



(RESEARCH ARTICLE)



CNFMD: Convolutional Network-based Face Mask Detection

Saksham Checker*

Student, Department of Applied Physics, Delhi Technological University, New Delhi, India.

World Journal of Advanced Research and Reviews, 2022, 13(02), 232–238

Publication history: Received on 07 January 2022; revised on 09 February 2022; accepted on 11 February 2022

Article DOI: <https://doi.org/10.30574/wjarr.2022.13.2.0142>

Abstract

After the rise of the coronavirus, every country has made a compulsion on wearing face masks in public places. Even though a few countries have reached a good number of vaccinations to date, the human body is still not immune to the new variants of the virus. It will take a few more years till everyone becomes mask-free. Thus, authorities need a system to keep a proper check on the discipline whether everyone is wearing a mask in a public place or not. This paper proposes a model with 99.5% accuracy which can be deployed and thus can monitor public places. This will help the authorities to control the spread of the virus. The model proposed is a fast model which took only 38.03 seconds on average per epoch while training. Using the Kaggle dataset, the CNFMD model is trained and tested. The dataset, pre-split into three parts is used from training, validation as well as testing of the model.

Keywords: Machine Learning; Deep Learning; COVID-19; Face Mask

1. Introduction

Face masks have become a need for human life after the breakout of SARS-Cov2 also known as the Coronavirus. Virus and Bacteria transmission can be prevented with the use of face masks along with social distancing, and proper sanitation [1]. Many studies thus suggest that during transportation, one must wear a mask to prevent the transmission of disease from themselves to others and vice-versa [1]. Studies suggest that the COVID-19 can spread with immediate contact between two people (nearly 6 feet). It can also spread to people who do not have symptoms [2]. Further, many reasons show that the virus could be air-borne, making Face Masks a must for everyone while traveling or working [3].

Face mask detection has become the latest field of study such that it can be assured that people are wearing face masks in public areas. This paper proposes a new CNFMD model to detect whether a person is wearing a face mask or not. For training the model, a dataset having around twelve thousand images is used divided into three parts—train, validation, and test [4].

In this paper, section 2 summarizes the previous work done by other authors in the field of face mask detection. Section 3 and section 4 give the proposed methodology and experimental results of the model which are all concluded in section 5.

2. Related Works

Since the outbreak of the virus, many pieces of research have been carried out to detect face masks and have also been applied to assure that the spread of the virus is limited reducing the number of cases as well as deaths among the people.

* Corresponding author: Saksham Checker

Student, Department of Applied Physics, Delhi Technological University, New Delhi, India.

The letter [1], shows a case study when a patient can infect 39 passengers during bus travel of 2 hours and 10 minutes whereas in a mini-bus, the patient can affect 14 other passengers for 50 minutes because the patient was not wearing a mask. The authors also refer to the study where 65.7% of healthcare workers agree that face masks can prevent the transmission of the virus.

Further, the study [2] proposed a residual network(ResNet) for extracting features, followed by classic machine learning techniques like SVM, Decision trees, and ensemble to detect the face masks. The researchers removed the last layer of ResNet and replaced it with three Machine learning classification layers. They used three datasets, RMFD, SMFD, and LFW where the dataset was split into train, validation, and testing as 70%, 10%, and 20% respectively. They also explained the Machine learning Layers which were applied in detail. Their proposed hybrid work achieved 99.64% accuracy using SVM on the RMFD dataset, whereas 99.49% and 100% accuracy on SMFD and LFW datasets.

A real-time study, SSDMNv2 [5] proposed a face mask detection model where the DNN module of OpenCV was used with ResNet-10 as the backbone. Further, the pre-trained model of MobileNetV2 was used to predict. They used an artificially created dataset by Prajna Bhandary. They explain the building of the MobileNetV2 model layer by layer. Their proposed model had a better F-1 score compared to the ResNet model. While detection of Faces using OpenCV DNN, it was established on the Single Shot Multi-box detector which has two versions- CaffeImplementation and Original TensorFlow Implementation.

The authors of [6] proposed an object detection method, which combined the one-stage and two-stage detectors. They thus worked on cropping facial areas which can be used in their proposed model. The ResNet-50 pre-trained model was fine-tuned by the researchers to check the performance. They used the images available at Kaggle to train and test their model. The train, test, and validation set was split up into a 64:40:16 ratio. They compared their model to AlexNet and MobileNet pre-trained models.

The latest study, [7], proposed a two-stage model. The first stage had the Face detector system where RGB image was taken as an input and a few pre-trained models were tested to detect the faces of people within the images. These models were trained on big datasets. Introducing stage two, which had the Face mask classifier. The detected faces from stage one were used in three known classification models- MobileNetV2, DenseNet121, and NASNet. After training, the best model was turned out to be the NASNetMobile model with 99.45% test accuracy.

Using the TensorFlow and Keras algorithms applied on datasets collected from Kaggle, authors of [8] proposed a system for face mask detection. They used 3918 images which were divided into two parts, with masks, and without masks. For classification, they implemented MobileNet and OpenCV. Their dataset had 3918 images which trained and tested the model.

3. Proposed Methodology

To predict if a person is wearing a mask or not, a proper dataset of faces with and without a mask was used. The dataset was thus used to train the proposed CNFMD model. This section explains the contents of the dataset, the building blocks of the CNFMD model, and also the algorithm of the model proposed.

3.1. Dataset used

As there has been only a year and a half since the outbreak of the virus and the importance of masks to stay protected from the deadly virus, there are not many datasets available compared to other fields of studies. The dataset published on Kaggle [4] is used to train and test our model.

The dataset consists of a total of 11,792 images as shown in Figure 1. The images are pre-divided in the dataset with 10000 images equally divided into the two-class labels – WithMask and WithoutMask, in the train set, 992 images in the test set with 483 as WithMask and 509 as WithoutMask, and the last set, Validation, having 800 images equally divided in the two labels. The images in all the sets are real-life images such that the model thus formed can be used in the future.

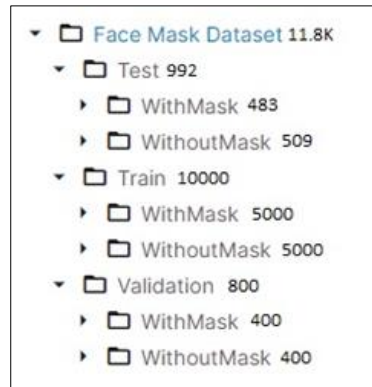


Figure 1 Data Directory

Figure 2 shows the WithMask images from the Train dataset, whereas Figure 3 shows WithoutMask images from the same dataset. It can be observed that there are images from different shapes

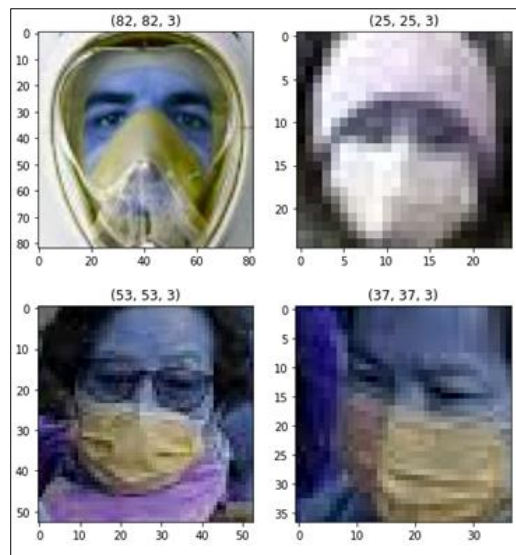


Figure 2 With Mask Images

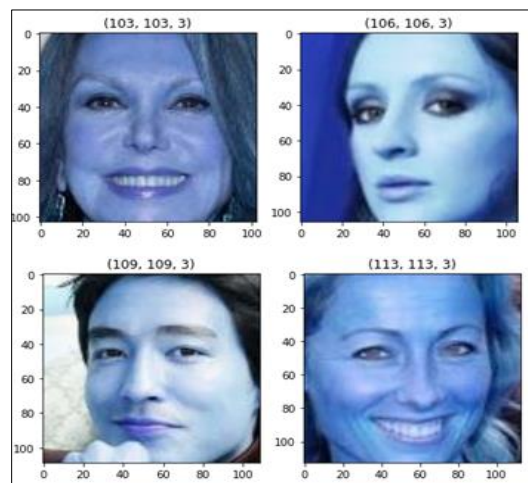


Figure 3 Without Mask Images

3.2. Building Blocks of CNFMD

3.2.1. Convolutional Layer

This is the fundamental block of CNN networks. Convolution is a combination of two functions to get a new function. Features are generated from an image using this layer. Figure 3 shows how the convolution process is done. The generated feature maps are thus processed further. Equation 1 shows the function with A being input matrix, and B as kernel to give output C [5][9][10].

$$C(T) = (A \times B)(x) = \int_{-\infty}^{\infty} A(T) \times B(T - x) \times dT \tag{1}$$

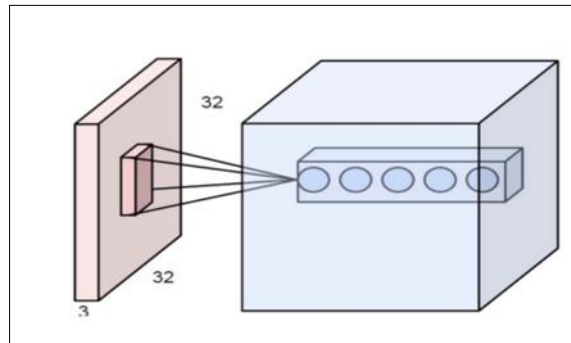


Figure 4 Convolutional Process

3.2.2. Non-Linear Layer

These layers follow the convolutional layers and are known as activation functions. In the CNFMD, ReLU activation is applied and is represented in Equation 2 [5][10].

$$ReLU: f(x) = \max(0, x) \tag{2}$$

3.2.3. Pooling Layer

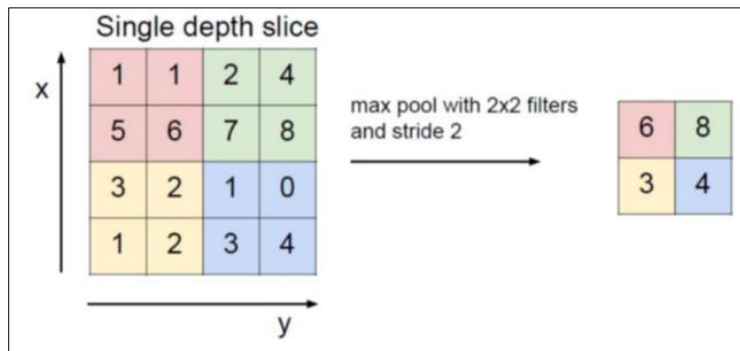


Figure 5 Max Pooling Process

Pooling works on the concept of down-sampling. It reduces the complexity of further layers. It can be compared to the process of reducing the resolution in images. Several filters are not affected by the pooling methods. MaxPooling is the most common type and it is the one that is applied in the proposed CNFMD model. Figure 5 shows the process of Maxpooling with 2x2 filters and stride as 2 [10] [5].

3.2.4. Data Augmentation

In data augmentations, images are applied with a few techniques to generate a larger dataset from the given one and thus for getting a better model for the problem [11]. In this study, augmentation techniques used are - rescale, rotation_range, width_shift_range, height_shift_range, shear_range, zoom_range and horizontal_flip.

3.2.5. Model Visualization

Using the layers mentioned above, the CNFMD model is formulated as shown in figure 6 below. The model takes an input of shape 82,82,3 and at the output, classifies the image as WithMask or WithoutMask image.

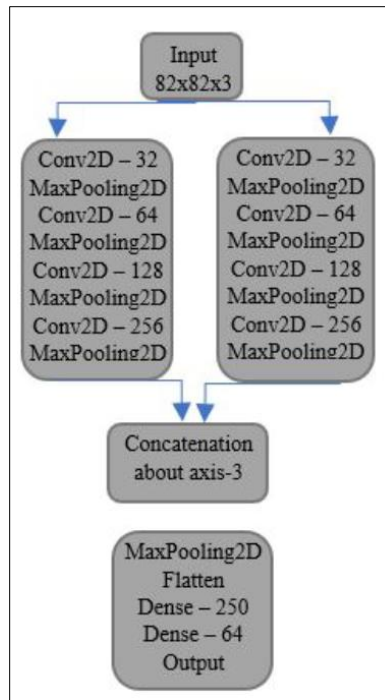


Figure 6 Model Design

With the generated model and defined layers, a final algorithm is thus proposed mentioned as algorithm 1.

ALGORITHM: Pre-processing and Training on Dataset

INPUT: Images along with their pixel values.

OUTPUT: Trained Model

- **STEP 1:** Loading images from the train folder.
- **STEP 2:** Load labels from the folder.
- **STEP 3:** Process the images, i.e., resizing, normalization of images.
- **STEP 4:** Perform Data augmentation and split the data into training and validation batches.
- **STEP 5:** Train the formulated CNFMD model on training batches.
- **STEP 6:** Save the model for the future.

4. Experimental Results

The generated model was trained on Kaggle Notebook Kernel having 13 GB of RAM and 16 GB of GPU memory. The images were loaded using the ImageDatagen where data augmentation was also done, for train dataset as follows- rescale=1. / 255, rotation_range = 10, width_shift_range = 0.2, height_shift_range = 0.2, shear_range = 0.2, zoom_range = 0.2. Whereas, test and validation set was just rescaled with 1./255.

The model was thus compiled using Adam as an optimizer. Metrics used to check the performance were Accuracy (Equation 3) and Binary cross-entropy /Log loss (Equation 4).

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (3)$$

$$\log loss = -\frac{1}{N} \sum_{i=1}^N -(\log(p_i)) \tag{4}$$

Here, TP is True Positive, FP is False Positive, TN is True Negative and FN is False Negative from equation 3 and N is the number of data entries, and p_i is corrected probability.

The model was trained with 50 epochs. It took a total time of 1904 seconds with an average of 38.08 seconds per epoch. As seen in figures 7 and 8 model accuracy for both Train and Validation increased and Loss was decreased. Table 1 shows the experimental results of the model after 50 epochs

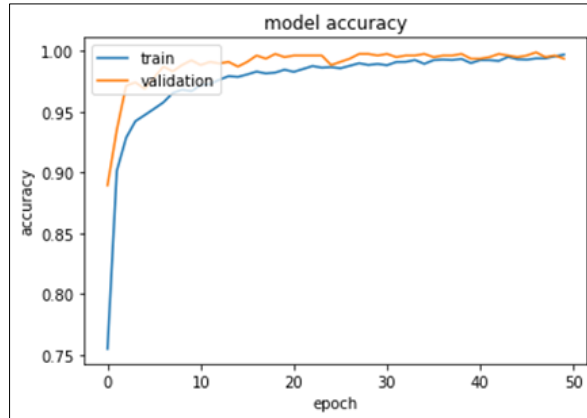


Figure 7 Model Accuracy

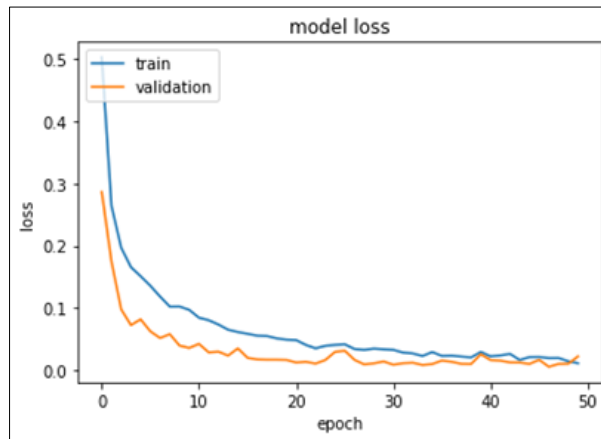


Figure 8 Model Loss

Table 1 Experimental results of CNFMD model

SN.	Dataset	Accuracy	Loss
1	Train	0.9970	0.0106
2	Validation	0.9935	0.0221
3	Test	0.9950	0.0262

5. Conclusion

As face masks will be an important part of human life for the next few years, authorities can use these models for real-time tracking of people without wearing masks and could educate them with the same. It was observed that the CNFMD

model gave good accuracy and minimal log loss on all three datasets. Test accuracy was turned out to be 99.50% which is very good on a real-life dataset

In the future, face recognition with face masks can be implemented and this model can be deployed using real-time cameras and thus can be used in public places. The research can further be broadened to detect people with face masks not worn properly.

Compliance with ethical standards

Acknowledgments

There is no external funding or institution involved in this study.

Statement of informed consent

The study doesn't include any work that requires consent.

References

- [1] X. Liu and S. Zhang, "COVID-19: Face masks and human-to-human transmission," *Influenza Other Respi. Viruses*. 2020; 14(4): 472–473.
- [2] M Loey, G Manogaran, MHN Taha, NEM Khalifa. "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic," *Meas. J. Int. Meas. Confed.* July 2020; 167: 108288.
- [3] T Greenhalgh, JL Jimenez, KA Prather, Z Tufekci, D Fisman, R Schooley. "Ten scientific reasons in support of airborne transmission of SARS-CoV-2," *Lancet*. 2021; 397: 10285 1603–1605.
- [4] A Jangra, "Face Mask Detection ~12K Images Dataset." 2020, [Online].
- [5] P Nagrath, R Jain, A Madan, R Arora, P Kataria, J Hemanth. "SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2," *Sustain. Cities Soc.* Mar. 2021; 66: 102692.
- [6] S Sethi, M Kathuria, T Kaushik. "Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread," *J. Biomed. Inform.* September 2020; 120; 103848.
- [7] A Chavda, J Dsouza, S Badgujar, A Damani. "Multi-Stage CNN Architecture for Face Mask Detection," *2021 6th Int. Conf. Converg. Technol. I2CT. 2021*; 1–8.
- [8] K Suresh, MB Palangappa, S Bhuvan. "Face Mask Detection by using Optimistic Convolutional Neural Network," *Proc. 6th Int. Conf. Inven. Comput. Technol. ICICT. 2021*; 1084–1089.
- [9] Y Li *et al.* "Robust detection for network intrusion of industrial IoT based on multi-CNN fusion," *Meas. J. Int. Meas. Confed.* 2020; 154: 107450.
- [10] S Albawi, TA Mohammed, S Al-Zawi. "Understanding of a convolutional neural network," *Proc. 2017 Int. Conf. Eng. Technol. ICET. 2017, 2018-Janua*, 1–6.
- [11] C Shorten, TM Khoshgoftaar. "A survey on Image Data Augmentation for Deep Learning," *J. Big Data.* 2019; 6(1).