Face identification for Masked and Unmasked Faces using an automated masked training set

Rahul Dhandhukia*
Monika Shah*

ABSTRACT

Facial recognition is a popular biometric authentication technique that is adopted everywhere. For the sustainability of the societal community, the COVID-19 pandemic and following viral attacks have mandated the wearing of masks at all public events and organizations. It set new standards for how existing facial recognition-based biometric identification systems must function in order to identify faces hidden behind masks. To meet this need, the researchers have attempted fixing the current algorithms and constructing new ones. Each has one or more drawbacks, such as the difficulty of identifying faces in masked images of various orientations, the expense of building a masked training set, the need to train systems using various mask colors and shapes, etc. Most tedious part here is to re-generate a dataset with different masks. This work uses machine learning and deep learning algorithms to automate the development of masked training sets and propose a model to recognize a person regardless of whether they are wearing a mask or not. With the aid of this function, industries may easily transition to new face recognition effectively.

Keywords - *Masked Face Recognition, Computer vision, Mask augmentation, Face recognition, Face detection, Biometric, Occlusion.*

Introduction

Face recognition is a popular biometric technique for identification and authentication in public and private organizations. Face recognition is used to verify people's identities at homes and businesses, record employees' in-and-out times, track students' attendance, and verify travelers' identities at border crossings.

The COVID-19 pandemic and subsequent viral attacks require masks at all public events and organizations for community sustainability. It defined new requirements for facial recognition-based biometric identification systems to identify faces behind masks. Researchers have fixed and created new algorithms to satisfy this demand. Each has drawbacks, such as the difficulty of recognizing faces in masked images of varied orientations, the expense of developing a masked training set, the requirement to train systems using multiple mask colors and forms, etc. Regenerating a dataset with different masks is tiresome. This paper automates training dataset generation of masked images from a legacy database and recognizes faces with or without masks.

The major contribution of this paper are : i) Summary of datasets, training set and algorithms used in related work, ii) Generating masked face images for training from the legacy non-masked dataset, iii) Augmenting mask on existing face image of a different orientation, iv) Face detection and recognition to the Masked and Non-Masked

Manuscript received March 30, 2023.
Rahul Dhandhukia was a student at Institute of Technology, Nirma University, Gujarat, India (e-mail: 16bit047@nirmauni.ac.in).
Monika Shah is with Institute of Technology, Nirma University, Gujarat, India. (e-mail: monika.shah@nirmauni.ac.in).

dataset both with better accuracy.

The paper continues like this. Section II describes masked face recognition methods. The datasets and training face images utilized in model training are also summarized. Section III discusses the proposed method to automate masked face image generation and masked face recognition. Section IV details the test setup and methodology outcomes. Section V concludes with the results of the masked face recognition experiment.

II. Related Works

Face recognition relies on occlusion. Caps, eyeglasses, veils, and other face-obscuring things can cause occlusion. Occlusion includes masks. The masked face covers almost half the face, making facial recognition difficult. This issue has many solutions. Fig. 1 demonstrates three categories of masked face recognition algorithms on paper [2]: Local matching, restoration, and discard occlusion-based methods.

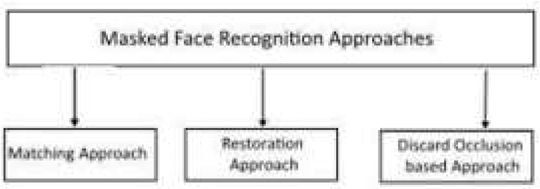


Fig. 1. Various Masked Face Recognition

Local-matching algorithms identify facial image similarities with and without masks. It analyses several areas of a similar-sized facial image to avoid obstructed regions affecting other aspects. Restoration uses the original image to restore the occluded face. Threshing 3D image face map values finds occluded regions. After reconstructing the occluded face using Principal Component Analysis (PCA) [7], iteration close point approach, and statistical curve estimation. Discard occlusion techniques remove occluded regions from images. Segmentation-based approaches can discover occluded sections and classify individual face images based on extracted attributes.

Mask-recognition research is abundant. Two steps: Face Detection finds the occluded face in an image, and Face Recognition matches extracted features to identify the person. Some of the examples based on this approach are [32], [33], [35], [41].

Retaking subject face images with and without masks for masked face recognition is difficult and time-consuming. Geng et al. [36], and Li et al. [38] have shown the use of GAN [13] to generate synthetic masked faces using Segmentation with a multi-level guided identity preserve module.

New approaches are suggested. Hariri [34] used VGG-16 and BoF [11] to crop the viewable face. Cropping-based techniques with the newly built module CBAM focused around ocular regions [43]. EUM employs facial recognition models to reveal masked faces [49]. Unmasking is very complex. Article [71] uses PCA, dilation, skin color painting, and inpainting to detect and rebuild occluded facial regions. Many frameworks start with ResNet. David Montero [52], GuiLing

TABLE I: Dataset, Training set, and algorithms

Qef	C`s`rds	Sq`hm Set	@kfnqhsgl r
Z32\	B@RH@-WebFace [25], Yela-B [31]	M, NM	BA@L Z2/\+L SBMMZ3\+ ResNet50 [12]
Z38\	L R0L U1 Z4/\	L + NoM	Ckha Z7\+Dl adcchmf Unmasking Mode [49], SRT Loss [49]
Z40\	QL EQC Z18\	L +ML	QdrMds-50 [12]
Z41\	L R0L U1 Z4/\	L +ML	QdrMds-50 [12],ArcFace [50], MaskTheFace [53]
Z43\	F dmdq`sdc eqnl uhcdnr BLP, M1P, M2P	L +ML	L SBMMZ3\+@qbE`bd Z4/\+ VGGFace [26],COTS [56] SphereFace [55]
Z46\	QL EQC Z18\+ SMFRD [29]	L +ML	L nchehdc QdrMds Z01\
Z47\	B@RH@MHQ-VIS 2.0 [60], Oulu-CASIA NIR-VIS [61], BUAA-VisNir [62]	L +ML	Gdsdqnf dmdnt r rdl h-siamese training (HSST) [59]
Z52\	QL EQC Z18\+ SMFRD [29]	L +ML	QdrMds Z01\
Z53\	QL EQC Z18\+ SMFRD [29], LFW dataset [27]	L +ML	Sqhokds enql t k`shnm Z54\
Z55\	L ECC Z56\+EL C [68], RMFRD [29], MaskedFace-Net [69]	L +ML	F`anqv`udkds Z6/\
Z23\	QL EQC Z18\	L % NM	A`f neEd`st re [11], VGG-16 [9],MLP classifier
Z24\	L EU c`s`rds 1// people with 400 images each, MFI for identification	ML	L SBMMZ3\+KOC Z24\+ ResNet [12]

Z25\	L `rj dc E`bd Segmentation & Recognition dataset	ML	Cnl `hmBnmrsq`hmdc Ranking Loss [36],GAN [13]
Z27\	Bdkeb-A [44], LFW dataset [27]	MnL + M, NM	F @MZ02\+QOB@Z03\+ GL [22], GFC [23], GA [25]
Z31]	F dmdq`sdc e`bd database	L +ML	L SBMMZ3\+QdrMds-50 [12], AlignedReID [21]
Z31\	K`adkdc E`bdr hmsgd Wild (LFW) [27]	L	E`bdL `rj Mds-21 [42], HOG [28]
Z21\	L `rj dc E`ce Database	MnL	L SBMMZ3\+E`bdMds Z2\
Z22\	NQK e`bd c`s`a`rd Z5\	L % NM	Uhnk`-Jones [5], PCA [7]
Z26\	Odnokd v hsg u`qhdc disguises in different backgrounds and varied illuminations	MnL + M, NM	Ro`sh`kEt rhmmBMM+ SVM [48], Disguised face identification
Z3/ \	BL T-PIE database [19], The AR database [20], KACST database	ML	L `rj dc Bnqqdk`shnm Filters (MCF) [40], CNN
Z28\	EQF B u-1 Z06\+TL A-DB [15] Bosphorus [16]	MnL	Hsdq`shud Bknrdrs Onhms+ CNN [18]

Wu [57], and Saumya Kumaar [37] have improved ResNet-50 and CNN-based algorithms. Face detection has improved, even if the face is partially obscured or oriented, because of the COVID-19 epidemic. DSFD [1], RetinaFace [45], and TinaFace [47] work well with occluded faces. Current approaches have drawbacks. For instance, [32],[39],[42],[43],[50], and [64] requires images of all persons with variable occlusion, alignment, and lighting within a threshold. Changing these thresholds can hurt outcomes. Masked face recognition accuracy decreases [33],[52]. Real-time face recognition requires speedier results. [41] requires degrading images for faster calculation at 64.23% accuracy. Some approaches only work on non-masked images [51]. Infrared images are difficult to make visible and require specific equipment [58]. The model training procedure at [57],[63] yields better results but takes longer. Bag of features [11] might provide two people with the identical patch [34]. [40] requires prior knowledge of the occlusion location.

Table I summarizes Masked Face Recognition datasets, training set, and method. Masked and non-masked face images were used to train the model. This work develops a masked face recognition system without requiring the retaking of training facial images with and without masks at varied angles. It should work on the NoM dataset, and masked facial pictures can greatly improve the model. The proposed method also intends to improve real-time systems for assessing facial image at varied orientations.

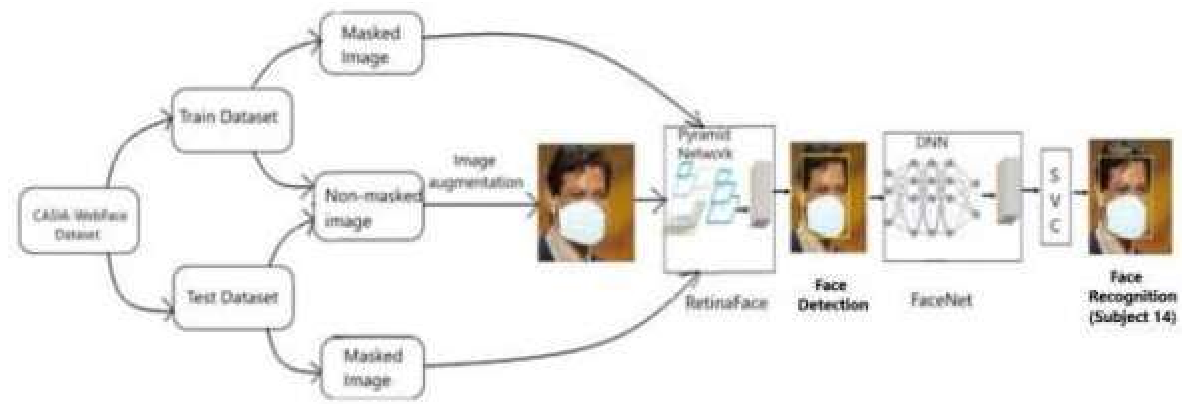
III. Proposed Method

Fig. 2 represents the basic workflow of the proposed masked face recognition solution. The input image can be a masked image or a non-masked facial image. If the input is a non-masked image, a mask image will be augmented to the input image, and then face detection and face recognition will be performed in sequence to recognize the identity.

Fig. 2. Workflow

RetinaFace is known for faster performance and better results in performing face localization on various scales of faces [45]. On the other hand, DSFD [1] is also known for being superior to other face detector algorithms. FaceNet [3] provides a better outcome for feature extraction-based face recognition even if the face is occluded. Therefore, the proposed solution suggests using RetinaFace [45] or DSFD for face detection, FaceNet [3] for feature extraction, and the Support Vector Classifier for face recognition. For the dataset, the CASIA-WebFace [24] dataset is used, which contains lots of images of more than 10k identities. Fig. 3 presents the proposed system architecture. It starts with building a database from CASIA-WebFace images for 20 identities, and the 21st-person pictures were wearing masks.

An image of a mask is augmented on non-masked images, which is done through face recognition [46] library. This library provides facial landmarks like the nose and different chin landmarks. Then, based on the landmarks received, resize the mask image and paste it onto the face. With this method, the process of generating masked images from non-masked images is more straightforward and works on any image, irrespective of orientation, illumination, or background. In the second phase, augmented images are passed to the face detection module. Here, DSFD [1], and RatinaFace [45] are tested for face detection. RatinaFace provides



similar results in less time. So that is finalized for detecting the face in the image. The third and final stage involves face recognition and classification. Face recognition identifies the person in the image. Based on the location of the face, features are extracted using FaceNet [3]. The vectors received in previous steps are normalized, and Support Vector Machine (SVM) [48] are used for classification. In-depth work is explained below in each separate sub-section.

A. Dataset

The CASIA-WebFace [25] dataset has around 494414 face images of 10575 different subjects. However, for the sake of simplicity, only 20-person images are used. However, there are a total of 3320 total images, 2510 for training and 810 for testing. Refer to Table II for a description of the training and testing datasets.

TABLE II: Data Set

	Training Images	Testing Images
Masked Images	1687	613
Non-Masked+ Masked	2510	810

This dataset contains images of a single person in a single folder, and the number of images per person also varies. Images have faces at different angles and orientations, which is a plus for the task as it should be able to detect the person from various angles. None of these images has a mask on it, so for generating masked face images, the face recognition [46] library is used, which provides fast detection and location of facial landmarks. So nearly 75% of the Non-Masked images of each person are converted into Masked images and split into training and testing images accordingly.

Landmarks are the critical locations in the face that help to differentiate one person from another, for example, the eyes, ears, nose, chin, and mouth. As explained earlier, based on the location of landmarks achieved, the mask image size is adjusted accordingly and augmented on top of the face to generate masked images. This method of generating masked images is more straightforward and can be applied to any image. Some of the examples of mask augmented images are shown in Fig. 4.

B. Face Detection

As discussed earlier, RetinaFace is selected here for face detection. RetinaFace [45] functionalities are divided into three parts:

- First part is Feature Pyramid Network which takes an image as an input and gives five feature maps calculated using ResNet. As explained in [45] it adds lateral connection with each layer of ResNet. This allows detection of smaller objects with better accuracy.
- Second part is Context Module, where Deformable Convolution Network (DCN) is used, similar to that of CNN, but with fewer offset parameters to help reduce constraints on the kernel. DCN applies the context module on all five feature pyramid levels independently. DCN allows the model to increase context modeling capacity and helps detect faces at different orientation.
- Third part is Cascade Multi-Task Loss. After getting the bounding box in the context module, cascade regression is used with Cascade Multitask Loss. It is used to improve the face-localization execution. The first context head uses regular anchor to predict the bounding box and second context head uses regressed anchor

to predict the bounding box with more accuracy.

The loss equation 1 computes the sum of the losses that occurred in all steps (classification, landmarks, and 3D pixel losses). i represents the person's face, Pi is the anchor probability, and Pi* is 1 for the positive anchor and 0 for the negative anchor. The loss function has four parts: softmax loss for binary classes, regression loss of the bounding box, regression loss of five landmarks, and regression loss of 3D points for better localization. RetinaFace also gives the person's coordinates in a masked image, and it is the latest methodology proposed for faster and more accurate results than other methods. Once it provides the face's location, it can pass it to FaceNet [3], by which we can extract features from face.

$$L = Lc\,(p_i + p_i^*) + \gamma_1 p_i^* L_{BOX}(t_i + t_i^*) + \gamma_2 p_i^* L_{pts}(l_i + l_i^*) + \gamma_3 p_i^* L_{pixel} \;(1)$$

C. Face Recognition

As shown above, in Fig. 3, the model takes the result images from RetinaFace and passes them to the deep convolutional network. The output is then L2-normalised, and a feature vector called embedding is created that contains features of the images given. Then triplet loss is used for calculating the loss, which enforces the difference between each pair of one person and all other faces. The triplet loss function is defined in Equation 2.

$$\sum_i^N \left[\|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha \right] \;(2)$$

Here, x represents the image, f(x) denotes the image embedding, and alpha denotes the margin between positive and negative pairs acting as a threshold value. xia is an example of an anchor image, xip is an example of a positive image, i.e., a picture of the same person as the anchor image, xin is an example of a negative image, i.e., the image of a different person than the anchor image. Then save the output in a.npz file, a compressed NumPy array. These extracted feature vectors can be used in any classification method to classify the person. Here, linear Support Vector Machine is used as they are very effective at separating face embedding vectors. Here, the SVC class in the SKLEARN [10] library for the classification.

IV. Testbed And Results Discussion

After training the model on the dataset, the system can use the model to evaluate a random person's test data images. Masked Images of a person are used for testing, containing 20-45 images. Then, it uses face embedding of that image, passes it to the model, and makes the person's prediction. A few test cases and their accuracy are shown in Fig. 5. The experiment was tested for six scenarios presented in Table III.

TABLE III: Testing Scenarios

Training Images	Testing Images	Train Result	Test Result
Masked	Masked	92.7	87.11
Masked	Non-Masked	92.6	86.3
Non-Masked	Masked	91.4	89.7
Non-Masked +Masked	Non-Masked	94.58	94.39
Non-Masked +Masked	Masked	94.58	88.69
Non-Masked +Masked	Non-Masked + Masked	94.58	90

The main objective of the work is to design an efficient masked face recognition model with an automated masked training set. Therefore, testing the masked face is essential here.

Fig. 5. Sample Output; *s represents subject

The result table III shows 89.7% and 87.11% accuracy for non-masked and masked training images, respectively, for testing of masked images. As non-masked images do not have occluded images, it should help improve face recognition as more features are extracted and compared. So randomly, 50 images from the original dataset are used in training, resulting in an accuracy of 94.58% for training and 90% for testing.

A. Results and Discussion

Upon further investigation, another observation is that when similar sets of training and testing images, for example, masked or masked + non-masked, are used, they show better training results than opposite sets of images, for instance, masked in training and non-masked in testing.

Also, when both training and testing contain masked images, the model's accuracy drops as features extracted from masked images may not be sufficient for comparison, and the model can fail to identify a person. Scenario 4 of testing non-masked images shows the best results. However, including non-masked images in the training and testing sets gives the best result. The similar observation of using non-masked images on a FaceNet based model shows better accuracy [2]. More feature extraction from non-masked images may be the reason for it. Table III suggests including non-masked during the training phase as it provides the best results among all six scenarios.

The confusion matrix gives us an inside look at the trained model. Therefore, the paper presents the confusion matrix for some cases in Fig. 6. Upon further testing, some blurry images gave a false-positive result. So, a suggestion is to use high-resolution images, but feature extraction will take a long time. Even with a total of 21 subjects and a resolution of 250x250, the feature extraction part took around 50 minutes. It would take more time if more images were used and require higher-end computer hardware. Another suggestion is to use the upscaling image feature, which might be helpful for blurry images but will take time too.

V. Conclusion

Pandemic situations like COVID-19 taught new practices like wearing masks in public gatherings to avoid any viral affection. Hence, a new face recognition algorithm must be kept in practice to test masked faces as well as non-masked face images with better accuracy and real-time speed. And automated masked training set generation is suggested for scalable face recognition systems, and this paper uses that model to automate masked training set generation and face recognition. RetinaFace [45] provides better accuracy in detecting a person's face with less time, even if it is occluded by a mask. RetinaFace for face detection and FaceNet for face recognition perform better when non-masked images are included in the training set along with masked images. The method provides relatively better results compared to the surveyed methods. We want to extend the work in the future to different types and sizes of masks and propose a generic occlusion solution.

References

J. Li, Y. Wang, C. Wang, Y. Tai, J. Qian, J. Yang, C. Wang, J. Li, and F. Huang, "Dsfd: dual shot face detector," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 5060–5069.

H. Hamdi and K. Yurtkan, "Masked face recognition based on facenet pre-trained model".

F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.

K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks", IEEE signal processing letters, 23 (2016), pp. 1499–1503.

P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001, vol. 1, Ieee, 2001, pp. I–I.

"The database of faces". https://cam-orl.co.uk/facedatabase.html, 2002. Accessed:2022-0926.

J. Yang, D. Zhang, A. F. Frangi, and J.-y. Yang, "Two-dimensional pca: a new approach to appearance-based face representation and recognition," IEEE transactions on pattern analysis and machine intelligence, 26 (2004), pp. 131–137.

D. E. King, "Dlib-ml: A machine learning toolkit", The Journal of Machine Learning Research, 10 (2009), pp. 1755–1758.

K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, (2014).

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python, Journal of Machine Learning Research", 12 (2011), pp. 2825–2830.

S. O'Hara and B. A. Draper, "Introduction to the bag of features paradigm for image classification and retrieval", arXiv preprint arXiv:1101.3354, (2011).

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, 2016, pp. 770–778.

I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks", Communications of the ACM, 63 (2020), pp. 139–144.

J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, "Robust principal component analysis:Exact recovery of corrupted low-rank matrices via convex optimization," Advances in neural information processing systems, 22 (2009).

A. Colombo, C. Cusano, and R. Schettini, "Umb-db: A database of partially occluded 3d faces", in 2011 IEEE international conference on computer vision workshops (ICCV workshops), IEEE, 2011, pp. 2113–2119.

A. Savran, N. Alyuz, H. Dibeklioglu, O. C¸eliktutan, B. Gokberk, B. Sankur, and L. Akarun, "Bosphorus database for 3d face analysis", in European workshop on biometrics and identity management", Springer, 2008, pp. 47–56.

P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), vol. 1, IEEE, 2005, pp. 947–954.

K. O'Shea and R. Nash, "An introduction to convolutional neural networks", arXiv preprint arXiv:1511.08458, (2015).

T. Sim, S. Baker, and M. Bsat, "The cmu pose, illumination, and expression (pie) database", in Proceedings of fifth IEEE international conference on automatic face gesture recognition, IEEE, 2002, pp. 53–58.

A. Martinez and R. Benavente, "The ar face database:" Cvc technical report, 24, (1998).

X. Zhang, H. Luo, X. Fan, W. Xiang, Y. Sun, Q. Xiao, W. Jiang, C. Zhang, and J. Sun, "Alignedreid: Surpassing human-level performance in person re-identification," arXiv preprint arXiv:1711.08184, (2017).

Y. Deng, Q. Dai, and Z. Zhang, "Graph laplace for occluded face completion and recognition", IEEE Transactions on Image Processing, 20 (2011), pp. 2329–2338.

C. Y. Wu and J. J. Ding, "Occluded face recognition using low-rank regression with generalized gradient direction", Pattern Recognition, 80 (2018), pp. 256–268.

J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative image inpainting with contextual attention", in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 5505–5514.

D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning face representation from scratch", arXiv preprint arXiv:1411.7923, (2014).

Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "Vggface2: A dataset for recognising faces across pose and age", in 2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018), IEEE, 2018, pp. 67–74.

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

"Histograms of oriented gradients, matlab central file exchange". https://www.mathworks.com/matlabcentral/fileexchange/33863-histograms-of-oriented-gradients. Accessed:2022-09-26.

Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H. Wu, P. Yi, K. Jiang, N. Wang, Y. Pei, et al., "Masked face recognition dataset and application", arXiv preprint arXiv:2003.09093, (2020).

S. Woo, J. Park, J.-Y. Lee, and I. S. "Kweon, Cbam: Convolutional block attention module", in Proceedings of the European conference on computer vision (ECCV), 2018, pp. 3–19.

A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose", IEEE transactions on pattern analysis and machine intelligence, 23 (2001), pp. 643–660.

M. S. Ejaz and M. R. Islam, "Masked face recognition using convolutional neural network", in 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), 2019, pp. 1–6.

M. S. Ejaz, M. R. Islam, M. Sifatullah, and A. Sarker, "Implementation of principal component analysis on masked and non-masked face recognition", in 2019 1st international conference on advances in science, engineering and robotics technology (ICASERT), IEEE, 2019, pp. 1–5.

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

F. Ding, P. Peng, Y. Huang, M. Geng, and Y. Tian, "Masked face recognition with latent part detection", in Proceedings of the 28th ACM international Conference on multimedia, 2020, pp. 2281–2289.

M. Geng, P. Peng, Y. Huang, and Y. Tian, "Masked face recognition with generative data augmentation and domain constrained ranking", in Proceedings of the 28th ACM International Conference on Multimedia, 2020, pp. 2246–2254.

S. Kumaar, A. Dogra, A. Majeedi, H. Gani, R. M. Vishwanath, and S. Omkar, "A supervised learning methodology for real-time disguised face recognition in the wild", arXiv preprint arXiv:1809.02875, (2018).

C. Li, S. Ge, D. Zhang, and J. Li, "Look through masks: Towards masked face recognition with de-occlusion distillation", in Proceedings of the 28th ACM International Conference on Multimedia, 2020, pp. 3016–3024.

N. Alyuz, B. Gokberk, and L. Akarun, "3-d face recognition under occlusion using masked projection", IEEE Transactions on Information Forensics and Security, 8 (2013), pp. 789– 802.

E. J. He, J. A. Fernandez, B. V. Kumar, and M. Alkanhal, "Masked correlation filters for partially occluded face recognition", in 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2016, pp. 1293–1297.

Q. Hong, Z. Wang, Z. He, N. Wang, X. Tian, and T. Lu,"Masked face recognition with identification association", in 2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI), IEEE, 2020, pp. 731–735.

R. Golwalkar and N. Mehendale, "Masked-face recognition using deep metric learning and facemasknet-21", Applied Intelligence, (2022), pp. 1–12.

Y. Li, K. Guo, Y. Lu, and L. Liu, “Cropping and attention based approach for masked face recognition”, Applied Intelligence, 51 (2021), pp. 3012–3025.

Z. Liu, P. Luo, X. Wang, and X. Tang, “Large-scale celebfaces attributes (celeba) dataset”, Retrieved August, 15 (2018), p. 11.

J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, “Retinaface: Single-shot multilevel face localisation in the wild”, in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 5203–5212.

“face-recognition 1.3.0”. https://pypi.org/project/face-recognition/. Accessed:2022-09-26.

Y. Zhu, H. Cai, S. Zhang, C. Wang, and Y. Xiong, “Tinaface: Strong but simple baseline for face detection”, arXiv preprint arXiv:2011.13183, (2020).

T. Evgeniou and M. Pontil, “Support vector machines: Theory and applications, in Advanced Course on Artificial Intelligence”, Springer, 1999, pp. 249–257.

F. Boutros, N. Damer, F. Kirchbuchner, and A. Kuijper, “Unmasking face embeddings by self-restrained triplet loss for accurate masked face recognition”, arXiv preprint arXiv:2103.01716, (2021).

J. Deng, J. Guo, N. Xue, and S. Zafeiriou, “Arcface: Additive angular margin loss for deep face recognition”, in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 4690–4699.

B. Mandal, A. Okeukwu, and Y. Theis, “Masked face recognition using resnet-50”, arXiv preprint arXiv:2104.08997, (2021).

D. Montero, M. Nieto, P. Leskovsky, and N. Aginako, “Boosting masked face recognition with multi-task arcface”, arXiv preprint arXiv:2104.09874, (2021).

A. Anwar and A. Raychowdhury, “Masked face recognition for secure authentication”, arXiv preprint arXiv:2008.11104, (2020).

N. Damer, F. Boutros, M. Sußmilch, M. Fang, F. Kirchbuchner, and A. Kuijper¨, “Masked face recognition: Human vs. machine”, arXiv preprint arXiv:2103.01924, (2021).

W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, “Sphereface: Deep hypersphere embedding for face recognition”, in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 212–220.

“Neurotechnology megamatcher 12.1 sdk”. https://www.neurotechnology.com/mmsdk.html. Accessed:2021-03-02.

G. Wu, “Masked face recognition algorithm for a contactless distribution cabinet”, Mathematical problems in engineering, 2021 (2021).

H. Du, H. Shi, Y. Liu, D. Zeng, and T. Mei, “Towards nir-vis masked face recognition,” IEEE Signal Processing Letters, 28 (2021), pp. 768–772.

H. Du, H. Shi, Y. Liu, J. Wang, Z. Lei, D. Zeng, and T. Mei, “Semi-siamese training for shallow face learning”, in European Conference on Computer Vision, Springer, 2020, pp. 36–53.

S. Li, D. Yi, Z. Lei, and S. Liao, “The casia nir-vis 2.0 face database”, in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2013, pp. 348–353.

J. Chen, D. Yi, J. Yang, G. Zhao, S. Z. Li, and M. Pietikainen, “Learning mappings for face synthesis from near infrared to visual light images,” in 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2009, pp. 156–163.

Y. W. D Huang, J Sun, “The buaa-visnir face database instructions”, School Comput. Sci. Eng., Beihang Univ., Beijing, China, Tech. Rep. IRIP-TR-12-FR-001, 3 (2012), p. 3.

W. Zheng, L. Yan, F.-Y. Wang, and C. Gou, “Learning from the web: Webly supervised meta-learning for masked face recognition, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition”, 2021, pp. 4304–4313.

Z. Wang and T. S. Kim, “Learning to recognize masked faces by data synthesis”, in 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), IEEE, 2021, pp. 036–041.

E. Hoffer and N. Ailon, “Deep metric learning using triplet network”, in International workshop on similarity-based pattern recognition, Springer, 2015, pp. 84–92.

P. Durga, G. Divya, D. Ayeshwariya, P. Sivakumar, et al., “Gabor-deep cnn based masked face recognition for fraud prevention”, in 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), IEEE, 2021, pp. 990–995.

“Masked face detection dataset (mfdd)”. https://zhuanlan.zhihu.com/p/107719641?utmsource=com.yinxiang. Accessed:2022-09-26.

“Face mask detection”. https://www.kaggle.com/datasets/andrewmvd/face-mask-detection. Accessed:2022-09-26.

A. Cabani, K. Hammoudi, H. Benhabiles, and M. Melkemi, “Maskedface-net–a dataset of correctly/incorrectly masked face images in the context of covid-19”, Smart Health, 19 (2021), p. 100144.

K. O. M. Aarif and P. Sivakumar, “Cursive script identification using gabor features and svm classifier”, International Journal of Computer Aided Engineering and Technology, 12 (2020), pp. 328–335.

M. Shah, D. Shukla, and D. Pandya, “Smart gate: Intelligent security system based on face recognition”, International Journal of Innovative Technology and Exploring Engineering (IJITEE), (2020).