International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 03 | Mar 2021 www.irjet.net p-ISSN: 2395-0072

Face Mask Detection using MTCNN and MobileNetV2

Vansh Gupta[1], Rajeev Rajput[2]

Electronics and Communication Engineering Department[1,2]
Dr. Akhilesh Das Gupta Institute of Technology & Management, New Delhi, India[1,2]

Abstract *- The COVID-19 pandemic is causing a worldwide wellbeing emergency so the viable security technique is wearing a face cover in open regions as indicated by the World Health Organization (WHO). As the world recuperates from the pandemic and plans to get back to a condition of routineness, there is an influx of tension among all people, particularly the individuals who expect to continue face to face movement. Studies have demonstrated that wearing a face veil essentially diminishes the danger of viral transmission just as gives a feeling of assurance. Notwithstanding, it isn't plausible to physically follow the execution of this strategy. Technology holds the key here. We will use the dataset to build a COVID-19 face mask detector with computer vision using Python, OpenCV, and Tensor Flow, and Keras. Our goal is to create a model that can detect face masks in crowded public places.*

Keywords: Computer Vision, COVID-19, Face Masks in Public Places, Safety Improvement, MTCNN, MobileNetV2

I. INTRODUCTION

The pattern of wearing face covers openly is ascending because of the COVID-19 pestilence everywhere in the world. Before Covid-19, People used to wear covers to shield their wellbeing from air contamination. While others are reluctant about their looks, they conceal their feelings from general society by concealing their countenances. Researchers sealed that wearing face covers deals with blocking COVID-19 transmission. COVID19 (known as Covid) is the most recent pandemic infection that hit human wellbeing in the only remaining century. In 2020, the quick spreading of COVID-19 has constrained the World Health Organization to pronounce COVID-19 as a worldwide pandemic. Man-made brainpower (AI) in light of Machine learning and Deep Learning can assist with battling Covid-19 from multiple points of view. Machine learning allows researchers and clinicians to evaluate vast quantities of data to forecast the distribution of COVID-19, to serve as an early warning mechanism for potential pandemics, and to classify vulnerable populations. The arrangement of medical services needs financing for arising innovation, for example, man-made reasoning, IoT, huge information, and AI to handle and anticipate new infections. To more readily comprehend contamination rates and to follow and rapidly distinguish diseases, the AI's force is being abused to address the Covid-19 pandemic. Individuals are constrained by laws to wear face veils out in the open in numerous nations. These standards and laws were created as an activity to the remarkable development in cases and passing in numerous zones. In any case, the way toward observing enormous gatherings of individuals is getting more troublesome. The checking cycle includes the location of any individual who isn't wearing a face cover. Here we present a veil face discovery model that depends on PC vision and profound learning.

II. LITERATURE SURVEY

The proposed model can be incorporated with observation cameras to block the COVID-19 transmission by permitting the discovery of individuals who are wearing veils not wearing face covers. The model is integration between deep learning and classical machine learning techniques with OpenCV, TensorFlow, and Keras. We have used deep transfer learning for feature extractions and combined it with three classical machine learning algorithms. We acquainted an examination between them with locating the most appropriate calculation that accomplished the most noteworthy exactness and burned-through minimal time during the time spent preparing and identification.

Machine Learning

ML is the study of computer algorithms that improve automatically through experience. It is seen as a subset of AI. AI calculations construct a numerical model dependent on example information, known as "preparing information", to settle on expectations or choices without being expressly modified to do as such. AI calculations are utilized in a wide assortment of utilizations, for example, email sifting and PC vision, where it is troublesome or infeasible to create traditional calculations to perform the needed tasks.

Computer Vision

PC vision is an interdisciplinary logical field that manages how PCs can acquire undeniable level comprehension from

computerized pictures or recordings. From the viewpoint of designing, it tries to comprehend and computerize assignments that the human visual framework can do, Computer vision errands incorporate strategies for securing, handling, examining, and understanding advanced pictures, and extraction of high dimensional information from this present reality to create mathematical or emblematic data.

Deep Learning

Profound learning techniques target taking in component pecking orders with highlights from more elevated levels of the progression framed by the structure of lower-level highlights. Consequently learning highlights at various degrees of reflection permits a framework to learn complex capacities planning the contribution to the yield straightforwardly from information, without relying totally upon human-created highlights. Profound learning calculations look to misuse the obscure design in the information dissemination to find great portrayals, regularly at numerous levels, with more elevated level learned highlights characterized as far as lower-level highlights.

OpenCV

Opencv (Open Source Computer Vision Library) is an open-source computer vision and machine learning software library. OpenCV was worked to give a typical foundation to PC vision applications and to quicken the utilization of machine discernment in business items. The library has in excess of 2500 streamlined calculations, which incorporates an extensive arrangement of both works of art and cutting edge PC vision and AI calculations. These calculations can be utilized to distinguish and perceive faces, recognize objects, group human activities in recordings, track camera developments, track moving articles, separate 3D models of items, produce 3D point mists from sound system cameras, line pictures together to create a high-goal picture of a whole scene, find comparable pictures from a picture data set, eliminate red eyes from pictures taken utilizing streak, follow eye developments, perceive view and build up markers to overlay it with enlarged reality, and so on.

TensorFlow

It is a free and open-source programming library for dataflow and differentiable programming across a degree of assignments. It is a representative mathematical library and is likewise utilized for AI applications, for example, neural organizations. It is utilized for both examination and creation at Google, TensorFlow is Google Brain's second-age framework. Keras is an API intended for people, not machines. Keras follows best practices for decreasing intellectual burden: it offers steady and straightforward APIs, it limits the number of client activities needed for basic use cases, and it gives clear and significant mistake messages.

III. PROPOSED SYSTEM

The proposed framework centers around how to recognize the individual on picture/video transfer wearing a face veil with the assistance of PC vision and profound learning calculation by utilizing the OpenCV, Tensorflow, Keras.
Our primary focus is to make a system work in crowded places.

Approach

1. Train Deep learning model (MobileNetV2) using Transfer Learning
2. Find ROI using MTCNN
3. Apply mask detector

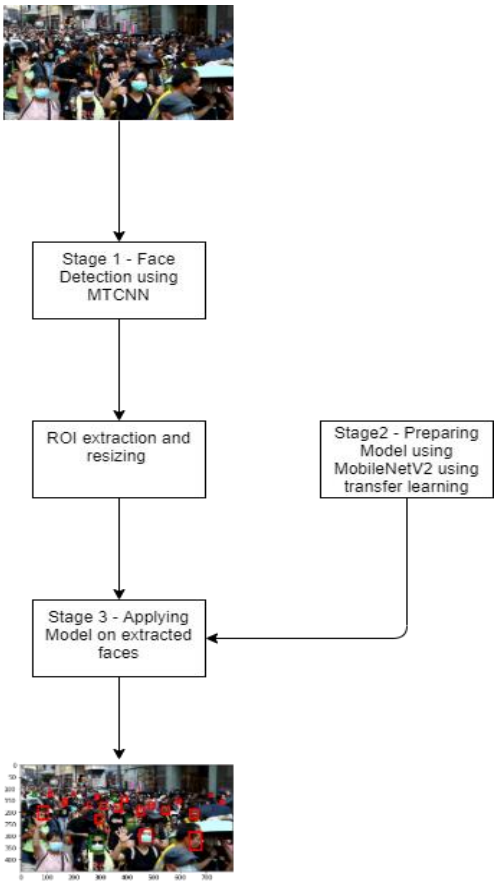

Fig.1: Flowchart of Face Mask Detection using MTCNN and MobileNetV2

Most of the pictures were increased by OpenCV. The set of images were already labeled “mask” and “no mask”. The pictures that were available were of various sizes and goals, likely separated from various sources or from machines

(cameras) of various goals. Data preprocessing steps as mentioned below was applied to all the raw input images to convert them into clean versions, which could be fed to a neural network machine learning model.

1. Resizing the input image (256 x 256)
2. Applying the tone separating (RGB) over the channels (Our model MobileNetV2 upholds 2D 3 channel picture)
3. Scaling / Normalizing images
4. Focus trimming the picture with the pixel estimation of 224x224x3
5. Finally Converting them into tensors (Similar to NumPy cluster)

Profound Learning Frameworks to actualize this profound learning network we have the accompanying choices.

1. Tensor Flow
2. Keras
3. PyTorch
4. Caffee
5. MxNet
6. Microsoft Cognitive Tool Kit

TensorFlow is chosen for this proposed model.

MobileNetV2

MobileNetV2 builds upon the ideas from MobileNetV1, using depth wise separable convolution as efficient building blocks. In any case, V2 acquaints two new highlights with the design:

1. Linear bottlenecks between the layers, and
2. Shortcut connections between the bottlenecks.
3. Typical MobilenetV2 architecture has as many layers,

The weights of each layer in the model are predefined based on the ImageNet dataset. The loads show the cushioning, steps, portion size, input channels, and yield channels. MobileNetV2 was picked as a calculation to construct a model that could be sent on a cell phone. A tweaked completely associated layer which contains four consecutive layers on top of the MobileNetV2 model was created. The layers are

1. Average Pooling layer with 7×7 weights
2. Linear layer with relu activation function
3. Dropout Layer
4. Straight layer with Softmax initiation work with the aftereffect of 2 qualities. The last layer softmax work gives the aftereffect of two probabilities every one addresses the characterization of "cover" or "not veil".

MTCNN

MTCNN (Multi-task Cascaded Convolutional Neural Networks) is a calculation comprising of 3 phases, which distinguishes the bounding boxes of appearances in a picture alongside their 5 Point Face Landmarks. Each stage gradually improves the detection results bypassing its inputs through a CNN, which returns candidate bounding boxes with their scores, followed by nonmax suppression. In stage 1 the information picture is downsized on numerous occasions to fabricate a picture pyramid and each scaled form of the picture is gone through its CNN. In stages 2 and 3 we extract image patches for each bounding box and resize them (24x24 in stage 2 and 48x48 in stage 3) and forward them through the CNN of that stage. Other than bounding boxes and scores, stage 3 moreover figures 5 face milestones focuses for each bounding box.

Face Mask Detection in CCTV camera

The stream to recognize the individual in the webcam wearing the face cover or not. The process is two-fold.
To distinguish the countenances in the webcam
Classify the faces based on the mask. Identify the Face in the Webcam.
To distinguish the countenances a pre-prepared model given by the OpenCV system was utilized. The model was trained using web images. OpenCV provides MTCNN which is a very powerful algorithm to detect faces in the photo. The face veil recognition model is a mix of face identification models to distinguish the current appearances from camera feeds and afterward running those countenances through a cover discovery model.

IV. CONCLUSION

As the technology is blooming with emerging trends availability so we have a novel face mask detector that can possibly contribute to public healthcare. The architecture consists of MobileNet and MTCNN as the backbone it can be used for high and low computation scenarios. In order to extract more robust features, we utilize transfer learning to adapt weights from a similar task face detection, which is trained on a very large dataset. The models were tried with pictures and continuous video transfers. This particular model could be utilized as a utilization case for edge investigation. Moreover, the proposed technique accomplishes cutting edge results on a public face cover dataset. By the development of face mask detection, we can detect if the person is wearing a face mask and allow their entry would be of great help to society.

V. FUTURE SCOPE

There are various perspectives we intend to deal with in a matter of seconds:

- The utilization of Machine Learning in the field of versatile organization is rising quickly. Henceforth,

we intend to port our models to their particular TensorFlow Lite renditions.

- Our engineering can be made viable with TensorFlow RunTime (TFRT), which will expand the surmising execution of anxious gadgets and make our models effective on multithreading CPUs.
- Stage 1 and Stage 2 models can be effectively supplanted with improved models, later on, that would give better exactness and lower inactivity.

REFERENCES

[1]. N. Zhang, J. Luo and W. Gao, "Research on Face Detection Technology Based on MTCNN," 2020 International Conference on Computer Network, Electronic and Automation (ICCNEA), Xi'an, China, 2020, pp. 154-158, doi: 10.1109/ICCNEA50255.2020.00040.

[2]. M. S. Ejaz and M. R. Islam, "Masked Face Recognition Using Convolutional Neural Network," 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 2019, pp. 1-6, doi: 10.1109/STI47673.2019.9068044.

[3]. A. K. Vani, R. N. Raajan, D. Haretha Winmalar. and R. Sudharsan, "Using the Keras Model for Accurate and Rapid Gender Identification through Detection of Facial Features," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2020, pp. 572-574, doi: 10.1109/ICCMC48092.2020.ICCMC-000106.

[4]. F. Saxen, P. Werner, S. Handrich, E. Othman, L. Dinges and A. Al-Hamadi, "Face Attribute Detection with MobileNetV2 and NasNet-Mobile," 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), Dubrovnik, Croatia, 2019, pp. 176-180, doi: 10.1109/ISPA.2019.8868585.

[5]. M. Ma and J. Wang, "Multi-View Face Detection and Landmark Localization Based on MTCNN," 2018 Chinese Automation Congress (CAC), Xi'an, China, 2018, pp. 4200-4205, doi: 10.1109/CAC.2018.8623535.

[6]. I. B. Venkateswarlu, J. Kakarla and S. Prakash, "Face mask detection using MobileNet and Global Pooling Block," 2020 IEEE 4th Conference on Information & Communication Technology (CICT), Chennai, India, 2020, pp. 1-5, doi: 10.1109/CICT51604.2020.9312083.

[7]. P. Viola and M. 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. CVPR 2001, Kauai, HI, USA, 2001, pp. I-I, doi: 10.1109/CVPR.2001.990517.

[8]. O. Cakiroglu, C. Ozer and B. Gunsel, "Design of a Deep Face Detector by Mask R-CNN," 2019 27th Signal Processing and Communications Applications Conference (SIU), Sivas, Turkey, 2019, pp. 1-4, doi: 10.1109/SIU.2019.8806447.

[9]. G. T. S. Draughon, P. Sun and J. P. Lynch, "Implementation of a Computer Vision Framework for Tracking and Visualizing Face Mask Usage in Urban Environments," 2020 IEEE International Smart Cities Conference (ISC2), Piscataway, NJ, USA, 2020, pp. 1-8, doi: 10.1109/ISC251055.2020.9239012.

[10]. A. S. Joshi, S. S. Joshi, G. Kanahasabai, R. Kapil and S. Gupta, "Deep Learning Framework to Detect Face Masks from Video Footage," 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), Bhimtal, India, 2020, pp. 435-440, doi: 10.1109/CICN49253.2020.9242625.

[11]. R. Moorthy H., V. Upadhya, V. V. Holla, S. S. Shetty and V. Tantry, "CNN based Smart Surveillance System: A Smart IoT Application Post Covid-19 Era," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2020, pp. 72-77, doi: 10.1109/I-SMAC49090.2020.9243576.