Research Article

Impact on Biometric Identification Systems of COVID-19

Yuheng Guo

Electrical & Computer Engineering, University of California–Los Angeles, Los Angeles 90095, CA, USA

Correspondence should be addressed to Yuheng Guo; yuhengguo@g.ucla.edu

Received 13 September 2021; Revised 1 November 2021; Accepted 19 November 2021; Published 2 December 2021

Academic Editor: Rahman Ali

Copyright © 2021 Yuheng Guo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

COVID-19 has had an inevitable impact on the daily life of people in 2020. Changes in behavior such as wearing masks have a considerable impact on biometric systems, especially face recognition systems. When people are aware of this impact, a comprehensive evaluation of this phenomenon is lacking. The purpose of this paper is to qualitatively evaluate the impact of COVID-19 on various biometric systems and to quantitatively evaluate face detection and recognition. The experimental results show that a real-world masked face dataset is essential to build an effective face recognition-based biometric system.

1. Introduction

The COVID-19 global pandemic may be one of the most persistent and deadly airborne viruses to exist in human history. People exposed to the virus have the potential to be infected at will, and everyone in the world can feel the momentum of how a pandemic can change the way we live. Carriers of the virus may be asymptomatic but contagious at the same time, which allows the virus to spread silently to humans. According to the World Health Organization, as of 10:27 a.m. EDT on Oct. 6, 2020, there have been 35,274,993 confirmed cases of COVID-19 worldwide, including 1,038,534 deaths. The new coronavirus has become the number one obstacle to human evolution.

The virus can certainly be transmitted by touching objects, so keeping social distance and staying at home are now the new common sense. Handshakes and other social activities are currently prohibited. In addition, most public regulations require people to wear masks before showing their faces. However, wearing a mask hides key information for face recognition, which makes it more difficult for normally trained models to process.

Face recognition using partially covered faces has not been fully addressed so far. Several approaches have been proposed to solve this problem. According to [11], to solve the technical challenge of recessed face recognition, Amazon has developed palm print recognition technology that can read palm prints and internal structures of blood vessels, tissues, or bones. Moreover, at the algorithmic level, other hard-coded methods, including printing faces on masks, are not ideal solutions. Therefore, there is an urgent need for a new deep learning model for partial coverage face recognition.

Biometric systems are necessary during COVID-19 and become even more important when we want to track the spread of the virus. The impact of COVID-19 is not limited to face recognition. Different biometric systems are affected to different degrees. As discussed in [2], noncontact technologies such as face and iris recognition are now being pushed to new heights to expand government protection, surveillance, and screening. In contrast, applications that rely on fingerprint and vein recognition modalities are suffering significant losses.

In this paper, we first provide a comprehensive review of the impact of COVID-19 on different biometric systems. The performance of existing solutions and new potential solutions is then discussed, using the masked face recognition task as an example.

Our contributions can be summarized as follows:

1. We present a state-of-the-art review of the impact of COVID-19 on different biometric systems
2. We evaluate the performance of recent deep learning models in masked face detection and recognition tasks
In this section, we briefly review the different biometric systems. We also summarize the advantages and disadvantages of different biometric systems under COVID-19 in Table 1. We will leave the discussion of face recognition to the next section.

2.1. Sclera Vasculature and Periocular Recognition. Scleral recognition is an emerging field of recognition that allows machines to determine the identity of a person by the tendency of veins in the sclera. According to [3], the research team collected a new dataset of ocular biometric features through a high-resolution camera SBVPI for model training. After segmentation of each region of the sclera, the UNet, RefineNet-50, RefineNet-101, and SegNet models were used for training, with accuracies of 0.938, 0.960, 0.955, and 0.949, respectively. In addition, the team divided SBVPI into different groups by age and gender to investigate whether these factors affect scleral vessel identification.

Similarly, periocular recognition applies to the recognition process around the eye. According to [4], a dual-stream convolutional neural network model can be applied to the periocular region to obtain more color information, which means that the network can accept RGB images. In addition, a new descriptor, the orthogonal combinatorial-local binary coding model, is proposed. The research team also proposed a new database, the Ethnic-Eye database, for the future development of periocular recognition. The accuracy of Rank 1 was 84.79 ± 1.9% and Rank 5 was 94.23 ± 1.3%.

According to the theory of [5], learning label smoothing regularization (L2SR) is another new approach to improve the performance of convolutional neural networks. Combined with cross-entropy and KL divergence loss function, this model can reduce intra class variance better. This model tested three datasets, Ethnic, Pubfig, and IMDB Wiki, among which Pubfig database has the highest accuracy of 97.53 ± 2.15%. Among other regularization methods, L2SR regularization has a clear advantage.

2.2. Finger Vein Recognition. Finger vein recognition has the advantages of security and privacy. There is no need for people to expose too much personal information. Collectors are usually portable and practical. According to [6], CPBFL works in two phases. Both the training phase and the testing phase use pixel difference vector (PDV) to extract features from the images. In the training phase, the PDV is directly transformed into binary features by category-preserving binary feature learning (CPBFL); while in the testing phase, CPBFL is used to generate mappings that convert the PDV into binary features. During the training process, the model uses binary codebook learning (BCL) to cluster binary features into codebooks for the histogram representation of finger veins in the testing phase. The accuracy of CPBFL-BCL is 98.48%, 98.95%, and 99.30% for the number of samples per category of 1, 2, and 3, respectively.

According to [7], Bosporus is another finger vein recognition dataset. In the preparation stage, before finger recognition, the entire binary hand image is hand normalized to generate the left finger contour and the right finger contour. To improve the accuracy, the research team collected 30 features for each finger and applied a greedy algorithm. The features were then classified using k-nearest neighbor and random forest. An accuracy of 96.56% was achieved for the right-side images and 95.92% for the left-side images.

According to the theory of [8], FPAD is a physical fingerprint acquisition method consisting of near-infrared (NIR), short-wave infrared (SWIR), laser scatter contrast imaging, and NIR backlight. The datasets used for recognition are the PADISI-Finger and LivDet datasets, and a fully convolutional neural network model is used for recognition. Among all combinations of physical fingerprint acquisition methods, visible light, NIR, and SWIR have the highest accuracy, all of which can reach the current level.

2.3. Gait Recognition. In order to recognize gait, both gait features and temporal features need to be extracted. Therefore, according to [9], the research group used sensors such as accelerometers and gyroscopes embedded in cell phones to collect information. Convolutional neural networks and recurrent neural networks were used to continuously extract gait-related features while maintaining the time series, and then, a long short-term memory network (LSTM) was used for gait recognition. The research team conducted experiments on 118 subjects by collecting information on their cell phones and achieved an accuracy of 93.5% in recognizing people.

2.4. Handwritten Signature Biometrics. Handwritten signature recognition faces a shortage of databases [10], and the DeepSignDB database provides researchers with more handwritten signatures available for training deep learning models. The research team used the signatures in DeepSignDB and passed them to the DTW and RNN models. After training, the model will be able to identify random forgeries. For skilled forgeries, bidirectional GRU detection is used. Accuracy rates of up to 8% are achieved for skilled forgeries and 1-2% for random forgeries.

2.5. Palm Recognition. To generate more realistic palm images by machine, the research team used a regular generative adversarial network. The innovation of this work is to use TV regularized GAN as a loss function to help the model...
learn more features of palm line connectivity [11]. Over 500 iterations, the discriminator helped the generator improve its ability to simulate palm images, and the generator’s loss function went from 17.5 at the beginning to about 5.0 at the end. As the number of each iteration increases, the palm images become clearer and more realistic.

2.6. Wrist Vein Recognition. With the NIR camera and NIR LED embedded in modern cell phones, the structure of wrist groove lines can be observed. In [12], more than 100 data feature points of the wrist grooves were extracted from the images and 2400 images of wrist veins were acquired. Three algorithms, SIFT, SURF, and ORB, were used in the data feature point acquisition process. After data extraction, two algorithms were used to process the dataset. tgs-cvbr algorithm was mainly used in the preparation phase, where the machine needs to identify the exact location of the wrist vein. pi-cvbr algorithm extracts features from the images and produces the correct decision output. Among the three data extraction algorithms, SIFT has the highest EER extraction accuracy between 6.82%~18.72%.

3. Masked Face Recognition

In this section, we use masked face recognition as a case study. We first outline the work on processing masks. Then, we present the experiments we conducted to evaluate state-of-the-art face detection/recognition models when masked faces are present.

3.1. Overview of Masked Face Recognition-Related Work. Partial face recognition has emerged as the new hotspot for biometrics, while other biometrics requiring physical contact, including fingerprint recognition and vein recognition, are set to face a general decline in biometrics. As potential opportunities for partial face recognition emerge, it is necessary to summarize the efforts to handle masked faces [2].

Residual neural networks are best suited to detect and extract features from a given database, and in [13], the authors implemented ResNet50 as a tool for extracting features from masked face images. The parameters of the deep migration learning model are trained by images and associated labels. To improve the accuracy, other components of the model include support vector machines, decision trees, and integration methods, which are used in the classification process. The entire model is trained based on three databases of masked face image sets RMFD, SMFD, and LFW. After completing the training process of the whole model, decision tree classification can achieve the highest test accuracy of 99.89%, support vector machine can achieve 100% test accuracy, and integration method achieves more than 99% accuracy.

During the COVID-19 city lockdown, many government decision makers proposed policies, including staying in restricted areas and limited shopping hours. In order to implement this decision, the help of technical methods is needed to identify individuals and change local authorities. In [14], the design of a three-layer transformation system IoT is proposed to provide the necessary technical power to the government. The whole IoT system includes physical layer, edge layer, face detection, bounding box, and edge computing. The physical layer is implemented by hardware devices such as cameras and smartphones to capture faces, each with a microchip, a Raspberry Pi, to update the data into a database for recording. The edge layer is then put in place to continuously collect all data from various device.

| Biometric identification method | Advantages during pandemic | Drawbacks during pandemic |
|--------------------------------|-----------------------------|---------------------------|
| Fingerprint recognition        | More accurate compared to other methods; easier for information gathering | The touching collection machine allows the virus to pass along crowds; to achieve high accuracy, facilities are in need, which adds up the complexity |
| Palm recognition               | Provide a contactless way to identify individuals | Variation of the palm line could easily cause face forgery of GAN generated patterns |
| Face recognition               | Nontouching methods that are convenient for information gathering | Wearing a mask could cause a recognition error |
| Sclera recognition             | Nontouching method identifies an individual by only a picture of the sclera | Need high-resolution cameras, which mean higher facility expense; pictures of different angles are needed for recording, which means time-consuming |
| Stride recognition             | Only need the embedded sensor in the phone to identify gait | Relatively low accuracy rate |
| Voice recognition              | The noncontact method which allows the machine to identify people without physically in touch | The environmental condition would largely affect the accuracy of the collection machine |
| Periocular recognition         | Similar to face recognition, a smartphone with a camera embedded could identify individuals | To successfully perform recognition, a higher resolution of the camera and a closer distance might be required |
| Handwritten signature          | Provide another way to solve the identification problem, and it does not need the actual bio-feature of individuals to process, which may have less discomfort | Needs extra handwritten input devices, and the collection of the handwritten script would not be as convenient as other biometric methods |
| Wrist vein recognition         | Provide real-time recognition for noncontact human wrist recognition | Relatively low recognition accuracy |

Table 1: The summary of the pros and cons of different biometric systems under COVID-19.
types, including cameras and smartphones. An application server will decide who are the residents eligible to go out shopping while keeping others in a restricted area. For face recognition, the group implemented a multitask cascaded convolutional network using the OpenCV library to successfully identify key face feature points within the bounding box. The enclosing box with low confidence level is calculated by cross-entropy loss regression and eliminated by a nonmaximal suppression process. The model is trained on P-Net, R-Net, and O-Net, and the database used is WIDER FACE with an accuracy of 0.607.

In addition to extracting face feature points as a preprocessing method for face recognition, combining the h-channel of HSV color space, face portrait, and grayscale images, i.e., the HGL method, yields better preprocessing results. According to [15], the research group used the HGL method to preprocess the MAFA database and then applied a convolutional neural network. After the color partially covered faces were imported into the model, they were converted into grayscale line portraits and color spatial images, and then feature extraction was performed using convolutional neural networks. Overall, the correct rate of frontal face recognition reached 93.64%, and the correct rate of side face recognition reached 87.17%.

Analyzing smaller parts of the face can also be used for recognition instead of using the entire unexposed face as input. In [16], regions of the face, including the left eye, right eye, and mouth, were collected from the LFW face database. To extract these specific facial features, various pretrained models were used, including AlexNet, ResNet50, ResNet101, DenseNet201, VGG-Face, and MobileNetV2. After applying the convolutional neural network to the soft biometric features classification, the optimal layer of CNN features was classified using support vector machines, such as gender, age, and race. Using only eye features as input, the accuracy rates of 92.6%, 60.2%, and 82.9% were achieved for gender, age, and race, respectively, which is an acceptable decrease in accuracy compared to the whole face recognition.

To comprehensively compare the effect of partially masked faces on face recognition, in [17], the research group analyzed the current state-of-the-art face recognition methods, ArcFace, SphereFace, and COTS, using normal and masked faces as inputs. In BLR-BLP, BLR-M1P, BLR-M2P, and BLR-M12P data segmentation, baseline true, baseline impostor, masked true, and masked impostor were compared, and the results showed that the true scores of baseline true, baseline impostor, masked true, and masked impostor had a tendency to be closer to the distribution of impostors.

3.2. Evaluation of Existing Face Detection Models. In this section, we aim to evaluate the performance of state-of-the-art deep learning models for face detection. The models we evaluate include CenterFace, LFFD, LibFaceDetection, SSD, RetinaFace, and these models are described in detail below.

With the increasing demand for facial recognition on devices with less storage space, CenterFace [18] provides researchers with an efficient and accurate method to find facial features for individuals. The main idea of CenterFace is to transform face recognition into key point prediction and build models through classification, box regression, and milestone regression. To recognize faces faster, CenterFace also includes a mobile feature pyramid network, which consists of existing models such as MobileNetv2 and feature pyramid network (FPN). CenterFace is then trained on FDDB and WIDER FACE datasets. CenterFace can achieve faster processing speed compared to DSFD, PyramidBox, S3FD, LFFD, etc. In terms of accuracy, the CenterFace model is tested on WIDER FACE on easy, medium, and hard classes, respectively, and all results can reach about 90% accuracy.

To accommodate the emerging need of processing face recognition on various portable devices with limited storage space, more and more models are designed to process faces with faster running time and higher accuracy, namely, light and fast face detector (LFFD) [19]. The main idea of LFFD is to implement RF and ERF with Gaussian characteristics, where the center of the image is clearer than the edges sharper than the edges, rather than predesigning various input anchors of different sizes. As the contextual information increases, the image grows from the middle of the image to contain more and more environment. For large size faces, the model only needs to retain face information, while for small size faces, the model needs to include more surrounding information. The main model of LFFD is a convolutional neural network, which consists of 25 layers divided into small, small, medium, and large parts. For different sizes of faces, different loss function branches are used, and there are 8 branches. The loss function of the model is softmax. The dataset used to evaluate the model is WIDER FACE. On the test set, the accuracy of easy, medium, and hard classification reaches 0.896, 0.865, and 0.770, respectively.

LibFace Detection [20] is an open source face recognition model using convolutional neural networks, with code based on FaceBoxes and SSD models. The website is attached to the reference section for reference.

YOLO has the advantage of faster processing speed compared to other two-stage detectors. However, it also has a drawback that YOLO is difficult to recognize accurately for faces with relatively small size. To address this drawback, YOLO-face [21] was designed to help YOLO recognize faces of different sizes by refining the loss function and selecting more appropriate anchor boxes. YOLO is a single-stage detector that uses a convolutional neural network to recognize the bounding box of a face. To improve the accuracy of recognizing faces of different sizes, YOLO-face uses darknet-53 as the backbone. Two types of anchor boxes are used, one keeps the general shape of YOLO anchor box but changes the anchor box to a higher shape, and the other anchor box is generated by performing k-clustering on WIDER FACE. For the loss function, YOLO uses confidence loss, regression loss, classification loss, and nonobject loss. By testing on a wider FACE and FDDB dataset, the detector can achieve higher accuracy while maintaining its processing speed.
For general target detection, the single-shot detector shows the possibility of scaling the wraparound box according to different target sizes. Single shot multibox detector (SSD) [22] has three important features, namely, a multiscale feature map for detection, a convolutional predictor for detection, and a default box and aspect ratio. With these available features, SSD can automatically scale the size of the object’s bounding box and match multiple feature maps.

RetinaFace [23] is a multilevel, single-shot face detection model. The main structure of RetinaFace consists of two parts, a feature pyramid network, and a cascaded multiscale contextual head module. For the feature pyramid network, a multiscale feature map is implemented inside the network; for the cascaded multiscale context head module, multitask losses are embedded in the model, including face classification loss, face five-marker regression loss, and face box regression loss. Experimenting with the model on a wider FACE dataset, training and testing can achieve state-of-the-art performance.

In this paper, we use two datasets to evaluate the above models.

(1) AFDB and afdb_masks: a real-world masked face recognition dataset in which face images are crawled from the Internet and further processed by reorganizing, cleaning, and adding symbols. The whole dataset includes 5000 masked faces and 90,000 unmasked normal faces.

(2) LFW_mask: a dataset of masked faces from normal faces, simulating the masked face scene by adding an overlay at the bottom of the face.

We evaluate these models based on the detection rate and detection time for each image. The results are shown in Figures 1 and 2. In these figures, we use AFDB and afdb_mask to distinguish the cases without and with masked face images.

We have two main observations from the results.

(1) Figure 1 shows the detection ratios of different datasets with different models. However, the accuracy of face recognition will decrease with the presence of masks, especially for face pictures with masks taken in real scenes. However, the accuracy of the LFW dataset with all masked faces generated by manually covering some faces decreases significantly. This shows that there is a huge difficulty difference between real masked face datasets and simulated masked face datasets in face recognition. Therefore, with sufficient budget and resources, real masked faces are more suitable for training than simulated faces.

(2) As can be seen in Figure 2, the time required to examine faces is prolonged in the case of masked faces, but the increase is within a negligible range.

3.3. Evaluation of Existing Face Recognition Models. In this part, we aim to evaluate the performance of state-of-the-art deep learning models for face recognition. We formulate this problem as a classification problem and evaluate the following models with the AFDB_masked datasets:

(1) MobileNet [24] was proposed on 17 Apr 2017 and uses depthwise separable convolutions to conduct object recognition.

(2) Xception [25] was proposed on 7 Oct 2016 and uses a structure that takes the middle point between the convolutional neural network and depthwise separable convolutions.

(3) ResNet152 [26] was proposed on 10 Dec 2015 and uses a residual learning framework to make the training for deep neural networks easier.
InceptionV3 [27] was proposed on 2 Dec 2015 to improve the computing efficiency and make the model more suitable for a bigger dataset.

EfficientNet [28] was proposed on 28 May 2019, and it is used for scaling the model ubiquitously to achieve higher performance.

We conducted four series of comparisons:

1. No mask (womasked): we train the model without the masked face images. We evaluate the model on the normal face images.

2. Masked (womasked): we train the model without the masked face images. We evaluate the model on the masked face images.

3. Both (womasked): we train the model without the masked face images. We evaluate the model on both the normal and masked face images.

4. Both (w masked): we train the model with the normal and masked face images. We evaluate the model on both the normal and masked face images.

We evaluate these comparisons in terms of accuracy. The results are shown in Table 2.

From Table 2, we have the following two observations:

1. By observing the second column, which is the accuracy of conducting human face recognition on a masked human face with models that are trained on human faces without masks. The result shows an extremely low accuracy.

2. The third column shows the resultant accuracy of training models solely on human faces without masks, and the fourth column shows the resultant accuracy on masked human faces. By comparison between the third column and the fourth column, we propose that, with the help of more masked human face pictures taken under real scenarios, models would be expected to generate better results.

4. Conclusion

In this paper, we first qualitatively evaluate the impact of COVID-19 on various biometric systems and summarize its strengths and weaknesses. This may be difficult for some systems, but it is also an opportunity for others. We then present a quantitative evaluation of face detection and recognition, with advanced deep learning models. We find that real masked faces are more suitable for training than simulated faces. We also propose that deep learning models are expected to produce better results with the help of more pictures of masked faces in real scenarios [29].

Data Availability

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] D. Kumar, “Introducing Amazon One—a new innovation to make everyday activities effortless,” 2020, https://blog.aboutamazon.com/innovation/introducing-amazon-one-a-new-innovation-to-make-everyday-activities-effortless.

[2] S. Carlaw, “Impact on biometrics of covid-19,” Biometric Technology Today, vol. 2020, no. 4, pp. 8-9, 2020.

[3] M. Vitek, P. Rot, V. Štruc, and P. Peer, “A comprehensive investigation into sclera biometrics: a novel dataset and performance study,” Neural Computing & Applications, vol. 32, pp. 1–15, 2020.

[4] L. C. O. Tiong, Y. Lee, and A. B. J. Teoh, “Periocular recognition in the wild: implementation of RGB-OCLBCP dual-stream CNN,” Applied Sciences, vol. 9, no. 13, p. 2709, 2019.

[5] Y. G. Jung, J. Park, L. C. O. Tiong, and C. Y. Low, “Periocular recognition in the wild with learned label smoothing regularization,” in Proceedings of the International Society for Optics and Photonics Twelfth International Conference on Digital Image Processing (ICDIP 2020), Article ID 115190T, Osaka, Japan, May 2020.

[6] H. Liu, G. Yang, and Y. Yin, “Category-preserving binary feature learning and binary codebook learning for finger vein recognition,” International Journal of Machine Learning and Cybernetics, vol. 11, pp. 1–14, 2020.

[7] A. Bera and D. Bhattacharjee, “Human identification using selected features from finger geometric profiles,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 50, 2017.

[8] L. Spinoulas, H. Mirzaalian, M. Hussein, and W. AbdAlmageed, “Multi-modal fingerprint presentation attack detection: evaluation on A new dataset,” 2020, https://arxiv.org/abs/2006.07498.

[9] Q. Zou, Y. Wang, Q. Wang, Y. Zhao, and Q. Li, “Deep learning-based gait recognition using smartphones in the wild,” IEEE Transactions on Information Forensics and Security, vol. 15, pp. 3197–3212, 2020.

[10] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and O. G. Javier, “Do you need more data? the DeepSignDB online handwritten signature biometric database,” in Proceedings of the 2019 International Conference on Document
Analysis and Recognition (ICDAR), pp. 1143–1148, IEEE, Sydney, Australia, September 2019.

[11] S. Minaee, M. Minaei, and A. Abdolrashidi, "Palm-GAN: generating realistic palmprint images using total-variation regularized GAN," 2020, https://arxiv.org/abs/2003.10834.

[12] R. Garcia-Martin and R. Sanchez-Reillo, "Vein biometric recognition on a smartphone," IEEE Access, vol. 8, 2020.

[13] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic," Measurement, vol. 167, Article ID 108288, 2020.

[14] M. Kolhar, F. Al-Turjman, A. Alameen, and M. M. Abualhaj, "A three layered decentralized IoT biometric architecture for city lockdown during COVID-19 o," IEEE Access, vol. 8, pp. 163608–163617, 2020.

[15] S. Li, X. Ning, L. Yu et al., "Multi-angle head pose classification when wearing the mask for face recognition under the COVID-19 coronavirus epidemic," in Proceedings of the 2020 international conference on high performance big data and intelligent systems (HPBDâ€-IS), pp. 1–5, IEEE, Shenzhen, China, May 2020.

[16] F. Alonso-Fernandez, K. H. Diaz, S. Ramis, F. J. Perales, and J. Bigun, "Soft-biometrics estimation in the era of facial masks," in Proceedings of the 2020 international conference of the biometrics special interest group (BIOSIG), pp. 1–6, IEEE, Darmstadt, Germany, September 2020.

[17] 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 Proceedings of the 2020 International Conference of the Biometrics Special Interest Group (BIOSIG), pp. 1–6, IEEE, Darmstadt, Germany, June 2020.

[18] Y. Xu, W. Yan, H. Sun, G. Yang, and J. Luo, "CenterFace: joint face detection and alignment using face as point," 2019, https://arxiv.org/abs/1911.03599.

[19] Y. He, D. Xu, L. Wu, M. Jian, S. Xiang, and C. Pan, "LFFD: a light and fast face detector for edge devices," 2019, https://arxiv.org/abs/1904.10633.

[20] libfacedet: Online: https://github.com/ShiqiYu/libfacedetection.

[21] W. Chen, H. Huang, S. Peng, C. Zhou, and C. Zhang, "YOLOface: a real-time face detector," The Visual Computer, vol. 37, pp. 1–9, 2020.

[22] W. Liu, D. Anguelov, D. Erhan et al., "Ssd: Single Shot Multibox detector," in Proceedings of the European Conference on Computer Vision, pp. 21–37, Springer, New York, NY, USA, October 2016.

[23] J. Deng, J. Guo, E. Vereras, I. Kostia, and S. Zafeiriou, "RetinaFace: single-shot multi-level face localisation in the wild," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5203–5212, Seattle, WA, USA, June 2020.

[24] A. G. Howard, M. Zhu, B. Chen et al., "Mobilenets: efficient convolutional neural networks for mobile vision applications," 2017, https://arxiv.org/abs/1704.04861.

[25] F. Chollet, "Xception: deep learning with depthwise separable convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1251–1258, Honolulu, HW, USA, July 2017.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778, Las Vegas, NV, USA, June 2016.

[27] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818–2826, Las Vegas, NV, USA, June 2016.

[28] M. Tan and Q. V. Le, "Efficientnet: rethinking model scaling for convolutional neural networks," 2019, https://arxiv.org/abs/1905.11946.

[29] WHO coronavirus disease (COVID-19) dashboard, https://covid19.who.int/, 2020.