International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT1217519

Face Mask Recognition Using MobileNetV2

Vatsal Patel*, Dhruti Patel
Devang Patel Institute of Advance Technology and Research, Charusat University, Gujarat, India

Article Info
Volume 7, Issue 5
Page Number : 35-42

Publication Issue :
September-October-2021

Article History
Accepted : 15 Sep 2021
Published : 26 Sep 2021

ABSTRACT

The pandemic of Corona Virus Disease is generating a public health emergency. Wearing a mask is one of the most efficient ways to combat the infection. This paper presents the detection of face masks, through mitigating, evaluating, preventing, and preparing actions regarding COVID-19. In this work, face mask identification is achieved using Machine Learning technique and the Image Classification algorithms are MobileNetV2 with major changes which includes Label Binarizer, ImageNet, and Binary Cross-Entropy. The methods involved in building the model are collecting the data, pre-processing, image generation, model construction, compilation, and finally testing. The proposed method can recognize people with and without masks. The training accuracy of the proposed method is 98.5% and the testing accuracy is 99%. This model is implemented in an image or video stream to detect faces with mask.

Keywords : Deep Learning, CNN, MobileNetV2, Face Mask, COVID-19

I. INTRODUCTION

As of July 2020, the Corona virus infection had infected 214 countries in just a few months, resulting in nearly five lakh deaths and affecting nearly every sector of life in nearly every corner of the world. School activities, jobs, travels, supply chains, and other activities are all paused or adjusted in an effort to reduce the pandemic's impact. The Corona virus was first reported in Wuhan, China, in December 2019. Dry cough, Fever, and tiredness are the most typical symptoms of COVID-19. These signs and symptoms are usually subtle and develop over time. Some people get infected but don't show symptoms that are comparable to COVID-19.

By following the WHO's advice, we can all lower our chances of contracting COVID-19. Hands should be washed and sterilised with soap and an alcohol-based sanitizer on a regular basis. According to several studies, keeping a 1-meter barrier between oneself and anybody who has a cold or cough is beneficial. Avoid contacting our eyes, mouth, and nose because touching different surfaces can lead to the spread of viruses. When you have the COVID-19 symptoms or are caring for someone who has the Corona virus, you must wear a mask. And, only a disposable face mask should be used.

In this paper, we used a Machine Learning algorithm to recognize face masks using photo classification approaches such as MobileNetV2. MobileNetV2 is a Convolution Neural Networks method with good performance and is more adequate. Label Binarizer, ImageNet, and Binary Cross-Entropy are some further image classification approaches. We can detect people with and without masks using these techniques. CNN can assign values to numerous objects in an image and distinguish between them using an assigned value, input image [8][9].

In section II, we addressed our views of various pieces of literature on Face mask recognition. A picture categorization approach based on MobileNetV2 CNN is also proposed. We put alternative classification algorithms to the test in section III and analyse the results in section IV.

II. RELATED WORK

Samuel Ady Sanjaya, Suryo Adi Rakhmawan [1]. The authors did research on face mask detection in this publication. Authorities can utilize the data to plan for COVID-19 mitigation, prevention, evaluation, and response against COVID-19. Face mask identification is accomplished in this work using a Machine Learning system called MobileNetV2. The image classification algorithm MobileNetV2 is based on CNN and is used to classify images (Convolution Neural Networks). In this case, the implemented model detects whether people are wearing masks or not with a 96.85 percent accuracy.

Pranav KN, Pramod Mahajan H, Praveen V, Sai Kiran, Dr. M Vinoth Kumar [2]. According to the authors' work in this study, some rudimentary Machine Learning libraries such as Tensorflow, OpenCV, and Keras are used. The writers' major goal is to wear masks that are appropriate for using their facilities. The suggested technique in this study consists of a pre-trained CNN and a cascade classifier with two 2D-Convolution layers that are connected to dense neuron layers.

Sanket Shete, Kiran Tingre, Ajay Panchal, Vaibhav Tapse, Prof. Bhagyashri Vyas [3]. Due to the increased challenges in detecting face details when wearing a mask, the authors have created a novel face detection task in this paper. They employed CNN-based Deep Learning algorithms to detect a face, gender, and visual recognition, and they also suggested a novel CNN approach to detect face masks that are made up of 3 Convolutional Neural Networks. Face mask detection is based on the experimental outcomes for face and mask kind. It also identifies the wearer's face and informs them if they are not wearing the mask.

Anushka G. Sandesara, Dhyey D. Joshi, Shashank D. Joshi [4]. The authors of this research suggested a stacked 2D-Convolution model for detecting facial masks that is highly effective. They used Gradient Descent for training and Binary Cross-Entropy as a loss function to create the proposed technique, which is a stack of 2D-Convolutional layers with ReLU activations and Max Pooling. Overall, they got a 95 percent testing accuracy and a 97 percent training accuracy.

Marielet Guillermo, Athena Rosz Ann Pascua, Robert Kerwin Billones, Edwin Sybingco, Alexis Fillone, Elmer Dadios [5]. The main goal of this research is to emphasize the importance of disease control and prevention in the use of face masks in densely populated areas. The authors of this article utilized an Artificial Neural Network (ANN), it is a system capable of identifying whether people in a crowd are wearing masks or not. In all of their testing and training, they received a perfect score of 99 percent.

III. METHODS AND MATERIAL

The main purpose of this proposed model is to detect the mask, which is necessary for this current pandemic situation. It first starts with collecting the data from various people with or without mask images. The trained model is able to differentiate between people with or without a mask. The proposed method uses 1915 images with masks and 1918 images without masks. At every first step, to initialize the Batch Size, Epochs and LR. The next approach is to label the data into two groups without a mask and with a mask Fig. 2 and 1. The process continues along with the conversion of the image to an array and Label Binarizer to encode the labels.

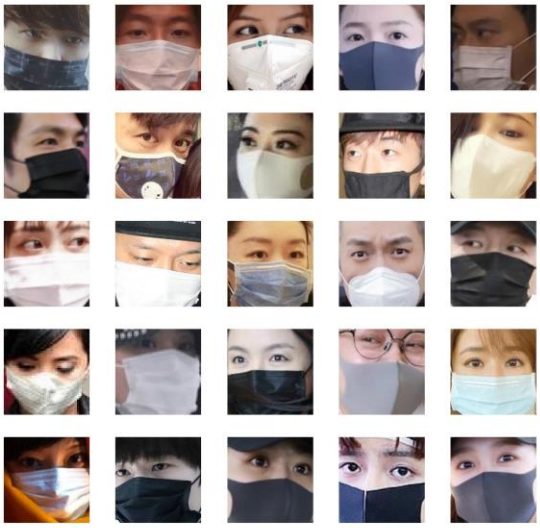

Figure 1. With Mask

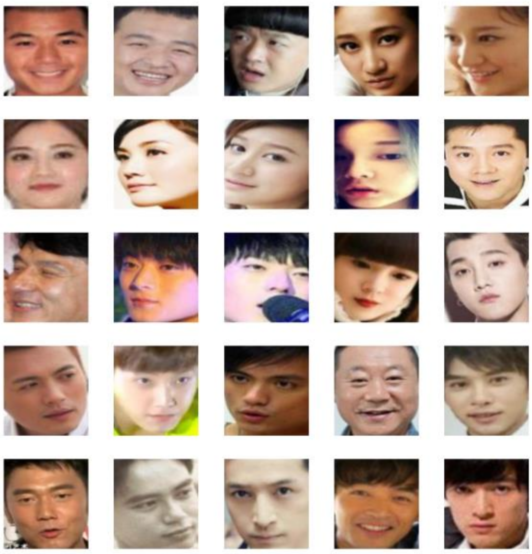

Figure 2. Without Mask

The pre-processing is phased along with train-test splitting. These four steps involved are Converting Image to Array, Resize the Data, Label Binarizer, Image Data Generator. Because of the efficacy of learning algorithms, picture scaling is a key pre-processing step in object recognition. The model will score higher if the image is lower in size. In this analysis, scaling the image results in the image are divided into 224*224 pixels. The following step is to convert all of the pictures in the dataset into an array. The picture is transformed into an array so that the repeat function may call it. Then, the picture will be utilized as pre-process data using MobileNetV2.

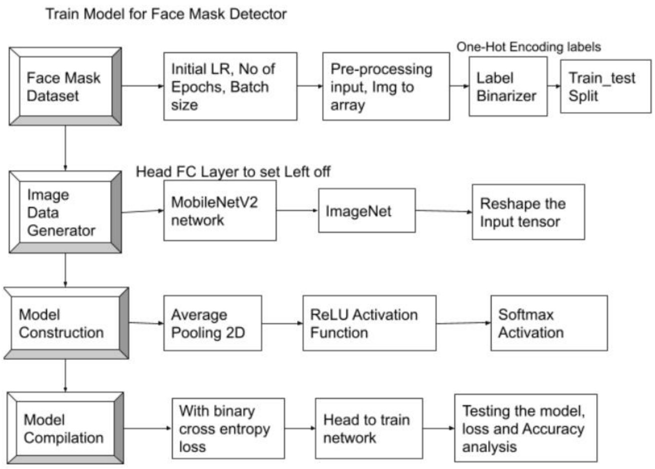

Figure 3. Block Diagram for Train Model

The final step in this stage is to execute Label Binarizer, although many Machine Learning algorithms cannot function immediately on data labelling. They demand that all incoming and outgoing parameters, including this procedure, be numeric. The tagged data will be converted into a numeric label so that the program can comprehend and analyse it.

Following the pre-processing step, the data is set into two batch sizes: training data (80%) and testing data (the remaining 20%). Every sample includes either with or without mask images.

MobileNetV2: This method is more effective for segmentation and object detection. It is a well-designed architecture in the Convolutional Neural Network. Mobile Net is a class of CNN that was openly built by Google, and thus, this offers us a good place to begin when it comes to training our incredibly compact and quick classifiers [11].

Label Binarizer: The transform method in the Label Binarizer makes this operation simple. When it comes to prediction, The matching model provides the highest level of performance by assigning the class. The inverse transform method in Label Binarizer makes this simple. Negative labels must be encoded with this value [13].

ReLU Activation: For short, ReLU is a piece-wise linear transformation that outputs the input directly if it is positive; else, it outputs zero. It became the standard activation function for several Neural Network models since it is willing to implement and typically results in high quality. A rectified linear activation unit, or ReLU for short, is a cluster or component that performs the activation function. Resolved networks are networks that employ the rectifier function for their hidden neurons. Acceptance of ReLU is readily regarded as one of the few watershed moments in the Deep Learning renaissance, alongside approaches that today allow for the everyday construction of incredibly Deep Learning techniques. The functions involved in the ReLU activation function are Computational Simplicity, Linear Behaviour, Representational Sparsity, and Train Deep Networks [15][16].

```
Algorithm 1: Proposed Model using CNN with MobileNetV2 with Major Changes

Input: RMFD Kaggle Dataset
Output: Recognition of Mask in Video stream
Initialization:
Labels and data
for category in Dataset:
        Path of the dataset
        for image in path:
                Read Image
                Convert 224*224
                Image to Array
                Preprocess the image
        End
train_test_split(data, labels,  test_size=0.20, stratify=labels,
random_state=42)
Data Augmentation
MObileNetV2 Network
Model Building
Compiling Model
Training the network
model.predict(testX, batch_size=BS),np.argmax(pred Idxs, axis=1)
Classification report
Plotting the Loss, Accuracy, val_accuracy, and val_loss
```

Figure 4. Pseudocode for Training Model

The testing process is as follows, the implementation of the proposed model in the real-world video stream is described in Fig. 5. At every first step, loading the face mask classifier from the model. And, applying the model in various ways like in images and recorded videos. We can predict the accuracy of the person with or without a mask in the frame. The usage of this model gives accurate results with better performance compared to the past. The testing algorithms is available in Fig. 6.

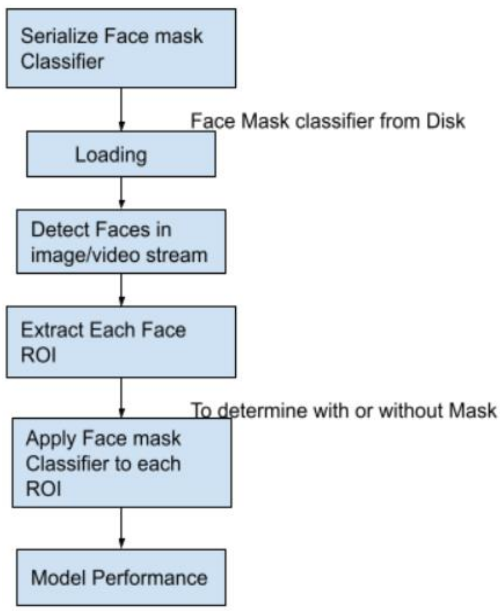

Figure 5. Block Diagram of Testing

```
Algorithm 2: Detection of Mask in video Stream
Function Detect and Predict mask
    Dimension of frame and Construct a bod
    Initialize faces,locations, and Prediction
    for i=0 to detection shape:
        confidence = detections[0, 0, i, 2]
        Filter the weak portion
        If confidence greater 0.5:
            (startX, startY, endX, endY) = box.astype("int")
            Bounding box fall in dimensions of frame
            Face ROI, BGR to RGB, Resize 224*224
            faces.append(face)
            locs.append((startX, startY, endX, endY))
    End
    Detect at least on face
While true: (Loop over the Frame from the video stream)
    Read, resize
    For box, pred in zip:
        Bounding box, prediction
        Color the portion
        Probability label
        Display the identification
    Press "q" to quit
```

Figure 6. Pseudocode for Testing

IV. RESULTS AND DISCUSSION

The performance and results are explained as follows. The proposed model was discovered by using MobileNetV2 with major changes to perform well when the model is ready to be tested in the real world. Our model gives the best results as compared to the latest models. This model uses techniques called Label Binarizer, ImageNet, Binary Cross-Entropy, ReLU activation function, and some other techniques to improve the model performance. The classification report is given for training and as well for testing with the best performance.

In this proposed method, we did not just evaluate the contained box of the photos generated, but we also focused on separating information about faces from pictures with varying backdrops and perspectives. We worked on the amalgamation of Kaggle and RMFD dataset. The train and test split, which is 80% and 20%, used 10 epochs with a batch size of 64. The table shows the performance of each epoch with the proposed model.

Epoch	Step-loss	Accuracy	Val-loss	Val-accuracy
1/10	0.4781	0.8175	0.2446	0.9544
2/10	0.2093	0.9587	0.1341	0.9713
3/10	0.1382	0.9707	0.0929	0.9778
4/10	0.1017	0.9770	0.0762	0.9817
5/10	0.0877	0.9780	0.0631	0.9857
6/10	0.0745	0.9830	0.0613	0.9804
7/10	0.0716	0.9810	0.0533	0.9883
8/10	0.0630	0.9823	0.0520	0.9857
9/10	0.0596	0.9823	0.0471	0.9883
10/10	0.0510	0.9877	0.0432	0.9896

Table 1. The Performance of Accuracy and Loss of the Iterations

From the above table, we can observe the relationship between train accuracy and loss for each epoch. If the epoch number is increased, the rise in accuracy and decrease in loss will be felt. The effective results were obtained when there is a loss rise in accuracy gives better performance. In the first epoch, the accuracy was 81%, but as the epoch value increased, the accuracy performance was 98.77% at the 10th epoch. All the performances of each epoch are shown in the graph Fig. 7.

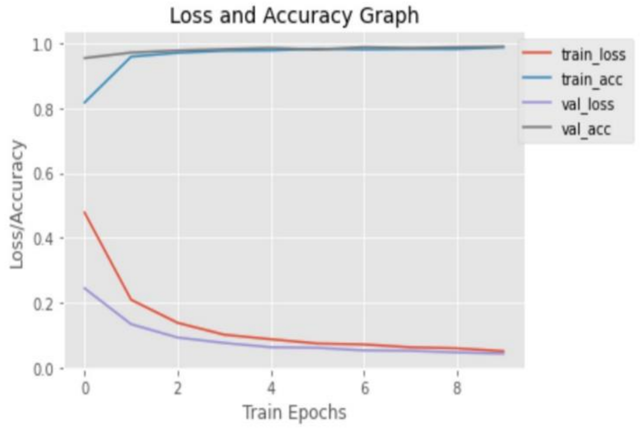

Figure 7. Train and Val Performance in Each Epoch

The training and testing performance of the model is characterized as a classification report in Fig. 8 and Fig. 9. The stable accuracy is represented in the graph. It indicates that without the need for further iterations to improve the performance of the models. So, as illustrated, the next step is to evaluate the model.

Train-loss	Train-accuracy	Test-loss	Test-accuracy
0.04178731	0.9849	0.04323413	0.9895

Table 2. Train and Test Performance

	precision	recall	f1-score	support
with_mask	0.978	0.993	0.985	1532
without_mask	0.993	0.977	0.985	1534
accuracy			0.985	3066
macro avg	0.985	0.985	0.985	3066
weighted avg	0.985	0.985	0.985	3066

Figure 8. Classification Report of Training Model Performance

	precision	recall	f1-score	support
with_mask	0.982	0.997	0.990	383
without_mask	0.997	0.982	0.990	384
accuracy			0.990	767
macro avg	0.990	0.990	0.990	767
weighted avg	0.990	0.990	0.990	767

Figure 9. Classification Report of Testing Performance

The model shown in the picture is the image which is read scene by scene, and then the facial recognition technology is used. If a face is spotted, the procedure moves on to the next step. Cognitive information processing will be performed on identified frames containing faces, including reducing the picture size, transforming to an array, and preprocessing input using MobileNetV2.

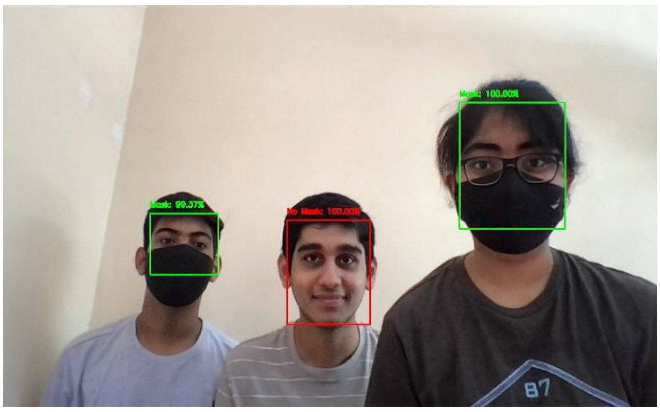

Figure 10. Recognition of Mask in Multiple Frame

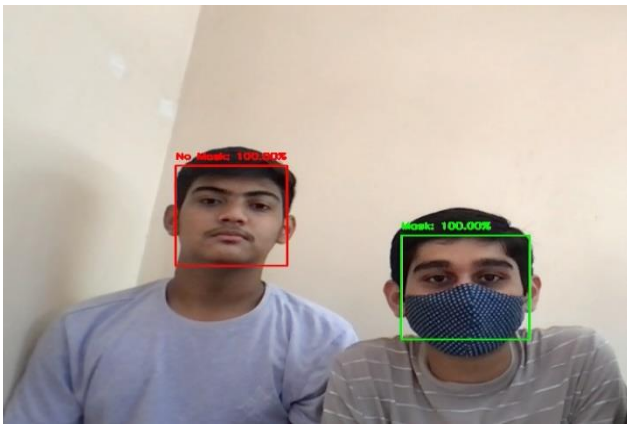

Figure 11. Accurate Recognition of Mask and Without Mask

The stored algorithm is then used to forecast data input. Estimate the processing input picture using a pre-existing model. Furthermore, the test image will be tagged with whether or not the individual is wearing a mask, as well as the predicted percent. Fig. 10 and 11 depicts one method of putting the concept into action.

V. CONCLUSION

To sum up, we discussed the importance of detecting whether or not people are wearing masks by implementing Machine Learning (ML) methods such as CNN's, MobileNetV2. We can predict how required Precautions should be taken in the future from the COVID-19 situation by using these strategies. This model may be improved further to determine whether the mask is virus-sensitive and to identify what type of mask it is, such as surgical, N95, etc.

VI. REFERENCES

[1] Sanjaya, S. A., & Rakhmawan, S. A. (2020, October). Face Mask Detection Using MobileNetV2 in The Era of COVID-19 Pandemic. In *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)* (pp. 1-5). IEEE.

[2] Shete, S., Tingre, K., Panchal, A., Tapse, V., & Vyas, B. (2021). Mask Detection and Tracing System.

[3] Sandesara, A. G., Joshi, D. D., & Joshi, S. D. (2020). Facial Mask Detection Using Stacked CNN Model.

[4] Guillermo, M., Pascua, A. R. A., Billones, R. K., Sybingco, E., Fillone, A., & Dadios, E. (2020). COVID-19 Risk Assessment through Multiple Face Mask Detection using MobileNetV2 DNN.

[5] Jiang, M., Fan, X., & Yan, H. (2020). Retinamask: A face mask detector. *arXiv preprint arXiv:2005.03950*.

[6] Balaji, S., Balamurugan, B., Kumar, T. A., Rajmohan, R., & Kumar, P. P. (2021). A brief Survey on AI Based Face Mask Detection System for Public Places. *Irish Interdisciplinary Journal of Science & Research (IIJSR)*.

[7] Haider, W., Bashir, H., Sharif, A., Sharif, I., & Wahab, A. (2014). A survey on face detection and recognition approaches. *Research Journal of Recent Sciences ISSN*, *2277*, 2502.

[8] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” in Advances in neural information processing systems, 2015, pp. 91–99.

[9] Rosebrock, A., 2020. COVID-19: Face Mask Detector With Opencv, Keras/Tensorflow, And Deep Learning - Pyimagesearch. [online] PyImageSearch. Available at: https://www.pyimagesearch.com/2020/05/04/covid19-face-mask-detector-with-opencv-keras-tensorflow-and-deep-learning/ .

[10] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520).

[11] Liu, J., & Wang, X. (2020). Early recognition of tomato gray leaf spot disease based on MobileNetv2-YOLOv3 model. *Plant Methods*, *16*, 1-16.

[12] Raghavan, A. K., Umaashankar, V., & Gudur, G. K. (2019). Label Frequency Transformation for Multi-Label Multi-Class Text Classification. In *KONVENS*.

[13] Ho, Y., & Wookey, S. (2019). The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling. *IEEE Access*, *8*, 4806-4813.

[14] Agarap, A. F. (2018). Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*.
[15] Yarotsky, D. (2017). Error bounds for approximations with deep ReLU networks. *Neural Networks*, *94*, 103-114.
[16] Lin, R. 2020. Face Detection and Segmentation Based on Improved Mask R-CNN. Discrete Dynamics in Nature and Society, 2020, p.9242917.

Cite this article as :

Vatsal Patel, Dhruti Patel, "Face Mask Recognition Using MobileNetV2", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7 Issue 5, pp. 35-42, September-October 2021. Available at
doi : https://doi.org/10.32628/CSEIT1217519
Journal URL : https://ijsrcseit.com/CSEIT1217519