Advances in Artificial Intelligence, Big Data and Algorithms
G. Grigoras and P. Lorenz (Eds.)

doi:10.3233/FAIA230860

A Survey on Masked Face Recognition Amid the COVID-19 Pandemic

Shanmei ZHONG[1]
Business school, Central South University, Changsha, Hunan, 410083, China

Abstract. Against the backdrop of the COVID-19 pandemic, the ubiquitous adoption of masks has caused significant hurdles to conventional face recognition technologies and their applications. In order to address these problems, a great deal of research related to masked face recognition has been put forward. Although masked face recognition has made great progress, the review or summary about this field is scanty. Therefore, this paper proposed a survey on masked face recognition. Firstly, the datasets of MFR and the mainstream deep learning methods are introduced. In addition, some common evaluation metrics are explained. Finally, this paper discusses the potential challenges of MFR and make conclusion.

Keywords: masked face recognition; face recognition; deep learning.

1. Introduction

With the emergence and spread of COVID-19 pandemic, wearing masks has proved a well-established preventive measure, which lessens the spread of infectious droplets, halts virus spread and decreases the transmissibility per contact in public situations effectively. Besides, while the cost of the intervention is cheap, the lowered transmissibility could significantly reduce the death toll and economic burden [1-3].

However, applications relying on conventional facial recognition algorithms, which have advanced significantly in recent years, were severely hampered by the use of masks [4]. For example, tremendous government buildings, schools, companies using facial recognition to secure their premises are faced with threats. Facial recognition-based immigration checkpoints that impose more intelligent border control are also coming under increased problems in pandemic. Many border controls already employ face recognition technology in conjunction with INTERPOL's information databases to identify people against a scale of accuracy. Processing face data via the cloud also provides plenty of opportunity for predictive algorithms to be run over the footage to take more complicated thing into account, such as aging, plastic surgery, cosmetics, and even the impacts of drugs [5]. Some ride-sharing companies ensure the right passengers are picked up by the right drivers based on face recognition. But this technology and algorithm are faced with difficulties when people are wearing masks [6]. IoT (the Internet of Things) improves from face recognition by enabling improved security measures and automated access management at home [7]. Retailers employ facial recognition to personalize their physical products and connect purchase

[1] Corresponding Author: Shanmei ZHONG, *Business school, Central South University;*
e-mail: zsmemail@csu.edu.cn

behaviors of customers online to their offline ones [8]. These widely used application scenarios are all affected seriously by masked face. Because face recognition relies on extracting human facial features and information, including eyes, mouth, nose, beard, skin etc., mask occlusion almost means only using the features above the nose, which significantly reduces the amount of information available, enabling the accuracy of facial recognition greatly reduced [9].

Masked Face Recognition (MFR) refers to detecting and identifying individuals wearing masks by their facial characteristics. The typical steps involved in masked face recognition are demonstrated in Figure 1. First, Masked Face Recognition Datasets are prepared, which usually comprises various face photos of a single individual with and without a mask. Second, masked face detector is used to localize faces in images. Third, two different approaches are used to solve the issue of face occlusion: (1) mask-robust feature extraction, (2) recovering unmasked faces for feature extraction and face identification. Fourth, some metrics are calculated to evaluate the performance of MFR models. Finally, to determine if this individual has been confirmed or recognized, the predicted face is compared with the original face without a mask.

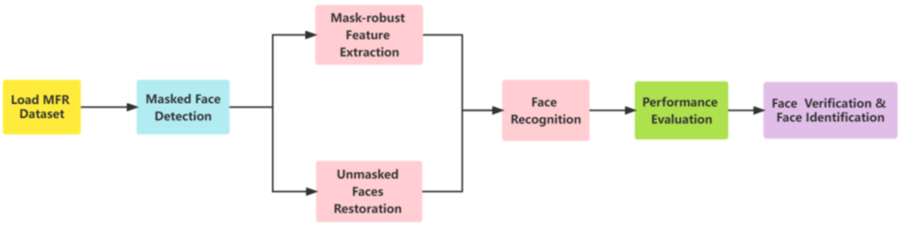

Figure 1. The MFR Pipeline.

Initially, the approaches for occluded face recognition or masked face recognition mainly focus on machine learning. Principal component analysis (PCA) was used in [10] to improve non-masked face recognition. The experiment showed that the accuracy of masked face picture identification is only 72% on average, which is subpar when compared to non-masked face recognition, which has an accuracy of 95% on average. In [11], the sparse coding was modeled as a sparsity-constrained robust regression task, namely robust sparse coding (RSC) method, and achieved a recognition rate of 97% in FR with scarf occlusion. In [12], support vector machine was applied in facial recognition with occlusion by defining an affine subspace of k dimensions (k is the number of features missing in a sample vector). Machine learning techniques utilize artificially constructed handcrafted features of occluded faces, followed by the application of various machine learning algorithms to classify or regress them. However, there are two main drawbacks of this approach: Firstly, feature construction is subjective and reliant on human input, causing uncertainty and instability in feature quality. Secondly, these superficial features primarily rely on statistical methods, leading to difficulties in accurately describing complex semantic information. Fortunately, the advancement of computing power and deep learning model has partially tackled the problems mentioned above. Deep learning can learn the relevant semantic information of the task automatically through multiple layers of nonlinear transformation. Therefore, a large amount of scientists have already used deep learning to perform masked facial recognition (MFR) tasks nowadays [13].

Although deep neural networks have already been applied in masked face recognition and helped solve many problems, there is still few reviews for this fast update task [13, 14]. As a result, many beginners cannot have a rapid and sound understanding of recent advances and novel methods of MFR. Besides, this review can be seen as a phased summary that contributes to the development of subsequent research. Therefore, this review paper reviewed make contributions to the development of MFR tasks in recent years.

This paper introduced the frequently-used datasets of MFR firstly and illustrate the mainstream deep learning methods used for MFR. In addition, some common evaluation metrics are presented. Finally, this paper discusses the potential challenges of MFR and make conclusion.

2. Public Datasets

This section provides several benchmark datasets for MFR training and testing that have been suggested by academics from across the globe.

Huang et al. [15] proposed three MFR datasets, which is claimed to be the first public dataset targeted to MFR globally, namely the Real-world Masked Face Recognition Dataset (RMFRD), Masked Face Detection Dataset (MFDD) and Synthetic Masked Face Recognition Dataset (SMFRD). RMFRD dataset are made up by 7,178 masked and unmasked sample pairs, 3589 of which are connected with the same identity and 3589 with different identities. 426 individuals are represented by RMFRD, whose faces are connected to both masked and unmasked faces. With order to aid in masked face identification, MFDD comprises 24,771 masked face photos from publicly available image sets [16] and web crawling. In order to enhance the diversity and capacity of MFR datasets, this paper put on masks on the existing unmasked faces and constructed a synthetic dataset SMFRD, which consists of 536,721 face images corresponding to 16,817 individuals.

In order to recognize masked faces in the wild, Kai et al. [17] introduced the Masked Aware Network (MAN) model, the first phase of which is to produce masked faces using facial landmarks. They did experiment on dataset Glink360K [18] and WebFace260M [19] with 17 million images and 265 M images respectively, obtaining

masked Glink360K and masked WebFace260M based on a mask generation module. They used masked Glink360K, masked WebFace260M and both of them to do experiments. The results achieve 0.1221 error rate of wild-masked MFR.

A user-friendly technique for face de-occlusion was presented by NIZAM et al. [20], where the user may choose which item to delete. By altering photos from publicly accessible pictures [21] and CelebA-HQ dataset [22], this paper produced a matched synthetic face-occluded dataset.

The MFSR (Masked Face Segmentation and Recognition) Dataset, which consists of two portions, was introduced by Geng et al. firstly [23]. 9,742 mask-enclosed images of faces with annotations for the mask area make up the first section. In the second segment, there are 11,615 photographs totaling 1,004 identities. These images include full-face and masked shots of each identity in a variety of orientations, lighting setups, and mask kinds.

3. Masked Face Recognition Methods

In this section, masked face recognition methods are broken down into two categories, containing mask-robust methods and unmasked face restoration methods. Nowadays, MFR techniques nearly exclusively rely on deep learning. Therefore, the focus of the following methods are masked face recognizers using deep learning models.

3.1 Mask Robust Methods

The difficulty in MFR caused by a lack of features may be greatly reduced if the features exhibit strong resilience to masks. Chang et al. [36] developed a ResSaNet feature extraction backbone. It combines the Self-attention module and CNN, especially Residual Block, into one network. By continually recording the local and global information of the same facial area, ResSaNet can provide impressive performance of more than 78% TAR on both masked and non-masked testing data. [36] To fast decrease the input resolution, a new stem unit drop block [37] was introduced. This unit contributed to the improvement of the ResNet backbone, which is helpful for effective MFR feature extraction. The experiment's baseline adding DropBlock accuracy in a mask MFR condition was 79.21%. In order to improve the efficiency of the feature extraction process, Wang et al. suggested the MaskOut data augmentation method [35]. A framework that incorporates the ArcFace [39] and pairwise loss are presented in [38] and help Masked Face Recognition operate better. This model aims to close the gap between a face's features while it is wearing a mask and when it is not. Additionally, the space between photos of various people is widened. Due to its greater adaptability and resistance to face masks, the paired loss-based technique on the WebFace42M dataset obtained an All-MFR rate of 14.38%. For integrating the mask into the facial image and controlling noisy training datasets, Tao et al. [40] suggested a UV-texture-mapping-based method as well as a self-learning-based cleaning procedure. The incorporation of the Balanced Curricular Loss (BCL), in addition to clever calculations that take the impact of long-tail distribution and challenging facial samples into account, has resulted in impressive experimental findings. Specifically, this recommended approach achieved values of 84.528% Mask and 88.355% MR-ALL in InsightFace ms1m Track. Feifei et al. [42] introduced a novel model called Latent Part Detection(LPD), which can locate the latent robust facial part based on a synthetic masked face dataset, and this part is further used to obtain robust features. Experiments showed LPD's superiority of the generalization on both realistic and synthetic datasets.

3.2 Unmasked Face Restoration Methods

In contrast to mask robust models, returning faces to an unmasked state is a simple concept. Unmasked face recognition methods employ standard face recognition techniques after recovering the whole face from the masked photos. In [41], a brand-new Gan-based network is suggested for unmasking disguised faces. In this study, two discriminators were utilized, one to learn the general structure of faces and the other to learn about the significant missing areas. As a consequence, both numerically and qualitatively, our picture inpainting model beats other cutting-edge methods. How to exclude user-selected foreground items from face pictures is described by Din et al. [20]. A unique GAN-based network was also used in this article. There are two stages in the model. It gains object detection skills in the first stage. With the use of the object detection information from the first step, the item is subsequently eliminated in the second stage. GAN-based networks were used in both phases. This article specifically incorporates partial and pure convolution techniques into the GAN network's generator step. Feihong et al.[43] proposed a Face-Mask Occluded Detection and Restoration framework(FMODR),which contains a unique self-adaptive contextual module, tackling the problem of large visual variance of masks globally. State-of-the-art results are displayed in several mainstream datasets, including AR and CelebA.

4. Performance Evaluation

4.1 Recall Rate

Recall rate, often known as recall, is the proportion of relevant documents that were found to the total number of connected documents in the document library.

There are four possible outcomes for the data test results: TP is for the case where the forecast is positive but it turns out to be positive, and TN stands for the case where the forecast is negative but it turns out to be negative. FP stands for a positive forecast that is actually negative, and FN stands for a negative prediction that is actually positive. Whether the forecast result is accurate is shown by T/F. P/N shows whether the sample for the prediction is positive or negative.

4.2 Accuracy

In contrast to the recall rate, the accuracy varies only with the denominator. The denominator is the predicted positive class, and the accuracy rate is proposed to keep the existing predictions of the model as correct as possible.

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \tag{1}$$

4.3 Precision

Precision in the realm of information retrieval is the percentage of documents that are relevant to the search [34]. Additionally, in the realm of masked face recognition, accuracy is the percentage of true predictions that are positive.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

4.4 F1-score

Obviously, the model performs better when the accuracy and recall rate is high. However, the accuracy and the recall rate do not increase simultaneously. Therefore, we usually assess the model using the F1-score, which is equal to the harmonic average of the accuracy rate and recall rate times two. The harmonic mean's importance is in determining how well A, B, and C are distributed on an overall average (assuming that

B, C do not overlap). The accuracy and recall rate can be balanced out by the F1-score. Comparing the outcomes of several rates may be more equitable.

$$\mathrm{F1-score} = \frac{2\mathrm{precision} \times \mathrm{recall}}{\mathrm{precision} + \mathrm{recall}} \tag{3}$$

5. Challenges

This section summarizes three great challenging of masked face recognition.

5.1 Data Shortage

The majority of face recognition systems nowadays are built based on deep learning, which only work when used to very big datasets. Thus, acquiring masked face recognition datasets is the first step of MFR training. However, the benchmark dataset in real life is still deficient since datasets targeted to MFR need plentiful pairs of samples connected with identities in addition to photos. The masks that the existing algorithms often use on typical face recognition datasets [20] do not accurately reflect the masked faces that are really seen in the wild.

5.2 3D Face Reconstruction

3-Dimensional facial reconstruction is a hot issue attracting increasing attention and has a trend to replace 2-Dimensional facial recognition in the MFR tasks. Due to the data shortage and variations, 3D MFR is expected to have low sensitivity to noisy data and have higher robustness. Some researchers have already utilized 3D representations for MFR task, including WearMask3D [24], MaskTheFace [25] and 3D morphable models [26].

5.3 Discrimination

It has been generally believed that MFR mistake rates varied significantly across various racial and ethnic groups [27,28,33]. For instance, research conducted jointly by the MIT Media Lab and Microsoft showed a high correlation between face recognition accuracy and skin tone [29]. The accuracy is only 65% for ladies with darker skin tones, but it is more than 90% when the picture is white. As a result, the fundamental database utilized for alignment must take not only the ethnic sample balance, but also the sample validity into account. Additionally, factors like light, decorating, and posture may also affect the outcomes of identification.

Discrimination may also result from certain inappropriate applications. Facial recognition technology is now widely used in a variety of industries, including recruitment, education, and dating [30,31]. The personality, psychology, aptitude, and emotional intelligence of the individual are assessed via the analysis of facial data, and the associated recommendations are supplied based on face recognition models. However, these sorts of applications can increase some prejudice and lead to discrimination owing to a number of reasons, including technological level, original data correctness, value judgment suggested in algorithms, and the validity of the database samples [32].

Additionally, facial recognition technology has inherent flaws as a security measure. Faces have a greater exposure rate than other biological characteristics used to identify individuals, such as iris, fingerprints, voice prints etc. Therefore, passive private data collection is much simpler. Face information leakage might result in property losses, invasions of personal privacy, and potentially massive database leaks that pose security dangers to an ethnic group or nation.

6. Conclusion

This paper has presented a survey on masked face recognition. Multiple popular benchmark datasets of MFR are introduced firstly, including RMFRD, Masked Glink360K etc. Additionally, some mainstream deep learning methods used for MFR are explained. Obviously, the MFR task is expected to be studied for a prolonged time and increasing works and research will be presented in the papers continuously. In addition, some useful evaluation metrics are proposed. Finally, this paper discusses the potential challenges of MFR, which also indicate the possible development directions.

Reference

[1] Howard J, Huang A, Li Z, et al. An evidence review of face masks against COVID-19[J]. Proceedings of the National Academy of Sciences, 2021, 118(4): e2014564118.

[2] Leung, N.H.L., Chu, D.K.W., Shiu, E.Y.C. et al. Respiratory virus shedding in exhaled breath and efficacy of face masks. Nat Med 26, 676–680 (2020). https://doi.org/10.1038/s41591-020-0843-2

[3] Cheng K K, Lam T H, Leung C C. Wearing face masks in the community during the COVID-19 pandemic: altruism and solidarity[J]. The Lancet, 2022, 399(10336): e39-e40.

[4] Zhu Z, Huang G, Deng J, et al. Webface260m: A benchmark unveiling the power of million-scale deep face recognition[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 10492-10502.

[5] https://anyconnect.com/blog/facial-recognition-applications/

[6] Gupta S, Buriro A, Crispo B. DriverAuth: A risk-based multi-modal biometric-based driver authentication scheme for ride-sharing platforms[J]. Computers & Security, 2019, 83: 122-139.

[7] Masud M, Muhammad G, Alhumyani H, et al. Deep learning-based intelligent face recognition in IoT-cloud environment[J]. Computer Communications, 2020, 152: 215-222.

[8] Shri Bharathi S V, Aadithya Kiran T, Dileep Kanth N, et al. A Hybrid Approach to Facial Recognition for Online Shopping Using PCA and Haar Cascade[C]//International Conference on Image Processing and Capsule Networks. Springer, Cham, 2022: 284-295.

[9] Ngan, M.L.; Grother, P.J.; Hanaoka, K.K. Ongoing Face Recognition Vendor Test (FRVT) Part 6B: Face Recognition Accuracy with Face Masks Using Post-Covid-19 Algorithms;NIST Interagency/Internal Report (NISTIR); National Institute of Standards and Technology: Gaithersburg, MD, USA, 2020.

[10] Ejaz M S, Islam M R, Sifatullah M, et al. Implementation of principal component analysis on masked and non-masked face recognition[C]//2019 1st international conference on advances in science, engineering and robotics technology (ICASERT). IEEE, 2019: 1-5.

[11] Yang M, Zhang L, Yang J, et al. Robust sparse coding for face recognition[C]//CVPR 2011. IEEE, 2011: 625-632.

[12] Jia H, Martinez A M. Support vector machines in face recognition with occlusions[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2009: 136-141.

[13] Alzu'bi A, Albalas F, Al-Hadhrami T, et al. Masked face recognition using deep learning: A review[J]. Electronics, 2021, 10(21): 2666.

[14] Wang B, Zheng J, Chen C L P. A Survey on Masked Facial Detection Methods and Datasets for Fighting Against COVID-19[J]. IEEE Transactions on Artificial Intelligence, 2021, 3(3): 323-343.

[15] Huang B, Wang Z, Wang G, et al. Masked face recognition datasets and validation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 1487-1491.
[16] https://github.com/AIZOOTech/FaceMaskDetection
[17] Wang K, Wang S, Yang J, et al. Mask aware network for masked face recognition in the wild[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 1456-1461.
[18] An X, Zhu X, Gao Y, et al. Partial fc: Training 10 million identities on a single machine[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 1445-1449.
[19] Zhu Z, Huang G, Deng J, et al. Webface260m: A benchmark unveiling the power of million-scale deep face recognition[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 10492-10502.
[20] Din N U, Javed K, Bae S, et al. Effective removal of user-selected foreground object from facial images using a novel GAN-based network[J]. IEEE Access, 2020, 8: 109648-109661.
[21] Liu Z, Luo P, Wang X, et al. Deep learning face attributes in the wild[C]//Proceedings of the IEEE international conference on computer vision. 2015: 3730-3738.
[22] Karras T, Aila T, Laine S, et al. Progressive growing of gans for improved quality, stability, and variation[J]. arXiv preprint arXiv:1710.10196, 2017.
[23] Geng M, Peng P, Huang Y, et al. Masked face recognition with generative data augmentation and domain constrained ranking[C]//Proceedings of the 28th ACM International Conference on Multimedia. 2020: 2246-2254.
[24] Hong J H, Kim H, Kim M, et al. A 3D model-based approach for fitting masks to faces in the wild[C]//2021 IEEE International Conference on Image Processing (ICIP). IEEE, 2021: 235-239.
[25] Anwar A, Raychowdhury A. Masked face recognition for secure authentication[J]. arXiv preprint arXiv:2008.11104, 2020.
[26] Egger B, Schönborn S, Schneider A, et al. Occlusion-aware 3d morphable models and an illumination prior for face image analysis[J]. International Journal of Computer Vision, 2018, 126(12): 1269-1287.
[27] Bacchini F, Lorusso L. Race, again: how face recognition technology reinforces racial discrimination[J]. Journal of information, communication and ethics in society, 2019.
[28] Štruc V, Pavešić N. Gabor-based kernel partial-least-squares discrimination features for face recognition[J]. Informatica, 2009, 20(1): 115-138.
[29] Hutchinson B, Denton E, Mitchell M, et al. Detecting bias with generative counterfactual face attribute augmentation[J]. 2019.
[30] Krithika L B, Venkatesh K, Rathore S, et al. Facial recognition in education system[C]//IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2017, 263(4): 042021.
[31] Nawaz N. Artificial intelligence applications for face recognition in recruitment process[J]. Journal of Management Information and Decision Sciences, 2020, 23: 499-509.
[32] Ngan M L, Grother P J, Hanaoka K K. Ongoing face recognition vendor test (frvt) part 6b: Face recognition accuracy with face masks using post-covid-19 algorithms[J]. 2020.
[33] https://venturebeat.com/ai/nist-benchmarks-show-facial-recognition-technology-still-struggles-to-identify-black-faces/
[34] https://en.wikipedia.org/wiki/Precision_and_recall#Precision
[35] Wang W, Zhao Z, Zhang H, et al. MaskOut: a data augmentation method for masked face recognition[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 1450-1455.
[36] Chang W Y, Tsai M Y, Lo S C. ResSaNet: A Hybrid Backbone of Residual Block and Self-Attention Module for Masked Face Recognition[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 1468-1476.
[37] Qi D, Hu K, Tan W, et al. Balanced Masked and Standard Face Recognition[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 1497-1502.
[38] Qian H, Zhang P, Ji S, et al. Improving Representation Consistency with Pairwise Loss for Masked Face Recognition[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 1462-1467.
[39] Deng J, Guo J, Xue N, et al. Arcface: Additive angular margin loss for deep face recognition[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 4690-4699.
[40] Feng T, Xu L, Yuan H, et al. Towards Mask-robust Face Recognition[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 1492-1496.
[41] Din N U, Javed K, Bae S, et al. A novel GAN-based network for unmasking of masked face[J]. IEEE Access, 2020, 8: 44276-44287.

[42] Ding F, Peng P, Huang Y, et al. Masked face recognition with latent part detection[C]//Proceedings of the 28th ACM international Conference on multimedia. 2020: 2281-2289.
[43] Feihong L, Hang C, Kang L, et al. Toward High-quality Face-Mask Occluded Restoration[J]. ACM Transactions on Multimedia Computing, Communications and Applications, 2023, 19(1): 1-23.