the IEEE conference on computer vision and pattern recognition (2016) 770-778. <https://doi.org/10.1109/CVPR.2016.90>



10. Howard A. G., Zhu M., Chen B., Kalenichenko D., Wang W., Weyand T., Andreetto M., Adam H. - Mobilenets: Efficient convolutional neural networks for mobile vision applications, arXiv preprint arXiv:1704.04861 (2017).  
<https://doi.org/10.48550/arXiv.1704.04861>
11. Sandler M., Howard A., Zhu M., Zhmoginov A., Chen L. C. - Mobilenetv2: Inverted residuals and linear bottlenecks, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2018) 4510-4520.  
<https://doi.org/10.1109/CVPR.2018.00474>
12. Khandelwal P., Khandelwal A., Agarwal S., Thomas D., Xavier N., Raghuraman A. - Using computer vision to enhance safety of workforce in manufacturing in a post covid world, arXiv preprint arXiv:2005.05287 (2020). <https://doi.org/10.48550/arXiv.2005.05287>
13. Jagadeeswari C., Theja M.U. - Performance evaluation of intelligent face mask detection system with various deep learning classifiers, International Journal of Advanced Science and Technology (2020) 3074-3082.
14. Li C., Wang R., Li J., Fei L. - Face detection based on YOLOv3, In: Jain, V., Patnaik, S., Popențiu Vlădicescu, F., Sethi, I. (eds) Recent Trends in Intelligent Computing, Communication and Devices. Advances in Intelligent Systems and Computing **1006**. Springer, Singapore (2020) 277-284. [http://dx.doi.org/10.1007/978-981-13-9406-5\\_34](http://dx.doi.org/10.1007/978-981-13-9406-5_34)
15. Lin T. Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. - Feature pyramid networks for object detection, In Proceedings of the IEEE conference on computer vision and pattern recognition (2017) 2117-2125. <https://doi.org/10.1109/CVPR.2017.106>
16. Jiang M., Fan X., Yan H. - Retinamask: A face mask detector, arXiv preprint arXiv:2005.03950 (2020). <https://doi.org/10.48550/arXiv.2005.03950>
17. Jia Y., Shelhamer E., Donahue J., Karayev S., Long J., Girshick R., Darrell T. - Caffe: Convolutional architecture for fast feature embedding, In Proceedings of the 22<sup>nd</sup> ACM international conference on Multimedia (2014) 675-678.  
<https://doi.org/10.1145/2647868.2654889>.

## **ABSTRACT**

### **STUDY ON FACE MASK DETECTION ALGORITHM FOR ROBOTS IN DISEASE PREVENTION**

Tran Hoan\*, Hoang Dac Huy

*Ho Chi Minh City University of Industry and Trade*

\*Email: [hoant@huit.edu.vn](mailto:hoant@huit.edu.vn)

The mask detection feature has significantly advanced in image processing and computer vision, making a crucial contribution to combating the COVID-19 pandemic. This paper presents a method utilizing the Caffe model in combination with the MobileNetV2 architecture, designed to apply to embedded devices with limited computational capabilities such as Raspberry Pi, NVIDIA Jetson Nano, etc. This method enables real-time mask detection even when individuals are wearing hats or glasses. Experimental results were obtained using a custom-built dataset from students of the Faculty of Electrical and Electronics Technology at the Ho Chi Minh City University of Industry and Trade, achieving an accuracy rate of up to 93%.

*Keywords:* Face mask, detection, Caffe, MobileNetV2, Raspberry Pi, Covid-19.