

NHẬN DIỆN KHUÔN MẶT KHI CÓ/KHÔNG ĐEO KHẨU TRANG

Ngô Dương Hà, Nguyễn Nhật Trường, Trần Như Ý*

Trường Đại học Công nghiệp Thực phẩm TP.HCM

*Email: ytn@cntp.edu.vn

Ngày nhận bài: 10/6/2022; Ngày chấp nhận đăng: 13/7/2022

TÓM TẮT

Nhận diện khuôn mặt là một lĩnh vực nghiên cứu của ngành thị giác máy tính. Hệ thống nhận diện khuôn mặt có thể nhận diện được khuôn mặt có đeo hoặc không đeo khẩu trang. Trong bài báo này, chúng tôi áp dụng kỹ thuật MTCNN giúp nhận diện và căn chỉnh khuôn mặt. Kỹ thuật ArcFace để nhận diện những đặc điểm trên khuôn mặt và trích xuất đặc trưng khuôn mặt. Ngoài ra, chúng tôi đề xuất thêm MobileNet cho phép rút gọn lại vài triệu tham số nhưng vẫn giữ được độ chính xác ổn định kết hợp cùng với ResNet giúp tăng độ chính xác khi nhận dạng khuôn mặt có/không đeo khẩu trang. Kết quả bài toán đã được thực nghiệm trên cơ sở dữ liệu RMFRD cho kết quả với độ chính xác trên 72%.

Từ khóa: MTCNN, ResNet, MobileNet, ArcFace.

1. GIỚI THIỆU

Nhận diện khuôn mặt được áp dụng trong nhiều lĩnh vực thiết yếu [1, 2] như y tế, giáo dục, công ty, khu dân cư, sân bay,... Cụ thể, việc điểm danh sinh viên ở các trường đại học bằng cách quét qua khuôn mặt sẽ rút ngắn thời gian hơn cách gọi tên truyền thống. Việc quét qua khuôn mặt sẽ giúp hệ thống nhận diện được mặt sinh viên đó và truy cập vào kho dữ liệu để xem thử sinh viên có trong danh sách ở phòng học đó không và thực hiện điểm danh. Từ đó, nhà trường có thể giám sát được sĩ số cũng như thời gian ra vào lớp. Hoặc những biển quảng cáo thông minh trong các sân bay hiện có thể xác định giới tính, dân tộc và tuổi xấp xỉ của người qua đường, sau đó nhắm mục tiêu quảng cáo đến đối tượng đó.

Phương pháp Cascade Face tìm kiếm khuôn mặt được đề xuất bởi Viola và Jones sử dụng tính năng Haar-Like và AdaBoost để huấn luyện phân loại đạt được hiệu suất tốt với hiệu quả thời gian thực [3]. Tuy nhiên, có nhiều công trình chỉ ra phương pháp này có thể suy giảm đáng kể trong các ứng dụng thế giới thực với các biến thể hình ảnh khuôn mặt người lớn hơn [4-6]. Bên cạnh đó, theo tác giả Mathias giới thiệu mô hình Deformable Part đào tạo phân biệt các thông tin tiềm ẩn để nhận diện khuôn mặt và đạt được hiệu suất đáng kể. Tuy nhiên, chúng tốn nhiều chi phí trong giai đoạn đào tạo về mặt tính toán [7-9]. Gần đây, mạng CNN đạt được những tiến bộ đáng kể trong thị giác máy tính, chẳng hạn như phân loại hình ảnh và nhận dạng khuôn mặt [10, 11]. Yang và cộng sự sử dụng mạng CNN chuyên sâu về nhận dạng thuộc tính ở các vùng khuôn mặt, điều này giúp mang lại nhiều các ứng cử viên. Tuy nhiên, do cấu trúc CNN phức tạp nên cách tiếp cận này tốn kém thời gian trong thực tế [12].

Căn chỉnh khuôn mặt thu hút sự quan tâm nghiên cứu rộng rãi. Các công trình nghiên cứu trong lĩnh vực này có thể được chia thành hai loại gồm phương pháp dựa trên hồi quy và phương pháp tiếp cận phù hợp với khuôn mẫu [13-17]. Gần đây, Zhang và cộng sự đề xuất sử dụng nhận dạng thuộc tính khuôn mặt như một nhiệm vụ để nâng cao hiệu suất căn chỉnh khuôn mặt bằng cách sử dụng CNN học sâu [18].

Nhận dạng khuôn mặt với sự kết hợp của khuôn mặt được che và không được che [19, 20]. Theo tác giả Anwar và cộng sự đề xuất kết hợp tập dữ liệu VGG2 với các khuôn mặt che tăng cường và đào tạo mô hình được mô tả trong FaceNet [21, 22]. Bằng cách này, mô hình học cách phân biệt khi nào một khuôn mặt đang đeo mặt nạ và các đặc điểm của nửa trên khuôn mặt, nhưng vẫn trích xuất thông tin từ toàn bộ khuôn mặt [19]. Mặt khác, Geng và cộng sự xác định hai đặc trưng trung tâm cho mỗi danh tính tương ứng với hình ảnh khuôn mặt đầy đủ và hình ảnh khuôn mặt được che mặt tương ứng. Tác giả sử dụng giới hạn miền để buộc tính năng của khuôn mặt được che đi gần hơn với trung tâm khuôn mặt tương ứng và ngược lại [20].

Việc phát hiện và căn chỉnh khuôn mặt trong môi trường không hoặc ít bị hạn chế là một thách thức do các tư thế, độ sáng khác nhau. Bài toán nhận diện khẩu trang là bài toán không quá khó khi áp dụng các phương pháp học sâu để giải quyết vấn đề. Tuy nhiên, có nhiều yếu tố dẫn đến việc sai lệch trong nhận diện khuôn mặt đeo khẩu trang như nhận diện một cặp song sinh hay khi họ đeo khẩu trang sẽ gặp khó khăn hơn trong việc nhận diện. Ngoài ra, các yếu tố như độ sáng, góc chụp hay phụ kiện cũng phần nào làm giảm độ chính xác của mô hình.

Bài báo đề xuất kết hợp kỹ thuật MobileNet, ResNet và ArcFace giúp rút ngắn thời gian nhận diện mặt người và mặt người có đeo khẩu trang. Phần còn lại của bài viết gồm phần 2 trình bày cơ sở lý thuyết thuật toán, phần 3 trình bày phương pháp thực hiện, phần 4 sẽ đưa ra kết quả thực nghiệm và phần 5 kết luận.

2. CƠ SỞ LÝ THUYẾT

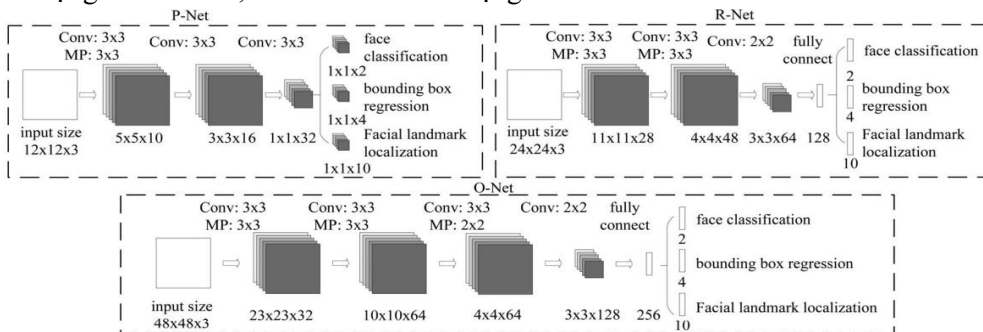
2.1. Kỹ thuật MTCNN

MTCNN được công bố năm 2016 bởi Zhang và cộng sự [23]. Một mạng nơ-ron có thể thực hiện nhận diện khuôn mặt và căn chỉnh khuôn mặt trên hình ảnh bằng mô hình MTCNN bao gồm 3 mạng riêng biệt như hình 1. Cụ thể,

P-Net (Proposal network) tìm tọa độ của các bounding boxes với độ tin cậy, tiếp đó xóa các bounding boxes có độ tin cậy thấp. Tuy nhiên, vẫn còn nhiều bounding boxes chồng lên nhau nên việc sử dụng non-maximum suppression sẽ giúp giải quyết vấn đề này.

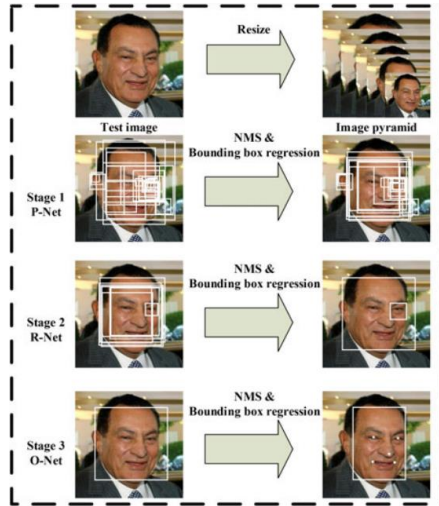
R-Net (Refine network) lấy các bounding boxes được cung cấp từ mạng P-Net, sau đó tinh chỉnh tọa độ để loại bỏ các vùng không phải khuôn mặt.

O-Net (Output network) lấy đầu ra của R-Net làm đầu vào và dùng để xác định tọa độ khuôn mặt gồm: hai mắt, mũi và hai khóe miệng.



Hình 1. Kiến trúc các mạng của MTCNN [23]

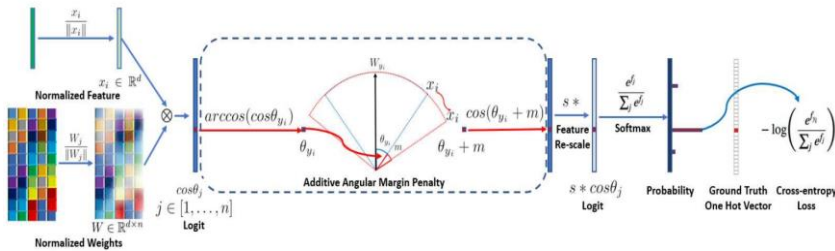
Với mỗi bức ảnh đầu vào, chúng sẽ tạo ra nhiều bản sao của hình ảnh đó với các kích thước khác nhau gọi là kim tự tháp ảnh [24]. Đầu tiên, các cửa sổ ứng viên được tạo thông qua một P-Net. Sau đó, lọc những ứng viên này trong giai đoạn tiếp theo thông qua một R-Net. Trong giai đoạn cuối, O-Net tạo ra hộp giới hạn cuối cùng và vị trí các điểm đặc trưng của khuôn mặt như Hình 2.



Hình 2. Ba giai đoạn của MTCNN [23]

2.2. Phương pháp ArcFace

ArcFace cho phép nhận dạng những đặc điểm sâu, có khả năng phân biệt cao và cho kết quả tối ưu với độ chính xác cao trong thử thách MegaFace Challenge [25]. MegaFace Challenge là một cộng đồng lớn với một triệu khuôn mặt được nhận dạng [26].

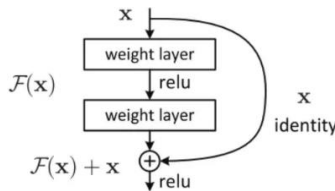


Hình 3. Xây dựng DCNN để nhận dạng khuôn mặt được giám sát bởi hàm lỗi trong mô hình ArcFace [25].

Theo Hình 3 mạng DCNN sử dụng tích chập giữa đặc trưng và trọng lượng bằng khoảng cách cosin. Nghĩa là sử dụng hàm arc-cosine để tính góc giữa đặc trưng và trọng lượng mục tiêu. Sau đó, thêm một hệ số góc phụ vào góc mục tiêu với hàm cosine.

2.3. Phương pháp ResNet

Mạng ResNet được giới thiệu vào năm 2015 đã giành được vị trí thứ 1 trong cuộc thi ILSVRC 2015 với tỷ lệ lỗi top 5 chỉ 3.57. ResNet là một mạng CNN hoàn chỉnh để giải quyết hiện tượng Vanishing Gradients [27]. Nghĩa là hiện tượng làm thay đổi nhiều weights của các tầng dẫn đến chúng không thể hội tụ và mạng sẽ không thu được kết quả tốt. Giải pháp mà ResNet đưa ra là sử dụng kết nối "tắt" đồng nhất để xuyên qua một hay nhiều lớp như trong Hình 4.



Hình 4. Cấu trúc kết nối tắt của mạng ResNet [27].

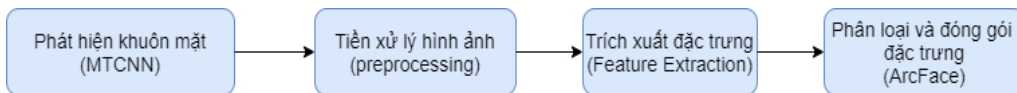
2.4. Phương pháp MobileNet

MobileNet được giới thiệu năm 2017, MobileNet có thể rút gọn lại vài triệu tham số nhưng vẫn giữ được độ chính xác ở mức ổn nhờ sử dụng một cơ chế gọi là Depthwise Separable [28]. Nghĩa là để giảm số lượng tính toán, giảm số lượng tham số, đồng thời có thể thực hiện trích xuất đặc trưng một cách tách biệt trên các kênh khác nhau.

3. XÂY DỰNG THỰC NGHIỆM

3.1. Bài toán nhận diện khuôn mặt không đeo khẩu trang

Hệ thống nhận dạng khuôn mặt không đeo khẩu trang kèm nhãn với việc kết hợp thuật toán MTCNN và ArcFace theo Hình 5.



Hình 5. Quá trình học mặt người kèm nhãn.

Giải thuật 1 (nomask):

Input: Tập hình ảnh người đã được dán nhãn (mỗi ảnh chỉ xuất hiện 1 mặt người).

Output: Danh sách đặc trưng đã được đóng gói theo nhãn của mỗi người trong tập dữ liệu.

Bước 1: Phát hiện khuôn mặt (MTCNN), tiền xử lý ảnh

Bước 1.1: Input tập hình ảnh (image)

Bước 1.2:

for image in images:

 bbox, landmark = MTCNN (image)

 aligned = preprocessing (bbox, landmark, image_size = "112,112")

 list_aligned.append(aligned)

return list_aligned (tập ảnh khuôn mặt đã căn chỉnh theo tỷ lệ 112x112)

Bước 1.3: Tạo dataset với tập ảnh vừa tạo ở bước 1.2.

Bước 1.4: Huấn luyện model ArcFace bằng dataset vừa tạo ở bước 1.3.

Bước 2: Trích xuất đặc trưng, phân loại và đóng gói với ArcFace

Bước 2.1: Input tập hình ảnh (image) có dán nhãn

Bước 2.2:

for folder(nhãn) in folders:

 images: danh sách hình ảnh trong mỗi folder

 for image in images:

 bbox, landmark = MTCNN (image)

 aligned = preprocessing bbox, landmark, image_size = "112,112")

 feature = modelArcFace.get_feature(aligned)

 list_feature.append(feature)

 list_label.append(folder)

return list_feature, list_label

Bước 2.3: Đóng gói danh sách đặc trưng (list_feature) và danh sách nhãn (list_label).

3.2. Bài toán nhận diện khuôn mặt đeo khẩu trang

Hệ thống nhận dạng khuôn mặt đeo khẩu trang kèm nhãn với kỹ thuật Resnet.

Giải thuật 2 (mask):

Input: Tập hình ảnh người đã được dán nhãn (mỗi ảnh chỉ xuất hiện 1 mặt người).

Output: Danh sách đặc trưng đã được đóng gói theo nhãn của mỗi người trong tập dữ liệu.

Bước 1: Phát hiện khuôn mặt (SSD) [29], tiền xử lý ảnh

Bước 1.1: Input tập hình ảnh (images).

Bước 1.2:

```
for image in images:
    image = preprocessing (bbox, landmark, image_size = "224,224")
    bbox, landmark = SSD (image)
    list_image.append(image)
return list_image (tập ảnh khuôn mặt đã căn chỉnh theo tỉ lệ 224x224)
```

Bước 2: Trích xuất đặc trưng, phân loại

Bước 2.1: Input tập hình ảnh (dataset) có dán nhãn.

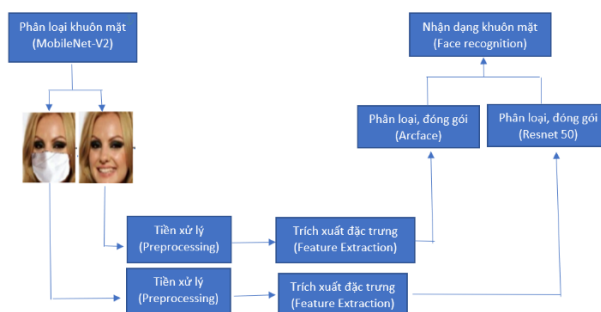
Bước 2.2:

```
for folder(nhãn) in dataset:
    images: danh sách hình ảnh trong mỗi folder
    for image in images:
        image = preprocessing (image_size = "224,224")
        bbox, pred = SSD (image)
        feature = modelResNet.get_feature(bbox)
        list_feature.append(feature)
        list_label.append(folder)
return list_feature, list_label
```

Bước 2.3: Đóng gói danh sách đặc trưng(list_feature) và danh sách nhãn (list_label).

3.3. Bài toán nhận diện khuôn mặt có/không đeo khẩu trang

Hệ thống nhận dạng khuôn mặt khi có/không đeo khẩu trang được trình bày trong Hình 6. Cụ thể, Hình ảnh đi qua mô hình MobileNet để cắt và phát hiện khuôn mặt có/không đeo khẩu trang. Nếu ảnh không có khẩu trang thì tiến hành tiền xử lý, trích đặc trưng và phân loại qua mô hình ArcFace. Ngược lại nếu là ảnh có đeo khẩu trang thì tiền xử lý, trích đặc trưng phân loại bằng mô hình ResNet để nhận dạng khuôn mặt đó là ai.



Hình 6. Quá trình các bước thực hiện bài toán nhận dạng khuôn mặt khi có/không đeo khẩu trang.

Giải thuật 3: MobileNet nhận diện khuôn mặt có/không đeo khẩu trang.

Input: Hình ảnh cần nhận diện khuôn mặt có/không đeo khẩu trang.

Output: Khuôn mặt đeo khẩu trang hoặc không đeo khẩu trang và tỷ lệ phân loại.

Bước 1: Với hình ảnh đưa vào

Bước 1.1: Tiền xử lý ảnh

```
image = cv2.cvtColor(face, cv2.COLOR_BGR2RGB)
```

```
image = cv2.resize(image, (224, 224))
```

```
image = img_to_array(image)
```

```
image = preprocess_input(image)
```

Bước 1.2: Dùng phương pháp SSD [29] cho phép phát hiện và trả về danh sách khung hình phát hiện khuôn mặt được xử lý.

```
faces.append(face)
```

```
locs.append((startX, startY, endX, endY))
```

Bước 1.3: Dùng model đã được huấn luyện để tiến hành predict từng khung hình phát hiện được, trả về danh sách vị trí khung hình và tỷ lệ (khẩu trang, không khẩu trang).

```
faces = np.array(faces, dtype="float32")
```

```
preds = maskNet.predict(faces, batch_size=32)
```

```
return (locs, preds)
```

Bước 2: Hiện thị box tương ứng vừa tìm được lên ảnh với khung xác định vị trí khuôn mặt và tỷ lệ xác định phân loại của khung hình đó.

Giải thuật 4 (MobileNet + ArcFace + ResNet): Nhận dạng khuôn mặt người khi có/không đeo khẩu trang.

Input: Hình ảnh khuôn mặt đeo có/không đeo khẩu trang.

Output: Nhận dạng được người hoặc người mà hệ thống chưa biết.

Bước 1: Dùng kết quả của giải thuật 3 để phân loại khuôn mặt, kết quả trả về danh sách gồm tọa độ các khuôn mặt và tỷ lệ phân loại mask/nomask của các khuôn mặt đó.

Bước 2: Kiểm tra từng khuôn mặt có trong ảnh

Bước 2.1: Nếu kết quả trả về là khuôn mặt không đeo khẩu trang: Sử dụng ArcFace xác định danh tính khuôn mặt bằng giải thuật 1.

Bước 2.2: Ngược lại, nếu kết quả trả về là khuôn mặt đeo khẩu trang: Sử dụng ResNet để tiến hành xác định danh tính khuôn mặt bằng giải thuật 2.

4. KẾT QUẢ THỰC NGHIỆM

Nhóm tác giả thực nghiệm trên máy tính sử dụng hệ điều hành Windows 10 Pro bản 64 bit, RAM 8 GB, Chip Intel Core (TM) 5i-3210M CPU @ 2.5GHz. Ngôn ngữ lập trình Python phiên bản 3.9.

Bộ dữ liệu được sử dụng trong thực nghiệm là Real World Fasked Face Recognition Dataset (RMFRD) [30]. Bộ dữ liệu gốc gồm 5000 khuôn mặt đeo khẩu trang và 90000 khuôn mặt không đeo khẩu trang của 525 người.

Để chạy thực nghiệm 2 bộ dữ liệu trên, chúng tôi đề xuất chia bộ dữ liệu thành 2 tập train và test theo tỷ lệ 8/2: 80% để train và 20% test.

Bảng 1. Kết quả thực nghiệm độ chính xác

Kỹ thuật	RMFRD
ArcFace	61,5%
MobileNet + ArcFace + ResNet	72,2%

Nhận xét: Dựa vào bảng 1 cho thấy kỹ thuật kết hợp MobileNet + ArcFace + ResNet cho kết quả tốt hơn kỹ thuật ArcFace là 10,7% trong việc nhận dạng khuôn mặt có/không đeo khẩu trang với tập dữ liệu RMFRD.

5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Việc thêm kỹ thuật mạng MobileNet để phân loại khuôn mặt đeo/không đeo khẩu trang kết hợp mạng ResNet để nhận diện người đeo khẩu trang vào mô hình ArcFace giúp nhận diện khuôn mặt đeo/không đeo khẩu trang một cách ổn định và chính xác.

Ngoài việc kết hợp MobileNet, ArcFace và ResNet để nhận diện khuôn mặt đeo/không đeo khẩu trang. Trong thời gian tới, Chúng tôi đề xuất việc kết hợp các kỹ thuật khác với ArcFace và đánh giá so sánh với kết quả hiện tại.

Hướng phát triển của bài toán từ những đặc trưng của gương mặt có thể truy xuất được giới tính, chủng tộc, độ tuổi,... Ngoài ra, nhận diện khẩu trang có thể cải tiến các thuật toán để nhận diện được các cặp song sinh, sinh ba hay người có khuôn mặt giống nhau.

Với sự phát triển của kỹ thuật hiện nay, việc thu thập được nhiều dữ liệu hơn từ khuôn mặt và khuôn mặt đeo khẩu trang là khả thi trong tương lai không xa. Ngoài ra, chúng tôi xây dựng dữ liệu khuôn mặt sinh viên của Khoa CNTT trường Đại học Công nghiệp Thực phẩm TP.HCM cho phép xây dựng ứng dụng hỗ trợ điểm danh sinh viên trong phòng học có sức chứa trên 200 sinh viên.

TÀI LIỆU THAM KHẢO

1. Silverstein E., Snyder M. - Implementation of facial recognition with Microsoft Kinect v2 sensor for patient verification, *Medical Physics* **44** (6) (2017) 2391-2399. <https://doi.org/10.1002/mp.12241>.
2. Zhao Q., Ye M. - The application and implementation of face recognition in authentication system for distance education, *International Conference on Networking and Digital Society* (2010). <https://doi.org/10.1109/ICNDS.2010.5479246>.
3. Viola P., Jones M.J. - Robust real-time face detection, *International Journal of Computer Vision* **57** (2) (2004) 137-154.
4. Yang B., Yan J., Lei Z., Li S.Z. - Aggregate channel features for multi-view face detection, *IEEE International Joint Conference on Biometrics* (2014) 1-8. <https://doi.org/10.1109/BTAS.2014.6996284>.
5. Pham M., Gao Y., Hoang V.D., Cham T.- Fast polygonal integration and its application in extending Haar-like features to improve object detection, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2010) 942-949. <https://doi.org/10.1109/CVPR.2010.5540117>.

6. Zhu Q., Yeh M.C., Cheng K.T., Avidan S. - Fast human detection using a cascade of histograms of oriented gradients, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2006) 1491-1498. <https://doi.org/10.1109/CVPR.2006.119>.
7. Mathias M., Benenson R., Pedersoli M., Gool L.V.- Face detection without bells and whistles, European Conference on Computer Vision **8692** (2014) 720-735. https://doi.org/10.1007/978-3-319-10593-2_47.
8. Yan J., Lei Z., Wen L., Li S.Z.- The fastest deformable part model for object detection, the IEEE Conference on Computer Vision and Pattern Recognition (2014) 2497-2504.
9. Zhu X., Ramanan D. - Face detection, pose estimation, and landmark localization in the wild, the IEEE Conference on Computer Vision and Pattern Recognition (2012) 2879-2886. <https://doi.org/10.1109/CVPR.2012.6248014>.
10. Krizhevsky A., Sutskever I., Hinton G.E. - ImageNet classification with deep convolutional neural networks, Advances in Neural Information Processing Systems (2012) 1097-1105.
11. Sun Y., Chen Y., Wang X., Tang X.- Deep learning face representation by joint identification-verification, Advances in Neural Information Processing Systems (2014) 1988-1996.
12. Yang S., Luo P., Loy C.C., Tang X. - From facial parts responses to face detection: A deep learning approach, the IEEE International Conference on Computer Vision (2015) 3676-3684.
13. Artizzu X.P.B., Perona P., Dollar P. - Robust face landmark estimation under occlusion, the IEEE International Conference on Computer Vision (2013) 1513-1520.
14. Cao X., Wei Y., Wen F., Sun J. - Face alignment by explicit shape regression, International Journal of Computer Vision **107** (2014) 177-190. <https://doi.org/10.1007/s11263-013-0667-3>.
15. Zhang J., Shan S., Kan M., Chen X. - Coarse-to-fine auto-encoder networks (CFAN) for real-time face alignment, Computer Vision – ECCV'2014 **8690** (2014) 1-16. https://doi.org/10.1007/978-3-319-10605-2_1.
16. Cootes T.F., Edwards G.J., Taylor C.J. - Active appearance models, IEEE Transactions on Pattern Analysis and Machine Intelligence **23** (6) (2001) 681-685. <https://doi.org/10.1109/34.927467>.
17. Yu X., Huang J., Zhang S., Yan W., Metaxas D.N.- Pose-free facial landmark fitting via optimized part mixtures and cascaded deformable shape model, the IEEE International Conference on Computer Vision (2013) 1944-1951.
18. Zhang Z., Luo P., Loy C.C., Tang X. - Facial landmark detection by deep multi-task learning, Computer Vision – ECCV'2014 **8694** (2014) 94-108.
19. Anwar A., Raychowdhury A. - Masked face recognition for secure authentication, arXiv:2008.11104 (2020). <https://doi.org/10.48550/arXiv.2008.11104>.
20. Geng M., Peng P., Huang Y., Tian Y. - Masked face recognition with generative data augmentation and domain constrained ranking, the 28th ACM International Conference on Multimedia (2020) 2246-2254. <https://doi.org/10.1145/3394171.3413723>.
21. Cao Q., Shen L., Xie W., Parkhi O. M., Zisserman A. - Vggface2: A dataset for recognising faces across pose and age, the 13th IEEE International Conference on

- Automatic Face & Gesture Recognition (2018) 67-74.
<https://doi.org/10.1109/FG.2018.00020>.
22. Schroff F., Kalenichenko D., Philbin J. - Facenet: A unified embedding for face recognition and clustering, the IEEE Conference on Computer Vision and Pattern Recognition (2015) 815-823.
 23. Zhang K., Zhang Z., Li Z., Qiao Y. - Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks, IEEE Signal Processing Letters **23** (10) (2016) 1499-1503. <https://doi.org/10.1109/LSP.2016.2603342>.
 24. Adelson E. H., Anderson C. H., Bergen J. R., Burt P. J., Ogden J. M. - Pyramid methods in image processing, RCA Engineer (1984).
 25. Deng J., Guo J., Xue N., Zafeiriou S. - ArcFace: Additive Angular Margin Loss for Deep Face Recognition, the IEEE Conference on Computer Vision and Pattern Recognition (2019) 4690-4699.
 26. Shlizerman I.K., Seitz S.M., Miller D., Brossard E. - The MegaFace Benchmark: 1 Million Faces for Recognition at Scale, the IEEE Conference on Computer Vision and Pattern Recognition (2016) 4873-4882.
 27. Wen L., Li X., Gao L. - A transfer convolutional neural network for fault diagnosis based on ResNet-50, Neural Comput & Applic **32** (2020) 6111-6124. <https://doi.org/10.1007/s00521-019-04097-w>.
 28. Yu J., Yin T., Li S., Hong S., Peng Y. - Fast Ship Detection in Optical Remote Sensing Images Based on Sparse MobileNetV2 Network, Genetic and Evolutionary Computing **1107** (2020) 262-269. https://doi.org/10.1007/978-981-15-3308-2_30.
 29. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.Y., Berg A.C. - SSD: Single Shot MultiBox Detector, Computer Vision – ECCV'2016 **9905** (2016). https://doi.org/10.1007/978-3-319-46448-0_2.
 30. Wang Z., Wang G., Huang B., Xiong Z., Hong Q., Wu H., Yi P., Jiang K., Wang N., Pei Y., Chen H., Miao Y., Huang Z., Liang J. - Masked Face Recognition Dataset and Application, arXiv:2003.09093 (2020). <https://doi.org/10.48550/arXiv.2003.09093>.

ABSTRACT

FACE RECOGNITION WITH/WITHOUT MASK

Ngo Duong Ha, Nguyen Nhat Trung, Tran Nhu Y*
Ho Chi Minh City University of Food Industry
*Email: ytn@cntp.edu.vn

Face recognition is a research area of computer vision. The facial recognition system can recognize faces, we propose to combine recognition of both mask-wearing and non-masking people. In this paper, we present the MTCNN technique to help identify and align faces. ArcFace technique to recognize facial features and extracts facial features. In addition, we propose to MobileNet which can reduce several million parameters but still maintain stable accuracy combined with ResNet to increase accuracy when recognizing faces with/without wearing a mask. The results of the problem have been tested on the RMFRD database, giving results with an accuracy of over 72%.

Keywords: MTCNN, ResNet, MobileNet, ArcFace.