# Masked Face Recognition Datasets and Validation

Baojin Huang, Zhongyuan Wang*, Guangcheng Wang, Kui Jiang, Zheng He, Hua Zou, Qin Zou
School of Computer Science, Wuhan University
{huangbaojin, kuijiang, hezheng, qzou}@whu.edu.cn
{wzy_hope, wangguangcheng0428}@163.com, zouhua08@gmail.com

## Abstract

*In order to effectively prevent the spread of COVID-19 virus, almost everyone wears a mask during coronavirus epidemic. This nearly makes conventional facial recognition technology ineffective in many scenarios, such as face authentication, security check, community visit check-in, etc. Therefore, it is very urgent to boost performance of existing face recognition systems on masked faces. Most current advanced face recognition approaches are based on deep learning, which heavily depends on a large number of training samples. However, there are presently no publicly available masked face recognition datasets. To this end, this work proposes three types of masked face datasets, including Masked Face Detection Dataset (MFDD), Real-world Masked Face Recognition Dataset (RMFRD) and Synthetic Masked Face Recognition Dataset (SMFRD). As far as we know, we are the first to publicly release large-scale masked face recognition datasets that can be downloaded for free at: https://github.com/X-zhangyang/Real-World-Masked-Face-Dataset.*

## 1. Introduction

Almost everyone wears a mask during the COVID-19 coronavirus epidemic. Face recognition techniques, as the most important means of identification, have nearly failed, which has brought huge dilemmas to authentication applications relying on face recognition, such as face access control,face gates, face authentication, etc. In particular, in the public security check in railway stations, the security check systems based on face recognition often reject the masked faces, but removing masks for passing authentication will increase the risk of virus infection. Because the COVID-19 virus can be spread through contact, the unlocking systems based on passwords or fingerprints are unsafe. It is much safer through face recognition without touching, but the existing face recognition is no longer reliable when wearing a mask. In view of the above difficulties, it is necessary to improve the existing face recognition approaches that do not take into account the specific requirements of masks covering the face, so that face identification can still be performed reliably when the face is not completely exposed.

The state-of-the-art face recognizers are almost developed based on deep learning, which depend on massive training samples [8, 11, 12, 13, 16]. Thus, above of all, developing face recognition approaches concerning masked faces requires a large number of masked face samples. A large dataset of 137,016 masked face images was proposed in [6], which is divided into two masked face categories: correctly worn and incorrectly worn. However, it is intended for mask detection rather than masked face recognition, and is created by defining a mask-to-face deformable model other than from realistic masked face samples. Some public masked face recognition challenges, such as ICCV 2021 [7, 18] and CSIG FAT-AI 2021 [1], conducted realistic masked face datasets, but they were not made public for the purpose of testing. Evidently, there are presently no publicly available masked face recognition datasets, especially real and non-synthetic masked face samples. This work constructs masked face datasets by different means and meanwhile validates their positive effects on masked face identification.

## 2. Proposed Datasets

There are two popular applications on masked faces, namely, facial mask detection task and masked face recognition task. The former needs to identify whether a person wears a mask as required, while the latter needs to identify the specific identity of a person with a mask. Each task has different requirements for the dataset. The former only requires masked face image samples, but the latter requires datasets which contain multiple face images of the same subject with and without a mask. Relatively, datasets used for the face recognition task are more difficult to construct.

This paper proposes three types of masked face datasets, including Masked Face Detection Dataset (MFDD), Real-world Masked Face Recognition Dataset (RMFRD) and

* Corresponding author

(a) (b)

Figure 1. Samples of masked face pairs in RMFRD (a) and SMFRD (b). The first row shows normal faces, and the second and third rows are their corresponding masked faces.

Synthetic Masked Face Recognition Dataset (SMFRD). Their details are presented as follows, respectively.

- MFDD: The source of MFDD mainly includes two parts: (a) Some of samples are from related release [2]; (b) The other part of MFDD is crawled from the Internet. We further label the crawled face images, performing annotations such as whether the face wears a mask and the position of the mask. This way, the dataset contains 24,771 masked face images. MFDD dataset can be used to train an accurate masked face detection model, along with the subsequent face recognition task. Additionally, it can also be used to determine whether a person wears a mask, as it is illegal without wearing a mask during coronavirus epidemic.

- RMFRD: A python crawler tool is used to crawl the front face images of public figures as well as their corresponding masked counterparts from massive Internet resources. Then, we manually remove the unreasonable face images resulting from the wrong correspondence. Similarly, we crop the accurate face areas with the help of semiautomatic annotation tools, like LabelImg and LabelMe [3]. The final dataset contains 4015 face images of 426 subjects in a size of 250 $\times$ 250 pixels, with each associated with a normal face and several masked faces. The dataset is further organized into 7178 masked and unmasked sample pairs, comprising 3589 pairs of the same identity and 3589 pairs of different identities. To the best of our knowledge, this is the world's first public masked face recognition dataset (Note that the previous masked face datasets are for face detection rather than face recognition). Fig. 1 (a) shows some examples.

- SMFRD: In order to expand the volume and diversity of the masked face recognition dataset, we have alternatively taken another means, which is to put on masks on the face images from the existing public face datasets. To promote data manipulation efficiency, we have developed an automatic mask-wearing tool based on Dlib library [4]. This software is then used to wear masks on face images in the popular face datasets, presently including CASIA-WebFace [17], LFW [10], CFP-FP [15] and AgeDB-30 [14] datasets. This way, we extra constructed a synthetic masked face dataset covering 536,721 face images of 16,817 subjects. In practice, SMFRD can be used together with their original normal counterparts. Fig. 1 (b) shows a set of synthetic masked face samples.

Table 1. Face datasets for training and testing.

| Dataset | #Identity | #Image/Video |
|---|---|---|
| WebFace [17] | 10K | 0.5M |
| SMFRD | 10K | 0.5M |
| MS1MV3 [8] | 9.3K | 5.1M |
| Glint360k [5] | 360K | 17M |
| RMFRD | 426 | 4,015 |
| ICCV2021-MFR-MASK [7] | 7K | 22K |
| ICCV2021-MFR-ALL [7] | 0.24M | 1.6M |

Table 2. Comparisons on face verification (%) on RMFRD, ICCV2021-MFR-MASK and ICCV2021-MFR-ALL.

| Dataset | Backbone | Method | Size / MB | RMFRD | ICCV2021-MFR-MASK | ICCV2021-MFR-ALL |
|---|---|---|---|---|---|---|
| WebFace | R50 | Arcface | 166 | 63.22 | 22.13 | 38.51 |
| SMFRD | R50 | Arcface | 166 | 71.13 | 39.03 | 39.16 |
| MS1MV3 | R18 | Arcface | 91 | 64.19 | 47.85 | 68.33 |
| Glint360k | R18 | Arcface | 91 | 64.59 | 53.32 | 72.07 |
| MS1MV3 | R34 | Arcface | 130 | 65.90 | 58.72 | 77.36 |
| Glint360k | R34 | Arcface | 130 | 68.61 | 65.10 | 83.02 |
| MS1MV3 | R50 | Arcface | 166 | 68.68 | 63.85 | 80.53 |
| Glint360k | R50 | Arcface | 166 | 71.40 | 70.23 | 87.08 |
| MS1MV3 | R100 | Arcface | 248 | 70.19 | 69.09 | 84.31 |
| Glint360k | R100 | Arcface | 248 | 72.74 | 75.57 | 90.66 |

# 3. Experiments

## 3.1. Experimental Settings

**Datasets**. As given in Table 1, we employ WebFace [17], SMFRD, MS1MV3 [8] and Glint360k [5] as our training sets. At testing stage, We extensively evaluate models on our built RMFRD and ICCV2021-MFR (Mask and All) [7] benchmarks. ICCV2021-MFR-MASK refers to the masked face testing set, which contains 6,964 masked images and 13,928 non-masked images of 6,964 identities. There are totally 13,928 positive pairs and 96,983,824 negative pairs. ICCV2021-MFR-ALL denotes the multi-racial testing set, covering 242,143 identities and 1,624,305 images.

**Models**. In experiments, Arcface [8] is used as the face recognizer. To provide evaluation reference for researchers using our datasets, we separately adopt the refined ResNet18, ResNet34, ResNet50 and ResNet100 in [8] as the backbones.

**Evaluation Metrics**. We evaluate all face recognition models by 1:1 face verification, explicitly referring to the true positive rate (TPR) or false rejection rate (FRR) under a certain false acceptance rate (FAR). For the ICCV2021-MFR set, true acceptance rate (TAR) is measured with mask-to-nonmask 1:1 protocol, with FAR less than $10^{-4}$. For other datasets, TAR is measured on all-to-all 1:1 protocol, with FAR less than $10^{-6}$. Unlike the above-mentioned single value evaluation protocols, we also compare the performance of the models using receiver operating characteristic (ROC) curve.

## 3.2. Implementation Details

For data prepossessing, we use the open-source RetinaFace [9] to detect and align raw face images, obtaining the normalised face crops ($112 \times 112$). Simultaneously, we perform data augmentation on the training set, such as flipping, to improve the generalization of the trained model. Our training experiments are implemented by Pytorch framework, running on two NVIDIA 3090 (24GB) GPUs. In particular, we set the batch size to 256 and the embedding dimension to 512 during the training process on SMFRD. Moreover, the learning rate starts from 0.1 and is divided by 10 at 20, 28, 32 epochs. The training process is finished at 34 epochs. Note that the models trained on MS1MV3 and Glint360k are all derived from ArcFace [1]. All evaluation results on ICCV2021-MFR are from the website [2].

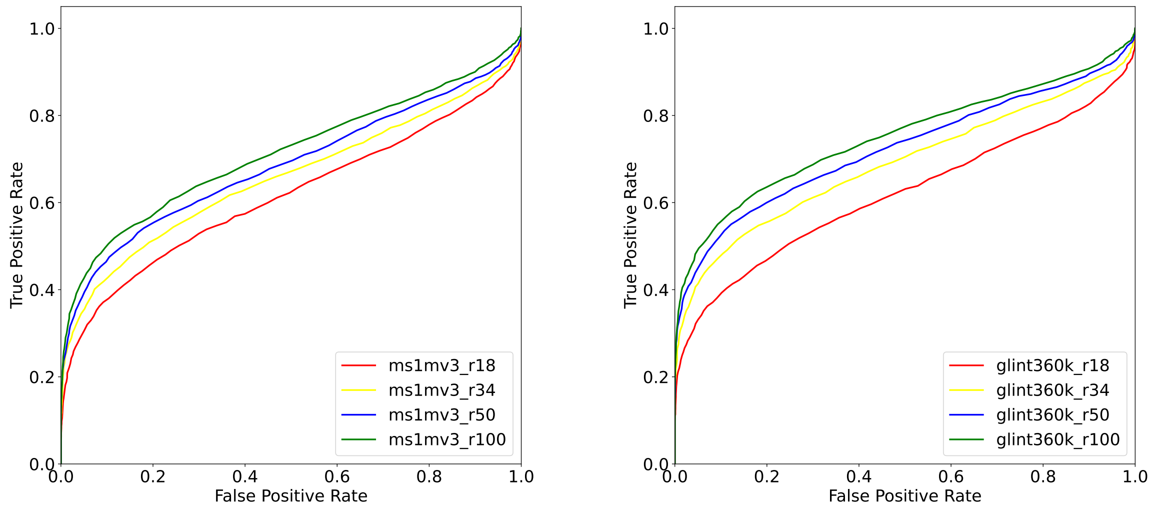
## 3.3. Results on Face Verification

We test all models on three masked face datasets: RMFRD, ICCV2021-MFR-MASK and ICCV2021-MFR-ALL by face verification. As tabulated in Table 2, ArcFace trained on SMFRD outperforms that trained on WebFace by a large margin on RMFRD, which confirms that the synthetic masked face images can notably improve the performance of masked face recognition. Under the same ResNet50, the model trained by Glint360k is slightly better than that trained by SMFRD, mainly because Glint360k is much larger than SMFRD in scale.

From the perspective of the testing set, the accuracy of multiple models (ArcFace trained on MS1MV3 and Glint360k) on RMFRD is positively correlated with their performance on ICCV2021-MFR-MASK. That is due to the fact that both RMFRD and ICCV2021-MFR-MASK are derived from real scenes. We should point out that although ICCV2021-MFR-MASK can effectively measure the masked face recognition model, its privacy (not publicly released) prevents it from being used as a validation set to optimize models during the training phase. Thus our built RMFRD can be used as a valid public validation set like LFW [10], CFP-FP [15], etc.

In Fig. 3, we show the full ROC curves of different models on RMFRD. It can intuitively see that TPR is significantly reduced when FPR is less than 0.1, which can be attributed to the variability facial poses, resolutions and masks of public celebrities collected from the website. As shown in Fig. 3, we also list some hard cases on RMFRD. In

[1] https://github.com/deepinsight/insightface/tree/master/recognition/arcface_torch
[2] http://iccv21-mfr.com/




(a) (b)

Figure 2. ROC curves of different models on RMFRD.

Figure 3. Some hard cases (positive pairs in the green box and negative pairs in the red box) on RMFRD. The first row shows masked faces, and the second row is their corresponding normal faces.

positive samples, the masked face image is usually accompanied by age and other occlusion factors, so that even the same face is actually different. In negative samples, because of the makeup of females, their important facial recognition features (such as eyebrows) become very similar. This fact implies that our built RMFRD is really a challenging dataset.

## 4. Conclusion

In this paper, we have proposed MFDD, RMFRD and SMFRD datasets to train and test deep learning based masked face recognition models. In particular, MFDD can be used to train detection models of wearing masks or masked faces. As the first publicly available realistic masked face recognition dataset, RMFRD can be used for validation or fine-tuning dataset in training, or testing dataset in testing faithful to the real situations. Due to the large scale of SMFRD, it is suitable for model training. In view of the scarcity of masked face datasets, and only a few have not yet been made public, our built datasets will be greatly beneficial to the masked face recognition communities and applications as well. At present, the prevention and control of COVID-19 is normalized, and considering the frequent haze weather in recent years, the demand for masked face recognition will persist for a long time.

## Acknowledgments

This work was supported by National Natural Science Foundation of China (U1903214, 62071339, U1736206), and Hubei Province Key RD Project (2020BAB018). The numerical calculations in this paper have been done on the supercomputing system in the Supercomputing Center of Wuhan University.

## References

[1] `http://fat.neliva.com.cn/CFAT2021/`. 1

[2] `https://aizoo.com/`. 2

[3] `https://tzutalin.github.io/labelImg/`. 2

[4] `http://dlib.net/`. 2

[5] Xiang An, Xuhan Zhu, Yang Xiao, Lan Wu, Ming Zhang, Yuan Gao, Bin Qin, Debing Zhang, and Fu Ying. Partial fc: Training 10 million identities on a single machine. In *Arxiv 2010.05222*, 2020. 2, 3

[6] Adnane Cabani, Karim Hammoudi, Halim Benhabiles, and Mahmoud Melkemi. Maskedface-net–a dataset of cor-

rectly/incorrectly masked face images in the context of covid-19. *Smart Health*, 19:100144, 2021. 1

[7] Jiankang Deng, Jia Guo, Xiang An, Zheng Zhu, and Stefanos Zafeiriou. Masked face recognition challenge: The insightface track report. In *Proceedings of the International Conference on Computer Vision Workshops*, 2021. 1, 2, 3

[8] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 4685–4694, 2018. 1, 2, 3

[9] Jiankang Deng, Jia Guo, Zhou Yuxiang, Jinke Yu, Irene Kotsia, and Stefanos Zafeiriou. Retinaface: Single-stage dense face localisation in the wild. *arXiv: Computer Vision and Pattern Recognition*, 2019. 3

[10] Gary B. Huang, Marwan Mattar, Tamara Berg, , and Eric Learned-Miller. 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*, 2008. 2, 3

[11] Bingyu Liu, Weihong Deng, Yaoyao Zhong, Mei Wang, Jiani Hu, Xunqiang Tao, and Yaohai Huang. Fair loss: Margin-aware reinforcement learning for deep face recognition. In *IEEE International Conference on Computer Vision*, pages 10052–10061, 2019. 1

[12] Weiyang Liu, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song. Sphereface: Deep hypersphere embedding for face recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 6738–6746, 2017. 1

[13] Weiyang Liu, Yandong Wen, Zhiding Yu, and Meng Yang. Large-margin softmax loss for convolutional neural networks. In *Conference on Machine Learning*, volume 2, page 7, 2016. 1

[14] Stylianos Moschoglou, Athanasios Papaioannou, Christos Sagonas, Jiankang Deng, Irene Kotsia, and Stefanos Zafeiriou. Agedb: The first manually collected, in-the-wild age database. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1997–2005, 2017. 2

[15] Soumyadip Sengupta, Juncheng Chen, Carlos D Castillo, Vishal M Patel, Rama Chellappa, and David W Jacobs. Frontal to profile face verification in the wild. In *Workshop on Applications of Computer Vision*, pages 1–9, 2016. 2, 3

[16] Anh Tuan Tran, Tal Hassner, Iacopo Masi, and Gerard Medioni. Regressing robust and discriminative 3d morphable models with a very deep neural network. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1493–1502, 2017. 1

[17] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z Li. Learning face representation from scratch. *arXiv: Computer Vision and Pattern Recognition*, 2014. 2, 3

[18] Zheng Zhu, Guan Huang, Jiankang Deng, Yun Ye, Junjie Huang, Xinze Chen, Jiagang Zhu, Tian Yang, Jia Guo, Jiwen Lu, Dalong Du, and Jie Zhou. Masked face recognition challenge: The webface260m track report. In *Proceedings of the International Conference on Computer Vision Workshops*, 2021. 1