Information System and Smart City 2023; 3(1): 467.
Original Research Article

Effective approach of face mask position detection and recognition

Om Pradyumana Gupta[1,*], Arun Prakash Agrawal[2], Om Pal[3]

[1] *Department of Computer Science & Engineering, Sharda University, Greater Noida, Uttar Pradesh 201 306, India*

[2] *Department of Computer Science & Application, Sharda University, Greater Noida, Uttar Pradesh 201 306, India*

[3] *Department of Computer Science, University of Delhi, New Delhi 110 007, India*

*** Corresponding author:** Om Pradyumana Gupta, op.gupta@nic.in

ARTICLE INFO

Received: 10 August 2023
Accepted: 18 September 2023
Available online: 9 October 2023

doi: 10.59400/issc.v3i1.467

ABSTRACT: During the recent COVID-19 pandemic across the world, face masks became necessary to stop the spread of infection. This has led to challenges with effective detection and recognition of human faces using the existing face detection systems. This paper proposes a Convolutional Neural Network (CNN)-based face mask recognition system, which offers two solutions—recognition of the person wearing the face mask and position of the face mask, i.e., whether the mask is correctly worn or not. The proposed model could play an instrumental role in face recognition. In the first stage, with the help of the Viola-Jones algorithm, the model detects the position of the face mask. In the second stage, we identify the person with a modified pre-trained face mask recognition that the that the DeepMaskNet model facilitates in identifying the person. The proposed model achieves an accuracy of 94% in detecting the face mask position and 99.96% in identifying the masked person. Lastly, a comparison with the existing models is detailed, proving that the proposed model achieves the highest and greatest performance.

KEYWORDS: DeepMaskNet; Viola-Jones; Convolutional neural network; Pooling layers; Softmax

1. Introduction

People must wear face masks in public places all around the world to reduce the impact of COVID 19. Controlling the spread of COVID-19 has been currently a major problem for WHO policymakers and the entire human race. Wearing a face mask lowers the spread of COVID-19 by lowering the likelihood of respiratory droplets being transmitted, according to most data from the WHO. Droplets can transmit coronavirus. The proper usage of face masks during the pandemic is essential to limiting and preventing the spread of COVID-19 amongst individuals[1,2]. To stop the spread of COVID-19, many nations mandate the use of face masks in public settings. It is necessary to develop automated techniques for spotting the face masks. Additionally, wearing a face mask presents unique difficulties for conventional facial recognition software, which is normally made for faces that are exposed. The failure of facial recognition techniques when used with face masks has created considerable difficulties for application verification and authentication. Additionally, the COVID epidemic has inhibited the adoption of several conventional biometric-based techniques, including fingerprint recognition. Despite the fact that recent research has suggested useful techniques for masked facial recognition, masks still greatly obscure a face[3] and research in the detection of proper use of face masks is limited. Although there has been tremendous research in the domain of face mask identification[4–7]. This paper presents a unique approach to not only identify the person wearing the face mask but also inform whether the mask is worn properly or not. One of the most effective ways to stop transmission is to correctly wear face masks. Such an automatic device may be installed at busy public gates to help monitor and prevent people from entering without masks or with the wrong kind of mask on.

The paper is structured as follows: In the next section, literature review is discussed. Methodology is detailed in Section 3. The dataset is described in Section 4. Section 5 is the Experimental Setup, Section 6 is Results, and Section 7 is Conclusion.

2. Literature review

In 2022, Rahman et al.[8] and Venkateswarlu et al.[9] developed a face mask position identification system using MobileNet. They have created a system that can identify different face mask placement mistakes. They achieved this by including additional layers, such as dropout, and making more advantageous choices of the number of thick layers, which significantly improved the classification accuracy. In 2021, Ullah et al.[10] are able to both detect face masks and recognize masked faces and suggest a revolutionary DeepMasknet framework that can do both. Because there is no dataset of masked people, they create their own. In 2020, Fan and Jiang[11] developed a model based on RetinaNet architecture to check whether a face mask is present. They extracted characteristics from photos using ResNet-50. In 2020, Vijitkunsawat and Chantngarm[12] determine the best model to normalize the face masks. Two conventional machine learning classifiers and one deep learning algorithm were discussed. Kumar et al.[13] and Mundial et al.[14] discuss how the tendency to fail in the detection of the methods used by the facial recognition systems has led to many challenges for the verification of true identities in many applications. However, Qin and Li[15] devised a method to automatically tell the placement of the mask on the face. Zhang et al.[16] prove that traditional techniques such as Support Vector Machines (SVM) take longer and consume more space than state-of-the-art models do. Chowdary et al.[17] discuss the robustness and effectiveness of the DeepMaskNet model in facial recognition.

3. Proposed methodology

The proposed methodology is divided in to two phases. The System flow chart is presented in **Figure 1**.

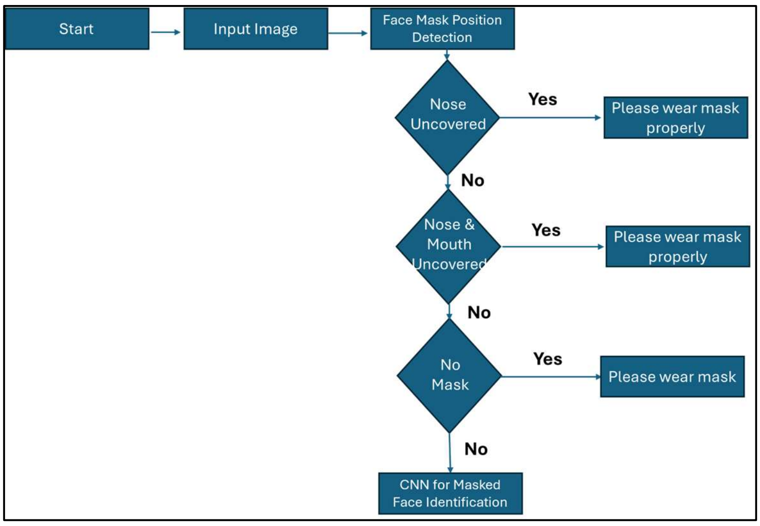

Figure 1. Process flow.

In the first phase, face mask position is detected with the help of the Viola-Jones algorithm[18,19], and an image is classified under i) Nose Uncovered (NU), ii) Nose & Mouth Uncovered (NMU), or iii) No Mask (NM). The classification is done sequentially, and a message is prompted in case the face mask is not properly worn. In case the mask is properly worn, the model enters the second stage, in which identification of the person is done. For this purpose, a modified DeepMaskNet CNN is deployed that helps in the identification of images of 50 persons.

3.1. Phase 1—Face mask position detection with Viola-Jones algorithm

Ensuring masks are properly worn has been a challenge. The proposed model categorizes images into the following:

1) Nose Uncovered (NU)
2) Nose and Mouth Uncovered (NMU)
3) No mask (NM)

The Viola Jones algorithm uses Haar features to detect the presence of a feature in that image. The Haar feature is calculated by running rectangular windows (with black and white rectangles) on an image. Haar features are shown in **Figure 2**. The black rectangle has a value of −1, and the white rectangle has a value of +1. The result is calculated by adding the pixels in both black and white rectangles and then subtracting the white rectangle from the black rectangle. This computation can be very slow and has around 160,000 features (with a 24 × 24 window) and can be reduced with the help of an integral image.

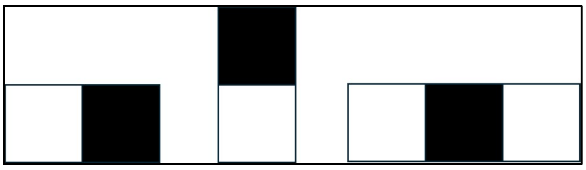

Figure 2. Different types of Haar features in Viola-Jones algorithm.

The integral image helps in a faster calculation of applying the Haar features. The number of pixels (on the left & above to the point (x, y)) in the integral image at one location is calculated as follows:

$$int\_img(x,y) = \sum_{x' \leq x, y' \leq y} img(x', y') \quad (1)$$

where $int\ img(x, y)$ is the integral image and $img(x, y)$ is the original image.

$$r(x,y) = r(x, y-1) + img(x,y) \quad (2)$$

$$int_{img(x,y)} = int_{img(x-1,y)} + r(x,y) \quad (3)$$

where $r(x, y)$ is the number of total rows.

In **Figures 3** and **4**, the calculation of the integral image is demonstrated pictographically.

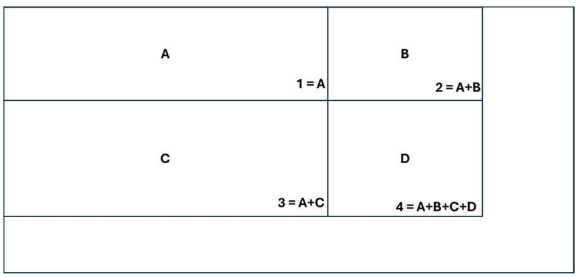

Figure 3. Integral image computation.

1	1	1		1	2	3
1	1	1		2	4	6
1	1	1		3	6	9

Figure 4. Input image (left) to integral image (right).

The cascade classifier algorithm was employed by Viola and Jones. Stratified classification is a feature of this cascade classification technique.

This algorithm is divided into multiple stages. To attain a low false positive rate, the first stage calls for the deployment of more sophisticated classifiers after a basic classifier has rejected the majority of the sub-windows. Because of this, the classifier set is a potent suite capable of classifying faces. As illustrated in **Figure 5**, the general form of the detection process is a degraded decision tree, sometimes known as a "cascade". When the first classifier yields a positive result, the second classifier—which has been proven to have an extremely high detection rate—is evaluated. The third classifier is triggered by a positive result from the second classifier, and so on. If the outcome is negative at any stage, the sub-window is rejected outright.

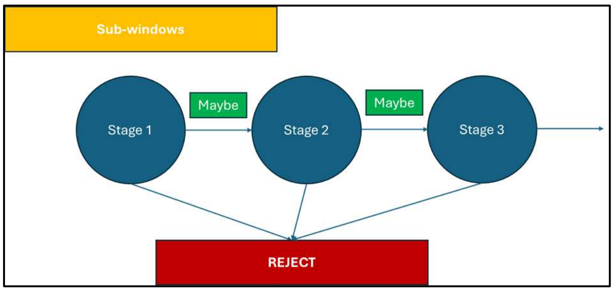

Figure 5. Cascading classifier.

Using AdaBoost machine learning to train classifiers and modifying the threshold value to reduce false negatives, the cascade stage is constructed. Keep in mind that the goal of this AdaBoost threshold is to minimize error rates in training data. **Figure 5** depicts the explanation of the cascade categorization. Adaboost constructs a strong classifier with a linear combination of weak classifiers as shown in Equation (4).

$$F(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) + \cdots \alpha_n f_n(x) \tag{4}$$

3.2. Phase 2—Recognition with modified DeepMaskNet

Once the image passes the face mask position check, it is evaluated for face mask person identification. A modified DeepMaskNet CNN model is deployed. Modifications to the size and stride were made to enhance the performance of the model. The model has ten layers and can process input images with a resolution of 256 × 256. The distribution of the input to each layer changes during the training. For instance, more surgical masks are present in the dataset as compared to other types of masks, which causes parameter training to become extremely time-consuming and requires better initialization. In the DeepMaskNet design, we applied batch normalization to address this covariate shifting and used the LeakyReLU activation function to overcome the dying ReLU problem. The details of the modified DeepMaskNet model are presented in **Table 1** and hyperparameter details in **Table 2**.

Table 1. Layers and architecture of modified DeepMaskNet model.

Sr. No.	Layer	Filters	Size	Stride
1	Input			
2	Convolutional-1	128	3 × 3	1 × 1
3	Max Pooling		3 × 3	2 × 2
4	Convolutional-2	512	3 × 3	1 × 1
5	Max Pooling		3 × 3	2 × 2
6	Convolutional-3	384	3 × 3	1 × 1
7	Max Pooling		3 × 3	2 × 2
8	Convolutional-4	256	3 × 3	1 × 1
9	Convolutional-5	256	3 × 3	1 × 1
10	Convolutional-6	256	3 × 3	1 × 1
11	Max Pooling		3 × 3	2 × 2
12	Fully Connected + LeakyRelu + Dropout			
13	Fully Connected + LeakyRelu + Dropout			
14	Fully Connected + LeakyRelu + Dropout			
15	Fully Connected			
16	Softmax			
17	Classification			

Table 2. Hyperparameters of DeepMaskNet Architecture.

Parameter	Values
Learning Rate	0.01
epoch	50
Optimizer	Adam

The output is passed to a 50-way softmax function for 50 persons.

4. Dataset

We created an internal dataset because there wasn't a uniform dataset for mask position detection and facial recognition covered by a mask. Different dataset requirements apply to each task. Face mask position requires photographs of several people wearing masks both correctly and improperly, whereas face recognition covered by a mask is trained on multiple images of the same person wearing a mask.

We have divided the data in the ratio of 80:20 for training and test data.

For face-mask position detection, since the Viola-Jones algorithm employed in this study has trained over 10,000 photos of human parts, including the face, nose, mouth, and eye, in a detection phase, training images was not needed[20].

For masked face recognition, we have developed a database of 10,397 images of 50 people using image augmentation[21]. **Table 3** details the image database.

Table 3. Details of the in-house database.

Sr. No	Mask Correctly Worn	Nose Uncovered	Nose & Mouth Uncovered	No Mask
Training	5822	832	832	832
Testing	1455	207	208	209

5. Experiment setup

As As part of the experimental setup, the Viola-Jones algorithm and DeepMaskNet are primarily used. Apart from DeepMaskNet, Python libraries such as Imutils, Numpy, dlib, random, OpenCV2, etc. are used for face detection, preprocessing, and face cropping.

A NVIDIA GTX 1070 GPU was used to develop the model.

6. Results

The combination of DeepMaskNet and skin detection methods helps us to reach a higher accuracy. Moreover, a skin detection method is used instead of a traditional face detection method to reduce the amount of computational processing required.

Table 4 and **Figure 7** depict the result on test data on our proposed model, DeepMaskNet only, ResNet 18, DenseNet, and DarkNet 53.

Table 4. Comparison of proposed model with standard models.

Model	Accuracy	Precision	Recall	F1 Score (%)
Proposed Model – Viola-Jones + modified DeepMaskNet	99.96	100	100	100
DeepMaskNet	98.08	98.51	99.25	98.87
ResNet 18	99.12	98.33	100	99.15
DenseNet	99.12	99.83	98.52	99.17
DarkNet 53	99.51	98.48	97.02	97.74

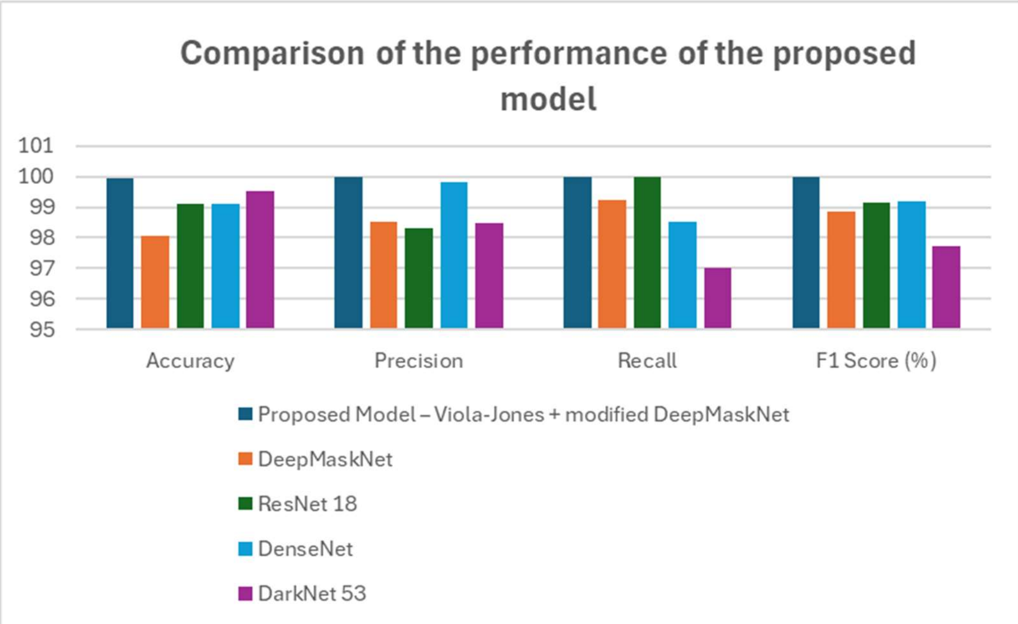

Figure 7. Confusion matrix for face mask position.

7. Conclusion

The proposed model achieves an accuracy of 94% in detecting the face mask position and 99.96% in identifying the masked person. Therefore, we conclude that with the help of the above model, we can improve the accuracy of the existing pre-trained models, which will help all the face recognition systems deployed. Automated face recognition systems have multiple use cases, such as attendance systems, authentication systems, accessibility systems, etc. Moreover, the proposed study is completed with the incorporation of face mask position detection that adds to the list of use cases—controlling entry to public places, narrowing the source of the spread of viruses, and identifying defaulters at schools, universities, offices, and other areas. The proposed model faces challenges in detecting different types of masks. This can be improved by adding more images of people wearing different types of masks in the training dataset. Secondly, computationally heavy but more accurate algorithms such as You-Only-Look-Once (YOLO)[22] can be deployed for face mask position detection.

Author contributions

Conceptualization of the whole research, OPG, APA and OP; methodology, software, OPG; validation part, APA; dataset, OPG, APA and OP; review and editing of the paper, OP. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

1. Jayaweera M, Perera H, Gunawardana B, et al. Transmission of COVID-19 virus by droplets and aerosols: A critical review on the unresolved dichotomy. Environmental Research. 2020; 188: 109819. doi: 10.1016/j.envres.2020.109819
2. Eikenberry SE, Mancuso M, Iboi E, et al. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. Infectious Disease Modelling. 2020; 5: 293-308. doi: 10.1016/j.idm.2020.04.001
3. Ejaz MdS, Islam MdR. Masked Face Recognition Using Convolutional Neural Network. In: Proceedings of the 2019 International Conference on Sustainable Technologies for Industry 40 (STI); 24-25 December 2019; Dhaka, Bangladesh. doi: 10.1109/sti47673.2019.9068044
4. Du H, Shi H, Liu Y, et al. Towards NIR-VIS Masked Face Recognition. IEEE Signal Processing Letters. 2021; 28: 768-772. doi: 10.1109/lsp.2021.3071663
5. Ding F, Peng P, Huang Y, et al. Masked Face Recognition with Latent Part Detection. In: Proceedings of the 28th ACM International Conference on Multimedia; 12-16 October 2020; Seattle, WA, USA. doi: 10.1145/3394171.3413731
6. Damer N, Boutros F, Süßmilch M, et al. Extended evaluation of the effect of real and simulated masks on face recognition performance. IET Biometrics. 2021; 10(5): 548-561. doi: 10.1049/bme2.12044
7. Deng J, Guo J, An X, et al. Masked Face Recognition Challenge: The InsightFace Track Report. In: Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW); 11-17 October 2021; Montreal, BC, Canada. pp. 1437-1444. doi: 10.1109/iccvw54120.2021.00165
8. Rahman MH, Jannat MKA, Islam MS, et al. Real-time face mask position recognition system based on MobileNet model. Smart Health. 2023; 28: 100382. doi: 10.1016/j.smhl.2023.100382
9. Venkateswarlu IB, Kakarla J, Prakash S. Face mask detection using MobileNet and Global Pooling Block. 2020 IEEE 4th Conference on Information & Communication Technology (CICT); 3-5 December 2020; Chennai, India. pp. 1-5. doi: 10.1109/cict51604.2020.9312083
10. Ullah N, Javed A, Ali Ghazanfar M, et al. A novel DeepMaskNet model for face mask detection and masked facial recognition. Journal of King Saud University - Computer and Information Sciences. 2022; 34(10): 9905-9914. doi: 10.1016/j.jksuci.2021.12.017
11. Fan X, Jiang M. RetinaFaceMask: A Single Stage Face Mask Detector for Assisting Control of the COVID-19 Pandemic. In: Proceedings of the 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC); 17-20 October 2021; Melbourne, Australia. pp. 832-837. doi: 10.1109/smc52423.2021.9659271
12. Vijitkunsawat W, Chantngarm P. Study of the Performance of Machine Learning Algorithms for Face Mask Detection. In: Proceedings of the 2020-5th International Conference on Information Technology (InCIT); 21-22 October 2020; Chonburi, Thailand. pp. 39-43. doi: 10.1109/incit50588.2020.9310963
13. Kumar RS, Rajendran A, Amrutha V, et al. Deep Learning Model for Face Mask Based Attendance System in the Era of the Covid-19 Pandemic. In: Proceedings of the 2021 7th International Conference on Advanced

Computing and Communication Systems (ICACCS); 19-20 March 2021; Coimbatore, India. pp. 1741-1746. doi: 10.1109/icaccs51430.2021.9441735
14. Mundial IQ, Ul Hassan MS, Tiwana MI, et al. Towards Facial Recognition Problem in COVID-19 Pandemic. In: Proceedings of the 2020 4rd International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM); 3-4 September 2020; Medan, Indonesia. pp. 210-214. doi: 10.1109/elticom50775.2020.9230504
15. Qin B, Li D. Identifying Facemask-Wearing Condition Using Image Super-Resolution with Classification Network to Prevent COVID-19. Sensors. 2020; 20(18): 5236. doi: 10.3390/s20185236
16. Zhang JP, Li ZW, Yang J. A parallel SVM training algorithm on large-scale classification problems. In: Proceedings of the 2005 International Conference on Machine Learning and Cybernetics; 18-21 August 2005, Guangzhou, China. Volume 3. pp. 1637-1641. doi: 10.1109/icmlc.2005.1527207
17. Chowdary GJ, Punn NS, Sonbhadra SK, Agarwal S. Face Mask Detection using Transfer Learning of InceptionV3. Available online: https://www.arxiv.org/pdf/2009.08369 (accessed on 10 May 2023).
18. Viola P, Jones M. 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; 8-14 December 2001; Kauai, HI, USA. doi: 10.1109/cvpr.2001.990517
19. Viola P, Jones MJ. Robust Real-Time Face Detection. International Journal of Computer Vision. 2004; 57(2): 137-154. doi: 10.1023/B:VISI.0000013087.49260.fb
20. Vikram K, Padmavathi S. Facial parts detection using Viola Jones algorithm. In: Proceedings of the 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS); 6-7 January 2017; Coimbatore, India. pp. 1-4. doi: 10.1109/icaccs.2017.8014636
21. Geng M, Peng P, Huang Y, et al. Masked Face Recognition with Generative Data Augmentation and Domain Constrained Ranking. In: Proceedings of the 28th ACM International Conference on Multimedia; 12-16 October 2020; Seattle, WA, USA. pp. 2246-2254. doi: 10.1145/3394171.3413723
22. Liu R, Ren Z. Application of Yolo on Mask Detection Task. 2021 IEEE 13th International Conference on Computer Research and Development (ICCRD); 5-7 January 2021; Beijing, China. pp. 130-136. doi: 10.1109/iccrd51685.2021.9386366