

# Efficient Two Stage Identification for Face mask detection using Multiclass Deep Learning Approach

**R Dhaya**

Professor, Department of Computer science and Engineering, King Khalid University, Kingdom of Saudi Arabia

E-mail: [dhayavel2005@gmail.com](mailto:dhayavel2005@gmail.com)

## Abstract

The World Health Organization (WHO) considers the COVID-19 Coronavirus to be a global pandemic. The most effective form of protection is to wear a face mask in public places. Moreover, the COVID-19 pandemic prompted all the countries to set up a lockdown to prevent viral transmission. According to a survey study, the use of facemasks at work decreases the chances of fast transmission. If the facemasks are not used or are worn incorrectly, it contributes to the third and fourth waves of the corona virus spreading throughout the world. This motivates us to conduct an efficient investigation of the face mask identification system and monitor people, who use suitable face mask in public places. Deep learning is the most effective approach for detecting whether or not a person is wearing a face mask in a crowded area. Using a multiclass deep learning technique, this research study proposes an efficient two stage identification (ETSI) for face mask detection. Whereas, the binary classification does not offer information about face mask detection and error. The proposed approach employs CNN's "ReLU" activation function to detect the face mask. Furthermore, in the current pandemic crisis, this research article offers a very efficient and precise approach for identifying COVID-19. Precision has increased as a result of the employment of a multi-class abbreviation in the final output.

**Keywords:** deep learning, face mask detection

## 1. Introduction

Since December 2019, the COVID-19 pandemic is creating a long-term impact on several countries throughout the world. On March 11, 2020, World Health Organization [WHO] classified it as a life threatening illness, with a pandemic spread to 114 countries by developing a significant population impact. Every medical expert, organization, medical practitioner, and researcher are working hard to develop appropriate vaccinations and medications to combat deadly diseases [1].

The global coronavirus pandemic caused by the COVID 19 virus has led to an increase in the use of face masks. Prior to Covid-19, people used masks to protect their health from air pollution risks. Some of them are self-aware, but they suppress their emotions in public so they don't look sad or depressed [2]. According to the World Health Organization, almost five million individuals were infected with COVID-19 in 188 countries in less than six months. The primary modes of virus transmission include close contact and crowded settings [3]. Figure 1 depicts the unawareness of individuals, who do not wear a mask in a crowded area.



**Figure 1.** Easy Transmission of Corona Virus in Crowded Places

Face mask detection encompasses detection and classification challenges since people must include their faces into digital images and select whether or not to wear a mask. Because facial detection methods were often employed, the first component of this issue, face detection, was extensively studied in computer vision literature [4]. Alternatively, the second pandemic component (whether or not the face is hidden) was recently drawn by the current study. While the component has done important work in the last year [5], it is often restricted to detect whether or not a mask is present in a picture.

It is not extremely important to assess whether the masks are properly positioned on the face and should therefore be used under medical supervision [6]. As a result, the use of existing visual mask detection techniques is restricted, and research into computer vision models is encouraged, with the purpose of not only identifying the presence of face masks in pictures but also determining whether the mask is correctly worn or not [7]. Figure 2 depicts an improperly worn face mask in a crowded location.



**Figure 2.** Incorrect Use of a Face Mask in a Crowded Area

This research work attempts to overcome this deficiency by focusing on the proper positioning of the face masks. Here, a two stage effective identification method is proposed with a large experimental dataset of faces from two publicly available donor datasets, MAFA [8] and Wider Face [9], for the utilization in research and development. Photographs from common and challenging donor datasets will be chosen and divided into groups of photos with correctly and badly placed facial masks. In addition, proper or improper and posture descriptors are provided. The built-in data set is initially used to conduct two stages of calculation (e.g. detection and classification) in the facial mask images for evaluating the performance of different current face detectors using masked face photos [10].

## **2. Organization of the Research**

This research article is arranged further as follows; Section 3 provides current research on facial mask recognition using various conventional methods. Section 4 covers the planned effort to efficiently identify the face mask. Section 5 describes the outcome of the proposed work. In section 6, the conclusion and future work will be addressed.

## **3. Preliminaries**

Zhu et al propose a CMS-RCNN, a multi-scale contextual neural network based on regions with a significant impact on the facial identification model [11]. Opitz et al have created a Grid Loss Layer to minimize CNN layer substitution errors and solve biased obstacles generated by mask detection models [12]. As technology advanced, more CNN-based 3D models started to appear. It was an apprenticeship to identify face masks. More research has been conducted on the identification of postures, gender evaluation, the placement of landmarks, etc. [13, 14].

Face mask detection models have many variations. These may be divided into many categories. The above-mentioned Viola-Jones (Jones, Viola, & Jones, 2001) was used to boost waterfalls with easy hair qualities. Then, a multiview face mask sensor was developed by inspiring the Viola-Jones detector concept. In addition, a face mask detector model was created by using decision-making methods. Face mask detectors were very efficient in recognizing facial masks in this category [15, 16].

However, these datasets have recently started to surface, owing largely to the COVID-19 pandemic. For instance, Wang et al., have presented three datasets to improve the performance of masked facial recognition systems. They include the masked face detection dataset (MFDD) and the real-world masked face recognition data set (RMFRD) (SMFRD). According to the authors, MFDD is a first-published enhanced dataset version with 24,771 masked faces [17, 18].

A picture containing numerous features in the face detection process recognizes a face. According to the author, face detection research involves facial monitoring, expression recognition, and posture assessment [19, 20]. For example, it is difficult to distinguish faces since photographs have varying forms, colors, sizes, and so on. The authors have identified two main occlusive facial recognition challenges in [21], they are:

- (1) The availability of large, masked and unmasked facial datasets
- (2) The exclusion from the covered area of facial expression.

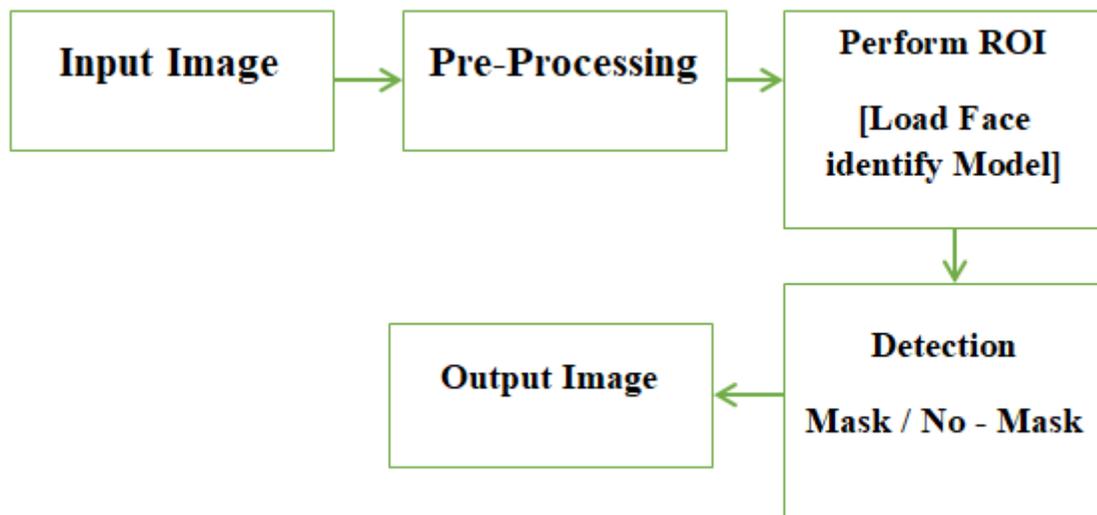
According to the S.Ge working report, the input image size of the computer-view convolution neural network (CNN) is severely restricted. The typical method reshapes the images to overcome the inhibition before they are incorporated into the network model [22]. The author S.Ghosh et al proposed an excellent technique in the article of identifying many faces with masks using a technology called deep learning to automatically identify masked faces [23]. Fasten RCNN

and YOLOv3 models were used in the experiment. Experimental results show that F-RCNN is more precise, but it is not used in security cameras in the current world. YOLOv3 is better if the tradeoff between speed and precision is the measure. The author suggests that R-CNN should be utilized faster if high-end GPUs are available on deployed devices [24].

#### 4. Methodologies

This project must be implemented by combining deep learning, machine learning, computer vision, and Python libraries. The proposal includes a mobile network backbone that may be used in both high and low-density settings. Moreover, the proposed method employs multi-class CNN algorithms [25].

##### Stage 1 - Facemask Detection



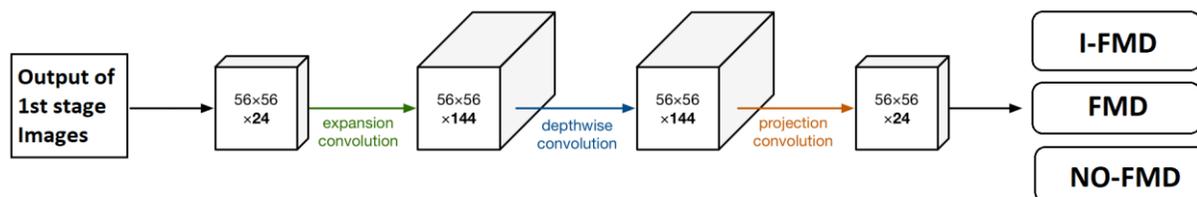
**Figure 3.** Stage 1 face mask identification

#### 4.1 Stage 2 - Effective Facemask detection

The classification of input image is considered as the output of stage one and it has taken into second stage for performing effective facemask detection. To determine whether or not facial masks are being worn correctly, all the faces will be identified in the initial step of our pipeline from where it should be classified into one of the two categories:

- (i) Compliant (complete wrong)
- (ii) Non-Compliant (partial error)

For the first class, we evaluate the mask's position as well as the nose, mouth, and chin covers. In the second session, we assess faces that do not have masks or have masks that have been improperly adjusted. Using this setup, the facial regions will be selected from the images we supply and later it will be sent to an identification procedure based on a standard size (provided by the proposed model which contains efficient two stage identification (ETSI)). Figure 4 shows the proposed MobileNet V2 architecture for stage 2 construction.



**Figure 4.** Stage 2 Face Mask Identification

#### 4.3 Two Stage Authentication

The complete experimental procedure utilizes the random pictures from the training sets and categorizes all the observed aspects as compliant or non-compliant with the COVID-19 mask placement criteria over the course of the experiment. The best-performing models from the first

and second phases are used to create an activity recognition system or to recognize the compliant face mask installations [27]. The system or installation is then tested against publically accessible facial mask detection models. For the sake of detection and classification, compatible and non-compliant faces are considered as two distinct classes (or objects) of faces.

## 5. Results Discussion

This research article comprises the ETSI face mask detection system, which is used to identify the mask pictures with and without mask data sets. More than 800 pictures from Kaggle were gathered and 1300 images from the FMLD benchmark data set were obtained for face mask recognition. The proposed algorithm has performed in two stage detection, initially, face has detected with or without mask for COVID-19 safety. After this procedure, extends our algorithm to detect the perfect or improper wearing mask or not with effective manner through multi class learning procedure [27]. Figure 5 shows some sample input test images.



**Figure 5.** Prediction on test image

The proposed algorithm has superior than binary class deep learning approach. Besides, it can detect perfect face mask detection (P-FMD) and improper FMD (I-FMD) in any crowd people for many reason now a day. Figure 6 shows the obtained results by the proposed two stage framework.



**Figure 6.** Results Obtained by our Proposed Algorithm

The dataset contains the selection of appropriate annotation images that are reasonably balanceable between different categories (e.g. unmasked vs. male vs. female) and represent real-world imaging conditions, where variability in appearance is expected over light, posture, image quality and other similar factors. Figure 7 shows the difference of perfect FMD (P-FMD) and improper FMD (I-FMD).

We effectively identify the MAFA and Wider Face datasets as appropriate donors and select images for labelling from these datasets. Table 1 contains the calculation of performance measures of various methods for detecting face mask.



**Figure 7.** Obtained Results from Various Complaints

**Table 1.** Overall performance measures of various methods

| S.No | Methods                          | Recall | Precision | Accuracy | Efficiency | Computation time |
|------|----------------------------------|--------|-----------|----------|------------|------------------|
| 1    | Image processing Method          | 80%    | 79%       | 81.5%    | 74.21%     | Low              |
| 3    | Single stage verification by CNN | 90%    | 88%       | 89%      | 89.23%     | Low              |
| 4    | Proposed ETSI                    | 97.24% | 92.5%     | 95.35%   | 99%        | Moderate         |

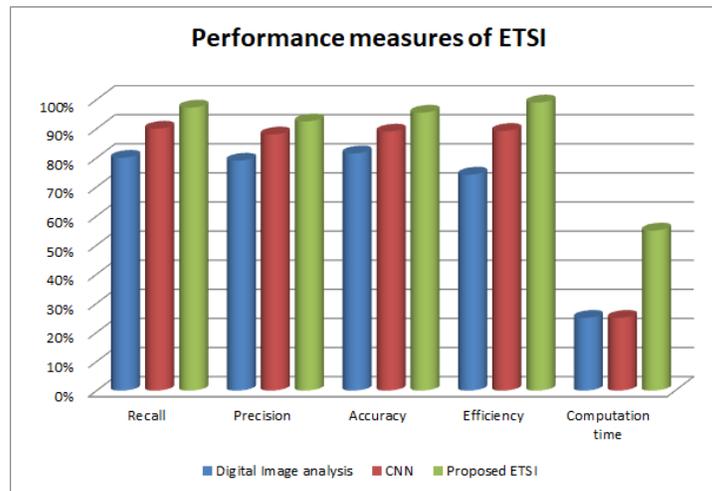
## 5.1 Performance Measures

The accuracy of classification is used to reflect the performance of the tested classification models. The ratio between the number of accurate predictions and the total number of input samples is determined. Precision may be calculated as,

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

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$



**Figure 8.** Overall Performance of Proposed Algorithm

Figure 8 shows the overall performance measures of proposed algorithm. The computation time is high when compared to other traditional methods.

## 6. Conclusion

Thus, the proposed ETSI facemask identification has been proved as a very efficient and good accuracy to other algorithm. Face-mask detection is critical for COVID-19 pandemic monitoring in many applications such as surgery room and so on, was the subject of this study, which looked at the problem further. Finally, this research work concludes that the proposed algorithm can use in any mass gathering public parks, examination hall and very sensitive area

such as hospitals. This research work has developed an annotated benchmark dataset for the investigation of face-mask detection issues, and an experimental research has been conducted to investigate:

- (i) The performance of current facial detectors using masked facial pictures
- (ii) The utility of current face-mask detection algorithms for COVID-19 surveillance applications
- (iii) The effectiveness of current face-mask detection algorithms.

Finally, the suggested method has been demonstrated to be successful, whereas existing face recognition algorithms are only moderately useful for real-world applications since they identify just facial masks in pictures but do not discriminate between where they are positioned. To the best of our knowledge, no literature study was conducted on this topic. The proposed next step is to train multiple classification models (hybrid) on our test data set to predict and evaluate the correct placement of facial masks. Future iterations of the proposed mask identification system will consider the candidate's age, gender, and geographic location.

## References

- [1] Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.
- [2] Tripathi, Milan. "Analysis of Convolutional Neural Network based Image Classification Techniques." Journal of Innovative Image Processing (JIIP) 3, no. 02 (2021): 100-117.
- [3] Grm, K.; Štruc, V. Deep face recognition for surveillance applications. IEEE Intell. Syst. 2018, 33, 46–50.

- [4] Prasad, Deepti, Rasika Gaiwal, Shivani Bharambe, Supriya More, and Ashwini Jarali. "Face-Based Alumni Tracking on Social Media Using Deep Learning." In *Innovative Data Communication Technologies and Application*, pp. 857-869. Springer, Singapore, 2021.
- [5] Adam, Edriss Eisa Babikir. "Evaluation of Fingerprint Liveness Detection by Machine Learning Approach-A Systematic View." *Journal of ISMAC* 3, no. 01 (2021): 16-30.
- [6] Shubha, P., and M. Meenakshi. "Human Face Recognition Using Local Binary Pattern Algorithm-Real Time Validation." In *International Conference On Computational Vision and Bio Inspired Computing*, pp. 240-246. Springer, Cham, 2019.
- [7] Kumar, T. Senthil. "Construction of Hybrid Deep Learning Model for Predicting Children Behavior based on their Emotional Reaction." *Journal of Information Technology* 3, no. 01 (2021): 29-43.
- [8] Balasubramaniam, V. (2021). Facemask Detection Algorithm on COVID Community Spread Control using EfficientNet Algorithm. *Journal of Soft Computing Paradigm*, 3(2), 110-122. doi:10.36548/jscp.2021.2.005
- [9] Manoharan, J. Samuel. "A Novel User Layer Cloud Security Model based on Chaotic Arnold Transformation using Fingerprint Biometric Traits." *Journal of Innovative Image Processing (JIIP)* 3, no. 01 (2021): 36-51.
- [10] Ranganathan, G. "Real time anomaly detection techniques using pyspark framework." *Journal of Artificial Intelligence* 2, no. 01 (2020): 20-30.
- [11] Zhu, C., Zheng, Y., Luu, K., & Savvides, M. (2017). Cms-rcnn: Contextual multi-scale region-based cnn for unconstrained face detection. *Deep learning for biometrics* (pp. 57–79). Cham: Springer.
- [12] Opitz, M., Waltner, G., Poier, G., Possegger, H., & Bischof, H. (2016). Grid loss: Detecting occluded faces. October. *European conference on computer vision* (pp. 386–402). Cham: Springer.

- [13] Smys, S., Abul Bashar, and Wang Haoxiang. "Taxonomy Classification and Comparison of Routing Protocol Based on Energy Efficient Rate." *Journal of ISMAC* 3, no. 02 (2021): 96-110.
- [14] Li, H., Lin, Z., Shen, X., Brandt, J., & Hua, G. (2015). A convolutional neural network cascade for face detection. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 5325–5334.
- [15] Vijayakumar, T., Mr R. Vinothkanna, and M. Duraipandian. "Fusion based Feature Extraction Analysis of ECG Signal Interpretation–A Systematic Approach." *Journal of Artificial Intelligence* 3, no. 01 (2021): 1-16.
- [16] Jones, P., Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. Charles Rich: University of Rochester.
- [17] Suma, V. "Community Based Network Reconstruction for an Evolutionary Algorithm Framework." *Journal of Artificial Intelligence* 3, no. 01 (2021): 53-61.
- [18] Wang, Z.; Wang, G.; Huang, B.; Xiong, Z.; Hong, Q.; Wu, H.; Yi, P.; Jiang, K.; Wang, N.; Pei, Y.; et al. Masked face recognition dataset and application. *arXiv* 2020, arXiv:2003.09093.
- [19] D. Meena and R. Sharan, "An approach to face detection and recognition," 2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE), Jaipur, 2016, pp. 1-6, doi: 10.1109/ICRAIE.2016.7939462.
- [20] Dhaya, R. "Analysis of Adaptive Image Retrieval by Transition Kalman Filter Approach based on Intensity Parameter." *Journal of Innovative Image Processing (JIIP)* 3, no. 01 (2021): 7-20.
- [21] S. Ge, J. Li, Q. Ye and Z. Luo, "Detecting Masked Faces in the Wild with LLE-CNNs," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 426-434, doi: 10.1109/CVPR.2017.53.

- [22] S. Ghosh, N. Das and M. Nasipuri, "Reshaping inputs for convolutional neural network: Some common and uncommon methods", *Pattern Recognition*, vol. 93, pp. 79-94, 2019. Available: 10.1016/j.patcog.2019.04.009.
- [23] Sunil Singh, Umang Ahuja "Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment" <https://doi.org/10.1007/s11042-021-10711-8>
- [24] Kumar, Jannaikode Yashwanth, and Nenavath Srinivas Naik. "COVID-19 Risk Management: A Mobile Application to Notify Alerts and Provide Real-Time COVID-19 Updates." In *Innovative Data Communication Technologies and Application*, pp. 507-518. Springer, Singapore, 2021.
- [25] Priyanka, K., and T. G. I. Fernando. "Mobile Application to Identify Fish Species Using YOLO and Convolutional Neural Networks." In *Proceedings of International Conference on Sustainable Expert Systems: ICSES 2020*, vol. 176, p. 303. Springer Nature, 2021.
- [26] Athreya, Shashank M., S. P. Shreevari, BS Aradhya Siddesh, Sandeep Kiran, and H. T. Chetana. "A Survey on Partially Occluded Faces." In *Evolutionary Computing and Mobile Sustainable Networks*, pp. 61-67. Springer, Singapore, 2021.
- [27] D'Souza, Rupal Mayo Diline, S. R. Deepthi, and K. Aarya Shri. "Detection Classification and Cutting of Fruits and Vegetables Using Tensorflow Algorithm." In *Evolutionary Computing and Mobile Sustainable Networks*, pp. 713-720. Springer, Singapore, 2021.

### Author's Biography

**R Dhaya** is currently a Professor, in the Department of Computer science and Engineering at King Khalid University, in the Kingdom of Saudi Arabia. His major area of research includes Image and Video Processing Algorithms, Computer Vision, Motion Analysis, Stereo Vision, Object Recognition, computer graphics, photo interpretation, image retrieval, Embedded Image Processing and Real-time image and video processing applications.