Face Mask Recognition based on Facial Feature and Skin Color Conversion

Md. Ziaur Rahman
Dept. of Computer Science and Engineering
International Islamic University Chittagong
Chattogram-4318

ABSTRACT
In current decays, COVID-19 became a great issue to all. Face mask detection is an important and active research area in the field of computer vision and human computer interaction. To get rid from COVID-19 everyone's needs to wear mask. So, it is an important task to detect the face mask in the open place to stop the spread of COVID-19. In this regard, in this work, a system is proposed to detect the face mask from an input image through the facial feature and skin color conversion. For that, initially, the facial region is extracted from the input image. From the extracted facial region 68 shape predictor facial landmarks are extracted. Based on these landmarks, the facial features of jaw, nose and mouth region are extracted. As the face mask covers the facial region of jaw, nose and mouth, so, whether, a person wear the mask or not can be recognize by analyzing the skin color of jaw, nose and mouth region. Finally, the proposed method is justified with different images i.e., mask or non-mask that are captured from different environmental conditions. The proposed method shows significant improvement with present state of the art.

General Terms
Pattern Recognition, Image Processing, Human Computer Interaction.

Keywords
Facial feature, COVID-19, face mask, facial landmark.

1. INTRODUCTION
In order to confine the expansion of COVID-19 virus, everybody wears a mask during corona virus pestilence. World health organization (WHO) submitted that COVID-19 has globally transited over 47.5M people where caused death over 1.21M and recovered 31.6M. The people of whole world are concerned about their death and also public heath came to the top priority for the government. So, it was considered that surgical face mask can cut the spread of the corona virus. After that WHO declared to wear mask the people who have respiratory symptoms. Now it became compulsory for all to wear mask to save own from corona affected people.

At first face mask is not mandatory for all people, however, for the restricted to spread of corona virus it is mandatory to wear mask by doctors and scientists. Face mask detection is that which refers to detect whether the people wear masks or not. In this work, we focus on the process to detect the faces those wear the mask or not.

In this work, the process of face mask detection is done based on the nose, jaw, and mouth regions detection. In this regard, the human face is initially detected. After that, the facial landmarks are extracted. Based on the landmark features the nose, jaw and mouth region are extracted. Finally, the face mask is recognizing based on the skin color of jaw, nose and mouth region.

To detect the face mask, many researchers perform many research works in the last few years. Such as, in [8], authors recognize the masked faces which addresses deep face recognition (FR) problem, having prior information of the face model. It worked with two perspectives. One is the created dataset and the other one is to use of open useful face features. They develop the recognition accuracy of masked faces from the initial 50% up to 95%. Therefore, it is very insistent to enhance the recognition representation of the remaining face recognition technology on the masked faces. These papers work on to detect face and avail public health care. In these papers, they propose Retina Face Mask, which have high-accuracy and effective facemask detector. The mentioned Retina Face Mask is a one-stage detector, which is forms of a feature pyramid network to dissolve high-level semantic information with multiple features of maps, and a novel context outlook module to concentrate on detecting face masks in introduced in [9].

In [7], it is founded on Mask-R CNN; their scheme is to Generalized Intersection over Union (GIoU) as the loss function for bounding box regression to improve accuracy of face detection. Specifically, in this proposed method, ResNet-101 is used to extract features, RPN is utilized to generate region of interest (RoI), and RoI align faithfully that preserves the exact spatial locations for generating binary mask through Fully Convolution Network (FCN).

The paper usedintroduced in [6] proposed Facemask net model. By using this it takes an input of an image, classifying the image as with a mask, and improperly worn mask and without a mask. They designed their projects into two phases such as, training face mask detector and implementing face mask detector. Using deep learning method called Facemask net, they got an accuracy of 98.6%.

In [4], authors introduce three types of masked face datasets that are, the correctly masked face dataset, incorrectly masked face dataset and their combination. And this provides masked to face deformable model to detect the faces if worn the mask or not and also to detect the mask is correctly worn or not.However, this work globally presents the applied mask-to-face deformable model for approving the generation of other masked face images, markedly with specific masks.

The paper is arranged into the following Sections. The proposed architecture is explained in Section II with the processing examples. The experimental results are denoted in the next Section. In the end, Section IV encloses the conclusions of the paper.

2. PROPOSED ARCHITECTURE
The proposed face mask detection architecture has two major portions. i.e., face detection using landmark feature and face recognition using YCbCr model. The architecture has six main steps, i.e., 1) pre-processing, 2) face detection, 3) facial

landmark detection, 4) jaw, nose and mouth detection, 5) extraction of skin color, 6) skin color for jaw, nose and mouth. The proposed architecture is shown in Fig.1.

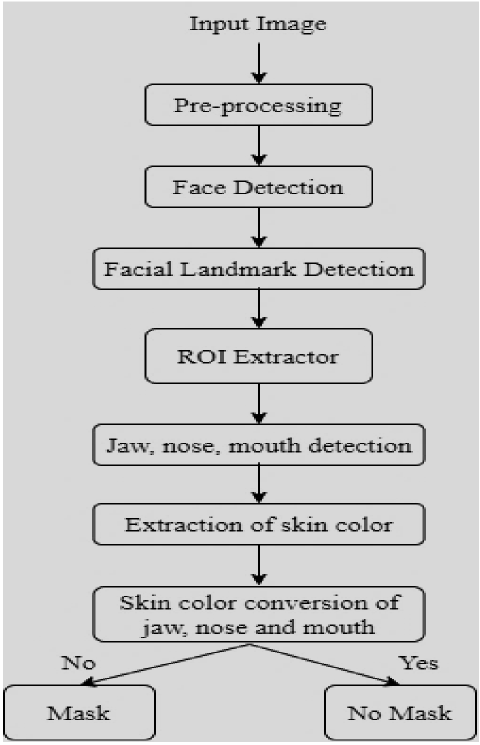


Fig.1. Proposed architecture of face masks detection.

2.1 Pre-processing

At first, the face image is resized into a standard shape, i.e., 256x256. Types of mask and non-mask images are liberated from the different sector. The images are resized, as the mask images are located in different environmental conditions with different background, consequently, the images are captured from different distances with different lighting conditions.

(a) (b)

Fig. 2.Processing example of input image resize for (non-mask and mask): a) input image, b) resized image.

2.2 Face Detection

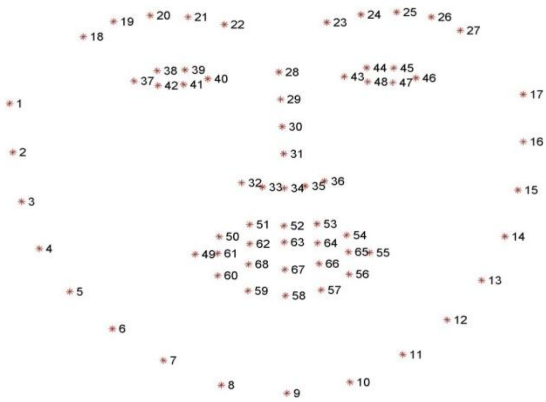
The face detector works as the second stage of our architecture. A raw RGB image is obtained as the input image which is passed at this stage. With their bounding box coordinates the face detector landmark algorithm extracts the facial regions. Detecting faces accurately is the important procedure of our architecture. Without detecting face correctly the system does not able to recognize the face mask. The processing example is shown in Fig. 3.

(a) (b)

Fig. 3.Processing example of face detection for (non-mask and mask): a) input image, b) Face detection with landmark.

2.3 Facial Landmark Detection

Facial landmark is a technique which is applied on applications such as face alignment. In this behalf of face landmarks which play a vital role. Here our aim is to detect facial structures on a person's face by using a method called shape predictor. In a number of ways we can detect face features. We used here OpenCV built-in TensorFlow and Keras. Over here singularly, we applied the Histogram of Gradients (HOG) object detector particularly for the task to detect face features. We can attain face by using bounding box through some process for which we used the (x, y) coordinates of the face in the image gradually. The facial landmark detector which is processed inside the dlib library of python to detecting landmarks is used to inventory the location of 68 points or (x, y) coordinates which is chart to the face structures. These exponents of 68 shape predictor to coordinates or points that can be easily noticed on the image beneath. At first we take an input image. Then detect facial landmark from the input image. After detecting the facial landmarks, find out the plot image from the original image. Then visualize the facial landmark. After that, detect the facial landmark which is shown in Fig.4 both for mask and non-mask.



Facial Landmark

(a) (b)

Fig. 4.Processing example of facial region detection using landmark (non-mask and mask): a) Facial Landmark detection, b) visualize facial landmark.

2.4 Jaw Detection from Facial Image

After detecting the facial landmark, we detect jaw, nose and mouth from the input image. First of all, individually detect jaw, nose and mouth from the landmark detection. Here we detect the jaw with landmark and then clone the original image for finding the Region of Interest (ROI) of the jaw. By landmark we detect point of the jaw is (2-15) points. Jaw points detection example is shown in Fig. 5. Here, we see that the jaw points can be detected both from the mask and non-masked images.

(a) (b) (c)

Fig. 5. Processing example of jaw detection for (non-mask and mask): a) facial region with landmark,b) jaw with landmark region, and c) jaw region extraction.

2.5 Nose Detection from Facial Image

In this section Nose region is extracted from the facial image through the landmarks points. Here, face part nose is divided into three portions for detecting the nose. One part is full nose, second part is 50% of nose, and third part is 25% of nose. After that nose portion is detecting from the cloning image. If any person covers the nose 25% with mask then it is not the perfect process of wearing the mask. The perfect process of wearing the mask is if person's covers the nose with mask 50%. The process of nose detection is shown in Fig. 6.

(a) (b) (c)

Fig. 6.Processing example of nose detection for (non-mask and mask): a) facial region with landmark b) nose with landmark region, and c) nose region extraction.

2.6 Mouth Detection from Facial Image

Detecting mouth is another process of feature detection from the face. Clone the original image for finding the mouth. The ROI of the face region for mouth is separately extracted. Input image detect the mouth region through the ROI. Without Mask, face region detect mouth points from ROI. Mouth points are (49-60), inner mouth points are (61-68). With mask, face region find the mouth inner portion. Finally get the region of mouth. Processing example shown in Fig.7.

(a) (b) (c)

Fig. 7.Processing example of mouth detection for (non-mask and mask): a) facial region with landmark, b) mouth with landmark region, and c) mouth region extraction.

2.7 Extraction of Skin Color

This portion elaborately describe about detecting the skin color region as well as jaw, nose and mouth which are effective features for mask detection. To detect the jaw, nose and mouth the face color region is used; this region is extracted from facial Region of Interest (ROI) which is described in the previous section. To recognize the face mask, we have to confirm whether the jaw, nose and mouth are covered with mask or not. If these features are covered with mask then it we cannot find any skin color information from these features, otherwise, the skin color can be detected from these features.

To confirm the process we apply the skin color conversion on these facial features using YCbCr color space. The YCbCr signals are generated from the corresponding gamma-configured RGB source as follows (JPEG conversion):

Y = 0.299×**R** + 0.587×**G** + 0.114×**B**

Cb = — 0.1687×**R —** 0.3313×**G** + 0.5×**B** + 128

Cr = 0.5×**R—** 0.4187×**G —** 0.0813×**B** + 128

The range of each input (R, G, B) is the full 8-bit range of [0…255]

2.8 Detect Skin Color for Jaw

For detecting skin color we use YCbCr color space. This color space is use for non-linear RGB signal and weighted as a sum of RGB values. Detecting jaw maximum and minimum YCbCr is used to get an actual array value from the cloning jaw image. The jaw is basically located at the lower portion of face. We used the jaw points are (2-15). Extracting the lower portion we get the skin region. If the face is covered with mask then skin color of jaw cannot be found, for this reason jaw points are not found. So, the color conversion of jaw is turned into black. If the face is not covered by mask, then detect the jaw and turned into YCbCr color space, which is shown in Fig.8.

(a) (b)

Fig. 8.Processing example of skin color for jaw: a) non-mask, b) mask.

2.9 Detect Skin Color for Nose

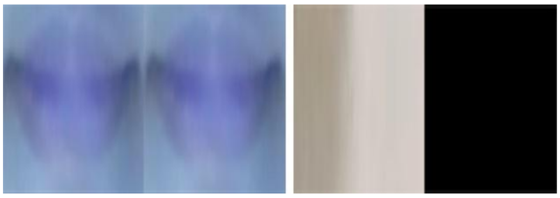
To detect nose is an essential part at the whole face. We know the points of nose abounding by points (28, 36). Nose portion is fully covered when mask wear properly that means covered the all landmark points of nose. If the point is (28-36) then the nose is fully covered, (28-30) for 50% covered the nose, and if the point is (31-36) the nose is 25% covered. By the points of nose we can detect that how the mask is covered the face. After detecting the skin color for nose, we can see that the nose is not getting skin color because of wearing mask. And which is not covered by maskcan detect the skin region of nose which is shown in Fig. 9.

(a) (b)

Fig. 9.Processing example of skin color for nose: a) non-mask, b) mask.

2.10 Detect Skin Color for Mouth

The efficient feature for color region is object detection. After detecting the skin color for mouth, the face which is covered can't be detected the skin color for the mask. And the face which is not covered by mask, can detect the mouth skin color. The mouth portion is divided in to two portions such as, inner portion and outer portion. Where the inner portion point is (61-68) and outer portion point is (49-60). The face which is uncovered can detect the point of the mouth so that it shows skin color. The face which is covered can't detect the points of mouth so that it can't recognize the skin color, which is shown in Fig. 10.



(a) (b)

Fig.10. Processing example of skin color for mouth: a) non-mask, b) mask.

2.11 Face Mask Detection

To detect the mask we used three parts from the facial part which are jaw, nose and mouth. If we can't detect the points of jaw, nose and mouth then we can say that the face is covered otherwise uncovered. Now, in our procedure we can see that if skin color of jaw, nose and mouth are detected then the face is detected as without mask. If we get the skin color of both mouth and nose then it can be say that the face is without mask. After that if skin color of mouth is detected and nose is not detected then the mouth is not covered by mask and nose is covered by mask. If we get the skin color of nose and mouth is not detected then we can say that the mouth is covered by mask and the nose is without mask. At last if we don't get the skin color for both mouth and nose then it is confirmed that the face is with mask. The face mask detection procedure isexplained in the procedure 1. The processing example for the face mask detection procedure is shown in Fig. 11.

Procedure 1: Face Mask Detection Procedure

if skin color of jaw region is detected
  if skin color of mouth and nose is detected
    decision='No Mask'
  else if skin color of mouth is detected but nose color is not detected
    decision='Mouth Uncovered Nose Covered'
  else if skin color of mouth is not detected but nose color is detected
    decision='Mouth Covered Nose Uncovered'
  else **decision** = 'Mouth and Nose are Covered'
else **decision** = 'Mask'

(a) (b)

Fig.11. Processing examples of face mask detection for: a) non-mask, b) mask.

3. EXPERIMENTAL RESULTS AND DISCUSSIONS

The practical implementation of our proposed method has used two libraries. We analyzed the recognition rate in different papers specifically for facial recognition. All experiments are performed on Intel(R) Core(TM) i3-6006U@2.00GHz with 4GB RAM in the COLAB environment.

Dataset images from different environmental conditions are given in Fig.12. The given dataset images are captured from the front side with different environmental conditions such as, noisy, background spot, low illumination condition and so on. By these conditions we able to detect the face mask.

(a) (b)

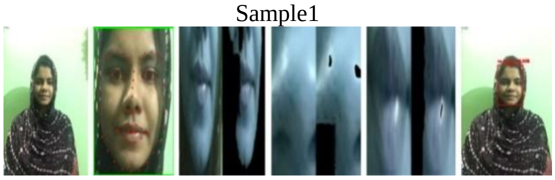
Fig.12. Dataset images from different environmental conditions: a) non-mask and b) mask.

Here we used 350 images with mask and 450 without mask images. Where from the mask images the number of female images are 150, for male is 120 images and for child is 50 images. After that for non-masked images for female we used 220 images, for male 170 images and for child 60 images.

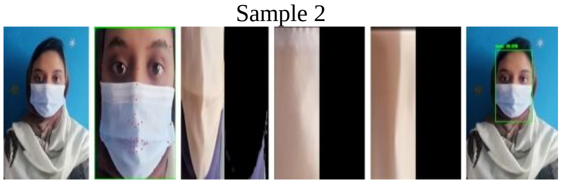
Some processing examples are shown in Fig. 13-15. That image liberated from different environment. The pixels of all images are 256x256. The processing example of the mask and non-mask of the input images is shown in Fig. 13 for female, Fig. 14 for male and Fig. 15 for child. Detect the input image with landmark, ROI image, jaw detection, nose detection, mouth detection and YCbCr color space conversion of jaw, nose and mouth. Which are shown in Fig.13 to Fig. 15.

The processing example for female is shown in Fig. 13. Where, Sample 1 and Sample 2 are not wear mask and Sample 3 for mask.

Sample1



Sample 2



Sample 3

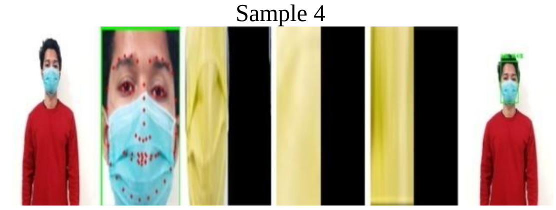
(a) (b) (c) (d) (e) (f)

Fig.13. Processing example of mask face detection for female: a) input image b) facial landmark detection c) extraction skin color for jaw d) extraction skin color for nose e) extraction skin color for mouth f) mask detection.

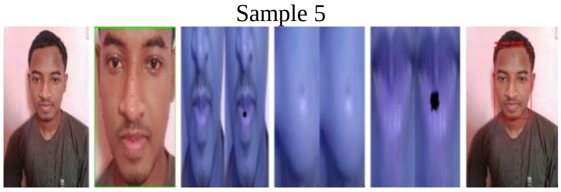
At Fig. 13 both for non-mask and mask images for female, where Sample 1 and Sample 2 are for non-mask images. After detecting these non-mask images we take an input image. From that we get a facial landmark images, where detected the 68 shape predictor for the whole face. We worked here with three portions such as jaw, nose and mouth by detecting their landmark points. After that skin color conversion applied for jaw, nose and mouth. By applying skin color if we get the jaw, nose and mouth extracted by skin color then we has that the face is not covered by mask which is shown in Sample 1 and Sample 2 at Fig. 13. And if we don't get the jaw, nose and mouth by extracting skin color then we get the face is covered with mask shown in Fig. 13 at Sample 3. For mask images we get the accuracy rate and 99.22% and for non-mask images 99.50%.

The processing example for female is shown in Fig. 14. Where, Sample 4 and Sample 5 are wear mask and Sample 6 for not wear mask.

Sample 4



Sample 5



Sample 6

(a) (b) (c) (d) (e) (f)

Fig.14. Processing example of mask face detection for male: a) input image b) facial landmark detection c) extraction skin color for jaw d) extraction skin color for nose e) extraction skin color for mouth f) mask detection.

At Fig. 14 is for both mask images non-mask images. Here first we take an input image for male. After that we used the same procedure which is similar to the images used for female. We detected the facial landmark on the input image then we applied skin color conversion on that image. If we get the detected face after applying skin color then it is uncovered by mask which is shown in Sample 6 and if we don't get the face then the image is covered by mask which is shown in Sample 4 and Sample 5 at fig. 22. However, we get the accuracy rate for male of non- mask images is 99.94% and mask images 98.94%.

The processing example for child is shown in Fig. 15. Where, Sample 7 is not wear mask and Sample 8 for wear mask.

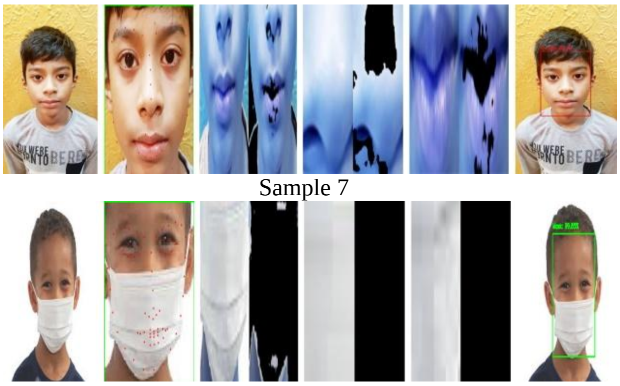


Fig.15. Processing example of mask face detection for child: a) input image b) facial landmark detection c) extraction skin color for jaw d) extraction skin color for nose e) extraction skin color for mouth f) mask detection.

At last, we take the image for child at Fig. 15. Here Sample 7 is for non-mask images and Sample 8 for mask images. Where same procedure is applied, which is used for both male and female. For child the accuracy rate for mask images is 99.85%and for non-mask images 99.31%.

The performance of our work is based on detection accuracy and masked, non-masked images position.Performance evaluation of our work is shown in Table I. In Table I, to evaluate the performance we have worked with two types of images that are mask and non-mask. Number of mask images are 350 where positively response 320 and detection accuracy 98.03%. And number of non-mask images are 450where positively response 450 and detection accuracy 99.99%.

Table 1. Performance evaluation of proposed framework

Image Type	No. of image	Positively Responded	Accuracy rate
Mask	350	320	98.03%
Non-Mask	450	450	99.99%

At our comparison table, we established detection rate for both mask and non-mask images. This detection rate is for female, male and child. Total number of images and how many of them are correctly detected are given in this table. By those images, we founded here accuracy rate for female, male and child. Between the mask and non-mask images, we get better accuracy rate for non-mask images. Comparison table is shown in Table II.

Table 2. Comparison table for face mask detection

SL No	Type	Detection rete of mask images			Detection rate for non-mask images		
		No of imag-es	Correc-tly detected	Accur-acy	No of imag-es	Correc-tly detected	Accur-acy
1	Male	120	112	98.94%	170	170	99.94%
2	Female	180	177	99.22%	220	219	99.50%
3	Child	50	48	99.85%	60	58	99.31%

For non-mask images, we accurately detected all images of male which is shown in our comparison table and the accuracy rate is 99.94%. When we detected the female images we get less accuracy from male detection rate for both mask and non-mask images. As like as female and male, child detection for non-mask is less than from mask images accuracy rate, i.e. 99.85%. Comparatively, we are able to detect the images for female more than male and child.

4. CONCLUSION

In this work, we have proposed a face mask detection system from the input image. For that, the facial image is detected using the image processing techniques. After that, the facial feature of jaw, nose and mouth region is extracted to recognize the face mask. Finally, the face mask region is recognize through the analyzing the skin color conversion. The system shows better performance on the different environmental images and shows the accuracy of 98.03% on the mask image and 99.99% on the non-mask images. The proposed method is least complicated in structure and gives instantaneous results.

5. REFERENCES

[1] Huang, G.B., Mattar, M., Berg, T. and Learned-Miller, E., "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," In Workshop on faces in 'Real-Life' Images: detection, alignment, and recognition, October, 2008.

[2] Zhao, Z.Q., Zheng, P., Xu, S.T. and Wu, X., "Object detection with deep learning: A review," IEEE transactions on neural networks and learning systems, 30(11), pp.3212-3232, 2019.

[3] Grassi, M. and Faundez-Zanuy, M., "Face recognition with facial mask application and neural networks," In International Work-Conference on Artificial Neural Networks, pp. 709-716, Springer, Berlin, Heidelberg, June, 2007.

[4] Cabani, A., Hammoudi, K.,Benhabiles, H. and Melkemi, M., "MaskedFace-Net–A dataset of correctly/incorrectly masked face images in the context of COVID-19," Smart Health, 19, p.100144, 2020.

[5] Meenpal, T., Balakrishnan, A. and Verma, A., "Facial mask detection using semantic segmentation," In 2019 4th International Conference on Computing, Communications and Security (ICCCS), pp. 1-5, IEEE, October, 2019.

[6] Inamdar, M. and Mehendale, N.,"Real-Time Face Mask Identification Using Facemasknet Deep Learning Network," Available at SSRN 3663305, 2020.

[7] Lin, K., Zhao, H., Lv, J., Li, C., Liu, X., Chen, R. and Zhao, R., "Face detection and segmentation based on improved mask R-CNN," Discrete Dynamics in Nature and Society, 2020.

[8] Wang, Z., Wang, G., Huang, B., Xiong, Z., Hong, Q., Wu, H., Yi, P., Jiang, K., Wang, N., Pei, Y. and Chen, H., "Masked face recognition dataset and application," arXiv preprint arXiv:2003.09093, 2020.

[9] Jiang, M. and Fan, X., "RetinaMask: a face mask detector," arXiv preprint arXiv:2005.03950, 2020.

[10] Hariri, W., "Efficient masked face recognition method during the covid-19 pandemic," *Signal, image and video processing*, *16*(3), pp.605-612,2022.

[11] Li, S.Z. and Lu, J., "Face recognition using the nearest feature line method," IEEE transactions on neural networks, 10(2), pp.439-443, 1999.

[12] Qin, B. and Li, D., "Identifying Facemask-Wearing Condition Using Image Super-Resolution with Classification Network to Prevent COVID-19", Sensors, 20(18), p.5236, 2020.

[13] Kortli, Y., Jridi, M., Al Falou, A. and Atri, M., "Face recognition systems: A Survey," Sensors, 20(2), p.342, 2020.

[14] Schroff, F., Kalenichenko, D. and Philbin, J., "Facenet: A unified embedding for face recognition and clustering," In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 815-823, 2015.

[15] Deng, J., Guo, J.,Xue, N. and Zafeiriou, S., "Arcface: Additive angular margin loss for deep face recognition," In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4690-4699, 2019.

[16] Liu, W., Wen, Y., Yu, Z. and Yang, M., "Large-margin softmax loss for convolutional neural networks," In ICML, Vol. 2, No. 3, p. 7, 2016.

[17] Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B. and Song, L., "Sphereface: Deep hypersphereembedding for face recognition," In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 212-220, 2017.

[18] Ye, Q., "Masked Face Detection Via a Novel Framework," In 2018 International Conference on Mechanical, Electronic, Control and Automation Engineering (MECAE 2018), Atlantis Press, March, 2018.

[19] Chavda, A., Dsouza, J., Badgujar, S. and Damani, A., "Multi-Stage CNN Architecture for Face Mask Detection," arXiv preprint arXiv:2009.07627, 2020.