Proceedings of the 3rd International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/6/20230761

# Face recognition with masks in the lasting era of COVID-19

**Longjie Fang**

Faculty of Information Technology, Beijing University of Technology, No.100 Pingleyuan, Chaoyang District, Beijing, China

fanglongjie@emails.bjut.edu.cn

**Abstract.** At the end of 2019, the pandemic began to have a huge impact on people's lives around the world. In order to control and prevent the virus from spreading, world-wide governments have encouraged all residents to wear masks in their daily lives. Since the 1960s, face recognition technology has gradually emerged and developed. In the 21st century, it has been widely used in security, finance, transportation, and even retail, advertising, smart devices, education, health care, entertainment and other fields, playing a pivotal role in everyone's life. In the era of COVID-19, the public wearing face masks has undoubtedly hindered the smooth technology application of face recognition. The technology of recognizing human face in the case of wearing masks should be developed, followed by the adjustment of the original technology and data. The development of technology also brings certain social impact. Our research reviews the adjustment and progress of face recognition technology in the era of COVID-19, summarizes the impact of technological progress on society and people's lives, and provides help for the government to better prevent and control the epidemic.

**Keywords:** masked face recognition, neural networks, occluded face detection, deep learning

## 1. Introduction

qiFrom the end of 2019, the COVID-19 pandemic gradually began to affect the lives of people around the world. So far this year, more than 1 million deaths worldwide have been attributed to the coronavirus. At the same time, the novel coronavirus is prone to produce variants and recombination among variants. Five types of mutant viruses α, β, γ, δ, η [1] and three types of mutant recombination viruses have been produced. These mutant viruses have spread so rapidly from their origin to many other countries that rapid and effective prevention and control measures must be adopted to control them [1]. Governments have adopted a number of effective and aggressive measures, such as local, national and international travel restrictions, bans on dinner parties, distance learning, work, quarantine and wearing masks. Because of the new type of coronavirus main route of transmission is airborne, infection by viral particles in the air is inhaled into the lungs [1], wearing masks became one of the most important prevention measures, almost in many cases make traditional facial recognition technology failure, for example, community entrance guard, facial security of access control, facial attendance, the railway station, etc., Which has greatly affected people's work and life and the operation of society. Therefore, face recognition technology urgently needs to be upgraded and adjusted to cope with this change. Since early 2020, the face recognition research of having masks on the basis of the existing face recognition technology gradually developed, the main work related to face recognition algorithm and training

DOI: 10.54254/2755-2721/6/20230761

algorithm of data set for the corresponding adjustment and transformation, now wear masks face recognition technology has many relatively mature robust model is established and reliable data set, For scholars in the epidemic era to carry out more in-depth research on this basis, and the development of this technology has also brought a very positive impact on society and people's life.

This article summarizes the main implications of the pandemic for the composition of face recognition technology in masks and the potential for the technology itself to help fight the virus. Face recognition performance is affected by the adoption of face masks that cover the mouth and nose, with secondary effects of rigorous hygiene measures adopted to keep the spread of COVID-19 under control. Therefore, new research is being in demand in order to ensure high recognition accuracy. Masks block important features of a person's face, leading to poor recognition performance.

## 2. Impact on existent face recognition technologies

In the era severely affected by the coronavirus pandemic, masked face recognition technology is particularly important, especially in the hospital environment where strictly wearing masks is required. Face recognition technology is widely used in edge control. We do not expect that wearing masks will have great influence on the precision of face recognition. Adopting face masks has a severe impact on face recognition technology, because face masks cover very important parts of the human face, including the nose and lips, making face recognition extremely difficult. A study led by NIST illustrates this problem. The study tested a variety of existing face recognition models on a dataset of images of faces wearing masks of various colors and shapes. The results showed that even the most robust models were able to misidentify faces up to 50% of the time. At the same time, different mask colors and shapes make face recognition difficult, and there is a lack of face mask wearing data sets. There are also many types of masks, including medical masks, printed masks, cloth masks and transparent masks, as the examples listed in figure 1. Many synthetic virtual masks appear in many datasets designed for face recognition model training with masks. Whether masks can be detected or even the type of masks is an important issue for masked face recognition technology. Even before COVID-19 era, technology of face recognition has studied impact of wearing glasses, sunglasses, hats and other facial occlusions on the recognition accuracy. In many cases, the presence of a face cannot be detected directly by wearing a mask. Therefore, it is of great significance to detect masks and their types for face recognition technology to restore faces and detect whether masks are worn correctly.

DOI: 10.54254/2755-2721/6/20230761

**Figure 1.** Masks of different types[1234].

Extracting features is the key step in the face recognition progress, which is aimed at extracting several face features with sufficient recognition power to learn and represent key facial properties, like mouth, nose and eyes. With the advent of masks, the pipeline becomes more complex, requiring extracting more representative and robust facial features. In mask face recognition technology, methods of extracting features are segmented into shallow representation and deep representation, where Shallow feature extraction is a classic feature extraction way, which clearly generates groups of manual features with low learning ability or optimization mechanism. Some of the shallow representation methods use hand-crafted low-level features to find occluded local parts and remove them from recognition [2]. Lbp [3], SIFT [4] and HOG [5] are most frequently-used descriptors that can represent local features, holistic learning, and shallow learning methods. They achieve significant precision and robustness for face variations such as affine, illumination, rotation, scaling, and the translation in non-occluded face recognition tasks. However, the performance of shallow features deteriorates when coping with occludes ,masks included, and the deep representations obtained by deep learning models are largely superior to shallow features. Many methods have been designed and assessed to extract features from the face with the use of deep learning.

## 3. Adjustment of datasets

Because of the technology of masked face recognition is mostly depend on depth study and machine learning, need pictures of people wearing face masks as its training model and algorithm of data set, and that before the start of the outbreak of the COVID-19, this kind of data set is applied to a broader more subject - the mask cover the face recognition technology research [6], Therefore, there is not a large number of face image datasets with masks that can be used, as a result of which, obtaining such data integration is a very important topic for face recognition technology with masks. Many researchers who study face recognition technology with masks have compiled their own datasets, and others have designed techniques for synthesizing face images with virtual masks in order to quickly generate usable face samples with masks. Wang et al. [7] proposed three datasets to improve the precision of masked face recognition. They have the Mask Face Detection Dataset (MFDD) [8], the Simulated Mask Face Recognition Dataset (SMFRD) and also the Real Mask Face Recognition Dataset (RMFRD) included. As the authors stated, the MFDD contains 24771 masked faces. The specific dataset, which tags people

[1] https://www.flair.be/fr/chillax/creez-votre-masque-sourire-a-partir-dune-photo-de-vous/
[2] https://www.bbc.com/zhongwen/simp/world-52760335
[3] https://www.gvm.com.tw/article/79578
[4] https://lifestyle.bg/tendencies/apple-clearmask-i-kak-kompaniyata-zapochva-da-proizvezhda-i-predpazni-maski.html

based on whether they wear masks, is primarily used to train and evaluate mask detectors. The three datasets have The RMFRD dataset including 5000 photos of 525 masks-wearing people and 90000 photos of 525 masks-not-wearing people. The images marked as masks-wearing included faces with the correct mask, but also included faces whose masks were improperly worn and therefore did not comply with the stipulations stated in the coronavirus pandemic era. The SMFRD dataset, short for simulated masked face dataset, is a simulated dataset images of human face containing 500000 facial images from 10000 subjects, with masks manually added to simulate the face wearing a mask.

Ge et al. [9] proposed MAFA as a masked face detection dataset constructed depending on the resources on the Internet. On the other hand, MAFA contains 30811 images of appearance changes in reality. The images in the dataset are not annotated by if the mask is properly placed, but by the type and degree of occlusion. MAFA datasets are also often used in studies of mask detection along with unmasked faces datasets [8], but as the annotation issues spoken of above, this usually leads to detecting techniques that are able to detect ordinary occlusions of faces rather than masks.

Although the dataset above does not distinguish much between correctly and incorrectly masked faces, on the Kaggle portal, there are two other smaller datasets available, which can make this distinction [10,11]. The first is the Mask Detection (FMD) [10] dataset containing 853 photos segmented into 3 categories :(i) with a mask, (ii) without a mask, and (iii) wearing the wrong mask. A sum of 4072 facial images were annotated in the dataset (with mask: 3232, without mask :717, wearing the wrong masks :123). Nonetheless, the main challenge within the dataset is that a great number of the faces are not big enough, that limiting the employment of this dataset for the evaluation and development of mask detectors. And the second one contains 6,024 images [11] obtained in an unconstrained setting. The dataset was consisted of a specific focus on diversity, including people of diverse races, regions, and ages. Images in this dataset were labeled by 20 categories, faces with masks, faces with the wrong mask, covered faces, faces without masks and other categories of masks, scarves, glasses and other covering things included. 4326 out of the 6024 images are annotated and can be trained. The rest of the images were annotated later as separate, named the Medical Mask Dataset [12] (MMD). It has a sum of 9067 labeled faces associated within the study mask detection question, namely 224 faces with the wrong mask, 2085 faces without wearing a mask, and 6758 faces wearing a mask.

## 4. Emerging technologies

Few steps are involved in the progress of Masked Face Recognition (MFR). After pre-processing the images, deep learning models are used to cope with the masked face datasets. Features are extracted from masked faces, and the masked face must be able to be detected. Then the faces with masks are restored, matched and recognized. It has several ways different from the traditional Face Recognition (FR) technologies, the specific improvement of which are stated below.

### *4.1. Image pre-processing*

The accuracy of FR systems, wearing masks or not, is heavily impacted by the essence of the face images used in the training, validation and testing phases. Few publicly available datasets, including facial image pairs with and without masked objects, are available to adequately train MFR systems in a progressive way. Thus, it reinforces the demand to enrich the testbed by adding simulated images of different types of masks [13,14]. The most popular methods used for mask synthesis are deep convolutional neural network (DCNN)[15], identity mask GAN (IAMGAN) [16] and CYCLE-GAN [17]. Images are also widely pre-processed using data enhancement techniques, through which many operations can be applied to increase the number and variation of images, for example, image flipping, cropping, alignment and rotation. Other enhancement processes are used to improve the quality of the image representation as well, such as segmentation, image scaling, smoothing, or noise removal. In addition, the image is able to be adjusted to improve its sharpness.

## *4.2. Deep learning models*

*4.2.1. CNN (Convolutional neural network).* CNN is an efficient neural network, which has exposed its advantage within a great range of applications such as image classifying, object detecting and recognizing. CNNS is usually composed of the input layer, the convolution layer and output to control the degree of displacement, scaling and distortion [18]. They can effectively learn from the training data all kinds of differences within the group, such as lighting, posture, facial expression and age. One of the most familiar pre-training models used successfully in face recognition issues is AlexNet [19]. AlexNet reduces training time and minimizes errors due to the availability of an integrated graphics processing unit (gpu). VGG16 and VGG19 [20] are familiar CNN-based models that have been used in diverse computer vision applications as well, face recognition included. Although the precision of this method is high, there are still some problems in training time and complexity. If more complex image recognition tasks are encountered, deeper neural networks are needed to process them. At the same time, this will make the training more complicated, leading to accuracy attenuation. To conquer this problem, the Residual Network (ResNet) [21] stacking additional layers to achieve higher performance was proposed. However, the added layers only have the ability to learn complicated features on the basis of being intuitionally determined to control the model performance degradations. The example of CNN is shown in figure 2.

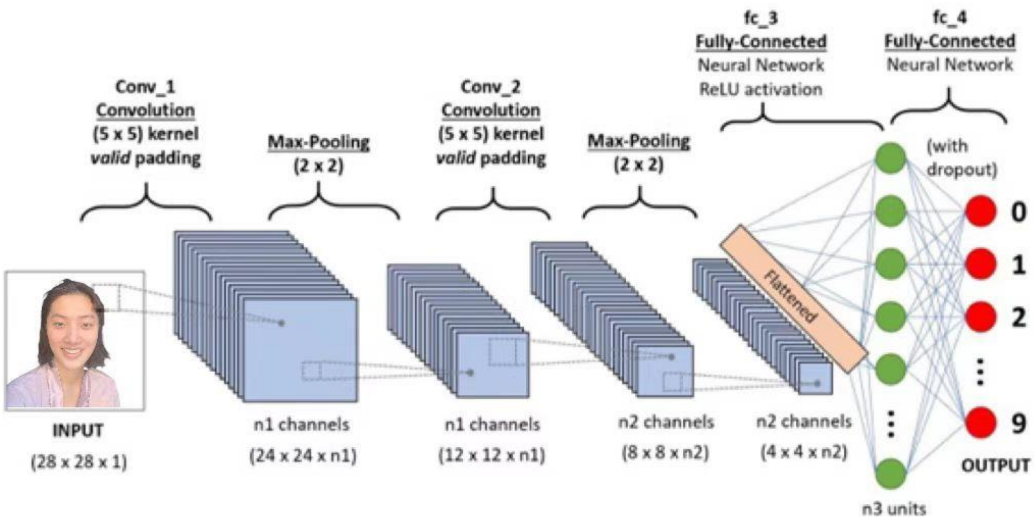


**Figure 2.** CNN [35].

*4.2.2. .Deep networks specifically for MFR*. A great number of deep learning architectures have been designed or tuned specifically for OFR or FR tasks, and they contribute significantly to the improvement in detection accuracy. FaceNet[22] maps images to Euclidean space through neural networks and constructs face embeddings according to triplet loss. When it comes to images belonging to the identical person, the distance in Euclidean space between them is small, while when images belong to different people, the distance between them will be large. This feature enables FaceNet to perform different tasks such as face detection, clustering, and recognition. Hereface [23] is another popular FR system that provides geometric interpretation and enables cnn to learn Angle recognition features, which makes it very effective in face recognition learning. ArcFace [24] is also an efficient FR network based on similarity learning, which replaces the soft maximum loss with the Angle loss. It uses cosine similarity to calculate the distance between images and find the minimum distance. Deng et al. [25] also proposed MFCosface as an MFR algorithm based on cosine loss. This method can effectively solve the problem of low face recognition rate under mask occlusion by detecting the key facial features of the face wearing mask. MFCosface relies on large cosine losses as well. By adding attention mechanism to the model, the representation of facial features is optimized.

### *4.3. Extraction of feature*

Song et al. [26] proposed a mask learning strategy with multiple steps mainly based on CNN in order to find and eliminate corrupted facial features from FR. Lots of other attention and context-aware methods use additional subnets to extract image features to capture important facial regions. The graphical image representation method of deep mapping convolutional network (GCN) is also applied in the field of mask face detection, reconstruction and recognition. GCNs has shown high capability in learning and processing face images using spatial or spectral filters built for shared or fixed graph structures, the example of which is as in figure 3. However, learning graphical representations is often limited by the number of GCN layers and unfavourable computational complexity.

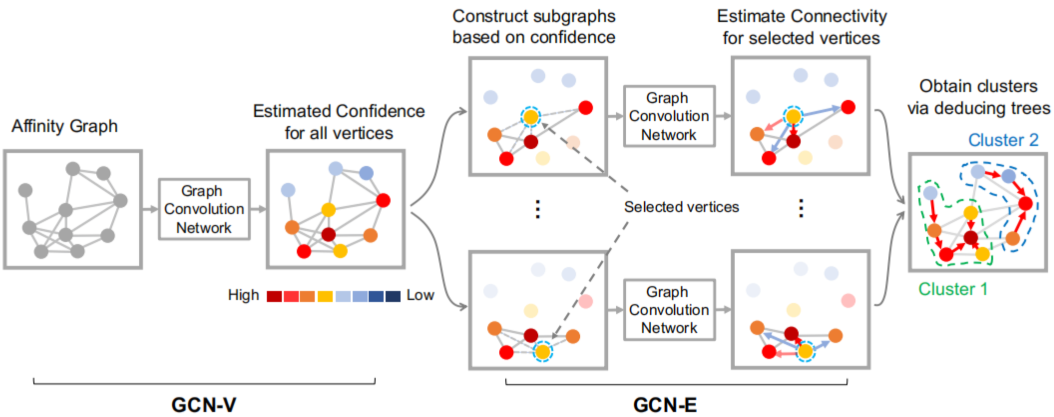


**Figure 3.** GCN [36].

### *4.4. Masks detecting*

Masks in the recent time have become one of the most quotidian things used to cover the face and come in different styles, textures, sizes and so on, which reinforces the need for deep learning model training to precisely detect masks covered on people's faces. Most of the exuberant detection methods are usually used for object detection, which are adjusted and studied in mask detection tasks. Regions with CNN features (R-CNN) [27] have been adopted globally in the field of object detection, where deep convolutional networks are used to classify object proposals. With mask-covered face image datasets, R-CNN extracts numerous facial regions by putting them into the CNN network and select an appropriate search algorithm for application, which gathers feature vectors for each region. Then objects in proposed candidate facial regions extracted from the extracted features will be classified by support vector machine (SVM). Fast R-CNN [28] and Faster R-CNN [29] were adopted to improve accuracy by changing the R-CNN architecture as well. Nonetheless, the methods' disadvantages is not oblivious, for example, the training process is a pipeline with multiple stages, so it would cost a great much time and space. In addition, Zhang et al. [30] came up with a context-focused R-CNN as a detection framework for mask wearing. The framework expands the intra-class distance and reduces the inter-class distance by extracting features, as a result of which, mask detection based on segment-based deep networks has become the focus of research. Fully Convolutional Neural network (FCN) [31] only performs convolution operations, which is consisted only of a simple layer that mainly used for CNN-based autoencoders. U-Net [32] was also first used for biometric division, but is largely used in many computer vision applications, including face detection. It has an encoder included that uses convolutions and maximum pooling layers to capture the image context, while a decoder that uses transposed convolution to upsample the encoded information. The feature map from the encoder is connected to the feature map of the decoder afterwards, which accelerates learning contextual information (the relationship between

DOI: 10.54254/2755-2721/6/20230761

image pixels) better. Other effective OFR or MFR methods have also been provoked in the research done before and mentioned in those paper works. Ge et al. [33] proposed an LLE-CNN to detect masked faces by extracting candidate facial regions combined with pre-trained cnn and representing them with high-dimensional descriptors. What comes next is that the local linear embedding module forms the face descriptors into weight vectors to recover the missing face cues in the occluded region. Ieamsaard et al. [34] designed the deep learning model that can work to detect masks, which was trained on YoloV5 for five separate periods. YoloV5 was used with CNN to testify the presence of masks and whether they were properly worn.

## 5. Conclusions

This article summarizes the main implications of the pandemic for the composition of face recognition technology in masks and the potential for the technology itself to help fight the virus. Face recognition performance is affected by the employment of face masks that cover the mouth and nose, with secondary effects of strict hygiene measures implemented to take control of the spread of virus. Therefore, new research is needed to ensure high recognition accuracy. Masks block important part of a person's facial features, contributing to poor recognition accuracy. It also has a lot to do with the type of mask; This applies not only to opaque masks, but to transparent masks as well, because the changes caused by the reflection of light are not trivial to the model. That said, the obstacles of ensuring dependable biometric accuracy have increased substantially in the era of COVID-19, requiring more renewed researches. Now with many people working from home or studying, some of the biggest problems are related to the employment of biometrics in unsupervised, remote verification scenarios. In turn, this makes continuous authentication, biometric template protection or presentation attack detection even more important to ensure privacy and security in this environment, thus defining the era of COVID-19.

## References

[1] Zhao, Y.; Huang, J.; Zhang, L.; Chen, S.; Gao, J.; Jiao, H. The global transmission of new coronavirus variants. Environ. Res. 2022, 206, 112240.

[2] Min, R.; Hadid, A.; Dugelay, J.-L. Improving the recognition of faces occluded by facial accessories. In Proceedings of the 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG), Santa Barbara, CA, USA, 21–25 March 2011; pp. 442–447.

[3] Ahonen, T.; Hadid, A.; Pietikainen, M. Face description with local binary patterns: Application to face recognition. IEEE Trans. Pattern Anal. Mach. Intell. 2006, 28, 2037–2041.

[4] Geng, C.; Jiang, X. Face recognition using sift features. In Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 7–10 November 2009; pp. 3313–3316.

[5] Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–26 June 2005; pp. 886–893.

[6] Wang, J.; Yuan, Y.; Yu, G. Face attention network: An effective face detector for the occluded faces. arXiv 2017, arXiv:1711.07246.

[7] Wang, Z.; Wang, G.; Huang, B.; Xiong, Z.; Hong, Q.; Wu, H.; Yi, P.; Jiang, K.; Wang, N.; Pei, Y.; et al. Masked face recognition dataset and application. arXiv 2020, arXiv:2003.09093.

[8] Daniell Chiang, Y.F. Detect Faces and Determine Whether they are Wearing Mask. Available online: https://github.com/ AIZOOTech/FaceMaskDetection (accessed on 26 November 2020).

[9] Ge, S.; Li, J.; Ye, Q.; Luo, Z. Detecting masked faces in the wild with LLE-CNNs. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2682–2690.

[10] Larxel. Face Mask Detection, 853 Images Belonging to 3 Classes. Available online: https://www.kaggle.com/andrewmvd/facemask-detection (accessed on 26 November 2020).

[11] Wobot Intelligence. Face Mask Detection Dataset. Available online:

https://www.kaggle.com/wobotintelligence/face-maskdetection-dataset (accessed on 26 November 2020).
[12] Humans in the Loop. Medical Mask Dataset. Available online: https://humansintheloop.org/medical-mask-dataset (accessed on 26 November 2020).
[13] Din, N.U.; Javed, K.; Bae, S.; Yi, J. A novel GAN-based network for unmasking of masked face. IEEE Access 2020, 8, 44276–44287. [CrossRef]
[14] Din, N.U.; Javed, K.; Bae, S.; Yi, J. Effective Removal of User-Selected Foreground Object from Facial Images Using a Novel GAN-Based Network. IEEE Access 2020, 8, 109648–109661.
[15] Hooge, K.D.O.; Baragchizadeh, A.; Karnowski, T.P.; Bolme, D.S.; Ferrell, R.; Jesu-dasen, P.R.; Castillo, C.D.; O'toole, A.J. Evaluating automated face identity-masking methods with human perception and a deep convolutional neural network. ACM Trans. Appl. Percept. (TAP) 2020, 18, 1–20.
[16] Zhu, J.-Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
[17] Wang, Z.; Wang, G.; Huang, B.; Xiong, Z.; Hong, Q.; Wu, H.; Yi, P.; Jiang, K.; Wang, N.; Pei, Y.; et al. Masked face recognition dataset and application. arXiv 2020, arXiv:2003.09093.
[18] Setiowati, S.; Franita, E.L.; Ardiyanto, I. A review of optimization method in face recognition: Comparison deep learning and non-deep learning methods. In Proceedings of the 2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE), Phuket, Thailand, 12–13 October 2017; pp. 1–6.
[19] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 2012, 25, 1097–1105.
[20] Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.
[21] 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, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
[22] Schroff, F.; Kalenichenko, D.; Philbin, J. Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 815–823.
[23] Liu, W.; Wen, Y.; Yu, Z.; Li, M.; Raj, B.; Song, L. Sphereface: Deep hypersphere embedding for face recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 212–220.
[24] Deng, J.; Guo, J.; Xue, N.; Zafeiriou, S. Arcface: Additive angular margin loss for deep face recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 4690–4699.
[25] Deng, H.; Feng, Z.; Qian, G.; Lv, X.; Li, H.; Li, G. MFCosface: A masked-face recognition algorithm based on large margin cosine loss. Appl. Sci. 2021, 11, 7310.
[26] Song, L.; Gong, D.; Li, Z.; Liu, C.; Liu, W. Occlusion robust face recognition based on mask learning with pairwise differential siamese network. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019; pp. 773–782.
[27] Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
[28] Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
[29] Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. Adv. Neural Inf. Process. Syst. 2015, 28, 91–99.
[30] Zhang, J.; Han, F.; Chun, Y.; Chen, W. A Novel Detection Framework About Conditions of

Wearing Face Mask for Helping Control the Spread of COVID-19. IEEE Access 2021, 9, 42975–42984.
[31] Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
[32] Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
[33] Ge, S.; Li, J.; Ye, Q.; Luo, Z. Detecting masked faces in the wild with lle-cnns. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2682–2690.
[34] Ieamsaard, J.; Charoensook, S.N.; Yammen, S. Deep Learning-based Face Mask Detection Using YoloV5. In Proceedings of the 2021 9th International Electrical Engineering Congress (iEECON), Pattaya, Thailand, 10–12 March 2021; pp. 428–431.
[35] Sumit Saha. A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. Towards Data Science,16 December 2018.
[36] Yang, L., Chen, D., Zhan, X., Zhao, R., Loy, C.C., Lin, D.: Learning to cluster faces via confidence and connectivity estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2020)