Face Mask Detection and Face Recognition Using Machine Learning

Prof. Jitendra Musale[1], Taher Travadi[2], Tabrez Shaikh[3], Shubham Thorve[4], Rutuja Kutwal[5]

[1,2,3,4,5] *Department of Computer Engineering, Anantrao Pawar College of Engineering and Research, Pune, India*

Abstract - *Face Recognition is a computer application that is capable of detecting, tracking, identifying, or verifying human faces from an image or video captured using a digital camera. The human face is a sophisticated multidimensional structure that shuttles a bunch of information including expression, feelings, facial features, etc. This research paper presents advanced image processing techniques. The growth in face mask detection algorithms in recent times has compelled humankind to come up with solutions to tackle the pandemic situation. We have designed a solution with a system that uses numerous algorithms and methods to identify unmasked individuals using several mask detections models and such people can be penalized for violating rules. The proposed system in this research adopts frameworks of OpenCV, Keras, TensorFlow, and deep learning libraries to detect face masks in real-time. To detect masks, initially, the model is trained using the Multi-Task Cascade Convolutional Neural Network (MTCNN) algorithm. The face is first detected using the OpenCV library and facial imagery is stored and passed as an input to the mask detection classifier for classification. For facial detection, the Viola-Jones algorithm is used which is also known as the Haar Cascade algorithm. Our experiment results show that our method is very accurate, reliable, and robust for a face recognition system that can be practically implemented in a real-life environment as a COVID-19 monitoring system.*

Key Words: Mask Detection, Face Recognition, Machine Learning, TensorFlow, OpenCV, Python, etc.

1. INTRODUCTION

In today's scenario, where the whole world has been under lockdown for more than a period of 1 and half years due to the worldwide spread of the deadly COVID-19 virus. All the institutes and organizations (Educational as well as Corporates) have been working remotely for this whole period from the safety of their home. The only weapon we had against this virus was Hand-Sanitizers and masks using these we have survived through these times. Now as the Antivirus Vaccines have been produced, we can see a decrease in the number of infections gradually and people have started to come into the organizations physically. But still, we know that even after taking shots of vaccine it is mandatory to follow the Covid norms of social distancing and wearing masks. The organizations have the responsibility to check whether the employees or students in the case of educational organizations are following these rules while entering the premises of the organization. Manually checking whether the person is wearing a mask or not is a hectic task and requires a lot of manpower. Here comes the use of our system of face mask detection and imposing a fine on the person who will violate the rules of wearing masks properly. Such a system can be installed at the entrance of the organization.

Computer vision is one of the emerging frameworks in the field of object detection and is widely being used in various aspects of research in artificial intelligence. There have been both supervised and unsupervised approaches to machine learning in the past for object detection in an image. An enhanced supervised machine learning technique of object detection has been deployed in this project. To detect masks, initially, a model is trained using CNN algorithms with datasets of faces provided. Face masks are first detected using OpenCV (Open-Source Computer Vision) and those image frames are stored and passed to the mask detection classifier for classification.

For facial detection, the Viola-Jones algorithm is used which is also known as the Haar-Cascade algorithm. We have used a mask classifier model to differentiate faces with masks and without masks. If a person without a mask is detected their face is recognized using face-recognition models and compared with existing databases and a corresponding fine is imposed on that person.

2. Proposed System

In the proposed system we are going to build a Face Recognition and Face-Mask detection ML model integrated with UI. detected without a mask, the name/id of that person is identified using face recognition and a corresponding fine will be charged to that person by the concerned authority. Using suitable UI the information of registered users and their corresponding fine can be seen. This data will help to control the misbehavior of people regarding wearing masks. The system is divided into the following parts:

i. Face Mask Detection

ii. Face Recognition

iii. Storage of data to impose fine

iv. Graphical User Interface (GUI)

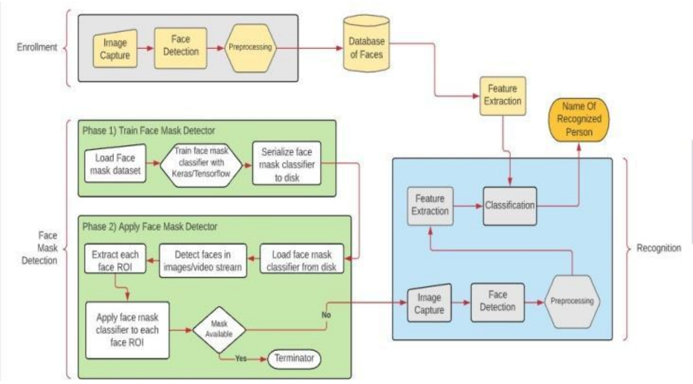


Figure 1. Proposed Architecture

3. Methodology

The System consists of various steps that are Image Capture, Face Mask Detection, Face Recognition, Comparison of Image, and Fine Implementation these can be achieved by ML techniques. The People in the organizations after being registered on the system their face image data is collected and stored in the database. At the entrance of organizations people entering in will be monitored using any suitable camera and if people without a mask are detected by a face mask detection model such people will be recognized by the face recognition model by matching their faces with an already available and fine will be charged for them by face identification Proposed model involves a two-stage process for detecting person wearing a face mask or not using webcam

i) Identifying faces from the input frame.

ii) Detecting mask on the recognized face region and classifying accordingly.

Dataset has been divided into the training and the testing datasets. To train efficiently and effectively, we have considered 80% of total images as a training dataset and 20% of total images as testing data sets to test the prediction accuracy. For simplicity, the images in our training data collection are classified into two categories "with mask" and "without mask". used a lightweight image classifier, MobileNetV2, which gives high accuracy and is well suited for mobile devices. In pre-processing steps, resize the image to 224 × 224 pixels to maintain consistency. The trained model is deployed using OpenCV. The model is applied to the real-time video frames captured using a webcam to detect people in the frame wearing a mask or not. For face recognition, following processes are involved:

i) Face Detection,

ii) Facial Feature Extraction from captured images,

iii) Classification & matching of images based on a database of face images.

Transferred learning using pre-trained convolution neural networks as feature extractors has been applied to face recognition. Two public well-known face databases will be used for face recognition of faces in this system. CNNs are supervised machine learning techniques that can extract deep knowledge from a dataset through rigorous example-based training. This machine learning approach mimics the human brain when learning. CNNs have been successfully applied to feature extraction, face recognition, classification, segmentation, and examination centers, where accuracy and precision are highly desired to serve the purpose. It can be used in smart city innovation, and it would boost the development process in many developing countries. Our framework presents a chance to be more ready for the next crisis or to evaluate the effects of the huge scope of social change in respecting sanitary protection rules.

4. Mathematical Model

A mathematical model is a description of a system using mathematical concepts and language. A model may help to explain a system and to study the effects of different components of a system to predict the behavior of a system. The mathematical modeling for our system is as follows

S= {∑, F, δ, C}

S = Face Recognition.

∑ = set of input symbols = {Video File, image, character information}

F = set of output symbol = {Match Found then notify to user, Not Found}

δ =

1. Start

2. Read training set of N*N images

3. Resize image dimensions to N2*1

4. Select training set of N2*M Dimensions, M: number of sample images

5. Find the average face, subtract from the faces in the training set, and create matrix A

$$\Psi = \frac{1}{M} \sum_{i=1}^{M} \Gamma i$$

Where, Ψ= average image,

M= number of images, and

Γi= image vector.

Φi = Γi – Ψ

Where, i = 1, 2, 3, …, M.

A = [Φ1, Φ2, Φ3…ΦM]

6. Calculate covariance matrix: AA'

7. Calculate eigenvectors of the c covariance matrix.

8. Calculate eigenfaces = No. of training images –no. of classes (total number of people) of eigenvectors.

9. Create reduced eigenface space. The selected set of eigenvectors is multiplied by the matrix to create a reduced eigenface

10. Calculate the eigenface of the image in question.

11. Calculate Euclidean distances between the image and the eigenfaces.

12. Find the minimum Euclidean distance.

13. Output: image with the minimum Euclidean distance or image unrecognizable.

C = {The system will not process the audio data, Eigenfaces will generate the grayscale images, and the algorithm will run only on keyframes.}

5. Experiment and Results

5.1 FACE RECOGNITION BASED ON LBPH ALGORITHM

In this research, identification of an individual for biometric monitoring has been made using Local Binary Patterns Histogram Algorithm. This monitoring system identifies a person from their front-facing and side-facing images and also functions well with low-quality pictures. The complete monitoring system follows a structure of steps from collecting the imagery to loading of the video camera. If the video camera is running, facial detection is initiated, then the processing of facial imagery is executed and special features of the face are extracted and identified. Using these extracted facial features, the model makes its recognition. The researchers have implemented this research using multiple datasets of facial images. Local Binary Patterns Histogram is considered to be one of the best suiting algorithms for face recognition.

Sample Images for face-recognition

5.2 FACE RECOGNITION BASED ON CONVOLUTION NEURAL NETWORK

CNN is primarily designed for multi-dimensional data processing and the main advantage is its automatic face-feature extraction, which is then forwarded to the classifier network. In this paper, We have tried to use the CNN algorithm for better human face recognition. The convolution Neural Net has been implemented to extract facial expressions and other features in a connected layer also the softmax classifier in use. A database of 50 images each of a person was tested for accuracy and by using this algorithm we found that this one of the finest face recognition algorithms but has some accuracy issues during dull lighting.

Sample images from all datasets: (a) sample images from ORL dataset; (b) sample images from GTAV face dataset; (c) sample images from Georgia Tech face dataset; (d) sample images from FEI dataset; (e) sample images from LFW dataset; (f) sample images from F_LFW dataset; (g) sample images from YTF dataset.

5.3 LBPH BASED FACE RECOGNITION AT LOW RESOLUTION

LBPH is used for face-recognition and it is known for its ability to recognize the face of a person from the front side as well as with the side face. In addition with LBPH, we have also used haar cascade classifier for extracting the face features like eyes,nose, smile, frontal face and side view of the face. The entire process consists of 3 steps as Collection of the image using a camera input stream followed by image detection and finally face recognition. In this paper, we have used the LR500 dataset which was designed to test LBPH. For detection purposes, haar classifier has used the adaboost algorithm.

After the extraction of features the training process is completed and the images stored are then saved in yml format. For the accuracy and reliability, LBPH algorithm was tested with the dataset and we have found the valid results but at some point the results were inaccurate.

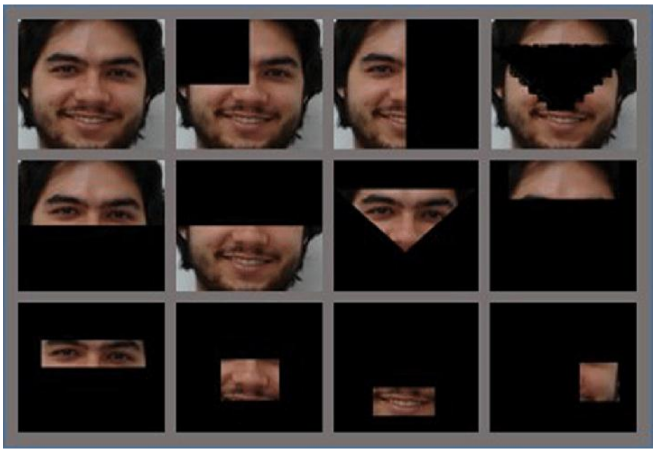
ACCURACY OF ALGORITHMS FOR DIFFERENT DATASETS

METHOD	ORL	FEI	DATASET CREATED
LBPH	97%	70.5%	80.3%
CNN	95.4%	94%	91%

5.4 Experiments on parts of the face using the FEI dataset

In this research, we have divided the face into various parts using FEI datasets. We generated twelve testing sets where each test corresponds to the part of the face. These parts were classified as forehead, eyes, mouth, nose, cheeks etc. Faces with eyes and nose, lower half of the face, upper half of the face, right and left half of the face, three quartered face and also full facial images were generated.

After generating images and extracting features from the images, the data was fed to VGGF model and then CS-Wo parts and linear SVM-Wo parts classifiers to check the rating of recognition for individual facial part separately. The results are summarized as; rate of recognition is highest on three quarters and full facial imagery with a rating of 100%. Later, It was noted that, for right half and top half facial images the rating falls down slightly with SVM-Wo classifier but with CS-Wo classifier, a rating of 100% is still maintained. For the bottom half of the face, the rating of recognition drops down to 50% with the SVM-Wo classifier and for CS-Wo classifier, a rating of 60% is obtained.



Parts of the face we have used for testing the recognition rates on the FEI dataset.

Classifier	Part of the face	Rate of Recognition (Accuracy)
CS-Wo	Three quarters and full facial imagery	98%
SVM-Wo		97.4%
CS-Wo	Right half and Top Half	98%
SVM-Wo		<98%
CS-Wo	Bottom Half	60%
SVM-Wo		50%

5.5 Experiments on rotated faces using the FEI dataset

In this experimental test, all the images were rotated by an increment of eighteen degrees, starting from ten degrees to one hundred and eighty degrees. All the rotated images were given as an input to VGGF model to extract facial features and then to the two classifiers mentioned above (CS-Wo & SVM-Wo). This experiment was carried out using both cases; with rotated and without rotated images.

It was quite adequate to obtain an accuracy of 100% with faces rotated at 10 and 20 degrees. At a rotation of about 30 to 40 degrees, the recognition rate fell down to 98% and 80% respectively for both classifiers. The worst rate of recognition was obtained at a rotation from 50 degrees and above. Resulting in almost 0% for both classifiers. In this experiment, the CS-W classifier recorded the highest accuracy in recognition rating in most of the cases. It was noted that as the rotation kept increasing, the recognition rate kept decreasing for SVM-W classifier while the CS-W classifier maintained its rate of recognition even at higher degree rotations. This indicates the CS-W is a better image classifier than SVM-W.

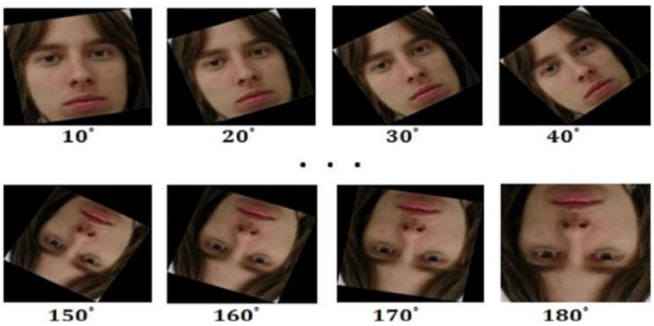


Illustration of face rotation (10°to 180°) on the FEI dataset.

Classifier	Rotation Level	Rate of Recognition
CS-Wo	10 degrees to 20 degrees	98%
SVM-Wo		98%
CS-Wo	30 degrees to 40 degrees	98% to 80%
SVM-Wo		98% to 80%
CS-Wo	> 50 degrees	~ 0%
SVM-Wo		~ 0%

5.6 Experiments on rotated faces using the FEI dataset

In this experimental test, the facial images were zoomed out at 10% to 90% to find the effects of zooming an image on the results of rating of face recognition. As experimented above, the zoomed out facial imagery was passed as an input to the VGGF model to extract unique facial features and then pass it to the above mentioned classifiers (CS and SVM).

As the first step, we evaluated the rate of recognition without zoomed images and trained the classifiers. It was noted that at 10% to 50% zoom, a rating of 100% was obtained using the SVM-Wo classifier while on the other hand, after 70% zoom the recognition rate dropped down significantly which eventually reached 0% as the zoom kept increasing. In contrary, while using the CS-Wo classifier, lower recognition rates were produced between 10% to 50% zoom when compared to SVM-Wo classifier. We noticed that as we added zoomed out facial images to the training dataset, the SVM-W classifier displayed an improvement in the rate of recognition at a zoom level of 40% and 50%. At 70%, 80% and 90% zoom levels, a steady rate of recognition was observed. In the case of the CS-W classifier, a recognition rate of 100% was obtained between 10% to 60% zoom level.

An example of zooming out (10% to 90%) of faces on the FEI dataset.

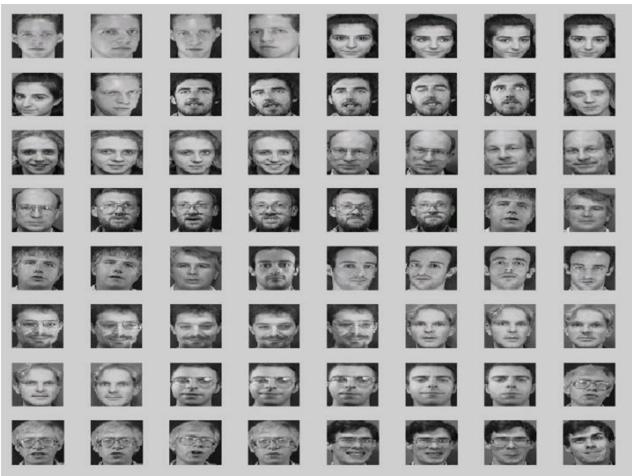
Classifier	Zoom Level	Rate of Recognition
CS-Wo	10% to 50%	< 10%
SVM-Wo		98%
CS-Wo	> 70%	~ 10%
SVM-Wo		< 10% to ~ 0%
CS-Wo	> 50%	~ 0%
SVM-Wo		~ 0%
CS-W	10% to 60%	~ 98%

6. DATASET DESCRIPTION

Here we are using three different generalized datasets namely ORL, FEI and LFW datasets.

6.1 ORL dataset:

The ORL database of faces is a collection of 10 different images of 40 individual persons which were collected between 1992 and 1994 at the laboratory. This dataset has images which were taken in bright to dark lightning areas. For some subjects the images were taken at different times also with varying facial expressions such as open/closed eyes, smiling face/not smiling with some facial details as glasses/no glasses. We have transformed these images into ORL.npz format which is given by numpy which stores these in an array format. Some plugins from image.io in the format .npz were also used.

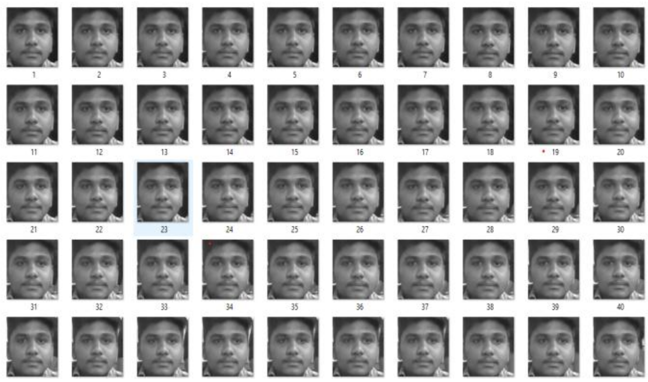


Sample face data from the ORL dataset.

6.2 FEI dataset:

The FEI face database is a Brazilian face database that has faces of 200 students with equal no. of males and females. This dataset was prepared by FEI university. For each person there are 14 images bringing the total to 2800 images in a dataset. All images are coloured and with a white homogenous background. The subjects are between the age group of 19 to 40 years old. These images have variations in facial expressions as well as with Posture. We have used 50% of the dataset for the training purpose and the other 50% for the testing purpose. We have also used our own database where we have 10 final year students and there are 100 images of each individual that sums to 1000 images. These images were taken with different lighting exposures and facial expressions like smile, seriousness, etc.

Sample face data from the FEI dataset.



6.3 The LFW dataset

Labeled faces in the wild (LFW) is a large dataset of face pictures designed for variable testing and studying the capability of face recognition algorithms in simulating uncontrolled scenarios. There are a total 13,233 images of 5749 people collected from the internet, all these images were detected and centered by viola jones face detector with the faces having large variations in expression, pose, illumination as well as resolution. Most of the subjects around 4070 person's have just a single image. The original database contained roughly four different sets of LFW images and also three different types of 'aligned' images. Also images have significant background clutter. We also had to do pre- processing of the images to extract the exact face dimensions from the original image.

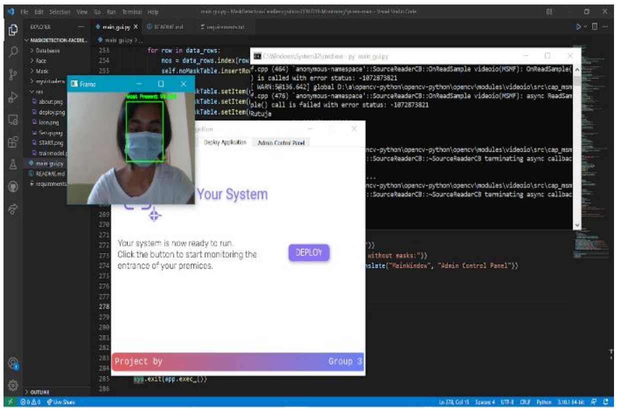
Sample face data from the LFW dataset.

7. EXPERIMENTAL OUTPUT

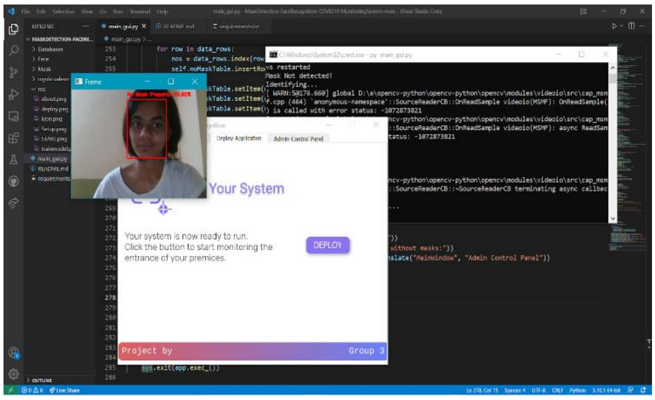
These experimental outputs were carried out on an Intel Core i5 8th gen laptop with 8GB RAM and 2GB integrated graphics. We made use of Google Collaboratory for testing of our face detection models. The datasets were taken from Kaggle which were classified as with and without masks. Some of the dataset was also prepared by us within our college students. This dataset was used to train the face detection model and the model was tested live using a laptop video camera. Our algorithm pre-processes the images and extracts unique facial features for making an accurate identification of faces.

The GUI of this desktop application was built using PyQT and the TKinter library. Our ML model was built using TensorFlow. The average accuracy of our ML model resulted in more than 90%. Calculation of Accuracy and Precision was measured for positive predicted values. We made a combination of models and classifiers with facial blending techniques to extract unique facial features even in low lighting conditions and low-resolution images. The accuracy of algorithms that we tested are mentioned below. Finally, after multiple testing, our face detection model was successfully able to detect faces against masked and unmasked faces.

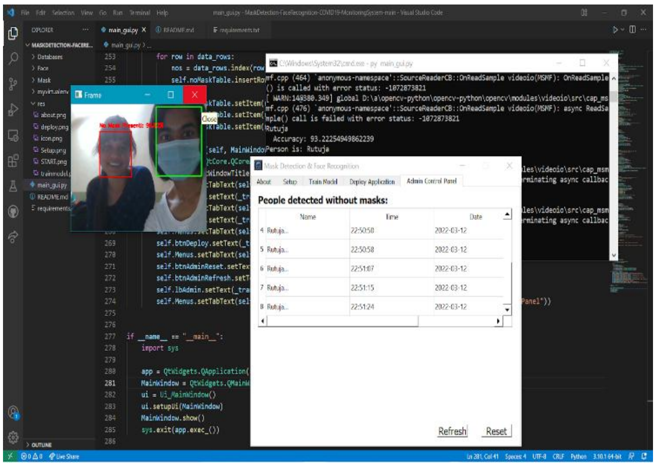
Dataset	***Classification Report***		
	Precision	***Recall***	***F1-score***
Without Mask	1.00	0.99	0.99
With Mask	0.99	1.00	0.99



Picture 1. Person With Mask Detected
Description: When the system comes across a person, it tries to identify them. Since, in the demo, the person is wearing a mask, the face mask model identifies that the person is wearing a mask and ignores their entry in the defaulter's database.



Picture 2. Person Without Mask Detected
Description: Similarly, if a person not wearing a mask is detected, then the face recognition algorithm is triggered to mark their entry into the defaulter's table for penalties.



Picture 3. Name of person without mask
Description: Using GUI, the admin can find the list of people detected without a mask which can be further used to verify the legitimacy of the system.

Conclusion

This project is focused on providing a better solution for regulating the use of a mask by the people using face mask detection and imposing a fine on the particular person who will violate the rule by their face recognition. This system will automate the hectic task of checking the mask manually thus reducing the efforts and increasing the time efficiency and accuracy. The system can work very efficiently using the power of Machine Learning techniques and algorithms like TensorFlow, Keras, CNN, OpenCV, etc.

The Computer Vision library was found to be a pioneer in the field of Image and Video Processing with the help of this system can be made very useful and robust in features. This embedded vision-based application can be used in any working environment such as public place, station, corporate environment, streets, shopping malls, and examination centers, where accuracy and precision are highly desired to serve the purpose. It can be used in smart city innovation, and it would boost the development process in many developing countries. Our framework presents a chance to be more ready for the next crisis or to evaluate the effects of the huge scope of social change in respecting sanitary protection rules.

REFERENCES

[1] Face Mask Detection System, Joby K. J.1, Priyanga K K2 | International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET) | ||Volume 10, Issue 6, June 2021|| DOI:10.15680/IJIRSET.2021.1006071

[2] Face MaskRecognition using TensorFlow a Deep Learning Framework | Bhagyashree Kulkarni1, Dr. Suma.V2 PG Scholar1, Professor2 Department of Information Science Dayananda Sagar College of Engineering, Bangalore, India | ISSN 2321 3361 © 2021 IJESC | Volume 11 Issue No.05

[3] CNN-BASED MASK DETECTION SYSTEM USING OPENCV AND MOBILENETV2 | Harriet Christa.G , Anisha.K , Jessica J, K Martin | 2021 3rd International Conference on Signal Processing and Communication (ICPSC) | 13 – 14 May 2021 | Coimbatore

[4] Face Recognition using Transferred Deep Learning for Feature Extraction | Amornpan Phornchaicharoen, Praisan Padungweang | The 4th International Conference on Digital Arts, Media and Technology and 2nd ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering

[5] Deep Convolutional Neural Network-Based Approaches for Face Recognition | Soad Almabdy 1, * and Lamiaa Elrefaei 1,2 | Appl. Sci. 2019, 9(20), 4397; https://doi.org/10.3390/app9204397

[6] Deep face recognition using imperfect facial data | Ali Elmahmudi, Hassan Ugail * | Future Generation Computer Systems Volume 99, October 2019, Pages 213-225

[7] Real-Time Detection and Recognition of Human Faces | Agnihotram Venkata Sripriya, Mungi Geethika, Vaddi Radhesyam| Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020) IEEE Xplore Part Number: CFP20K74-ART; ISBN: 978-1-7281-4876-2

[8] Image processing effects on the deep face recognition system | Jinhua Zeng1, *, Xiulian Qiu2, and Shaopei Shi1 | Mathematical Biosciences and Engineering Volume 18, Issue 2, 1187–1200 | 13 January 2021