Face recognition under mask-wearing based on residual inception networks

Warot Moungsouy, Thanawat Tawanbunjerd, Nutcha Liamsomboon
and Worapan Kusakunniran

*Faculty of Information and Communication Technology, Mahidol University,
Nakhon Pathom, Thailand*

Abstract

Purpose – This paper proposes a solution for recognizing human faces under mask-wearing. The lower part of human face is occluded and could not be used in the learning process of face recognition. So, the proposed solution is developed to recognize human faces on any available facial components which could be varied depending on wearing or not wearing a mask.

Design/methodology/approach – The proposed solution is developed based on the FaceNet framework, aiming to modify the existing facial recognition model to improve the performance of both scenarios of mask-wearing and without mask-wearing. Then, simulated masked-face images are computed on top of the original face images, to be used in the learning process of face recognition. In addition, feature heatmaps are also drawn out to visualize majority of parts of facial images that are significant in recognizing faces under mask-wearing.

Findings – The proposed method is validated using several scenarios of experiments. The result shows an outstanding accuracy of 99.2% on a scenario of mask-wearing faces. The feature heatmaps also show that non-occluded components including eyes and nose become more significant for recognizing human faces, when compared with the lower part of human faces which could be occluded under masks.

Originality/value – The convolutional neural network based solution is tuned up for recognizing human faces under a scenario of mask-wearing. The simulated masks on original face images are augmented for training the face recognition model. The heatmaps are then computed to prove that features generated from the top half of face images are correctly chosen for the face recognition.

Keywords Face recognition, Mask wearing, Partial occlusion, Heatmaps

Paper type Research paper

1. Introduction

Currently, there has been an outbreak of the COVID-19 pandemic [1, 2] that is the defining global health crisis and one of the most challenges that the world has faced since previous years. One way to slow down the spread of the disease is to wear a face mask in public areas. However, the masked faces become more challenging to the existing face recognition systems [3–5]. The adoption of face recognition in this pandemic situation could identify the main difficulty of masked face recognition, when compared with the mask removed. Moreover, several studies show that the effect of wearing a mask on face recognition points out a large drop in the recognition performance [6–9]. Therefore, developing and studying masked face recognition can beneficially enhance the potential of a facial recognition system to support any aspects of the situation. In addition, deep learning has certainly been one of the most successful techniques for the face recognition system [10–16].

Before a step of recognizing human faces under mask-wearing, the faces with or without the mask-wearing must be detected. For example, Loey *et al.* [17] proposed the networks that were able to detect masks on human face images with an average precision of up to 81% on a custom dataset combined from Medical Masks Dataset and Face Mask Dataset. In this model, the ResNet-50 was used for the feature extraction, where the YOLO v2 was deployed as the mask detector. Kumar *et al.* [18] proposed the mask detection system based on tiny YOLO v4. The network was improved by adding a spatial pyramid pooling module at the end of the feature extraction step. This was to improve the small-sized object detection. It was tested using the self-created face masks detection dataset. It achieved an average precision of up to 84%.

Moreover, there are several approaches introduced for the masked face recognition in recent studies. Mandal and Okeukwu [19] fine-tuned the pre-trained ResNet-50 model on their dataset of faces without masks. Then, the model was operated on the masked faces, with additional fine-tuning step based on the previous results of identifying individuals without masks.

They considered many alternative approaches such as cropping the occluded part and supervised domain adaptation to the resulting model. Li and Guo [20] proposed an attention-based approach to focus on regions around eyes by integrating a cropping-based approach with the Convolutional Block Attention Module. The cropping helped the model to get more attention on extracting features of face images. Then, an attention mechanism was embedded in every convolution block of ResNet-50 to refine feature maps. Boutros and Damer [21] presented the Embedding Unmasking Model operated on top of existing face recognition models, with the Self-Restrained Triplet loss function. Deng and Feng [22] proposed a masked-face recognition algorithm based on the large margin cosine loss (MFCosface). The restoration approach was applied to remove a mask from each face image. Then, such missing information was restored to complete the face.

Recently, Li and Ge [23] proposed an end-to-end de-occlusion distillation framework to migrate the mechanism of amodal completion for the task of masked face recognition. Din and Javed [24] employed a GAN-based network using two discriminators where one discriminator helped in learning the global structure of the face and another discriminator was used to learn the deep missing region. Based on our literature reviews, there have been many contributions to address this challenge. The restoration-based techniques were new approaches in the field of face recognition. However, the restoration approach was sensitive to a variety of conditions such as light, occluded items and segmentation results of detected masks. So, it led to an imperfect generated face image, which dropped the recognition accuracy. Then, the transfer learning approach focused on enhancing the existing face recognition models on different details of techniques and datasets. The main focus was to find the best setup of the model that can recognize the masked faces. Many researchers have been studying on the same challenge of finding the best setup based on their experiments.

This paper introduces a new solution to recognize human faces under mask-wearing with the Inception-ResNet-v1 and our simulated masked face dataset. The augmentation of simulated masked face images is applied to original face images without masks. Several experiments are conducted to find the best setup of the model. Details of the proposed method are explained in Section 2. The experiment and discussion are described in Sections 3 and 4, respectively. Then, the conclusion is drawn in Section 5.

2. Proposed method
This section explains details of the proposed method, where some related supplementary materials of additional figures and tables are located in https://github.com/mwarot1/fr-undermaskwearing.

2.1 Overview

This research project aims to modify the existing facial recognition model, to cover both scenarios of mask-wearing and without mask-wearing face images. It consists of three processes including data acquisition, data preprocessing and data modeling, as shown in Figure 1.

2.1.1 Data acquisition. The first step is to use the public face databases for the data modeling process. This paper uses the two well-known face datasets which are publicly available, including the CASIA-WebFace and LFW [25, 26] datasets. The CASIA-WebFace dataset is a collection of 10,575 unique identities of celebrities with 494,414 images. The data was collected from the IMDb website. In addition, the LFW is a public benchmark test set for the face verification. The dataset contains 13,233 images of 5,749 identities. The face images were also collected from the web. Both datasets are completely independent in terms of identities.

2.1.2 Data preprocessing. The data preprocessing step is to create a completed dataset for data modeling and model evaluation. Two sub-processes are used in the dataset. The first
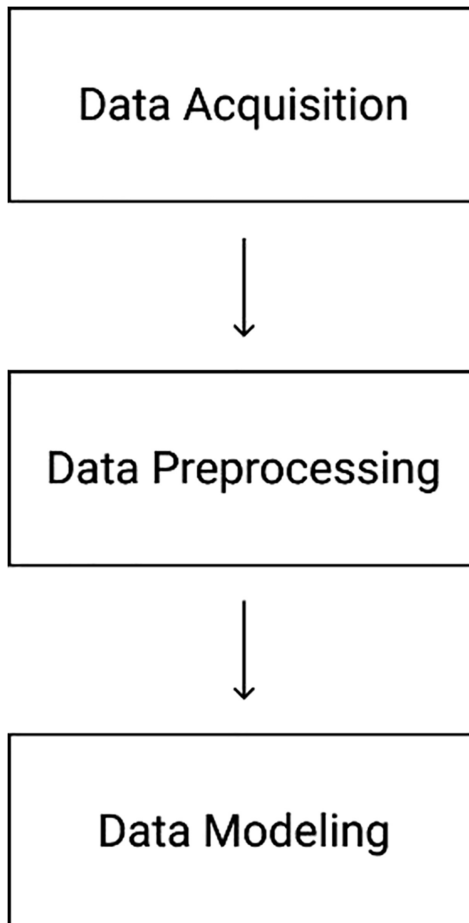

Figure 1.
Overview of the
proposed method

sub-process is to create simulated masked-face images using an open-source tool, namely MaskTheFace. Then, the Multi-task cascaded convolutional neural networks model is applied to crop the face images [27]. MaskTheFace is a computer vision-based script to generate a masked face from an original face image with extended feature supports. This process is used to create different variations of the simulated mask face dataset. The flow to create the masked face dataset is shown in Figure 2. The second sub-process is to split the dataset into two sets, which are the training set for 80% and the validation set for 20%. The training set contains the samples used to train the model for classifying individuals. The validation set is then used to provide an unbiased evaluation of a fitting model while tuning model's hyperparameters. These two sub-processes are repeatable, so the process can work iteratively to create various scenarios of datasets and test cases.

2.1.3 Model training. In the model training step, a convolutional neural network (CNN) [28–31] based approach is created for the face recognition task. The Inception-ResNet-v1 [32–34], a deep CNN architecture with a combination of Inception block and residual neural network, is deployed as our baseline network. The Inception-ResNet-v1 architecture is represented in Figure 3. In each training epoch, each training sample is parsed forward to fit and improve model's weights. Next, it is back-propagated for obtaining the minimum value of the error function in the weight space. The trained model is used for the feature extractor to validate the results from the validation dataset. Moreover, the callback function is set to monitor the validating loss. So, the training process will stop if validating loss starts to increase or is still the same as the last epoch.

2.2 Dataset

2.2.1 M-CASIA. M-CASIA is our custom dataset that was created based on the CASIA-WebFace dataset. The M-CASIA dataset consists of 689,686 images with 10,575 identities. Each identity can be divided into two subcategories. The first subcategory contains the normal face images from the CASIA-WebFace dataset. It contains 453,525 face images, which is roughly about 2:3 of the M-CASIA dataset. The next step is to compute the simulated masked-face images in the second subcategory. The masked part of the simulated face images is generated using the open-source tool MaskTheFace on the CASIA dataset. It consists of 236,161 simulated masked-face images, which is roughly about 1:3 of the M-CASIA dataset. The number of simulated face images for each identity is 50% on average. Further, this dataset includes only four variations of a mask, which are surgical green, surgical white, cloth black and cloth white.

2.2.2 LFW30. LFW30 is a subset of the LFW dataset. This dataset filtered only the identities that contain more than or equal to 30 face images. LFW30 has been used to create our custom dataset for the model testing process, which includes SMF-LFW30 and
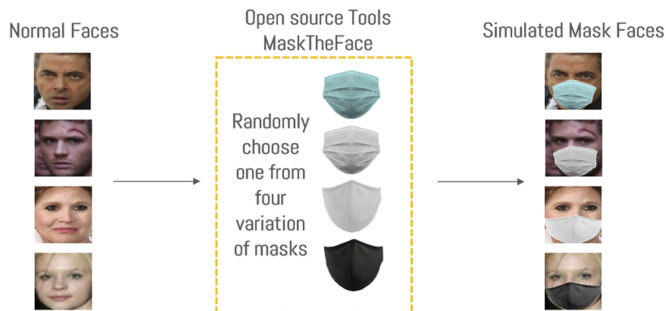

Figure 2.
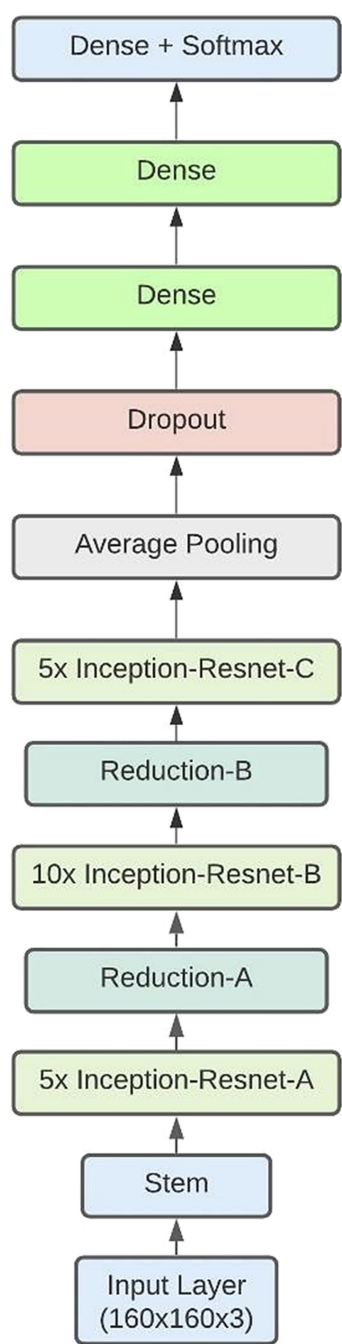Data preprocessing flow. Face images used in this figure are from the publicly available dataset, namely CASIA-WebFace dataset, at https://paperswithcode.com/dataset/casia-webface

M-LFW30. First, the SMF-LFW dataset consists of 125 simulated masked-face images with 32 identities. Second, the M-LFW30 dataset consists of 272 normal face and simulated masked-face images with 32 identities. Both datasets contain four variations of masks, which are surgical green, surgical white, cloth black and cloth white.

2.3 Experiment setup

The experiments are designed to create the optimal model for recognizing human faces from both mask-wearing and without mask-wearing scenarios by improving the performances with custom datasets and network tuning. To begin with, the M-CASIA dataset was prepared for the model training process. The step of fine-tuning the network [35–37] requires an appropriate dataset to shift the network's attention correctly. So, tuning the network with both mask-wearing and without mask-wearing face images could help the model to understand key features for recognizing both scenarios. Our adopted base network is the Facenet which uses the Inception-ResNet-v1 as the main architecture. Each part of Inception-ResNet-v1 is separated by the inception blocks including Block A, Block B and Block C, as shown in Figure 4. In between the connection of each block, there will be a reduction block which helps in reducing the dimension before being passed to the next Inception block. Moreover, two dense layers have been added as trainable layers on top of the original network.

With the transfer learning [38], we can transfer the initial weights and train the model using the M-CASIA dataset. The training process could converge faster than training the model from scratch. The next step is to fine-tune the model. This step is an iterative process to find the best setting for the model training. Finally, the Adam optimizer is used with the learning rate of 0.0001 and the categorical cross-entropy is applied as the loss function [39, 40]. The accuracy, precision and recall are used as the measurement metrics. During the training process, we set a callback function to save the best model based on the monitoring of the validation loss on each epoch. In the case of face verification, the better model means the better feature extraction it could perform on the face images. Therefore, we use a feature heatmap to explore and prove that the trained model could perform well in the feature extraction. The heatmap is created to represent the weights of pixels. Moreover, we examine the relationship between input data and face database by using gallery and probe evaluation experiments. In the real-world scenario, users must register their faces to the recognition system first. Then, the system can start to recognize individual identities by comparing input images with face images in the database. Therefore, the gallery and probe experiment can identify the best setup for the face database that covers as many input variations as possible.

3. Experiment

This section describes two main scenarios of experiments, where additional figures and tables of supplementary results are located in https://github.com/mwarot1/fr-undermaskwearing.

3.1 Experiment #1: network tuning

In this experiment, four setups based on the Inception-ResNet-v1, which consist of different unfreezing parts of the network, are evaluated. Each model is trained on the upper layers starting from 1) setup1: the last two dense layers, 2) setup2: Block C, 3) setup3: Block B and 4) setup4: Block A. All the models are trained using the same training parameters, initial weights and training set. After the training process, we evaluate each model with different combinations of gallery and prob. Then, the comparison graphs of the four models on the gallery sets of LFW30 and M- LFW30 are shown in Figures 5 and 6, respectively.
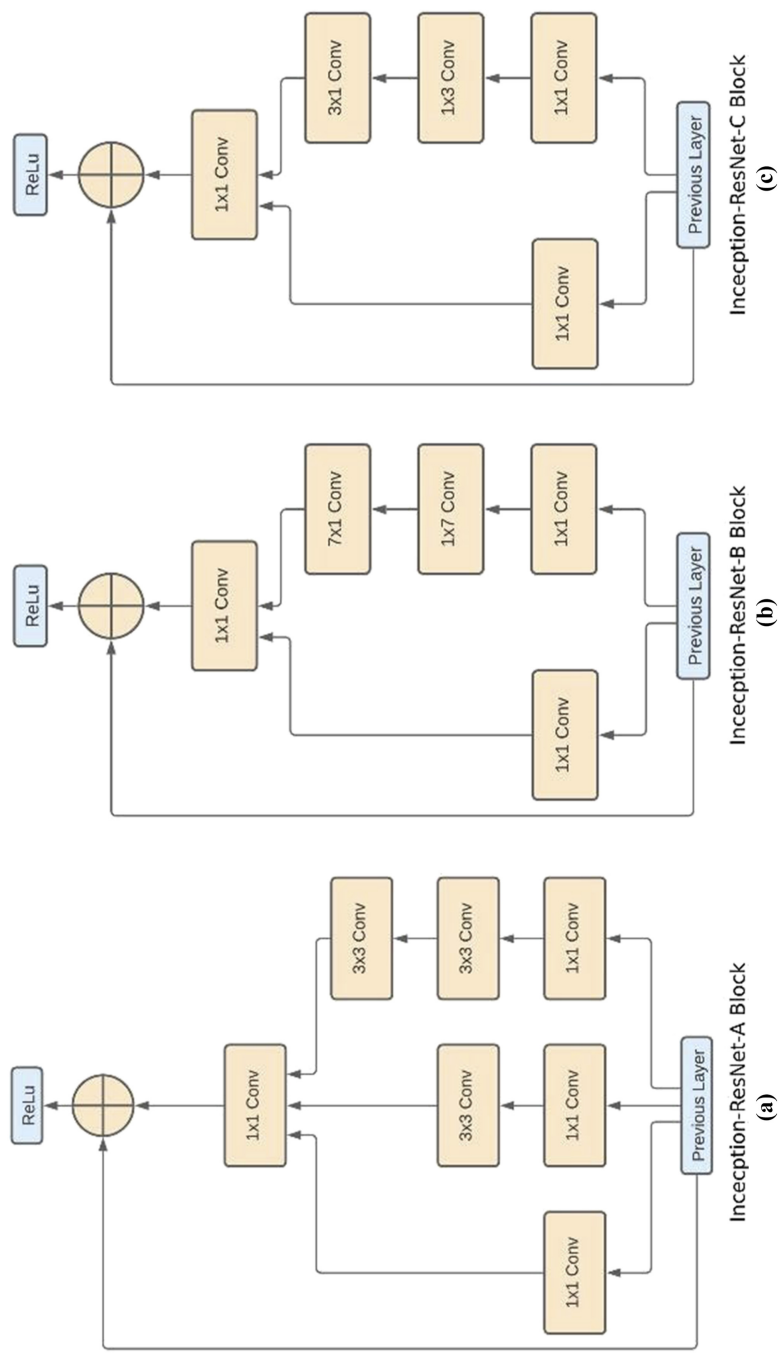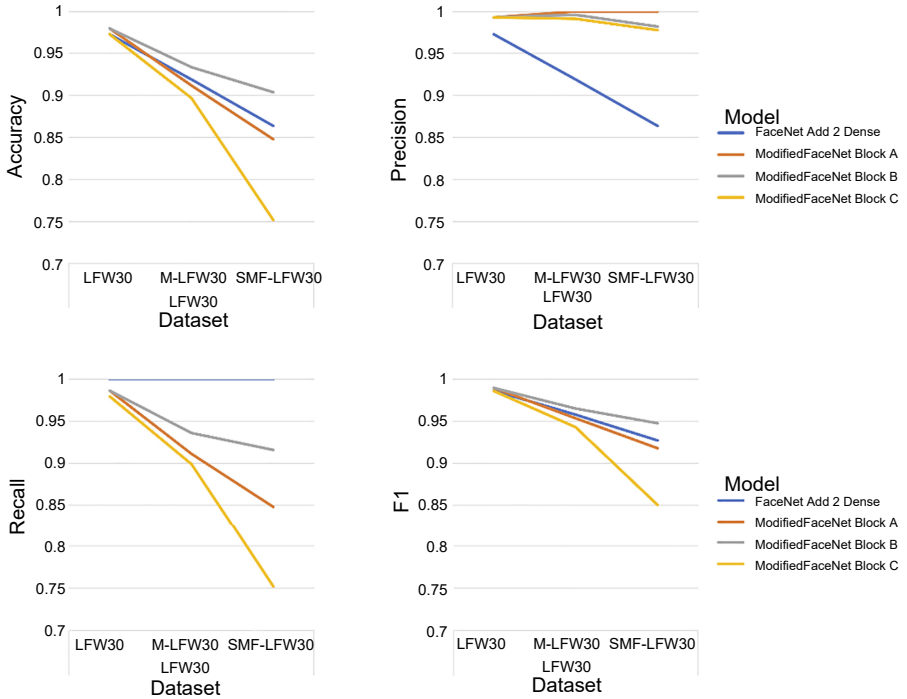
Figure 5.
Comparison of the four models (Experiment#1) on the gallery set of LFW30

3.2 Experiment #2: gallery and probe

Our best approach from experiment #1 is chosen as an experimental model, which is Modified FaceNet Block A, as demonstrated in Figures 5 and 6. The model is trained with the M-CASIA dataset, which focuses on four types of face masks. The result of experiment #1 shows that the gallery must consist of both masked and unmasked face images. For the gallery and probe, the LFW30 dataset has been used as an initial dataset to create many galleries and probes for the test scenario. The gallery is one of the data partitions, which act as a collection of database or search datasets. The gallery of LFW30 contains 3,384 images with 32 identities. Instead, a probe is a collection of data that needs to be recognized from our model by comparing it with all images in a gallery using a classification algorithm. The probe of LFW30 contains 125 images with 32 identities. With MaskTheFace, we create a simulated masked-face dataset with multiple combinations of four mask types and colors, including surgical green, surgical white, cloth black and cloth white. Besides, additional test problem that would be found in a real-world scenario, which is an out-scope color mask, is evaluated. All information about each dataset used in the gallery and probe experiment is shown in Table 1.

In this experiment, we set up different combinations of gallery and probe sets [41, 42] for evaluating the recognition system. The gallery set is a mix of unmasked face and masked face images, which contains some variations of masks' colors identified by the dataset codes shown in Table 1. The probe set consists of masked face images, such that the variations of masks' colors are also based on the dataset codes shown in Table 1. For each iteration, an image from the probe set is fed into the model for the feature extraction process. Then, it will be compared with every feature vector extracted from all data samples in the gallery set,
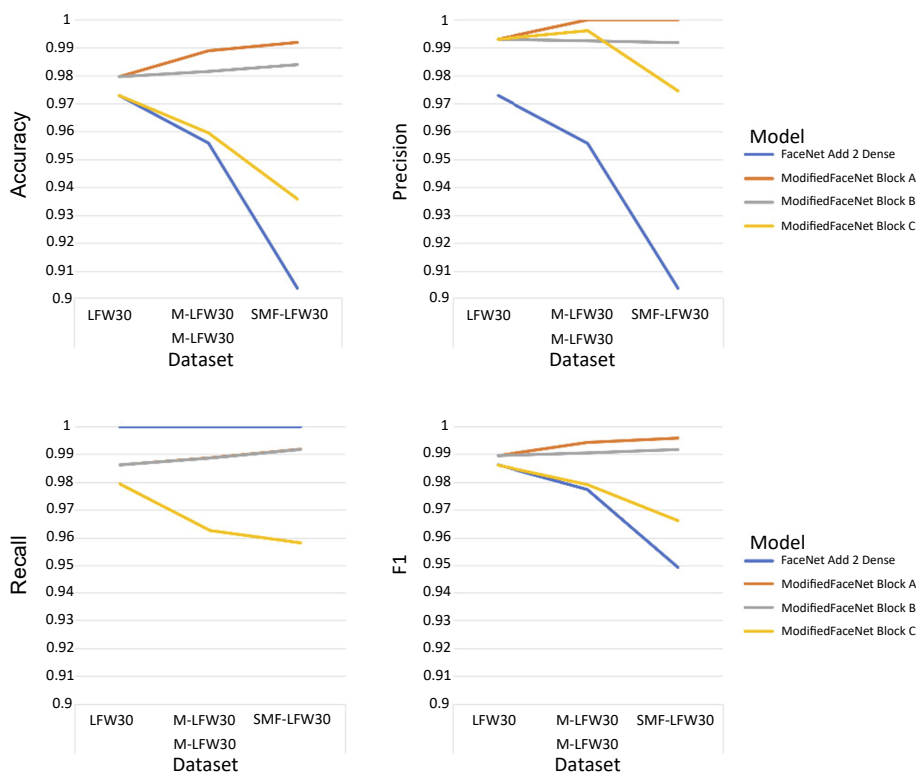
using the cosine similarity [43]. Finally, the K-nearest neighbor algorithm [44] is applied to get the top three matches from the gallery set. These three closest scored identities will be voted to return the final identity. The accuracy on each gallery set lies around 98% to 99% on average for all probe sets.

4. Discussion

4.1 Result

We perform two main experiments that aim to find the best model and setup for recognizing human faces under mask-wearing. In experiment #1, we have improved the performance of the Inception-ResNet-v1 with augmented data of simulated masked-face images and network tuning, in order to find the best setup for the masked-face recognition model. Moreover, we take our best approach to be evaluated with the real-world set of data, to seek out the limitations of our model in experiment #2.

In the first experiment, we improve the performance of FaceNet with the augmented data of simulated masked-face images and the network tuning, to recognize both mask-wearing and without mask-wearing faces. It is shown to improve the performance of the original FaceNet model. Our best approach is to fine-tune the network with the M-CASIA dataset starting from the last dense layer until the inception of Block A, which covers almost 80% of the entire Inception-ResNet-v1 network. Modified FaceNet Block A achieves the best accuracy among the other test cases. The accuracy increases by around 62.4% when

| | | Mask variation | | | For gallery |
Dataset	Cloth black	Cloth white	Surgical white	Surgical green	No mask
LFW-4	✔	✔	✔	✔	✔
LFW-31	✔	✔	✔		
LFW-32	✔	✔		✔	
LFW-33	✔		✔	✔	
LFW-34		✔	✔	✔	
LFW-21	✔	✔			
LFW-22	✔		✔		
LFW-23	✔			✔	
LFW-24		✔	✔		
LFW-25		✔		✔	
LFW-26			✔	✔	
LFW-11	✔				
LFW-12		✔			
LFW-13			✔		
LFW-14				✔	
LFW-00					
LFW-unknown	Consists of unknown mask color				

Table 1.
Information about each dataset used in the gallery and probe experiments

evaluating with the masked-face dataset and does not decrease the performance of the unmarked-face recognition. Overall, our best approach can improve the performance of the original FaceNet model on the same training parameters and pre-trained weights. The model should be fine-tuned with the balanced mask-wearing and without mask-wearing datasets. Examples of feature heatmaps from modified FaceNet Block A are shown in Figure 7.

In the second experiment, we have created multiple combinations of simulated masked-face datasets for gallery and probe evaluations. After the investigation, it is observed that registering a normal face along with its simulated masked-face in the database is the best setup for real-world usage. However, the accuracy of using the mixed database is only 0.6% higher than using the original unmasked-face database. Therefore, both mixed database and original unmasked-face database can be applied in a real-world application. It depends on the situation of the system or the organization if 0.6% higher accuracy could be fairly exchanged with a double space consumption of the computational resource. Next, the variation of the types of masks including colors and patterns does not affect the performance of the model. This is because the key value of the face feature is shifting to the upper part of the face, not on the masked area. For this reason, registered masked-face images in the database can be in any color or pattern.

4.2 Comparison with other approaches

Due to the variety of datasets used in the testing, results could be fluctuated based on the test datasets used. First, Anwar and Raychowdhury [45] used a similar approach to our proposed method. They used the existing FaceNet and retrained the network with the custom dataset generated using the MaskTheFace. This technique achieved an accuracy of 97.25% on simulated masked-faces of LFW dataset. They also reported that the model could achieve a roughly 38% increase in true positive rate, when compared with the original model. In addition, Ding *et al.* [46] applied the CNN and the latent part detection approach using two branches of CNN to separately learn from the global and the partial part of human faces. The global branch learned the full face with occlusion, while the partial branch learned the face without the occlusion. The model achieved 95.7% on the synthesized LFW dataset.

In addition, David [47] used an ArcFace model as a baseline and modified some of the backbones and loss functions. By using the LResNet-50 as a backbone and adding a newly

created dense layer, the method obtained two logits as the output. Adding them together created the MTArcFace loss function. Then, the total loss was created using the MTArcFace loss and the regularization. The evaluation of the MTArcFace on the masked-faced LFW dataset achieved up to 98.92% accuracy. The accuracy comparisons are shown in Table 2.

5. Conclusion

This research work developed the technique for recognizing human faces under both scenarios of mask-wearing and non-mask-wearing. The proposed method was based on the FaceNet model using the residual inception network of Inception-ResNet-v1 architecture. In addition, the simulated masked-face images were constructed on top of the original unmasked-face images from the publicly available face datasets. Both simulated masked-face images and original unmasked-face images were applied in the transfer learning process of the original FaceNet model. The best model based on our experiments was the fine-tuned FaceNet with the retraining from Inception Block A on the M-CASIA dataset. In the evaluation, this model achieved 99.2% accuracy on the masked-face test dataset. Despite the variety of masks in a real-world situation, the model could recognize faces with any type of mask, varying in colors and patterns. Also, from the experiments, we could conclude that having masked-face images along with the original unmasked-face images in the gallery

Paper	Our	Anwar A	Feifei D	David M
Test dataset	SMF-LFW	LFW-SM	Syn-LFW	Masked LFW
Technique	Transfer learning	Augmentation	LPD	MT ArcFace
Accuracy (%)	99.20	97.25	95.70	98.92

ACI

database could improve the accuracy of the model by 0.6%. However, this would consume a double space of the computation resource for storing the database. However, the proposed method also has the limitation. Since the dataset that we used was the simulated masked-face images, any unrealistic part in the simulated images might cause some inaccuracies in the recognition. Therefore, in the future work, to improve the recognition performance, the proposed model could be further trained with real masked-face images. Also, another attempt could be retraining the model with face images without a bottom part covered by a face mask. In terms of the application-based usage, the trained model could be plugged-in to a web application, as an example, with user-friendly interfaces.

References

1. Paules CI, Marston HD, Fauci AS. Coronavirus infections–more than just the common cold. JAMA. 2020; 323(8): 707-8. doi: 10.1001/jama.2020.0757.

2. How to protect yourself & others from coronavirus disease. 2020. Available from: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html.

3. Deng J, Guo J, An X, Zhu Z, Zafeiriou S. Masked face recognition challenge: the insightface track report. Proceedings of the IEEE/CVF International Conference on Computer Vision; 2021. p. 1437-44.

4. Zhu Z, Huang G, Deng J, Ye Y, Huang J, Chen X, Zhu J, Yang T, Guo J, Lu J, Du D. Masked face recognition challenge: the webface260m track report. arXiv preprint arXiv:2108.07189. 2021 Aug 16. p. 1-8.

5. Boutros F, Damer N, Kolf JN, Raja K, Kirchbuchner F, Ramachandra R, Kuijper A, Fang P, Zhang C, Wang F, Montero DMFR. Masked face recognition competition. 2021 IEEE International Joint Conference on Biometrics (IJCB), 2021 Aug 4. IEEE; 2021. p. 1-10.

6. Fitousi D, Rotschild N, Pnini C, Azizi O. Understanding the impact of face masks on the processing of facial identity, emotion, age, and gender. 2021 nov 3. Front Psychol. 2021; 12: 743793. doi: 10.3389/fpsyg.2021.743793.

7. Fu B, Kirchbuchner F, Damer N. .The effect of wearing a face mask on face image quality. 2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021). 2021 Dec 15. IEEE. p. 1-8.confproc.

8. Damer N, Grebe JH., Chen C, Boutros F, Kirchbuchner F, Kuijper A. The effect of wearing a mask on face recognition performance: an exploratory study. 2020 International Conference of the Biometrics Special Interest Group (BIOSIG); 2020. p. 1-6.

9. Freud E, Stajduhar A, Rosenbaum RS, Avidan G, Ganel T. The COVID-19 pandemic masks the way people perceive faces. Scientific reports. 2020 Dec 21; 10(1): 1-8.

10. Schroff F, Kalenichenko D, Philbin J. Facenet: a unified embedding for face recognition and clustering. Proceedings of the IEEE conference on computer vision and pattern recognition; 2015. p. 815-23.

11. Liu J, Deng Y, Bai T, Wei Z, Huang C. Targeting ultimate accuracy: face recognition via deep embedding. arXiv preprint arXiv:1506.07310. 2015 Jun 24.

12. Yu W, Yang K, Bai Y, Yao H, Rui Y. Visualizing and comparing convolutional neural networks. arXiv preprint arXiv:1412.6631. 2014 Dec 20.

13. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, inception-resnet and the impact of residual connections on learning. Thirty-first AAAI conference on artificial intelligence 2017 Feb 12.

14. Mi Q, Keung J, Xiao Y, Mensah S, Mei X. An inception architecture-based model for improving code readability classification. Proceedings of the 22nd International Conference on Evaluation and Assessment in Software Engineering, 2018 Jun 28; 2018. p. 139-44.

15. Yi D, Lei Z, Liao S, Li SZ. Learning face representation from scratch. arXiv preprint arXiv: 1411.7923. 2014 Nov 28. p. 1-9.

16. AbdAlmageed W, Wu Y, Rawls S, Harel S, Hassner T, Masi I, Choi J, Lekust J, Kim J, Natarajan P, Nevatia R. Face recognition using deep multi-pose representations. 2016 IEEE winter conference on applications of computer vision (WACV), 2016 Mar 7. IEEE. p. 1-9.

17. Loey M, Manogaran G, Taha MH, Khalifa NE. Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. Sustainable cities and society. 2021 Feb 1; 65: 102600: 1-8.

18. Kumar A, Kalia A, Sharma A, Kaushal M. A hybrid tiny YOLO v4-SPP module based improved face mask detection vision system. J Ambient Intelligence Humanized Comput. 2021 Oct 20; 1-14.

19. Mandal B, Okeukwu A, Theis Y. Masked face recognition using ResNet-50. arXiv preprint arXiv: 2104.08997. 2021 Apr 19.

20. Li Y, Guo K, Lu Y, Liu L. Cropping and attention based approach for masked face recognition. Appl Intelligence. 2021 May; 51(5): 3012-25.

21. Boutros F, Damer N, Kirchbuchner F, Kuijper A. Unmasking face embeddings by self-restrained triplet loss for accurate masked face recognition. arXiv preprint arXiv:2103.01716. 2021 Mar 2. p. 1-15.

22. Deng H, Feng Z, Qian G, Lv X, Li H, Li G. MFCosface: a masked-face recognition algorithm based on large margin cosine loss. Appl Sci. 2021 Jan; 11(16): 7310.

23. Li C, Ge S, Zhang D, Li J. Look through masks: towards masked face recognition with de-occlusion distillation. Proceedings of the 28th ACM International Conference on Multimedia, 2020 Oct 12. p. 3016-24.

24. Din NU, Javed K, Bae S, Yi J. A novel GAN-based network for unmasking of masked face. IEEE Access. 2020; 8 Mar 2: 44276-87.

25. Huang GB, Mattar M, Berg T, Learned-Miller E. Labeled faces in the wild: a database for studying face recognition in unconstrained environments. Workshop on faces in' Real-Life'Images: detection, alignment, and recognition 2008 Oct.

26. Huang GB, Learned-Miller E. Labeled faces in the wild: updates and new reporting procedures. Technical Report 2014 May, 14(003). Amherst, MA: Department of Computer Science, University of Massachusetts Amherst.

27. Zhang K, Zhang Z, Li Z, Qiao Y. Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Process Lett. 2016 Aug 26; 23(10): 1499-503.

28. Bansal A, Castillo C, Ranjan R, Chellappa R. The do's and don'ts for cnn-based face verification. Proceedings of the IEEE international conference on computer vision workshops. 2017 Oct 29. p. 2545-54.

29. Bell S, Bala K. Learning visual similarity for product design with convolutional neural networks. ACM transactions on graphics (TOG). 2015 Jul 27; 34(4): 1-0.

30. Lu P, Song B, Xu L. Human face recognition based on convolutional neural network and augmented dataset. Syst Sci Control Eng. 2021 May 3; 9(sup2): 29-37.

31. Hassan RJ, Abdulazeez AM. Deep learning convolutional neural network for face recognition: a review. Int J Sci Business. 2021; 5(2): 114-27.

32. Peng S, Huang H, Chen W, Zhang L, Fang W. More trainable inception-ResNet for face recognition. Neurocomputing. 2020 Oct 21; 411: 9-19.

33. Sharma V, Gangaraju S, Sharma VK. Masked face recognition. Stanford University; 2019. 1-7.

34. Gwyn T, Roy K, Atay M. Face recognition using popular deep net architectures: a brief comparative study. Future Internet. 2021 Jul; 13(7): 164.

35. Lu Z, Jiang X, Kot A. Enhance deep learning performance in face recognition. 2017 2nd International Conference on Image, Vision and Computing (ICIVC), 2017 Jun 2. IEEE. p. 244-8.

36. Wang YX, Ramanan D, Hebert M. Growing a brain: fine-tuning by increasing model capacity. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017. p. 2471-80.

37. Kim J, Jo H, Ra M, Kim WY. Fine-tuning approach to nir face recognition. ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019 May 12. IEEE. p. 2337-41.

38. Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, Xiong H, He Q. A comprehensive survey on transfer learning. Proceedings of the IEEE, 2020 Jul 7, 109(1). p. 43-76.

39. Srivastava Y, Murali V, Dubey SR. A performance evaluation of loss functions for deep face recognition. Revised Selected Papers 2020. Computer Vision, Pattern Recognition, Image Processing, and Graphics: 7th National Conference, NCVPRIPG 2019, December 22-24, 2019, Hubballi, India, 1249. Springer Nature. p. 322.

40. Gordon-Rodriguez E, Loaiza-Ganem G, Pleiss G, Cunningham JP. Uses and abuses of the cross-entropy loss: case studies in modern deep learning. ICBINB@NeurIPS, 2020 Dec 6. p. 1-10.

41. Banerjee S, Das S. Mutual variation of information on transfer-CNN for face recognition with degraded probe samples. Neurocomputing. 2018 Oct 8; 310: 299-315.

42. Shang K, Huang ZH, Liu W, Li ZM. A single gallery-based face recognition using extended joint sparse representation. Appl Math Comput. 2018 Mar 1; 320: 99-115.

43. Nguyen HV, Bai L. Cosine similarity metric learning for face verification. Asian conference on computer vision 2010 Nov 8, Berlin, Heidelberg. Springer. p. 709-20.

44. Li B, Han L. Distance weighted cosine similarity measure for text classification. International conference on intelligent data engineering and automated learning 2013 Oct 20, Berlin, Heidelberg. Springer. p. 611-18.

45. Anwar A, Raychowdhury A. Masked face recognition for secure authentication. arXiv preprint arXiv:2008.11104. 2020 Aug 25.

46. Ding F, Peng P, Huang Y, Geng M, Tian Y. Masked face recognition with latent part detection. Proceedings of the 28th ACM International Conference on Multimedia, 2020 Oct 12. p. 2281-9.

47. Montero D, Nieto M, Leskovsky P, Aginako N. Boosting masked face recognition with multi-task ArcFace. arXiv preprint arXiv:2104.09874. 2021 Apr 20.

Corresponding author

Worapan Kusakunniran can be contacted at: worapan.kun@mahidol.edu