International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 07 | July 2022 www.irjet.net p-ISSN: 2395-0072

FACE MASK DETECTION USING MACHINE LEARNING AND IMAGE PROCESSING

Siya Kamat[1], Akshata Karapurkar[2], Rinki Matonkar[3], Sanat Shirodkar[4], Rajesh Gauns[5], Diksha Prabhu Khorjuvenkar[6]

1,2,3,4 Department of Computer Engineering Agnel Institute of Technology and Design, Assagao, Goa
5,6 Assistant Professor, Department of Computer Engineering Agnel Institute of Technology and Design, Assagao, Goa

---***---

Abstract - *Abstract - In light of the current COVID-19 pandemic, many public places require that individuals wear face masks. However, few individuals don't consider it necessary. People use different excuses not to use a mask whenever in public areas and it is not always possible for the authorities to monitor whether people are wearing face masks or not. This is where machine learning-based face-mask detection can be helpful.*

Through machine learning, face-mask detection can be carried out to automatically determine if a face mask is being worn by someone or not. This can be achieved by utilizing a camera or a webcam to take a static image or video of the person.

With the help of this project, we will be able to detect whether or not a person is wearing a mask in both a static image and a real-time stream. This project can be used at the entrance of public places like shops, hospitals, banks, schools, colleges, malls, airports, etc. as a digital scanning tool. This will lessen human mistakes and instantly notify those without masks. Firstly, a dataset will be prepared to consist of 2 classes; with-mask and without-mask. A machine learning model shall be created and trained with the dataset gathered. After the training is completed, the model shall be assessed with a different dataset. Machine learning techniques will be used for classifying the detected people into the with-mask or without-mask category.

Key Words: Machine Learning, Image Processing, Face Mask Detection, MTCNN, EfficientNet, MobileNet

1. INTRODUCTION

Coronaviruses are a large family of viruses that infect humans, other mammals, and birds. These viruses can cause several respiratory, gastrointestinal, and neurological diseases. COVID-19 is a disease caused by a coronavirus, and its effects can possibly range from mild to fatality.

At times of a pandemic, the best prevention is to avoid contracting the virus. A few things one should follow are when coughing and sneezing covered with tissue which is then safely disposed of. Next, cleanse your hands routinely with a soap or a disinfectant containing at least 60% alcohol (in the absence of soap and water), avoiding direct contact with an infected person. It's also important to avoid using unwashed hands to contact the mouth, nose, or eyes.

Utilizing face masks is the most crucial step in stopping the spread of sickness. However, some people don't think it's necessary and give various justifications for not wearing a mask when in public places. Using image processing and machine learning, it is possible to identify people's faces and divide them into two groups: those who are wearing masks and those who are not.

2. LITERATURE SURVEY

2.1 Face Mask Detection using MTCNN

Vansh Gupta and Rajeev Rajput developed the face mask detection system using MTCNN. The pattern of wearing face covers openly is ascending because of the COVID-19 pestilence everywhere in the world. People used to wear coverings before Covid-19 to protect their health from air pollution.

The suggested model in their work can be combined with observation cameras to prevent COVID-19 transmission by allowing the identification of people who are not wearing facial coverings but are using veils.[1]

2.1.1 Machine Learning (ML)

In general, Machine Learning is the study of algorithms that can automatically improve given more data. It is often seen as part of AI since these algorithms can often make decisions without being specifically programmed to do so. They learn from data, building a model that allows them to make predictions or decisions.[1]

2.1.2 Image Processing

Image processing is a method of manipulating an image to either enhance it or extrapolate knowledgeable data from it. It's a form of signal processing where the input is an image and the output can be either another image or characteristics/features related to that image.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 07 | July 2022 www.irjet.net p-ISSN: 2395-0072

2.1.3 OpenCV

OpenCV is a computer vision and machine learning software library that gives computer vision applications a common framework and expedites the deployment of machine learning in commercial goods. The library includes a wide range of essential computer vision and AI computations, including more than 2500 simplified calculations.[1]

2.2 High-Precision Portrait Classification Based on MTCNN and Its Application on Similarity Judgment

Juan Du proposed a High-Precision MTCNN-Based Portrait Classification and Similarity Judgment Application. A complex course, portrait classification contains at least face detection, recognition, and comparison, each of which has several jobs. There are many difficult issues because of the skewed positions, illuminations, occlusions, blurred images, and small-scale faces in the photographs.

Multi-task convolutional neural network (MTCNN), a revolutionary face detection model, emerged in 2016 and quickly gained popularity. Real-time effects based on lightweight CNN, high efficiency and accuracy on face detection and face alignment tasks, and successful online hard sample mining all contribute to a large improvement.

MTCNN is a reframed CNN model composed of three layers of networks in the following sequence: P-net→R-net→O-net. It makes use of the candidate box and classifier concepts to carry out timely and efficient face detection: P-net is used to quickly create candidate boxes; R-net serves as a filter for accurately identifying candidate boxes, and O-net is used to automatically generate boundary boxes and essential aspects of the face.

MTCNN's thematic architecture is comparable to cascaded CNNs, but it addresses both face area and facial feature recognition simultaneously. MTCNN uses picture pyramid, bounding box regression, Non-Maximum Suppression (NMS), and several CNN technologies to solve image defects, just like many other CNN models.[2]

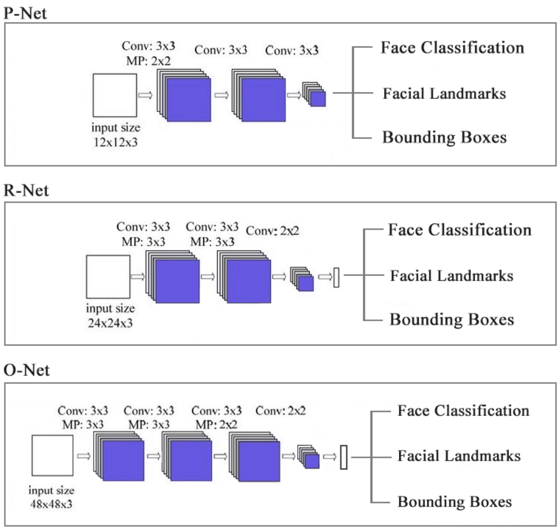

Fig -1: Structure of MTCNN

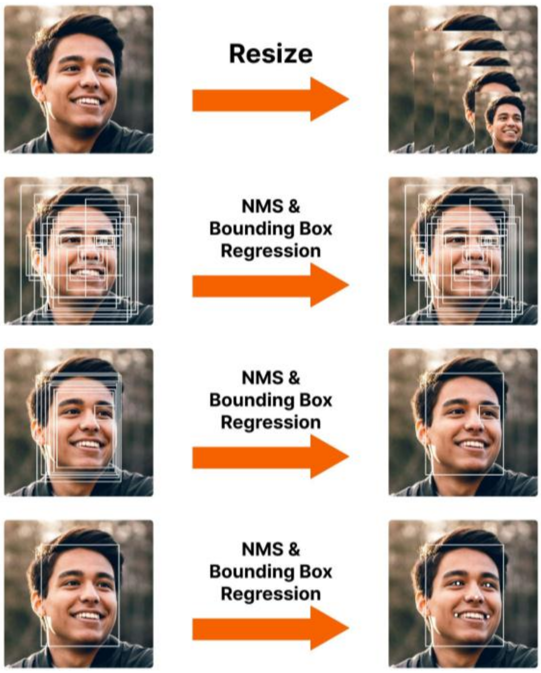

Fig -2: Stages of MTCNN

2.3 EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks

The development of Convolutional Neural Networks (ConvNets) typically begins with a limited resource budget and, if more resources are available, can be scaled up for better accuracy. Mingxing Tan has thoroughly investigated model scaling in this paper and found that optimizing network depth, width, and resolution can improve performance. Based on this insight, he has developed a new scaling technique that employs a straightforward but incredibly potent compound coefficient to equally scale all depth, width, and resolution parameters. Through the scaling up of MobileNets and ResNet, the usefulness of this strategy is shown. Scaling typically increases the model's accuracy for a particular task. According to the experiments done, the improvement in accuracy saturates pretty soon if only one way of scaling is done.

Further, a new baseline network is created via neural architecture search, and it is then scaled up to obtain a family of models, called EfficientNets, which outperform prior ConvNets in terms of accuracy and efficiency.

Stage i	Operator $\hat{\mathcal{F}}_i$	Resolution $\hat{H}_i \times \hat{W}_i$	#Channels $\hat{C}_i$	#Layers $\hat{L}_i$
1	Conv3x3	224×224	32	1
2	MBConv1, k3x3	112×112	16	1
3	MBConv6, k3x3	112×112	24	2
4	MBConv6, k5x5	56×56	40	2
5	MBConv6, k3x3	28×28	80	3
6	MBConv6, k5x5	14×14	112	3
7	MBConv6, k5x5	14×14	192	4
8	MBConv6, k3x3	7×7	320	1
9	Conv1x1 & Pooling & FC	7×7	1280	1

Fig -3: The architecture of baseline network EfficientNet-B0

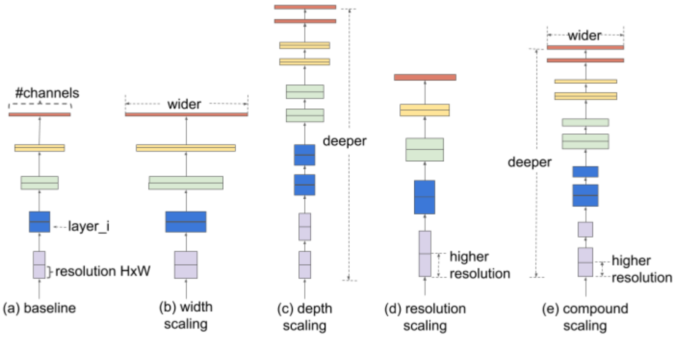

Fig -4: Comparison of different scaling methods. (e) shows Compound Scaling used in EfficientNet

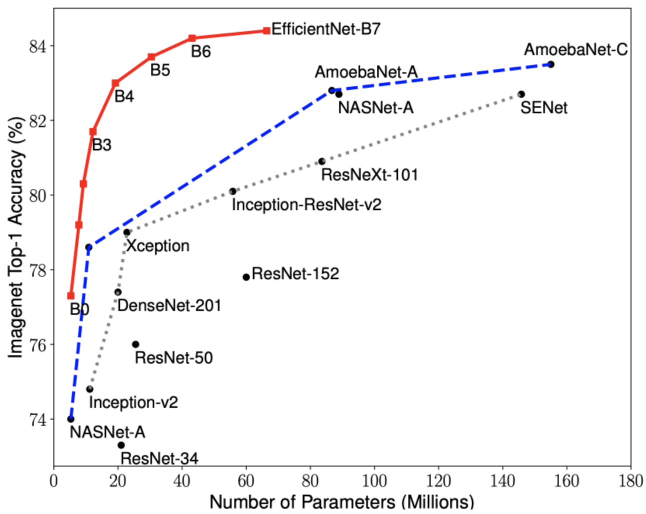

Fig -5: Performance comparison of EfficientNet vs other existing CNNs on ImageNet

The motivating insight of this project is that many individuals don't wear masks in public and therefore there is a need for a system that will identify whether or not someone is wearing a mask. Firstly, a dataset will be prepared to consist of 2 classes; with-mask and without-mask. A machine learning model will be created and trained with a dataset gathered. After the training is done, the model will be assessed with a different dataset. Machine learning techniques will be used for classifying the detected people into wearing masks or not wearing masks categories. Facial detection will be done using image processing. [3]

3. PROPOSED WORK

3.1 Problem Definition

This project's objective is to reduce the amount of labor required to manually find and identify individuals wearing or not wearing masks. This project uses face mask detection to automate the process of finding and identifying individuals who are wearing masks or not. This can replace the need for humans to do this manually, which can save time and labor.

Sometimes, humans can miss out on seeing a person without masks. For instance, a person could either not be wearing a mask or wearing a mask incorrectly, so it's not an easy task for the human to check everyone in a crowd unless they are standing in a queue. So therefore you can have an accurate system that can aid this work. Our system's goal is to identify mask wearers so that the coronavirus can be stopped from spreading further.

3.2 Project Specification

For face detection, we will be using the MTCNN algorithm which algorithm will perform its tasks in 3 parts ie. face classification, box bound regression, and facial landmark detection. The image classification i.e. whether the face in the image contains wear or not will be done by EfficientNet. The main goal of this project is to create a fully automated face-mask detection system.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 07 | July 2022 www.irjet.net p-ISSN: 2395-0072

3.3 Block Diagram

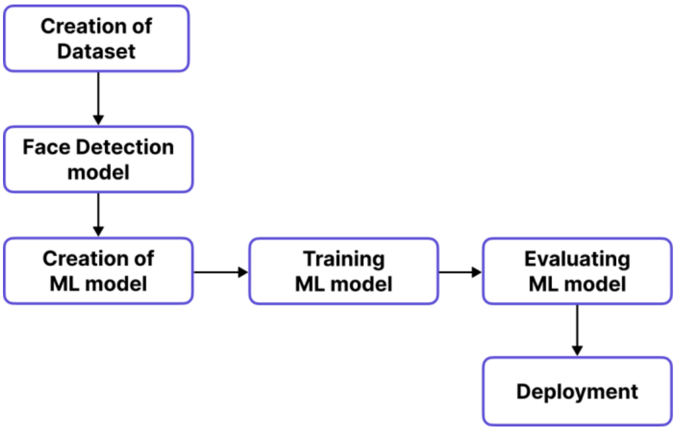

Fig -6: General project flowchart

3.4 Implementation

The objective of the project is to implement a face-mask detection model using MTCNN and EfficientNet. The algorithms are implemented using Python language. The first task that the model performs is detecting faces in the image or video stream. Face detection is carried out by the MTCNN algorithm. Now that we know the exact coordinates of the face, we can move on to further processing. The second task is to detect whether the detected face contains a mask or not, which is executed by the EfficientNet algorithm. The dataset used for the training of the model contains 20743 images belonging to 2 classes. The obtained model performs face-mask detection in real-time, even without GPU acceleration.

3.5 Results and Analysis

This section discusses the results obtained by implementing the model. The first version of the model made some false detections (Figure 7.1). The errors were eventually fixed by increasing the number of dataset images (from 10K to 20K+), improving the camera quality (from 720P to 1080P), and increasing the number of epochs (from 2 epochs to 7 epochs). Figure 7.1 shows results for various test cases obtained by running the latest version of the model.

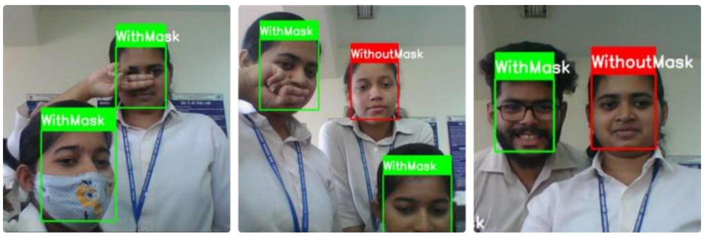

Fig -7.1: Errors found in initial versions

Fig -7.2: Test cases upon improvement of the model

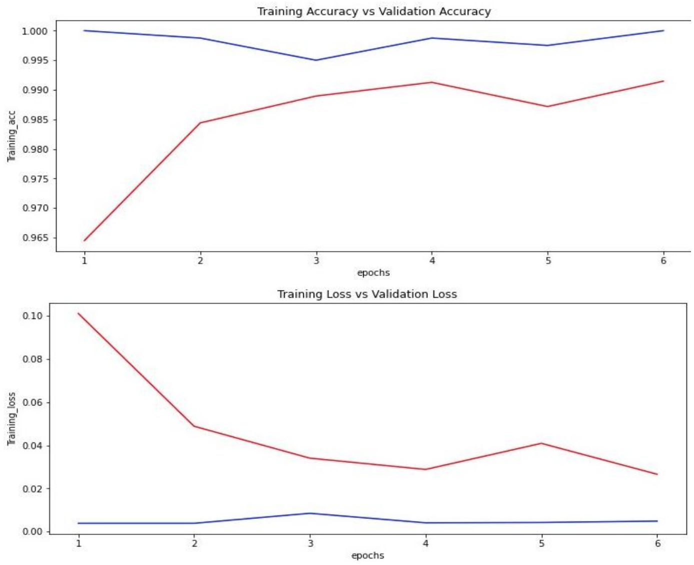

Fig -8: Training Accuracy v/s Validation Accuracy

REFERENCES

[1] Vansh Gupta and Rajeev Rajput (2021), "Face Mask Detection using MTCNN and MobileNetV2", International Research Journal of Engineering and Technology (IRJET) ,Vol: 08, page 309-310

[2] Juan Du (2020), "High-Precision Portrait Classification Based on MTCNN and Its Application on Similarity Judgment", Journal of Physics: Conference Series , page 1-9

[3] Mingxing Tan and Quoc V. Le (2019), "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", page 1-11

[4] Andrey V. Savchenko (2021), "Facial expression and attributes recognition based on multi-task learning of lightweight neural networks", HSE University, Laboratory of Algorithms and Technologies for Network Analysis, Nizhny Novgorod, Russia, Vol:03, page 1-15

[5] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li and Yu Qiao (2016), "Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks", The Institute of Electrical and Electronics Engineers(IEEE), Vol:01, page 1-5