Deep Learning Model Based on Mobile-Net with Haar-like Algorithm for Masked Face Recognition at Nuclear Facilities

Nadia.Nawwar, Hani.Kasban, May salama

Abstract: During the spread of the COVID-19 pandemic in early 2020, the WHO organization advised all people in the world to wear face-mask to limit the spread of COVID-19. Many facilities required that their employees wear face-mask. For the safety of the facility, it was mandatory to recognize the identity of the individual wearing the mask. Hence, face recognition of the masked individuals was required. In this research, a novel technique is proposed based on a mobile-net and Haar-like algorithm for detecting and recognizing the masked face. Firstly, recognize the authorized person that enters the nuclear facility in case of wearing the masked-face using mobile-net. Secondly, applying Haar-like features to detect the retina of the person to extract the boundary box around the retina compares this with the dataset of the person without the mask for recognition. The results of the proposed model, which was tested on a dataset from Kaggle, yielded 0.99 accuracies, a loss of 0.08, F1-score 0.98.

Keywords: COVID-19, Deep learning, Mobile-Net, and Haar-Like.

I. INTRODUCTION

In the computer vision field and due to COVID-19, masked-face recognition has a great interest of most researchers[1]. Therefore, the individuals in the authority must wear a mask to limit the spread of COVID-19. The tools to identify features of masked-face are ineffective in numerous circumstances, like Face access control at the authority’s gate attendance. [2] Consequently, it is very critical to develop an efficient frame-work recognition on the masked faces. The effort related to face features detection recognized many variants in scale, location, in-plane rotation, pose, facial appearance, camera lighting environments, etc.

The mobile-Net classifier uses deep intelligent separable convolution is an enhancement of Convolutional neural network (CNN) due to the limitation of the CNN in face recognition where features extracted from an image like an eye, eyebrow, mouth, and nose are more shrinking in size as a result of pooling and filtering and also, The non-linearity inside the narrow level is removed. Most researchers use pre-train models of CNN to enhance biometrics recognition.

Therefore, the motivation of this paper to use a Mobile-Nets classifier to improve feature extraction of masked-face best identification and authentication to introduce a better algorithm that can attain face features information. In [3] introduced how to implement a Haar-like algorithm to detect faces with superior performance in multi-condition.

In this research. We propose a framework recognition of masked-face using a Mobile-Nets network. The reset or research is organized as follows, sec. one related work, and sec. two the proposed method of the work used in detect and recognize the masked-face, finally sec. three the simulating results and conclusion.

II. RELATED WORKS

In [4] introduced an algorithm to detect face features for mobile applications for objects wearing a mask. [5] proposed the capability of dissimilar structures of the mask with and without operational Micro-Ventilator. [6] proposed Cascade-Classifier to concentrate the individual, the Convolutional Neural Network models to recognize the individual and non-individual image, and the Kalman filter technique to track individual motion with an accuracy range of 77.7%.[7] introduced deep-learning (DL) perceptions relevant to face appearance examination and face-recognition, and delivers a survey of readings on exact face recognition difficulties. [8] proposed a method using CNN by restricted training ratio, considerable progress in face-recognition amount. [9] introduced clustering on the sizes of bounding boxes to get a good image and speed up the divergence speed of the network using the YOLO algorithm.[10] suggested a face detection prototype to detect the faces and their consistent facial milestones current in the video frame using the Mobile-Net-V2 construction. [11] showed a non-masked-face recognition and also relate in the masked-face recognition method. Principal-Component-Analysis (PCA) got a more operative and popular numerical technique and generally used. [12] suggested by what means to form mobile semantic separation models concluded a reduced form of Deep-Labv3 which called Mobile Deep-Labv3.[13] showed a deep Meta Caps-net Equivariant Embedding Modal with three different innovations. [14] the proposed algorithm improves the face recognition accuracy from 50% - 95% without providing data on their starting point. [15] suggested a model called the Face-Net trained which is used for improving masked- face gratitude. [16] proposed a Multilayer-Perceptron (MLP) for the classification process and showed that that the middling size of the code-book (RBF neurons) provided a greater recognition-rate associating with the code-book of size 70.
[17] showed that the Viola-Jones algorithm gave the efficiency 12% on average. [18] suggested a deep-cascaded multi-task to detect the face. [19] suggested a modal called the anchor-free category and showed that it realized real-time speed and great correctness with a minor model-size.

[20] introduced precise face-segmentation masks from any arbitrary size input object using VGG-16 with malicious pixel-level accuracy of 93.884 % for the face-segmentation masks.[21] proposed a CNN prototype to be used in ensemble technique for detection.[22] used CNN models to progress the accuracy in face detection.[23] introduced Neural Networks for matching symmetrical structures of the social face.[24] used only one feature for each fragile classifier. Specified the same amount of features, this technique diminishes the error by 37% and 2.6 times as fast as Viola-Jones' face-detector.[25] suggested an algorithm for real-time human-robot communication uses, the modal was clever for detection, recognition, and tracking faces up to 24 frames per second.[26] illustrated global-dynamic global Haar-like features in which it was is faster than that of the customary method.

[27] suggested a four new Haar-Like feature in which it got improved consequences compared with other customary face detection classifiers like Haar-Like.[28] proposed a novel usual of ±26.565 ° haar-like features which can be designed rapidly to characterize the features of rotated faces.

III. THE PROPOSED METHOD

3.1 Masked face Data set

This paper used a large dataset associated with face patterns, which were images of faces wearing masks and without-mask. The images were collecting from Kaggle datasets. This dataset consists of 3835 images belonging to two classes:

- With mask: 1916 images (refer with: Figure.1).
- Without mask: 1916 images (refer with: Figure.1).

Figure.1 Dataset with-mask

Figure.2 Dataset without-mask

3.2 Data-Augmentation

In the next stage, to succeed the training and propose an efficient modal we must enlarge and develop the data-set to get a greater number of patterns for training by rotate, casual crop, and flip each image in the data-set in which, data-augmentation avoid over-fitting or training patterns. A purpose image data generation is generated for image-augmentation, which profits test and train collections of data-set. After that, this dataset splitting to 80% for training and 20% for test as shown in (refer with: Figure.1).

![Figure.3 Block diagram of used data-augmentation.](image)

3.3 Mobile-Net in use

Mobile-Net stayed practical as a detection network to detect the people-wearing mask or people without the mask, which was an inconsequential CNN and accomplished very high accuracy in image-classification. The highest structures of mobile-Net were residual blocks to the training of deep network and depth-wise to separable convolution (refer with: Figure.1). In this step, we use Conv2D, Batch Normalization, ReLU, expanded_conv_depthwise (Depthw), blocks expand, flatten, dense, dropout to build the proposed modal to improve the accuracy.

![Figure.4 Blocks of Mobile-Net.](image)

As shown in (refer with: Figure.1), illustrate the block diagram of the proposed modal, in this step will appear the motivation of our work and how to recognize the authorized person that enters the nuclear facility in case of wearing the masked-face. After pre-trained the dataset on mobile-net we detect if a person wears a mask or not, by applying the steps of planned algorithms illustrate in Table1 step by step from Image data generator to increase the dataset for training until reach to training the modal on Mobile-net to detect the masked-face and evaluate the performance of modal by accuracy/accuracy-validation, loss/loss-validation, and precision/recall/F-score.
The interesting step after that applying Haar-like features to detect the retina of the person to extract the boundary box around the retina compare this with the dataset of the person without the mask for recognition.

![Diagram](image)

**Figure 5** The block diagram of the proposed modal

(Refer with: Table 1) there are the following steps of our proposed algorithm:

1. Apply data augmentation method to increase the dataset.
2. Pre-processing method on the image (rescale).
3. Dataset splitting to 80% for training and 20%.

### Table 1. The algorithm of the planned modal

| Algorithm 1 proposed modal |
|----------------------------|
| **Input** | Face mask dataset |
| **Output** | Face recognition |
| **Initialize on** | |
| 1 | Import numpy, pandas, sklearn, matplotlib, Os, keras and tensorflow_gpu libraries |
| 2 | Set INIT LR=1e-4, EPOCHS=20, BS=32<br>aug=ImageDataGenerator()<br>rotation_range=20,<br>zoom_range=0.15,<br>width_shift_range=0.2,<br>height_shift_range=0.2,<br>shear_range=0.15,<br>horizontal_flip=True,<br>fill_mode='nearest' |
| 3 | Test_dataset=keras.preprocessing.image.ImageDataGenerator(rescale=1./255) |
| 4 | Training_set = Train_dataset.flow_from_directory(<br>‘training_folder’,<br>target_size=(image_width,image_high),<br>batch_size=BS,<br>class_mode='categorical') |
| 5 | test_set = test_dataset.flow_from_directory(<br>‘test_folder’,<br>target_size=(image_width,image_high),<br>batch_size=Batch_size,<br>class_mode='categorical') |
| 6 | Training_Model |
| 7 | Training_Model.compile(loss='categorical_crossentropy',optimizer='Adam(lr=INIT_LR, decay=INIT_LR / EPOCHS), metrics=['categorical_accuracy']) |
| 8 | Training_Model.fit(aug.flow((trainX,trainY),batch_size=BS),<br>steps_per_epoch=int(1/BS),<br>validation_data=(testX,testY),<br>validation_steps=10/len(testX)/BS,<br>epochs=EPOCHS) |

### 3.4 Haar-like proposed algorithms for training

Technology in the extent of computer vision for support and distinctive objects in an image. Individuals recognize a commonality of objects in images with brief attempts, spite the circumstance that the image of the objects may modify something in separate survey appoint [19], in many separate dimension and scales, or even when they are transfer or rotated. Objects can recognize when they are partly retard from view. This task is still a challenge for computer vision systems [1, 29].

Many approximate to the work has been accomplished over multiple decades. So, Haar-like is used in object-recognition. The important advantage of a Haar-like feature above maximum other features is its calculation rapidity[1, 17].

Owing to the use of a summed-area table, a Haar-like of any size can calculate in constant time. A Haar-like is characterizing by taking a boundary box of an image and separating that box into several parts [30]. They are visualizing as black and white adjacent boxes [31].

In the limited access area as it is within a designated zone comprising a nuclear facility to which access is limited and controlled for physical protection purposes. At this zone, the masked-face recognition to limit the spread of COVID-19 applies to all individuals that enter the facility.

In this interesting step of work, after we training the Mobile-Net to our data set and detect which person wears a mask, and he takes permission to enter the facility to limit the spread of COVID-19, it must recognize if this person in our data set.

(Refer with: Table 2), by applying the proposed modal for Haar-like detection and recognition algorithm as illustrated next, First, we determine the interesting features to detect in our case it is the retina and determines the (x,y,w,h) dimensions to extract the boundary box from the object which it is Region of interest to recognize the person by using Haar cascaded detect function. After that, each object assigns a label and saves it with this label for comparison with the dataset without the mask according to the following steps:

1. Extract feature using haar-cascaded function.
2. Labeling each object using people.index(person) function.
3. Determine the boundary box of right and left eyes.
4. Use face recognition function.
3.5 Haar-like proposed algorithms for recognition

(Refer with: Fig. 6) the proposed structure of face recognition classifier. After applying the retina detection algorithm and determine the boundary box of the right eye and left eye by using Haar-cascaded classifier the features were saving and labeling. So, in the final step applying the face recognition algorithm as illustrated in Table 3, by comparing the boundary box of the labeled object with the dataset of objects without the mask for recognition to take the permission to the person to access and enter the facility after following the next steps of the algorithm of face recognition:

1. Using haar-cascaded classifier function.
2. Apply face.LBPHF face recognition function.
3. Training the masked-images.
4. Put label using a put text function.
5. Finally, using imshow function to review the detected face.

| Table 2. Algorithm of retina detection modal |
|---------------------------------------------|
| **Input** | **Output** | **Initialization** |
| Face mask dataset | Retina detection | Initialize DIR, people /DIR directory of face mask data set, and people |
| 1: import numpy, sys, PIL,HaarlikeFace | Haar CascadeClassifier('haarcascade_righteye.xml') |
| 2: features=[] | labels=[] |
| def create_train(): | |
| for person in people do | |
| path=sys.path.join(DIR,person) | |
| label=people.index(person) | |
| for img in os.listdir(path) do | |
| img_path=os.path.join(path,img) | |
| img_array=cv.imread(img_path) | |
| gray=cv.cvtColor(img_array,cv.COLOR_BGR2GRAY) | |
| faces_rect=HaarCascadeDetecMultiScale(gray,scaleFactor=1.1, minNeighbors=2) | |
| for(x,y,w,h) in faces_rect do | |
| features.append(faces[x,y,x+w,y+h]) | |
| labels.append(label) | |
| 4: create_train() | |
| 5: features=np.array(features, dtype='float32') | |
| labels=np.array(labels) | |
| 6: face_recognizer=face.LBPHFFaceRecognizer_create() | |
| 7: face_recognizer.train(faces,labels) | |
| 8: pred,labels=face_recognizer.predict(faces) | |

| Table 3. Algorithm of face recognition |
|----------------------------------------|
| **Input** | **Output** | **Initialization** |
| Face mask dataset | Face recognition | |
| 1: Import numpy, sys, PIL, HaarlikeFace | Haar CascadeClassifier('haarcascade_righteye.xml') |
| 2: face_recognizer=face.LBPHFFaceRecognizer_create() | face_recognizer.train(faces, labels) |
| 3: gray=cv.Color(img,cv.COLOR_BGRGRAY) | |
| 5: faces_rect=HaarCascadeDetectorMultiScale(gray,scaleFactor=1.1, minNeighbors=2) | |
| 6: for(x,y,w,h) in faces_rect do | |
| faces=img[y:y+h,x:x+w] | |
| 7: cv.putText(img,people[labels][20],[20,20],cv.FONT_HERSHEY_COMPLEX,1.0,0.5, thickness=2) | |
| 8: cv.imshow('Detected Face',img) | |

(Refer with: Fig. 7) shows the recognized output person after applying the proposed algorithm for face recognition which indicates the label of the person with the mask.

3.6 Experimental Result.

The proposed methods are applied to employ a laptop background (Intel(R) Core (MT) i7-7700 HQ CPU @ 2.80 GHz, win10 64-bit operating method, RAM 16 GB, 256 GB SSD, NVIDIA GTX 1060 graphics card, and implemented in python 3.6 contains a deep learning library. The assessment metrics engaged here are accurate, (refer to Fig.8, Fig.9, Fig.10) precision, recall, and F1-score. accuracy in, accuracy – validation, loss, loss- validation.

(Refer with: Eq. (1), Eq. (2), Eq. (3), Eq. (4))

The evaluation metrics of the proposed modal is illustrating as follow:

1. Accuracy $\frac{Tp+Tn}{Tp+Pp+Fn+Tn}$
   Eq. (1).
2. Precision $\frac{Tp}{Tp+Pp}$
   Eq. (2).
3. Recall $\frac{Tp}{Tp+Fn}$
   Eq. (3).

![Figure.6 the proposed structure of face recognition classifier](image-url)

![Figure.7 the recognized output person](image-url)
4. F1 score

Eq. (4),

(\( Tp \) True positive, \( TN \) true negative), (\( Fp \) False positive, \( FN \) False negative)

(Refer with: Fig. 10), we only iterated 20 epoch, the data improved and accelerate the training.

(Refer to Table 7), which illustrates the Related research used different techniques CNN, Mobile-Net, Hyperface-ResNet, and cross-class object removal algorithm used in face detection and recognition with our proposed modal.

| Authors/year | Technique | Performance |
|--------------|-----------|-------------|
| (Singh et al., 2021) | Single Shot Multi-box Detector as a Face detector and MobileNetV2 | Accuracy of 93.96% and an F1 score of 0.93. |
| (Rahman, Alvi, & Kim, 2020) | C3DST | Accuracy of 98.7% |
| (Song & Pan, 2019) | cross-class object removal algorithm | 2.3% and 1.5% for face and mask detection precision |
| (Pottail, 2020) | multithreads | 2.3% and 1.5% for face and mask detection precision |
| (Razou, Budi, & Chellappa, 2019) | HyperFace-ResNet | |
| Our proposal model | Mobile-Net and Hare-like | An accuracy of 0.99 and an F1 score of 0.88.
IV. CONCLUSION

In our proposed modal, a novel face detection and recognition framework is implementing using mobile-net and haar-like algorithms applying in the dataset contain images with-mask and image without -mask in which, data augmentation is applying and after that, the data set is splitting into 80%for training and 20% for testing. The proposed techniques and algorithms give an excessive accuracy, the results of the proposed modal was tested on a dataset from Kagggle, yielded 0.99 accuracies, a loss of 0.08, F1.score 0.98 and it would be supportive to the facility to detect and recognize masked –faces in this great pandemic.

REFERENCES

1. M. Alsaawaf, Z. Chaczko, and M. Kulbacki, “In Your Face: Person Identification Through Ratios Between Facial Features,” 2020, pp. 527-536.

2. A. Bhadani and A. Sinha, “A FACEMASK DETECTOR USING MACHINE LEARNING and PROCESSING TECHNIQUE,” vol. 13, 11/09 2020.

3. Z. Zhu, T. Mecchio, H. Adachi, O. Kiyiyama, T. Koide, and H. J. Mattausch, “Multi-view Face Detection and Recognition using Haar-like Features,” 01/01 2004.

4. K. Hammoudi, A. Cabani, H. Benhabiles, and M. Melkemi, “Validating the Correct Wearing of Protection Mask by Taking a Selfie: Design of a Mobile Application “CheckYourMask” to Limit the Spread of COVID-19,” Computer Modeling in Engineering & Sciences, vol. 124, no. 3, pp. 1049–1059, 2020.

5. S. S. Zhou, S. Lulaka, C. Chissone, R. W. Nims, D. B. Suchmann, and M. K. Izaq, “Assessment of a respiratory face mask for capturing air pollutants and pathogens including human influenza and rhinoviruses,” (in eng), J Thorac Dis, vol. 10, no. 3, pp. 2059-2069, Mar 2018.

6. J. Angelico and K. Wardani, “Convolutional Neural Network Using Kalman Filter for Human Detection and Tracking on RGB-D Video,” CommDT (Communication and Information Technology) Journal, vol. 12, p. 105, 10/31 2018.

7. G. Guo and N. Zhang, “A survey on deep learning based face recognition,” Computer Vision and Image Understanding, vol. 189, p. 102805, 2019/12/01 2019.

8. U. Anman and V. P. Vishwakarma, “Face recognition using modified deep learning neural network,” in 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2017, pp. 1-5.

9. H. Huang, D. Sun, R. Wang, C. Zhu, and B. Liu, “Ship Target Detection Based on Improved YOLO Network,” Mathematical Problems in Engineering, vol. 2020, p. 6402149, 2020/08/17 2020.

10. A. S. Joshi, S. S. Joshi, G. Kanahusabai, R. Kapil, and S. Gupta, “Deep Learning Framework to Detect Face Masks from Video Footage,” in 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), 2020, pp. 435-440.

11. M. S. Ejaz, M. N. Islam, M. Sifatullah, and A. Sarker, “Implementation of Principal Component Analysis on Masked and Non-masked Face Recognition,” 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), pp. 1-5, 2019.

12. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510-4520.

13. F. Wu, J. Smith, W. Lu, and B. Zhang, “Pose-robust Face Recognition by Deep Meta Capsule network-based Equivariant Embedding,” 2021.

14. N. Damer, J. H. Grebe, C. Chen, F. Boutros, F. Kirchbuchner, and A. Kuijper, “The Effect of Wearing a Mask on Face Recognition Performance: an Exploratory Study,” in 2020 International Conference of the Biometrics Special Interest Group (BIOSIG), 2020, pp. 1-6.

15. M. Ejaz and M. Islam, Masked Face Recognition Using Convolutional Neural Network, 2019, pp. 1-6.

16. H. Wald, Efficient Masked Face Recognition Method during the COVID-19 Pandemic, 2020.

17. J. Huang, Y. Wang, and H. Chen, “Improved Viola-Jones face detection algorithm based on HoloLens,” EURASIP Journal on Image and Video Processing, vol. 2019, no. 1, p. 41, 2019/02/11 2019.

18. K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, “Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks,” IEEE Signal Processing Letters, vol. 23, no. 10, pp. 1499-1503, 2016.

19. Y. Xu, W. Yan, G. Yang, J. Luo, T. Li, and J. He, “CenterFace: Joint Face Detection and Alignment Using Face as Point,” Scientific Programming, vol. 2020, p. 7845384, 2020/07/02 2020.

20. T. Meenpal, A. Balakrishnan, and A. Verma, “Facial Mask Detection using Semantic Segmentation,” in 2019 4th International Conference on Computing, Communications and Security (ICCCS), 2019, pp. 1-5.

21. J. Lee, S. Lee, and S. Yang, “An Ensemble Method of CNN Models for Object Detection,” in 2018 International Conference on Information and Communication Technology Convergence (ICTC), 2018, pp. 898-901.

22. J. Lee, S.-K. Lee, and S.-I. Yang.

23. T. H. Le, “Applying Artificial Neural Networks for Face Recognition,” Advances in Artificial Neural Systems, vol. 2011, p. 673016, 2011/11/03 2011.

24. T. Mita, T. Kaneko, and O. Hori, “Joint Haar-like features for face detection,” in Tenth IEEE International Conference on Computer Vision (ICCV’05) Volume 1, 2005, vol. 2, pp. 1619-1626 Vol. 2.

25. P. Menezes, J. C. Barreto, and J. Dias, “Face tracking based on haar-like features and eigenfaces,” IFAC Proceedings Volumes, vol. 37, no. 8, pp. 304-309, 2004/07/01 2004.

26. M. Rezaei, H. Ziaei Nafchi, and S. Morales, “Global Haar-Like Features: A New Extension of Classic Haar Features for Efficient Face Detection in Noisy Images,” in Image and Video Technology, Berlin, Heidelberg, 2014, pp. 302-313: Springer Berlin Heidelberg.

27. M. Songyan and B. Lu, “A face detection algorithm based on Adaboost and new Haar-like feature,” in 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), 2016, pp. 651-654.

28. S. Du, N. Zheng, Q. You, Y. Wu, M. Yuan, and J. Wu, “Rotated Haar-Like Features for Face Detection with In-Plane Rotation,” in Interactive Technologies and Sociotechnical Systems, Berlin, Heidelberg, 2006, pp. 128-137: Springer Berlin Heidelberg.

29. M. Jiang and X. Fan, RetinaMask: A Face Mask detector, 2020.

30. S.-K. Pavani, D. Delgado, and A. F. Frangi, “Haar-like features with optimally weighted rectangles for rapid object detection,” Pattern Recognition, vol. 43, no. 1, pp. 160-172, 2010/01/01 2010.

31. Y. Li and F. Ren, “Light-Weight RetinaNet for Object Detection,” CoRR, vol. abs/1905.10011, 2019.

AUTHOR PROFILES

Nadia M. Nawar, was born in Cairo, Egypt, in 1984. She received the B.S. degree in 2006 from the University of Banha, Egypt, Nawwar, in 2014 and the M.S. degree from the University of monifia, Egypt. She is currently an Assistant Lecture Egyptian Atomic Energy Authority (EAEA), Cairo, Egypt.

Prof. Kasban, received B.Sc., M.Sc. and Ph.D. degrees in Electrical and Electronic Engineering from Menoufia University, Egypt. Currently, He is the Head of Engineering and Scientific Instruments Department, Nuclear Research Center (NRC), Egyptian Atomic Energy Authority (EAEA), Cairo, Egypt. He is a co-author of international journals papers. His research areas of interest are in, electronics, digital signal processing, communication systems and nuclear applications in industry and medicine.

May Salama, received B.Sc., M.Sc. from Ain shams University, Egypt and Ph.D. degrees from University of Banha, Egypt in computer science, machine learning and security.