ECE3502-IOT DOMAIN ANALYST
CONTACTLESS ENTRY SYSTEM WITH FACEMASK DETECTOR IN SURVEILLANCE FOR COVID-19

MANSI NARADE(19BML0022), RIDITA GARG(19BML0004), SARNITHA G U(19BEC057)

Vellore Institute of Technology , Vellore Campus

***Abstract*—After the breakout of the worldwide pandemic COVID-19, there arises a severe need of protection mechanisms, face mask being the primary one. The basic aim of the project is to detect the presence of a face mask on human faces on live streaming video as well as on images. We have used MobileNetV2 architecture to develop our face detector model. The architecture is computationally efficient, making it easier to deploy. Alongside this, we have dataset from Kaggle. Experimental results show that our model performs well 99% training accuracy and 98% validation accuracy.**

***Keywords*— face detection, real time facemask detection, Python, OpenCV, Keras.**

I. Introduction

To smother transmission and save lives, masks should be utilized as a piece of a system of measures. All-inclusive mask use will impressively downsize infection transmission inside the local area by forestalling anybody and those who unwittingly convey the infection and communicate it to other people. Infection displaying suggests that masks worn by a considerable population could prompt significant decreases in impacted case numbers and passings[1]. Mask may also downsize the dangerous effect of the pandemic, eminently for those that live in crowded environments, any place social removing is intense, and for those that add front line roles wherever, there is a more significant gamble of openness to the infection[2]. Coronavirus infected people have fever which can easily be measured but one has to consider that there are asymptomatic infected people and incubation periods. These seemingly unaffected people are free to walk in society and possibly further the spread of the infection[3]. The main purpose of our project is to prevent the further escalation of the virus due to not wearing masks in places accessed by the general public. Our objective is to reduce this transmission of pathogens on high-touch surfaces, like door handles, and to prevent entry in a community area without a face mask.by making the process contactless and reduce crowding at common areas.[4][5]

II. Literature Study

Face mask detection algorithms have become more topical recently, since masks can help control the spread of COVID 19 during the pandemic. The algorithmic task focuses only on detecting physical masks, as shown in [6], [7], [9]. Among these, YOLO based models are the most popular detectors. ResNet based YOLOv2 was used by [8] to improve feature extraction for face mask detection. To enhance the robustness of detection by YOLOv3, an image mix-up and multi-scale method was utilized in [10]. A distance intersection over union non-maximum suppression (DIOU-NMS) algorithm was used to improve the post-processing stage of YOLOv3 [11]. YOLOv3 achieved the highest mAP in a comparison of YOLOv3, YOLOv3-tiny, SSD, and Faster R-CNN on the newly-established Moxa3K face mask detection dataset [12]. A person tracking system with a three-part face mask recognition system, a person detector, a tracker, and a face mask classifier, was developed to facilitate face mask detection applications in smart cities. Face mask classification or recognition, assuming faces were detected, has also been studied [13], [14].

In the paper[1] they proposed a novel SL-FMDet, which is efficient and has low hardware requirements. To overcome the lower feature extraction capability caused by its light-weight backbone, they proposed RCAM and SGHR. RCAM extracts highly contextual information and focus on face mask related areas. And by using SGHR as an auxiliary task, the model is able to learn more refined features for faces with and without masks. The model with SGHR yielded a better attention map, which qualitatively supports the efficiency of this auxiliary task. The proposed model achieved state-of-the-art results on two face mask datasets Experimentally, the paper has shown that light-weight models can achieve similar or even better performance than heavy models by using RCAM and SGHR[14]. The results also show the model is capable of overcome the challenges present in face mask detection. Therefore, the face mask detector proposed by this paper, has a high potential to contribute to public health care to control the spread of COVID-19. One disadvantage of this method is that it requires extra computation for generating heatmaps and, due to lack of enough datasets, the method cannot distinguish between correct and incorrect method of wearing mask[15].

New deep learning detectors may be used to further improve the performance. Recently, advanced work on anchor-free deep learning detectors, such as CenterNet or CornerNet has appeared[16]. Anchor-free detectors operate more like how human beings detect objects than anchor-based methods such as proposed method. CenterNet first detects the center of the objects, and then regresses the coordinates of corners relative to the centers[17]. Detection Transformer (DETR) a newly-proposed transformer-based deep learning detector borrows advantages from language transformers to use patch-based sequential information, and shows the method does not require post processing[18]. In addition, can develop a real-world face mask detection system on high performance edge devices, and integrate it with the internet of things systems.

III. PROPOSED METHOD

Assuming an area like a store, with an entry, we will be making a mask detection system that opens a door, without the use of your hands, and keep a live count of the number of people in the store to prevent crowding. Cameras will be placed at the entry gate of the store[14]. When a person is ready to enter the store, they will need to stand in front of the camera. If the customer is wearing a mask, the camera will detect it and grant them entry into the store. If the customer is found not wearing a mask and not following safety protocol, they will not be granted permission to enter the store, unless they wear a mask.

Our face mask detector is accurate, and since we used the MobileNetV2 architecture, it's also computationally efficient, making it easier to deploy the model to embedded systems (Raspberry Pi , Google coral etc.,). Detecting Face masks using Python , Keras and OpenCV on real video stream. We have used Face mask detection dataset from Kaggle.

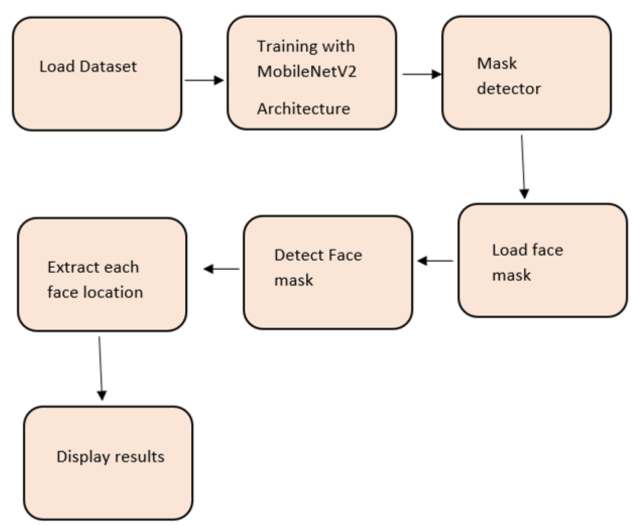

Fig 1 .Block Diagram

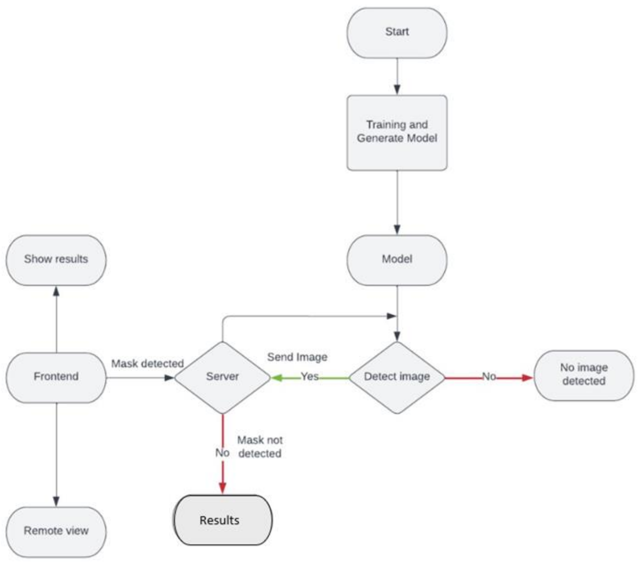

Fig 2. Flow Chart of Proposed Work

The design approach will be broken down into three main phases: Training, Deployment, and Alerts to train the face mask detector as shown in the block diagram [Fig 1]. Training will focus on loading the face mask detection dataset and then serializing the mask detector. The deployment will focus on executing actual detection of the mask and categorizing each face as with or without a mask. The Flow chart of the proposed method is given in the figure[Fig 2.]

A. *Training*

A dataset containing images will be used to train and test the code. 80% of the dataset is used for training the algorithm, and the remaining 20% is used for testing. In training, at first, we will be loading the complete dataset of images as a list. We will be using two arrays in which one of the arrays will contain the data (images), and the other one will contain the label (names) of the image. Then both arrays are changed into a NumPy array. Now, we split the data set into training and testing sets. We have to arrange every image in the dataset according to our conditions, so we will be using different syntax to vary the dimensions of every image now using the MobileNetV2 architecture, which was designed to enable hundreds or thousands of convolutional layers. We will focus on loading our face mask detection dataset from disk, training a model (using Keras/TensorFlow) on this dataset, and then serializing the face mask detector to disk. Using this MobileNetV2 we will create a base model used to train the algorithm. This base model predicts different images in the dataset, and the label used in the second array is matched with these notions. This will help in serializing the data and end the training part of the system.

B. *Deployment*

To test the application, we will use a Real-time video as our input. As a video is a combination of frames, we will be extracting the frames from the video in this test. For mask detection, every second a frame is extracted from the video, and the algorithm identifies the person in that frame by bounding the face of that person within a rectangular box. Using OpenCV, we are able to find the region of interest, that is the face of a person. The data extracted from a frame is converted into an array, which is then stored.[7] When the value of the array exceeds 0, the algorithm has identified the face of the person. Only when a face is detected, can the classifier detect a mask. According to the output, if the person is wearing a mask he will be bound with a green box. If not, he will be bound with a red box. Every image will be compared with a set of images in the dataset and the output is given according to the results. Thus, the value stored in the array will allow us to identify the face of the person in the image as soon as it appears.

C. *Alerts*

- The detection of the mask takes place, so if the person is wearing a mask, it is going to print 'Mask' along with its accuracy percentage.
- If the person is not wearing a mask, it is going to print 'No Mask' along with its accuracy percentage.

IV. Results And Disscussion

As shown in Fig.3 The training and validation accuracy rose until they reached their maximal peak at epoch 20, with 98% training accuracy and 95% validation accuracy.

The model's training and validation losses both reduced until they reached their lowest at epoch 20, with a validation loss of 0.30 and a training loss of less than 0.15. [Fig 3]

We demonstrate some qualitative results in Fig. 4 and Fig 5. The model can successfully distinguish some confusing occlusions, such as occlusion by hands, hair or other objects. and all diverse mask types were detected . Side views of faces with masks could be detected.

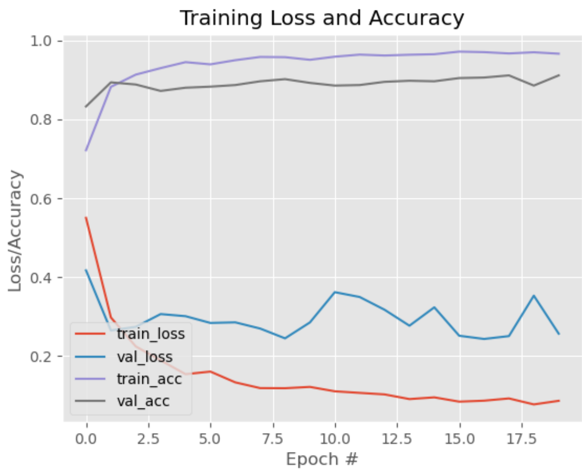

Fig 3. Training Loss and Accuracy

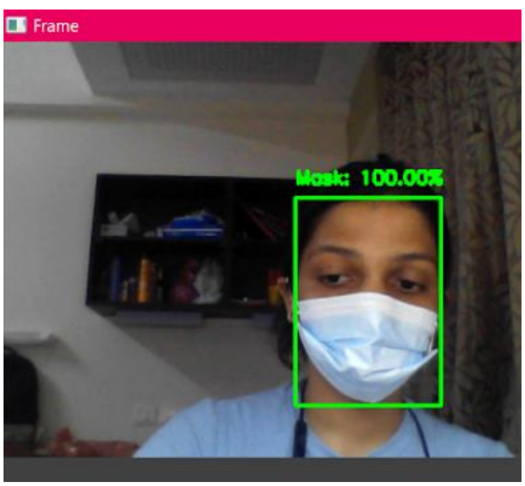

Fig 4 . Result of Mask Detected with Accuracy Percentage

The performance of our model on the Kaggle face mask dataset was compared with existing models used in face mask detection as shown in Table 1 . We compared our model with the best results reported by [1]. The light-weight MobileNetV2 model achieved better performance than heavy models like SSD.

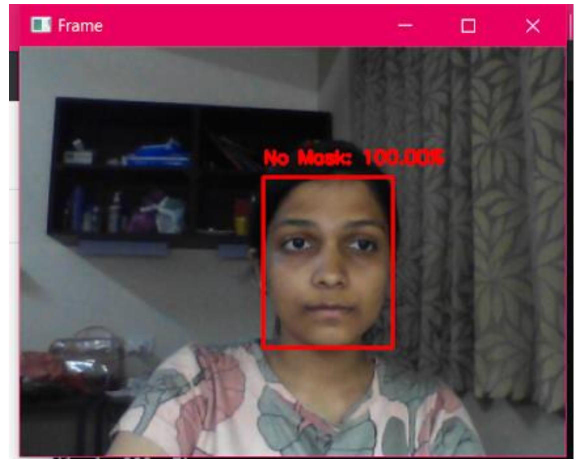

Fig 5. Result of No Mask Detected with Accuracy Percentage

Model	Dataset	Accuracy
SSD	AIZOO	89.6
YOLOv3	AIZZO	92.6
RetinaFace	AIZOO/ MOXA3k	92.8/53.7
RetinaFaceMask -M	AIZZO/ MOXA3k	93.6/54.21
MobileNetV2	Dataset from KAGGLE	91

Table 1 . Comparison with Other Models and Datasets

RetinaFace and RetinaFaceMask-M, and these models give 1-2% higher performance than MobileNetV2 model in terms of accuracy.[Table 1]. But MobileNetV2 can find most of the confusing occlusions, such as occlusion by hands, hair or other objects. and all diverse mask types in the wild. Although there are some failure cases, due to occlusions by people or objects, the result seems satisfactory.

V. CONCLUSION

As the technology are blooming with emerging trends the availability so we have novel face mask detector which can possibly contribute to public healthcare. The architecture consists of Mobile Net 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 adopt weights from a similar task face detection, which is trained on a very large dataset.[Table1]

We used OpenCV, keras, python to detect whether people were wearing face masks or not. The models were tested with images and real-time video streams. The accuracy of the model is achieved and, the optimization of the model is a continuous process[Fig4]. This specific model could be used as a use case for edge analytics. Furthermore, the proposed method achieves state-of-the-art results on a public face mask[Fig5]. This system can therefore be used in real-time applications which require face-mask detection for safety purposes due to the outbreak of Covid-19. This project can be integrated with embedded systems for application in airports, railway stations, offices, schools, and public places to ensure that public safety guidelines are followed.

This is a small step to ensure that people are protected in areas of mass crowding. The surveillance of face mask detection can be a leap in reducing the number of active cases of the virus in the long run.

References

[1] X. Fan, M. Jiang and H. Yan, "A Deep Learning Based Light-Weight Face Mask Detector With Residual Context Attention and Gaussian Heatmap to Fight Against COVID-19," in IEEE Access, vol. 9, pp. 96964-96974, 2021, doi: 10.1109/ACCESS.2021.3095191.

[2] M Inamdar, N Mehendale (2020). Real-Time Face Mask Identification Using Facemask Net Deep Learning Network. http://dx.doi.org/10.2139/ssrn.3663305

[3] M Loey, G Manogaran, MHN Taha, NEM Khalifa (2020). A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. Measurement, 2020 – Elsevier

[4] M Jiang, X Fan (2020). Retina Mask: A Face Mask detector. arXiv preprint arXiv:2005.03950

[5] S Peng, H Huang, W Chen, L Zhang, W Fang (2020). More trainable inception-ResNet for face recognition. Neurocomputing, 2020 – Elsevier

[6] CE Hwang, SH Lee, JW Jeong (2019). Web-based interactive IoT management with deep visual object detection. IRJET-2018

[7] A. S. Joshi, S. S. Joshi, G. Kanahasabai, R. Kapil, and S. Gupta, ''Deep learning framework to detect face masks from video footage,'' in Proc. 12th Int. Conf. Comput. Intell. Commun. Netw. (CICN), Sep. 2020, pp. 435–440

[8] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, ''Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection,'' Sustain. Cities Soc., vol. 65, Feb. 2021, Art. no. 102600

[9] M. Jiang, X. Fan, and H. Yan, ''RetinaMask: A face mask detector,'' 2020, arXiv:2005.03950. [Online]. Available: http://arxiv.org/abs/2005.03950

[10] M. Jiang, X. Fan, and H. Yan, ''RetinaMask: A face mask detector,'' 2020, arXiv:2005.03950. [Online]. Available: http://arxiv.org/abs/2005.03950.

[11] X. Ren and X. Liu, ''Mask wearing detection based on YOLOv3,'' J. Phys., Conf. Ser., vol. 1678, no. 1, pp. 1–6, 2020

[12] B. Roy, S. Nandy, D. Ghosh, D. Dutta, P. Biswas, and T. Das, ''MOXA: A deep learning based unmanned approach for real-time monitoring of people wearing medical masks,'' Trans. Indian Nat. Acad. Eng., vol. 5, no. 3, pp. 509–518, Sep. 2020.

[13] B. Roy, S. Nandy, D. Ghosh, D. Dutta, P. Biswas, and T. Das, ''MOXA: A deep learning based unmanned approach for real-time monitoring of people wearing medical masks,'' Trans. Indian Nat. Acad. Eng., vol. 5, no. 3, pp. 509–518, Sep. 2020.

[14] M. M. Rahman, M. M. H. Manik, M. M. Islam, S. Mahmud, and J.-H. Kim, ''An automated system to limit COVID-19 using facial mask detection in smart city network,'' in Proc. IEEE Int. IoT, Electron. Mechatronics Conf. (IEMTRONICS), Sep. 2020, pp. 1–5.

[15] J Chatterjee, T Debnath (2018). Environmental Monitoring Using Sense HAT based on IBM Watson IoT Platform. Journal of Sensor and Actuator Networks, 2019I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl,

[16] M. M. Rahman, M. M. H. Manik, M. M. Islam, S. Mahmud, and J.-H. Kim, ''An automated system to limit COVID-19 using facial mask detection in smart city network,'' in Proc. IEEE Int. IoT, Electron. Mechatronics Conf. (IEMTRONICS), Sep. 2020, pp. 1–5.

[17] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, and A. Desmaison, ''PyTorch: An imperative style, high-performance deep learning library,'' in Proc. Adv. Neural Inf. Process. Syst., 2019, pp. 8024–8035.

[18] R. Padilla, S. L. Netto, and E. A. B. da Silva, ''A survey on performance metrics for object-detection algorithms,'' in Proc. Int. Conf. Syst., Signals Image Process. (IWSSIP), Jul. 2020, pp. 237–242.