Face Mask Detection

(Sem 6 - Data Science Minor Project)

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Computer Science and Engineering Department

Thapar Institute of Engineering and Technology, Patiala

Submitted by: Chandan Taneja (101917048)

Submitted to: Dr. Sahil Sharma

Date of Submission: 18 May 2022

Group: CSE2(3rd year)

TABLE OF CONTENTS

1. Introduction of the Project

Spread of Covid-19 started from WUHAN, China in late December 2019 and the world is till fighting with this deadly disease. The World Health Organization (WHO) confirmed that this deadly virus could be transmitted to humans through droplets and airborne pathogens. According to data, by now 526 billion people got affected by this deadly virus all over the world. When it comes to prevention, wearing a face mask is important when going out or meeting others. With the outbreak of Covid-19 in 2019 and we humans are still unable to defeat this virus and that's all due to negligence of Government and common people who don't want to wear masks in spite of reminders, daily news, online campaigns and who can forget that **Caller Tune by Indian Government to weak mask and maintain social distancing.** Since manual monitoring of face masks is often infeasible in the crowd, automatic detection can be beneficial. Therefore, there was a need to develop some mechanisms to ensure that people wear a mask. As the numbers of infected patients per day can rise and there is risk of high alert. Even during the ongoing COVID-19 epidemic, there are no effective Rules, regularizations and automatic applications that are much needed in public transportation vehicles, densely populated areas, residential regions, major manufacturers, and other businesses to ensure safety.

Machine learning is a method of analyzing data using an analytical model that is built automatically, or 'learned', from training data. The idea is that the model gets better as you feed it more data points. Machine learning is important because it gives enterprises a view of trends in customer behavior and business operational patterns, as well as supports the development of new products. The absence of large datasets of Mask on images has made this task cumbersome and challenging.

This project uses a Deep Neural Network, more specifically a Convolutional Neural Network, to differentiate between images of people with and without masks. Labels in real time, using OpenCV.

This program can therefore be used for real-time applications that require face detection for security purposes due to the Covid-19 explosion. This project can be integrated with embedded application programs at airports, train stations, offices, schools and public places to ensure that public safety guidelines are followed.

2. Related Work

A Convolutional neural network (CNN) is a neural network that has one or more convolutional layers and are used mainly for image processing, classification, segmentation and also for other auto correlated data. Rather than looking at an entire image at once to find certain features it can be more effective to look at smaller portions of the image.

The most common use for CNNs is image classification, for example identifying satellite images that contain roads or classifying hand written letters and digits. There are other quite mainstream tasks such as image segmentation and signal processing, for which CNNs perform well at.

1. Decoding Facial Recognition

Facial recognition is broken down by a convolutional neural network into the following major components -

- Identifying every face in the picture
- Focusing on each face despite external factors, such as light, angle, pose, etc.
- Identifying unique features
- Comparing all the collected data with already existing data in the database to match a face with a name.
-

2. Analyzing Documents

Convolutional neural networks can also be used for document analysis. This is not just useful for handwriting analysis, but also has a major stake in recognizers. For a machine to be able to scan an individual's writing, and then compare that to the wide database it has, it must execute almost a million commands a minute. It is said with the use of CNNs and newer models and algorithms, the error rate has been brought down to a minimum of 0.4% at a character level, though it's complete testing is yet to be widely seen.

3. Historic and Environmental Collections

CNNs are also used for more complex purposes such as natural history collections. These collections act as key players in documenting major parts of history such as biodiversity, evolution, habitat loss, biological invasion, and climate change.

4. Understanding Climate

CNNs can be used to play a major role in the fight against climate change, especially in understanding the reasons why we see such drastic changes and how we could experiment in curbing the effect. It is said that the data in such natural history collections can also provide greater social and scientific insights, but this would require skilled human resources such as researchers who can physically visit these types of repositories.

5. Grey Areas

Introduction of the grey area into CNNs is posed to provide a much more realistic picture of the real world. Currently, CNNs largely function exactly like a machine, seeing a true and false value for every question. Allowing the machine to understand and process fuzzier logic will help it understand the grey area us humans live in and strive to work against. This will help CNNs get a more holistic view of what human sees.

6. Advertising

CNNs have already brought in a world of difference to advertising with the introduction of programmatic buying and data-driven personalized advertising.

7. Other Interesting Fields

CNNs are poised to be the future with their introduction into driverless cars, robots that can mimic human behavior, aides to human genome mapping projects, predicting earthquakes and natural disasters, and maybe even self-diagnoses of medical problems. So, you wouldn't even have to drive down to a clinic or schedule an appointment with a doctor to ensure your sneezing attack or high fever is just the simple flu and not symptoms of some rare disease. One problem that researchers are working on with CNNs is brain cancer detection. The earlier detection of brain cancer can prove to be a big step in saving more lives affected by this illness.

Table 1. Comparison of state-of-the-art

Citation	**Year**	**Paper Title**	**Methodology**	**Dataset Used**	**Evaluation**
[1]	2021	WearMask: Fast In-browser Face Mask Detection with Serverless Edge Computing for COVID-19	Converted the PyTorch model into an NCNN model with a model size of 581 KB. After compiling this C++ program into WASM format, the entire framework was executed as a function in JavaScript.	MAsked FAces (MAFA) dataset	AP (average precision)
[2]	2020	Masked Face Recognition with Latent Part Detection Feifei	Data augmentation is used to multiply images. Detect landmarks of the face image and locate the bounding box of the mask region. Then we affine the mask image to the mask region of the face image using the algorithm of Delaunay Triangulation.	MFV and MFI	Two standard evaluation metrics: rank scores (rank1, rank5 and rank10) obtained from Cumulated Matching Characteristics (CMC), and mean Average Precision (mAP).
[3]	2018	MobileNetV2: Inverted Residuals and Linear Bottlenecks Mark	Replace all the regular convolutions with separable convolutions (depthwise followed by 1 × 1 projection) in SSD prediction layers. The intermediate expansion layer uses lightweight depthwise convolutions to filter features as a source of non-linearity.	COCO dataset, PASCAL VOC 2012	Evaluation metric mIOU. (Intersection over union)

Table 2. Comparison of state-of-the-art: Pros &Cons

Citation	**Year**	**Pros**	**Cons**
[1]	2021	Compared our performance with the most similar settings among previous works, which (1) used real mask pictures instead of simulated mask pictures, (2) performed simultaneous object detection and classification. It is able to correctly identify the cases when a subject is not wearing the mask properly, such as when the nostrils or mouth are exposed Serverless edge-computing design, Easy deployments, Installation free. Low privacy risk are its other pros.	The existing dataset divides the no mask and the wearing masks incorrectly into one category Due to the system limitation in iOS, only Safari supports WebAssembly-related functions, Frames Per Second (FPS) is low on IOS It does not support parallel computing features such as SIMD (Single instruction, multiple data
[2]	2020	Outperforms Cosface, which is a typical model for common face recognition, by 50.58% (87.12-36.54) (rank1), 28.59% (49.08-20.49) (mAP) and 6.9% (94.34-87.44) (accuracy) on MFI. A novel latent part detection (LPD) model is proposed to locate the latent facial part which is robust to mask wearing	Proposed method focuses on a novel task of masked face recognition (MFR) where most facial cues are occluded by mask.
[3]	2018	Basic building block is a bottleneck depth-separable convolution with residuals. It allows very memory-efficient inference and relies utilize standard operations present in all neural frameworks. The proposed convolutional block has a unique property that allows to separate the network expressiveness (encoded by expansion layers) from its capacity (encoded by bottleneck inputs). Compared the needed sizes for each resolution between MobileNetV1, MobileNetV2 and ShuffleNe	Didn’t compare performance with other architectures such as Faster-RCNN and RFCN.

3. Methodology

The digitalization in almost every field has made our lives easy, for instance nowadays deep learning models like CNN, YOLO, etc. are used in object detection, handwriting detection, Car speed detection, Facial-Hand Gesture Recognition and many more.

With the outbreak of Covid-19 in 2019 and we humans are still unable to defeat this virus and that's all due to negligence of Government and common people who don't want to wear masks. Hence, we thought of developing a model that would intake live footage of public places and would detect people who are not wearing masks and would be able to alert the security guards (front-line workers) with a beep.

Our Mask Detector Model comprises of **MobileNet v2** which is a Convolutional Neural Network that is 53 layers deep. It is an architecture of bottleneck depth-separable convolution building of basic blocks with residuals. MobileNetV2 is very similar to the original MobileNet, except that it uses inverted residual blocks with bottlenecking features. It has a drastically lower parameter count than the original MobileNet.

It has two types of blocks. The first one is a residual block with stride of 1. Second one is also residual block with stride 2 and it is for downsizing.

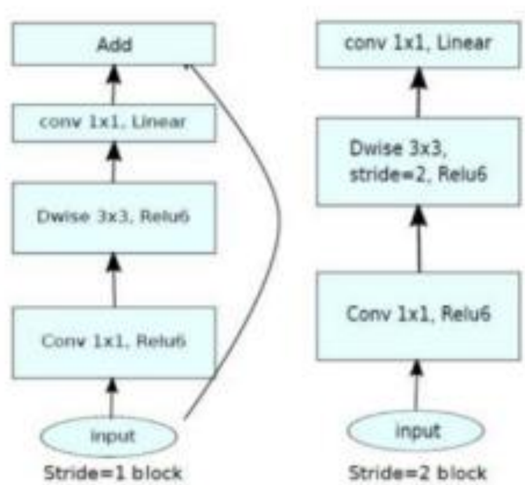


Fig. MobileNet V2

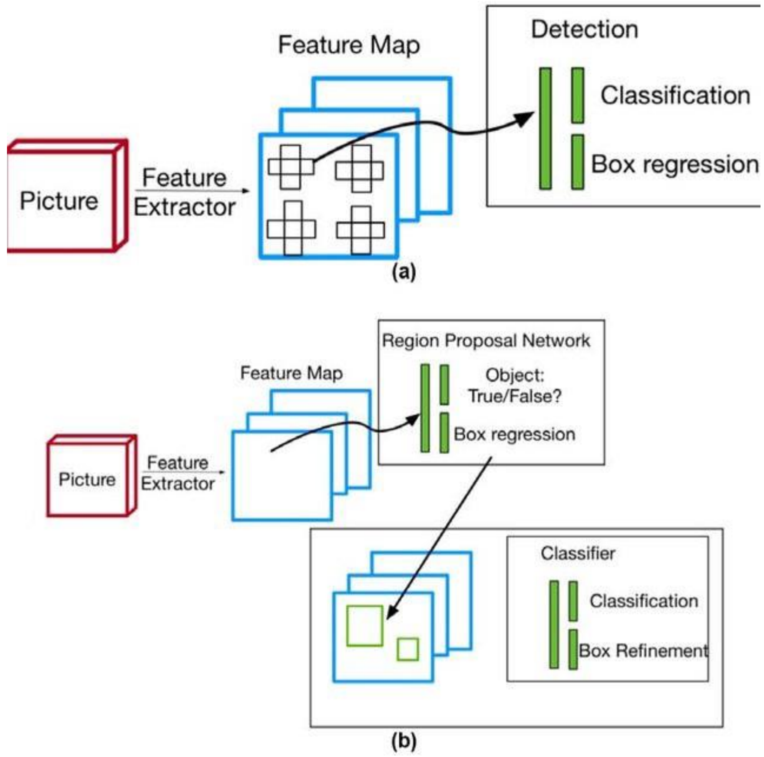


Fig. Object Detection using MobileNet V2

To increase the dataset an augmented image generator is created using ImageDataGenerator API in Keras to generates batches of image data with real-time data augmentation. And then dataset is preprocessed and labels are given to Mask & No-Mask. Dataset is then trained using MobileNetV2 on the ROI (Region of Interest) and hyper-parameterized using Adam to give output label as “MASK” and “No-Mask”.

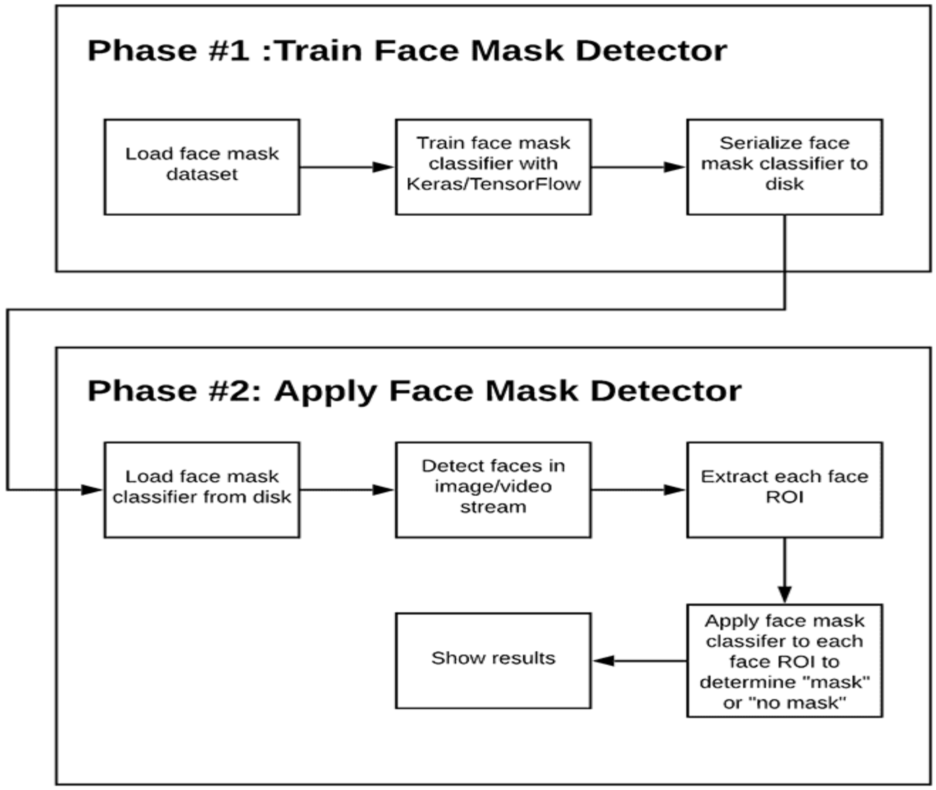


Fig. Working Flow

4. Proposed Architecture

Due to increase in covid-19 infected patients and continuous threat of high alert, prevention and safety comes. But most people would not wear a mask, which has caused the virus to spread widely. Since manual monitoring of face masks is often infeasible in the middle of the crowd, automatic detection can be beneficial. Therefore, there was a need to develop some mechanisms to ensure that people wear a mask.

In this model we will implement MobileNet v2 which is 53 layers deep because it enhances both the speed and accuracy of the model due to integration with embedded system.

It's a real-time software, which can be used in public places and feed the live camera footage. It will be able to detect multiple faces and would be able to label as "**MASK**" and "**NO MASK**". A **BEEP SOUND** for a short period of time (0.5 ms) will be heard when it will detect a person with no mask and continue detecting people in the live video.

The software would be able to

- Beep would alert front-line workers and other people in the particular area.
- People would feel the necessity of wearing masks after hearing the beep sound in public.
- Would surely decrease the spread of covid-19
- No need of Manual monitoring will be required
- Can further implementations could be made (discuss in Conclusions and Future Scope)

5. Experiments and Results

MobileNet V2 model is used as it enhances the speed and accuracy of the project. It is a type of Convolutional Neural Network (CNN) with 53 neural layers.

The data was fed into the Neural-Network with the following train, test and validation splits - 80:20 train and test, the validation was 20% of the train.

1) **Hyperparameter Tuning**

Learning Rate - Learning rate of 0.0001 is used for this model. (Tried 0.001,0.00001 but that worsened the accuracy)

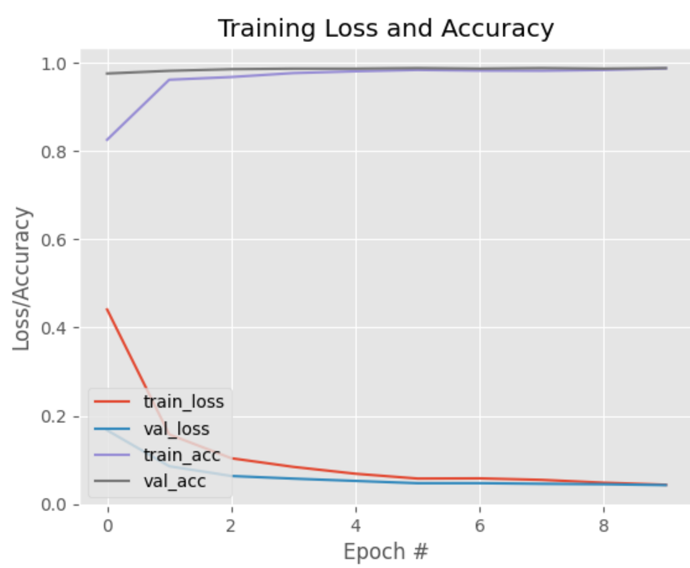
Batch Size - Batch size of 42.

Optimizer - Adam

Epochs - 10 epochs (also tested for 20, 25 epochs)

2) **Model Accuracy**

The performance of the model was judged based on accuracy and loss values for train, validation and holdout set. The loss on train and validation sets for twenty epochs is given below



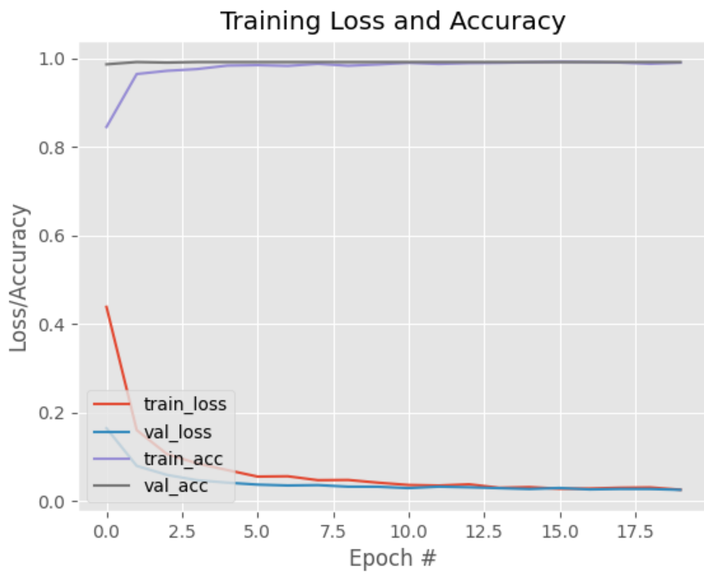


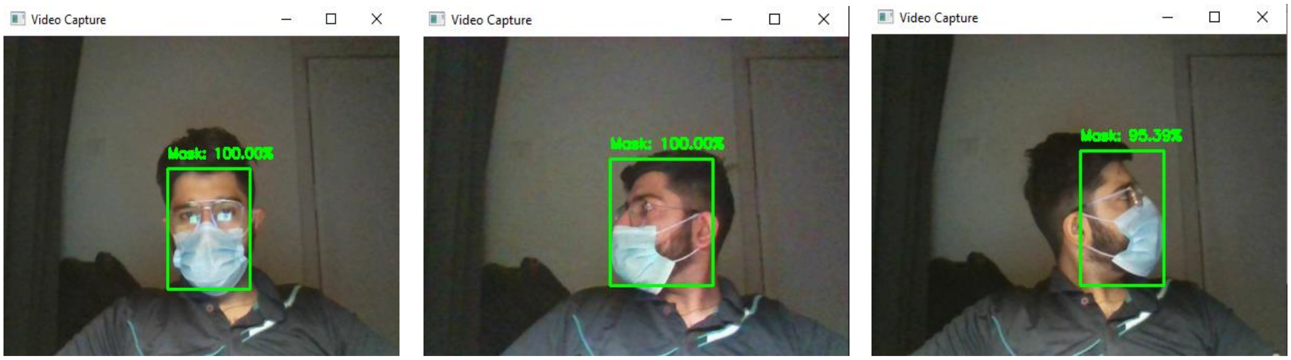
Fig. Loss and Accuracy w.r.t #epochs

Model trained for both 10 &20 epochs and does well on both train and validation loss. Loss, Accuracy, validation_loss, Validation_accuracy per epochs are as follows:

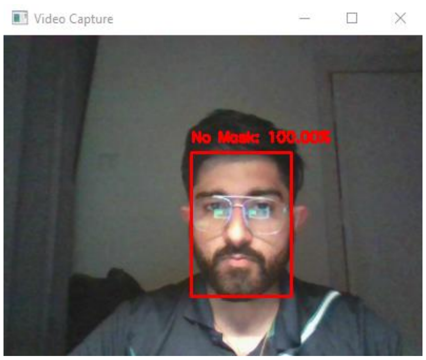
```
Anaconda Prompt (anaconda3)
Training the head
Epoch 1/10
2022-05-14 21:03:40.228971: W tensorflow/core/framework/cpu_allocator_impl.cc:82] Allocation of 51380224 exceeds 10% of free system memory.
2022-05-14 21:03:40.324810: W tensorflow/core/framework/cpu_allocator_impl.cc:82] Allocation of 51380224 exceeds 10% of free system memory.
2022-05-14 21:03:40.372110: W tensorflow/core/framework/cpu_allocator_impl.cc:82] Allocation of 51380224 exceeds 10% of free system memory.
2022-05-14 21:03:40.429065: W tensorflow/core/framework/cpu_allocator_impl.cc:82] Allocation of 154140672 exceeds 10% of free system memory.
2022-05-14 21:03:40.495906: W tensorflow/core/framework/cpu_allocator_impl.cc:82] Allocation of 156905472 exceeds 10% of free system memory.
102/102 [==============================] - 138s 1s/step - loss: 0.4409 - accuracy: 0.8255 - val_loss: 0.1672 - val_accuracy: 0.9756
Epoch 2/10
102/102 [==============================] - 136s 1s/step - loss: 0.1578 - accuracy: 0.9615 - val_loss: 0.0858 - val_accuracy: 0.9817
Epoch 3/10
102/102 [==============================] - 134s 1s/step - loss: 0.1039 - accuracy: 0.9676 - val_loss: 0.0636 - val_accuracy: 0.9853
Epoch 4/10
102/102 [==============================] - 134s 1s/step - loss: 0.0843 - accuracy: 0.9766 - val_loss: 0.0577 - val_accuracy: 0.9866
Epoch 5/10
102/102 [==============================] - 135s 1s/step - loss: 0.0687 - accuracy: 0.9806 - val_loss: 0.0524 - val_accuracy: 0.9866
Epoch 6/10
102/102 [==============================] - 134s 1s/step - loss: 0.0578 - accuracy: 0.9837 - val_loss: 0.0473 - val_accuracy: 0.9878
Epoch 7/10
102/102 [==============================] - 135s 1s/step - loss: 0.0582 - accuracy: 0.9821 - val_loss: 0.0473 - val_accuracy: 0.9866
Epoch 8/10
102/102 [==============================] - 134s 1s/step - loss: 0.0548 - accuracy: 0.9818 - val_loss: 0.0459 - val_accuracy: 0.9878
Epoch 9/10
102/102 [==============================] - 134s 1s/step - loss: 0.0484 - accuracy: 0.9837 - val_loss: 0.0450 - val_accuracy: 0.9866
Epoch 10/10
102/102 [==============================] - 133s 1s/step - loss: 0.0437 - accuracy: 0.9871 - val_loss: 0.0430 - val_accuracy: 0.9878
Evaluation of the model network
              precision    recall  f1-score   support

   with_mask       0.99      0.99      0.99       433
without_mask       0.99      0.98      0.99       386

    accuracy                           0.99       819
   macro avg       0.99      0.99      0.99       819
weighted avg       0.99      0.99      0.99       819

[INFO] saving mask detector model...

(faceo) C:\Users\lenovo\Desktop\Face-Mask-Detection-master>
```

Fig. Precision, Recall, f1-score used as Evaluation Matrices.

6. Screenshot of Working Model

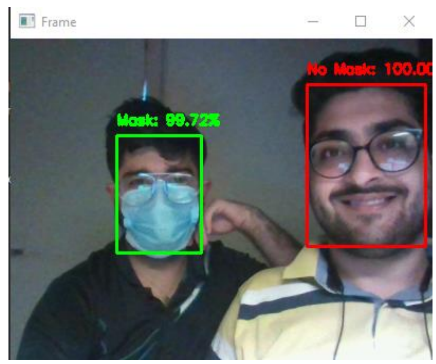


Mask Detection and labelled as "Mask" with prediction percentage



No mask detection

(With a beep sound)



Mask detection with multiple people

(With a beep sound)

Software Screen Recording & Google Drive link:

https://bit.ly/Mask_detection

7. Conclusion and Future Scope

In this work, a deep learning-based approach for detecting masks over faces in public places to curtail the community spread of Coronavirus is presented. The proposed technique efficiently handles occlusions in dense situations by making use of an ensemble of single and two-stage detectors at the pre-processing level. The ensemble approach not only helps in achieving high accuracy but also imp Augmented Images using Contrast Stretching (left), Histogram Equalization (middle) and Adaptive Histogram Equalization (right) roves detection speed considerably. Furthermore, the application of transfer learning on pre-trained models with extensive experimentation over an unbiased dataset resulted in a highly robust and low-cost system. The identity detection of faces, violating the mask norms further, increases the utility of the system for public benefits.

Both Biometric detection using face recognition and Mask detection can be used by Government for online E-Challan for not wearing masks at public places.

Finally, the work opens interesting future directions for researchers. Firstly, the proposed technique can be integrated into any high-resolution video surveillance devices and not limited to mask detection only. Secondly, the model can be extended to detect facial landmarks with a facemask for biometric purposes.

8. References

[1] Z. Wang, P. Wang, P. C. Louis, L. E. Wheless, and Y. Huo, "WearMask: Fast In-browser Face Mask Detection with Serverless Edge Computing for COVID-19," no. December, pp. 1–8, 2021, [Online]. Available: http://arxiv.org/abs/2101.00784

[2] F. Ding, P. Peng, Y. Huang, M. Geng, and Y. Tian, "Masked Face Recognition with Latent Part Detection," *MM 2020 - Proc. 28th ACM Int. Conf. Multimed.*, pp. 2281–2289, 2020, doi: 10.1145/3394171.3413731.

[3] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.