IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 7, July 2021

DOI 10.17148/IJARCCE.2021.10785

Face Mask Detection using OpenCV and Machine Learning

Honnaraju B[1], Suhana B U[2], Spoorthi Raghavendra Prasad[3], Pooja B[4], Gayathri K Raj[5]

[1] Assistant Professor, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Karnataka, India

[2,3,4,5] UG Student, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Karnataka, India

Abstract: COVID-19 is the greatest humanitarian challenge facing the world ever since World War II. World has been fighting the pandemic with great spirit, with the unlocking phases being in motion. Wearing mask properly has become an effective method to cut down the spread of virus transmission. In this work, we are targeting to build a system where we can detect fine-grained state of wearing mask. This work involves three major challenges: 1) different face orientations, 2) various mask types and patterns and 3) facial occlusions. To resolve these challenges we created a new practical dataset, which contains 4000 images. The proposed approach here uses OpenCV and Machine Learning with MobileNetV2 architecture as image classifier to perform mask detection in real time. This model achieves an accuracy of 97%. The new practical dataset created can be used for further advanced models for thermal screening and facial recognition.

Keywords: COVID-19, new practical dataset, OpenCV, Machine Learning, MobileNetV2.

I. INTRODUCTION

Outbreak of the COVID-19 pandemic has threatened human lives safety, and rapidly spread to the majority of countries in the world. On 30th January 2020, the World Health Organization (WHO) declared the outbreak a Public Health Emergency of International Concern and suggests that every individual should wear masks whenever they step out of home, in order to cut the spread of coronavirus. Hence, detection of face masks have become one of the significant task in Image Processing and Computer Vision domain.

Numerous researches and mask detection applications do not concern at problem of facial occlusions and detecting different various mask types and patterns. To address these limitations we are targeting to build a system where we can detect fine-grained state of wearing mask. For this purpose, we created an efficient practical dataset consisting of 4000 images. We collected images from various sources such as MAFA, RMFD, CelebA and some images are downloaded from internet and cleaned the dataset by manual inspection.

In this proposed work, we have developed an efficient model for the detection of face masks using OpenCV and Machine Learning with MobileNetV2 as image classifier. To hinder the transmission of COVID-19 virus this model can be incorporated with surveillance cameras to detect people wearing mask improperly or not wearing masks. This model intend towards the elimination of incorrect predictions in real world scenarios that transpired in other related models.

Detection of masked face is a challenging task as they have various accommodations, various mask types and patterns and also various degree of obstructions. Multiple reasons are noticed for the unsatisfactory achievement of the existing models two of them were lack of suitable dataset and the mask on faces produces an inevitable noise that lessen the detection process. Addressing these issues, an efficient model is constructed.

The significant contributions of proposed work are given below:

i. Creation of fine-grained practical dataset with facial occlusions and various mask types and patterns, which can be helpful for studying new models.

ii. Usage of OpenCV, allows us to perform mask detection in real time with limited resource. It also detects occluded faces and faces in different orientations with good accuracy.

 This work is licensed under a Creative Commons Attribution 4.0 International License

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 7, July 2021

DOI 10.17148/IJARCCE.2021.10785

The remaining part of the work is categorized into different sections as shown below: some of the related previous works done are discussed in Section 2. The detailed descriptions of the methods that are used in the construction of the model is reviewed in Section 3. Experimental results discussion together with the analogy table of proposed work and previous works are mentioned in Section 4. Finally, in Section 5, the proposed work with a concise summary is concluded.

II. RELATED WORK

Over the past few decades, numerous analysts and researchers worked on detection of face technology. One of the eminent face detector proposed by Viola-Jones, provided a breakthrough in real time face detection [1]. But it encountered few problems like face orientations, face brightness, making hard to intercept. In 2006, a random forest tree model approach in detecting masked faces was proposed by Ramanan, which predicts face poses and structures.

In 2015, Yang designed a face detection model with facial features aggregation method [2]. In the classification based on convolutional neural network, the facial detection models learns directly from the data given by the user and then various machine learning algorithms applied on it [3]. In 2016, Opitz-Waltner formulated a grid loss in order to resolve facial occlusion problems [4].

The masked face detection task become evident with the COVID-19 outbreak, which aims minimize the spread of virus by detecting whether an individual is wearing a mask or not. Jiang-Fan-Yan proposed single-stage generic detector based RetinaMask, which tries to extract robust features taking the context information into consideration [5]. Loey-Manogaran designed a hybrid model for detecting medical face masks using the combination of various conventional machine learning algorithms and deep neural network and an implementation is done using ResNet-50 and the Single Stage Detector YOLOv2 is made, which increases the performance of detection [6]. Militante-Dionisio proposed a model where the input images is divided into several segments and the prediction of final results using VGG16 models [7].

In the above face detection and face mask detection models aims at detecting face masks or clean faces but ignored different fine-grained mask wearing conditions. The proposed model targets to solve all the issues of previous works with high average precision on the practical dataset and proves that it has strong potential to derive the distinguish feature map.

III. METHODOLOGY

A. Dataset Visualization

For the face mask detection, only a few datasets are available and most of them are either created artificially or full of noise with wrong labels. Hence, we created a practical dataset to train the model which yielded fruitful results at the end. The dataset comprising of images that are obtained from abundant data sources including MAFA, RMFD, CelebA and some images are downloaded from internet, cleaned the dataset by manual inspection. Finally, a dataset having 4000 images was created, with the label “with_mask” and “without_mask” and the distribution is visualized in Figure 1.

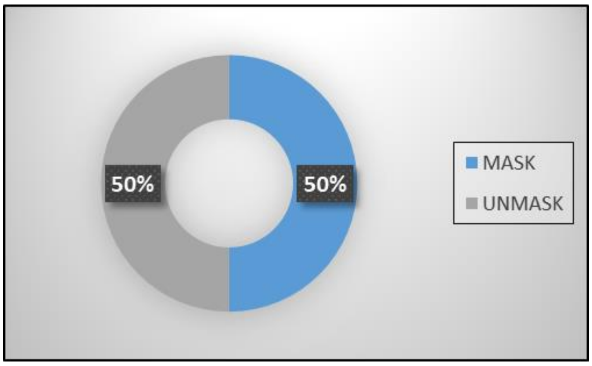

Figure 1: Dataset Visualization

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 7, July 2021

DOI 10.17148/IJARCCE.2021.10785

B. Work Flow

The initial stage of the process starts with training the model using the appropriate dataset to predict in real time whether an individual is wearing mask or not. After the classifier is trained, a precise face detector model is essential for detection the faces so that it can classify masked and unmasked faces. For this task, an OpenCV's Single Shot Multibox Detector with ResNet-10 as its backbone, helps in real time detection of faces. The classifier uses MobileNetV2 pre-trained model to do predictions. The proposed model's work flow is illustrated in Figure 2.

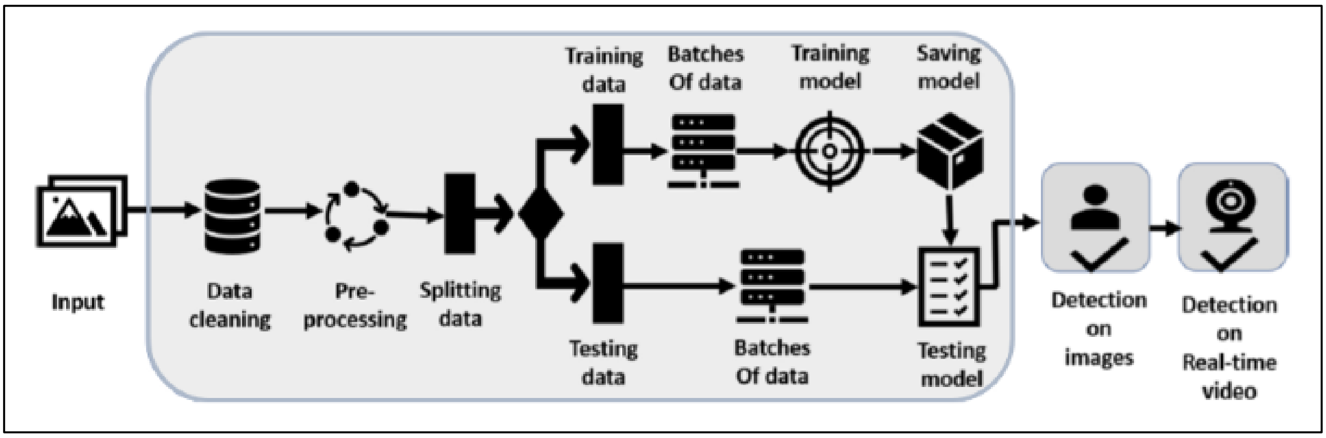

Figure 2: Proposed work flow

C. Detection of Faces

The model constructed on the 'Single Shot Detector' using ResNet-10 architecture as its base model, has the potential for detecting multiple objects in an image at only one shot. This model with two versions are made available in GitHub repository of OpenCV –

i. Caffe Implementation
ii. Original Tensorflow Implementation

We have used Caffe model for our proposed model to detect faces and face masks as it is more powerful, faster and efficient in contrast with other object detection models. On the application of facial detection model, the sum total of all the faces with their boundary box location and the confidence score of the predictions is obtained. These predictions are taken as the mask classifier's input that allows real time detection of faces at different orientations.

D. MobileNetV2 Classifier

MobileNetV2 is a convolutional neural network model, employed for image classification. Pre-trained ImageNet weights are added and to avoid learning features impairment, the layers of the base model are froze. Then to determine features to distinguish between the faces without masks from the faces with masks, new layers which are trainable are appended and trained on the newly created practical dataset. Finally, fine-tuning of the model done and saved their weights.

MobileNetV2 is a convolutional neural network architecture that utilizes the below mentioned layers and functions. The Figure 3 depicts the MobileNetV2 architecture.

- Convolutional Layer

It is the primitive block of Convolutional Neural Network, which is a mathematical combination of dual functions. It helps in feature extraction and generating feature maps.

- Pooling Layer

To make calculations faster by minimizing the input matrix size without dropping many features, pooling operations are applied. Some of the pooling operations are - Max Pooling and Average Pooling.

- Dropout Layer

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 7, July 2021
DOI 10.17148/IJARCCE.2021.10785

This helps in reducing a model from overfitting, by dropping random biased neurons, which can be a fragment of visible layers and hidden layers.

- Non-Linear Layer

It consists of an activation function that draws generated feature map from convolutional layer as input and creates activation map. Some of the most commonly used non-linear functions are various Rectified Linear Unit (ReLu), sigmoid functions and tanh functions.

- Fully-Connected Layer

These layer helps in classifying images in binary classification. These layers have complete connections to activation layers and are appended in the model. It yields the predicted output classes results.

- Linear Bottlenecks

As we cannot minimize the multiple matrix multiplication into a solitary numerical operation, some non-linear activation functions are implemented, thus allowing us to build multiple layers by discarding values which are less than zero. This information loss can be tackled by increasing the capacity of the network.

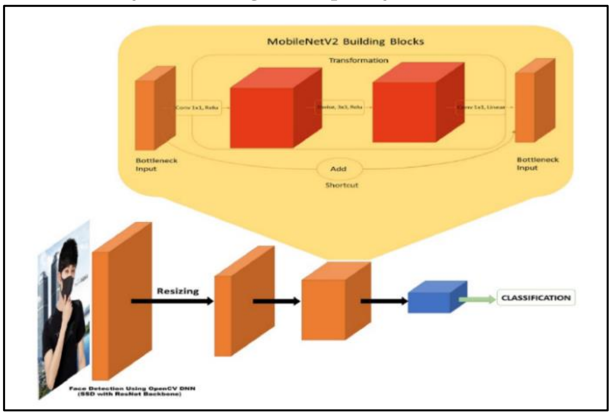

Figure 3: MobileNetV2 Architecture

E. Algorithms

ALGORITHM 1: PRE-PROCESSING AND TRAINING ON DATASET

INPUT: Images
OUTPUT: Trained model

Step1: Load images
Step2: Process the images (i.e., resizing, normalization and convert to 1D array)
Step3: Load the file names and their respective labels
Step4: Perform data augmentation and then split the dataset into train and test batches
Step5: Load MobileNetV2 model from Keras, train them on batches and compile it using Adam optimizer
Step6: Save the model

ALGORITHM 2: DEPLOYMENT OF FACE MASK DETECTOR

INPUT: Images or Live Stream
OUTPUT: Real time classification of images into Masked or Unmasked

Step1: Load face detector from OpenCV and saved model from disk
Step2: If input is Image:
Load image(s)
Step2.1: To detect the faces in an image apply face detection model
Step2.2: If the faces are detected:
Crop the faces to boundary box co-ordinates from face detection model
Receive predictions from the classifier model

 This work is licensed under a Creative Commons Attribution 4.0 International License

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 7, July 2021
DOI 10.17148/IJARCCE.2021.10785

Display the prediction
Else:
Display no output
Step3: If input is Live Stream:
Load live stream feed from OpenCV
Read the live stream feed frame by frame
Step3.1: To detect the faces in frames apply face detection model
Step3.2: If faces are detected:
Crop the faces to boundary box co-ordinates from face detection model
Receive predictions from the classifier model
Display real time output
Else:
Show feed normally
Step4: Stop video or live stream

IV. RESULTS AND DISCUSSION

To conduct experimental trails we have used a laptop customized with AMD Ryzen 5 3550H processor with 8GB RAM. The Jupyter Notebook Software with Python 3.8.5 kernel is employed for the experimental trail development and implementation.
For the evaluation of the proposed model, the following metrics are used.

$$Accuracy = \frac{TP + TN}{(TP + FP + FN + TN)} \quad (1)$$

$$Precision = \frac{TP}{(TP + FN)} \quad (2)$$

$$Recall = \frac{TP}{(TP + FN)} \quad (3)$$

$$F1\ Score = 2 * \frac{Recall * Precision}{(Recall + Precision)} \quad (4)$$

Where,

TP = True Positive
TN = True Negative
FP = False Positive
FN = False Negative

As the classes were balanced, the accuracy obtained was a satisfactory measure to start with. The predictive positive values measure was given by Precision. The potential of a classifier to acquire the possible positive samples was given by Recall and the test accuracy measure was given F1 score.

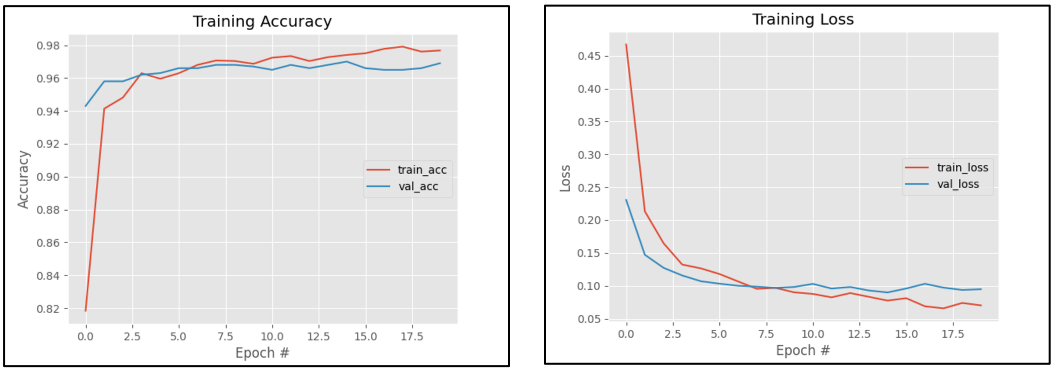

Figure 4: Training Accuracy and Training Loss Plots

© IJARCCE This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 7, July 2021

DOI 10.17148/IJARCCE.2021.10785

Table 1: Classification Report

	precision	recall	f1-score	support
with mask	0.95	0.99	0.97	500
without mask	0.99	0.95	0.97	500
accuracy			**0.97**	1000
macro avg	0.97	0.97	0.97	1000
weighted avg	0.97	0.97	0.97	1000

Figure 5 depicts the predictions made on few images by the proposed model. The green rectangular box with an accuracy score depicts that the person is wearing mask, while the rectangular box with an accuracy score depicts that the person is no wearing masks.

Figure 5: Visualization of Results

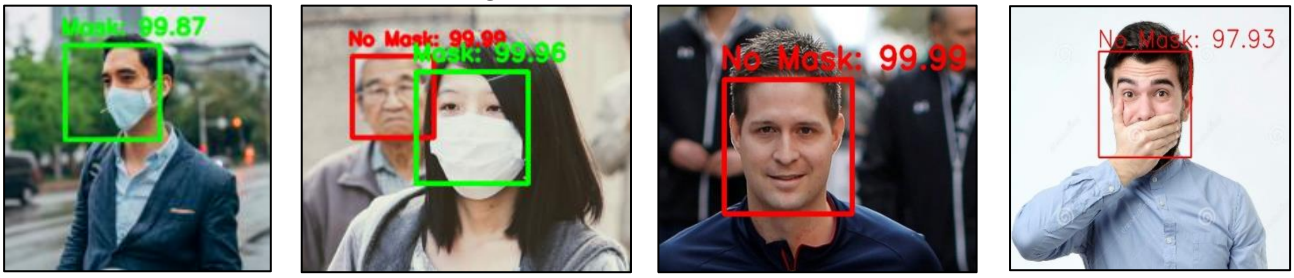

Table 2: Comparison of accuracy of previous work with our proposed work

Model	Accuracy
VGG-16	0.92
ResNet-50	0.91
Proposed Model	**0.97**

V. CONCLUSION

In this paper, our proposed method which is developed using OpenCV and Machine Learning with MobileNetV2 image classifier handled the limitations of the previous work and yielded fruitful results. The issues of wrong predictions have been successfully removed and thus we can notice that the model's accuracy has increased. This automated system will operate in an efficient manner to track people in public places where crowd monitoring is necessary. The new practical dataset created can be used for further advanced models for thermal screening and facial recognition.

REFERENCES

[1] Paul Viola and Michael Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, February – 2001, pp.I-511.

[2] Shou Yang, Ping Luo, Chen Change Loy, Xiaoou Tang, "From Facial Parts Responses to Face Detection: A Deep Learning Approach", www.arXiv.org, September – 2015, arXiv: 1509.06451.

[3] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", Advances in Neural Information Processing Systems, June - 2015, pp. 91 - 99.

[4] Michael Optiz, Georg Waltner, Georg Poier, Horst Possegger and Horst Bischof, "Grid Loss: Detecting Occluded Faces", European Conference on Computer Vision, October - 2016, pp.386 - 402.

[5] Mingjie Jiang, Xinqi Fan and Hong Yan, "RetinaMask: A Face Mask Detector", .org, May - 2020, arXiv: 2005.03950.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 7, July 2021

DOI 10.17148/IJARCCE.2021.10785

[6] Mohamed Loey, Gunasekaran Manogaran, Mohamed Hamed N Taha and Nour Eldeen M Khalifad, "A Hybrid Deep Transfer Learning Model with Machine Learning methods for Face Mask Detection in the era of the COVID-19 Pandemic", Measurement, Jan - 2021, Vol.178, ArtNo.108288.

[7] Sammy V. Militante and Nanette V. Dionisio, "Real-time Facemask Recognition with Alarm system using Machine Learning", 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC), August – 2020, pp. 106 - 110.

[8] Hoanh Nguyen, "Fast Object Detection Framework based on MobileNetV2 Architecture and Enhanced Feature Pyramid", Journal of Theoretical and Applied Information Technology, March – 2020, Vol.98, No.05.

[9] Haoxiang Li, Zhe Lin, Xiaohui Shen, Jonathan Brandt and Gang Hua, "A Convolutional Neural Network Cascade for Face Detection", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June - 2015, pp.5325 - 5334.

[10] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov and Liang-Chieh Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp.4510 - 4520.

[11] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu and Alexander C Berg, "SSD: Single Shot MultiBox Detector", European Conference on Computer Vision, Springer, 2016, pp. 21-37.

[12] Mohamed Loey, Gunasekaran Manogaran, Mohamed Hamed N Taha and Nour Eldeen M Khalifad, "Fighting against COVID-19: A Novel Machine Learning model based on YOLO-V2 with ResNet-50 for Medical Mask Detection", Sustainable Cities and Society, February - 2021, Vol. 65, Art No. 102600.

[13] Bosheng Qin and Dongxiao Li, "Identifying Facemask-Wearing Condition Using Image Super-Resolution with Classification Network to Prevent COVID-19", Sensors, September - 2020, Vol.20, No.18, pp.5236.

[14] Zhongyuan Wang, Guangcheng Wang, Baojin Huang, Zhangyang Xiong,, Qi Hong, Hao Wu, Peng Yi, Kui Jiang, Nanxi Wang, Yingjiao Pei, Heling Chen, Yu Miao, Zhibing Huang and Jinbi Liang, "Masked Face Recognition Dataset and Application", arXiv.org, March - 2020, preprint arXiv: 2003.09093.

[15] Rajeev Ranjan, Vishal M Patel and Rama Chellappa, "Hyperface: A Deep Multi-Task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), December – 2017, pp. 121 - 135.

AUTHOR DETAILS

1. Honnaraju B
Assistant Professor, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Karnataka, India.
2. Suhana B U
UG Student, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Karnataka, India.
3. Spoorthi Raghavendra Prasad
UG Student, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Karnataka, India.
4. Pooja B
UG Student, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Karnataka, India.
5. Gayathri K Raj
UG Student, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Karnataka, India.