International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 05 | May - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Real Time Face Mask Detection and Email Alert

Shekhar Nayak [1], Yogesh Dewangan[2], Shubham Verma[3]
Under graduate[1,2,3,]

Dr.Dolly Shukla
Associate Professor

Dept. Of Information Technology Engineering, Shri Shankaracharya Technical Campus, Bhilai, C.G., India

Abstract— Coronaviruses, a large family of distinct viruses, have recently become exceedingly prevalent, communicable, and harmful to the whole human race. It transmits from person to person by exhaling infected breath, which deposits virus droplets on various surfaces, which are subsequently breathed by another person, who contracts the illness. As a result, it is critical that we safeguard ourselves and others around us from this predicament. We may take safeguards like as keeping social distance, washing hands every two hours, using sanitizer, and, most importantly, wearing a mask. Wearing a mask in public has grown quite widespread in recent years all around the world. Due of its high population in a short space, India is the most afflicted and terrible situation. This study presents a technique for detecting whether or not a face mask is used in offices or any other workplace with a large number of individuals. For this, we employed a convolutional neural network. The model was trained on a real-world dataset and successfully tested using live video streaming. The model's accuracy is further tested using several hyper parameters and many persons at various distances and locations inside the picture .

Keywords— Face Mask Detection, Convolutional NeuralNetwork, MobileNetV2, Corona virus Precaution.

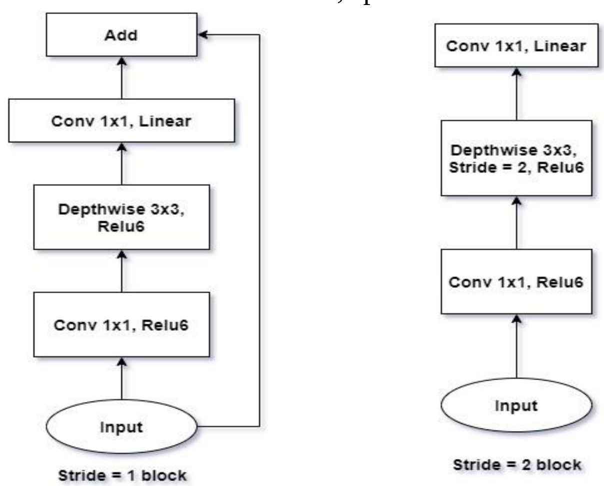
I. INTRODUCTION

Since the introduction of the new coronavirus epidemic, public use of face masks has become prevalent in China and other countries across the world. According to the Health Centre's alert, we now know that a considerable fraction of people with coronavirus have no symptoms ("asymptomatic") and that even those who later acquire symptoms ("pre-symptomatic") can spread the virus to others before they display symptoms..“This means that the virus can spread between people interacting in close proximity — for example, speaking, coughing, or sneezing — even if those people are not exhibiting symptoms”.

The latest information also points to a new corona virus strain, the mutant corona virus, in which the virus's structure has altered and it has become mutant. The novel strain is undetectable by the RT-PCR technique we now employ. As a result, it is unavoidable for the citizens of an overcrowded country like India to put on masks and continue working. Nobody can keep track of whether or not everyone entering the workplace is wearing a mask. As a result, the requirement for Face mask detection arose. The Convolutional Neural Network is used in this model. It's a deep neural network model that can analyses any type of visual imagery. It accepts picture data as input and captures all the data, and send to the layers of neurons. It contains a fully linked layer that processes the final output, which is the picture prediction. The MobileNetV2 architecture is the Convolutional neural network model employed here. The Mobile Net model is a network model in which the fundamental unit is depth wise separable convolution. It has two layers of depth wise separable convolution: depth wise convolution and point convolution [1] . It is based on an inverted residual structure where the residual connectionsare between the bottleneck layers. The intermediate expansion layer uses lightweight depth wise convolutions to filter features as a source of non-linearity. As which is used in the model discussed in this paper. The hyper parameters tried are learning rate, it is a tuning parameter that is used in optimization models which determines the step size of the model and helps to reduces the loss function. It is a very important hyper parameter as itresults in either convergence or overshoots the model. The other hyper

Figure1. MobileNetV2

parameters used are batch size, epochs etc . The model has



used OpenCV to fulfil the purpose of using thevideo stream for capturing the frames in the video stream

II. RELATED WORK

In [3] they have proposed a pre-trained Mobile Net with a global pooling block for face mask detection. The pre- prepared Mobile Net takes a shading picture and creates a multi-dimensional component map. The worldwide pooling block that has been used in the proposed model changes the element map into an element vector of 64 highlights. At long last, the SoftMax layer performs paired order utilizing the 64 highlights. We have assessed our proposed model on two openly accessible datasets. Our proposed model has accomplished 99% and 100% exactness on DS1 and DS2 separately. The worldwide pooling block that has been utilized in the proposed model dodges overfitting the model. Further, the proposed model beats existing models in the quantity of boundaries just as preparing time. But this model cannot detect face mask for multiple faces at a time. In [5] paper utilizes a proficient and strong item location calculation to naturally identify the appearances with veils or without covers, making the plague avoidance work cleverer. In particular, they gathered a broad data set of 9886 pictures of individuals with and without face covers and physically named them, at that point use multi-scale preparing and picture mistake techniques to improve YOLOv3, an article recognition calculation, to consequently distinguish whether a face is wearing a veil. Our analysis results show that the mean Average Precision (mAP) of the improved YOLOv3 calculation model came to 86.3%. This work can viably and naturally distinguish whether individuals are wearing veils, which decreases the pressing factor of conveying HR for checking covers openly puts and has high functional application esteem. In face detection method, a face is detected from an image that has several attributes on it. According to [21], research into face detection requires expression recognition, face tracking, and pose estimation. Given a solitary image, the challenge is to identify the face from the picture. Face detection is a difficult errand because the faces change in size, shape, color, etc. and they are not immutable. It becomes a laborious job for opaque image impeded by some other thing not confronting camera, and so forth. Authors in [22] think occlusive face detection comes with two major challenges: first, unavailability of sizably voluminous datasets containing both masked and unmasked faces, second, exclusion of facial expression in the covered area. Utilizing the locally linear embedding (LLE) algorithm and the dictionaries trained on an immensely colossal pool of masked faces, synthesized mundane faces, several mislaid expressions can be recuperated and the ascendancy of facial cues can be mitigated to great extent. According to the work reported in [11], convolutional neural network (CNNs) in computer vision comes with a strict constraint regarding the size of the input image. The prevalent practice reconfigures the images before fitting them into the network to surmount the inhibition. In [23], a robust and efficient technique for liveness detection was proposed. The authors used the deep learning DeBNet approach for feature extraction and classification. In [24], the authors used SVM for proposing a machine learning based face detection and recognition system. The proposed model was used to detect the faces of students for monitoring their activities during online examinations. The proposed system used feature vectors from the input images for detecting the faces in a faster manner. In [25], a multi-task deep learning method called F-DR Net for recognizing and detecting was used .

III. PROPOSED SYSTEM

TensorFlow, Keras, and OpenCV are among the Python libraries used to construct and model the model presented here. The convolutional neural network model we utilized was MobileNetV2. Transfer Learning is the way of employing MobileNetV2. Transfer learning is the process of utilizing a pre-trained model to train your current model and obtain a prediction, which saves time and simplifies the process of training various models. The hyper parameters: learning rate, number of epochs, and batch size are used to fine-tune the model. The model is trained on a set of photos divided into two categories: with and without mask. There are 993 photos with masks and 1918 images without masks in the collection.

(i) Training the model with the taken dataset.
(ii) Deploying the model

We used the above-mentioned libraries to create a model in the paper. We evaluated the model under various settings and with various hyper parameters, and the findings are shown in the next section. We feed the dataset into the model first, then execute the training algorithm to train the model on the provided data. Then, using the object detection method, we execute the detection software, which turns on the video stream and grabs the frames continually from the video stream with an anchor box. This information is sent through the MobileNetV2 model layers, which determine if the picture has a mask or not. A green anchor box is displayed if the person is wearing a mask, and a red anchor box is presented if the person is not wearing a mask, with the accuracy for the same tagged individual on the anchor

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 05 | May - 2022 Impact Factor: 7.185 ISSN: 2582-3930

box. The flow of the Face Mask Detection model utilized in this work is shown in Figure 2. The face mask identification system use artificial intelligence to determine whether a person is wearing a mask or not. It may be linked to any surveillance system you have placed on your property. Authorities or administrators can use the system to validate the person's identification. If someone enters the premises without wearing a face mask, the system sends an alarm message to the designated individual. Detecting a person wearing a face mask is 95-97 percent accurate, depending on the digital capabilities. The data has been automatically transmitted and kept in the system, allowing you to run reports whenever you wish .

IV. DATASETS

The model was built using two datasets. Dataset 1 [16] has 1376 photos, 690 of which have persons wearing face masks while the remaining 686 images do not. Figure 2 shows a front face posture with a single face and a mask of the same kind and colour (white only).

Fig. 2 Samples from Dataset 1 with faces with and without masks [16].

Fig. 3 Samples from Dataset 2 including faces without masks and with different types and colors of masks [17]

Dataset 2 from Kaggle [17] has 853 pictures, each of which is clarified with or without a mask. Head turn, tilt, and slant with several faces in the frame with varied types and colors of masks are some of the face collections shown in Fig. 3 .

V. RESULTS

We evaluated the model in three distinct settings, and the table below shows the outcomes of those scenarios with a consistent number of epochs (20) and batch size 32 for all three scenarios. For getting a smooth picture, we employed Average Pooling. Table 1 displays the outcomes of a comparison of several hyperparameters and scenarios.

Model	Learning rate	With mask distance	Without mask distance	Blur image quality	Multiple people capturing
1	1e-4	161 cm	190 cm	Good	4 people
2	1e-3	155 cm	187 cm	Average	3 people
3	1e-2	146 cm	179 cm	Bad	3 people

Table 1. Result Comparison Table

The first model, according to the given results, is the best of all the models. Below is a graphic of the best model from our study. It plots the number of epochs vs loss or accuracy for training loss, validation loss, training accuracy, and validation accuracy. The plot shows that the training and validation accuracy increases as the number of epochs grows, whereas the training and validation accuracy falls. Furthermore, the validation accuracy is better than the training accuracy, indicating that the model is not overfitted. Figure 4 depicts a plot of the number of epochs vs. accuracy for dataset 1. Figure 4 depicts a plot of the number of epochs vs. accuracy for the number of epochs corresponding to dataset 2.

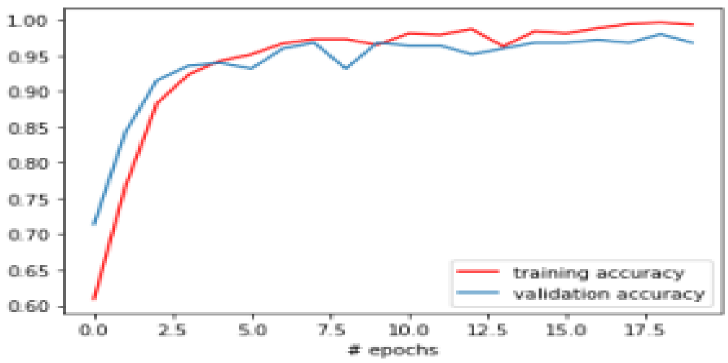


Figure4. epochs vs. accuracy corresponding to dataset 1.

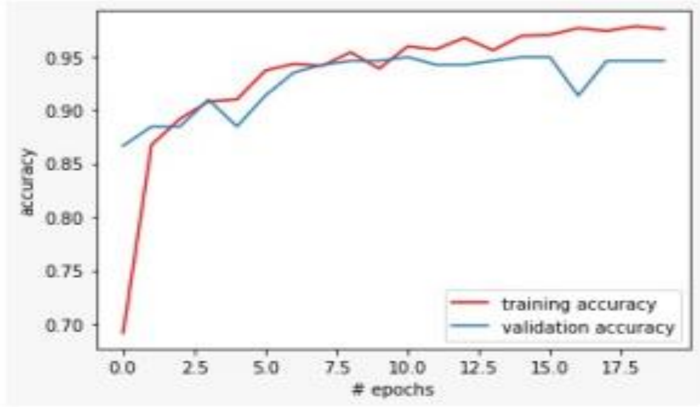


Figure5. epochs vs. accuracy corresponding to dataset 2.

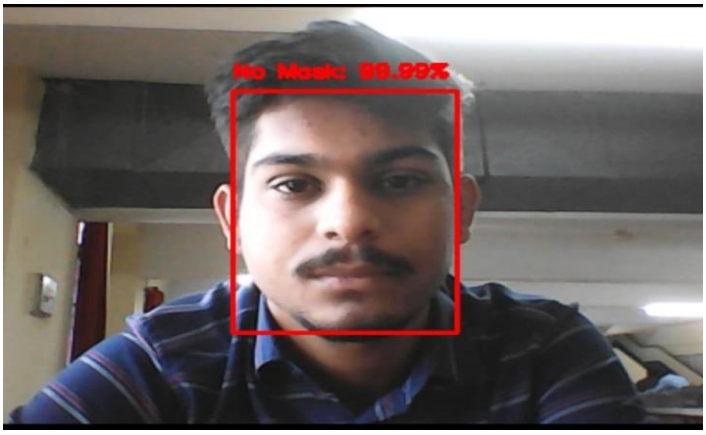


Figure6: Person without mask and its accuracy

Figure7: Person with mask and its accuracy

VI. FUTURE WORK

For single faces with and without masks, the suggested model provides excellent accuracy. It also has a high level of accuracy when dealing with several faces. It works on any mobile device simply by turning on the video stream; no further hardware is required. Further, we will work to improve the accuracy of multiple face mask detection by adding datasets with images of people wearing masks that do not cover their noses properly and also by detecting the masked face using the FaceNet model of convolutional Neural Network as described in [4] so that we can improve our model and add marking attestation in it by detecting the face even when the mask is on .

VII. CONCLUSION

Measures should be done to slow the spread of the COVID-19 pandemic. In neural organizations, we showed a facemask detector utilizing Convolutional Neural Networks and motion learning algorithms. We used a dataset with 993 masked faces photographs and 1918 exposed faces pictures to train, validate, and test the model. These images were compiled from a variety of sources, including Kaggle and RMFD databases. On photographs and live video transmissions, the model was induced. We evaluated metrics such as precision, accuracy, and recall to determine a base model, and the best exhibition was MobileNetV2 architecture, which had 99 percent precision and 99 percent recall. MobileNetV2 is also more computationally efficient, making it easier to integrate the model into other frameworks. This face mask detector may be deployed at a variety of locations, including shopping malls, airports, and other high-traffic areas, to screen individuals in general and prevent the spread of illness by determining who is following basic standards and who is not be dispatched to a variety of locations, including shopping malls, airports, and other high-traffic areas, to screen individuals in general and prevent the spread of the virus by determining who is following basic standards and who is not .

VIII. REFERENCE

[1]. G. Howard, M. Zhu, B. Chen et al., “Mobilenets: efficient convolutional neural networks for mobile vision applications,” 2017, https://arxiv.org/abs/1704.04861.

[2]. Wei Wang, Yutao Li, Ting Zou, Xin Wang, Jieyu You, Yanhong Luo, "A Novel Image Classification Approach via Dense-MobileNet Models", *Mobile Information Systems*, vol. 2020,ArticleID 7602384, 8 pages, 2020. https://doi.org/10.11 55/2020/7602384

[3]. Venkateswarlu, J. Kakarla and S. Prakash, "Face mask detection using MobileNet and Global Pooling Block,"

[4]. Liu, C., and Wechsler, H. (2002). Gabor feature

based classification using the enhanced fisher linear discriminant model for face recognition. IEEE Transactions on Image processing, 11(4), 467-476.

[5]. Kepenekci, B., and Akar, G. B. (2004, April). Face classification with support vector machine. IEEE 12th Signal Processing and Communications Applications Conference, 2004. (pp. 583-586). IEEE.

[6]. H. Anandakumar and K. Umamaheswari, A bio-inspired swarm intelligence technique for social aware cognitive radio handovers, Computers & Electrical Engineering, vol. 71, pp. 925–937, Oct. 2018. doi:10.1016/j.compeleceng.2017.09.016

[7]. R. Arulmurugan and H. Anandakumar, Early Detection of Lung Cancer Using Wavelet Feature Descriptor and Feed Forward Back Propagation Neural Networks Classifier, Lecture Notes in Computational Vision and Biomechanics, pp. 103–110, 2018. doi:10.1007/978-3-319-71767-8_9.

[8]. Savvides, M., Heo, J., Abiantun, R., Xie, C., and Kumar, B. V. (2006, May). Class dependent kernel discrete cosine transform features for enhanced holistic face recognition in FRGC-II. IEEE (Vol. 2, pp. II-II).

[9]. Vu, N. S., and Caplier, A. (2010, September). Face recognition with patterns of oriented edge magnitudes. (pp. 313-326). Springer, Berlin, Heidelberg.

[10]. Abusham, E. E., Jin, A. T., Kiong, W. E., and Debashis, G. (2008). Face recognition based on nonlinear feature approach.

[11]. G. K. Jakir Hussain and S. Natarajan, Subpixel Based Image Scaling Using Continuous Domain Analysis, International Journal of Computational Research and Development, Volume 1, Issue 2, Page Number 92-96, 2016.