FACE MASK AND SOCIAL DISTANCE DETECTION IN CCTV VIDEO STREAMS USING AI AND COMPUTER VISION

Senthil Kumar V[1], Joycema J[2], Benisha Joan M[3], Aathish Vicram J[4]

[1]*Assistant Professor, Department of Computer Science and Engineering, Kumaraguru College of Technology [anonymous], Coimbatore, Tamilnadu, India*
[2-4] *Department of Computer Science and Engineering, Kumaraguru College of Technology [anonymous], Coimbatore, Tamilnadu, India*

---***---

Abstract - *According to the data obtained by the World Health Organization, COVID-19, the global pandemic has severely impacted the world and infected more than hundreds of million people worldwide which includes more than three million deaths. This global pandemic enforced governments across the world to impose lockdowns in order to prevent the virus transmissions. Reports indicate that wearing face masks and following safe social distancing are two of the enhanced safety protocols need to be followed in public places in order to prevent the spread of the virus. To ensure the public safety in environment, we propose an efficient Computer Vision based approach that focused on the real-time automated monitoring of people to detect both safe social distancing and face masks in public places by implementing the model to monitor the activity and detect violations through camera. After detection of breach, the system sends to control center at state police headquarters and also give alarm to public. In this proposed system, modern deep neural network based model have been mixed with geometric techniques for building a robust model which covers three aspects of detection, tracking, and validation. Thus, the proposed system helps the society in saving time and reducing the spread of corona virus. It could be practiced effectively in current situation when lockdown is eased to inspect persons in public gatherings, shopping malls, etc. Automated inspection reduces manpower to inspect the public and can be used in any place to ensure safety.*

Key Words: Deep Learning, Computer Vision, Deep Neural Networks, World Health Organisation, YOLO, Face Mask, Social Distance, public Safety.

1. INTRODUCTION

The spread of COVID-19 has created the most crucial global health crisis all over the world which has had a deep impact on the way we perceive our world and our everyday lives. In December 2019, the spread of severe acute respiratory syndrome corona virus 2 (SARS- CoV-2), emerged in Wuhan, China, and has infected 7,711 people and 170 reported deaths in China before coronavirus was declared as a global pandemic, was named by the World Health Organization as COVID-19 (coronavirus disease 2019). This has resulted in person-to-person transmission but so far as we know, the transmission of the novel corona virus causing COVID-19 can also be from an asymptomatic carrier with no symptoms. Until now there is no report on any clinically approved antiviral medicine or vaccines which treats COVID-19. WHO recommends that people should wear face masks to avoid the risk of virus transmission and also recommends that a social distance of at least 2m be maintained between individuals to prevent the person-to-person spread of disease since virus spreads rapidly across the world, bringing massive health, economic, environmental and social challenges to the entire human population.

1.1 CONCEPTUAL STUDY OF THE PROJECT

Our proposed model describes an approach to prevent the spread of the virus by monitoring whether a person is following safe social distancing and wearing face masks in public places in real time. Our approach adopts the combination of lightweight neural network MobileNetV2 and Single Shot Detector (SSD) with transfer learning technique to achieve the balance of resource limitations and recognition accuracy so that it can be used on real-time video surveillance to monitor public places to detect if persons wearing the face mask and maintaining safe social distancing using YOLO object detection on video footage and images in real time. The experimental results infer that the detection of masked faces and human subjects based on YOLO has stronger robustness and faster detection speed.

Our solution uses deep neural networking models to analyze video streams using Open CV and MobileNetV2.We mix the approach of modern-day deep learning and classic projective geometry techniques which not only helps to meet the real-time requirements but also keeps high prediction accuracy.

1.2 OBJECTIVES OF THE PROJECT

Our project aims in proposing an efficient AI and Computer Vision based approach focused on real-time automated monitoring of people which can detect faces in real-world videos and perform tasks namely identify if the detected faces are wearing masks or not and monitor if proper social distancing measures are maintained.

1.3 SCOPE OF THE PROJECT

As future scope of the project, Temperature Screening can be performed in this approach since body temperature is another key symptom of COVID-19 infection. In the present scenario, thermal screening in done using handheld contactless IR thermometers where health-care workers need to come on close proximity with the person who needs to be screened. This makes the health-care workers vulnerable to get infected and requires a lot of manpower to capture for each person. This proposed use case could be equipped with thermal cameras for screening body temperature of the people in public places that can add another helping hand for enforcement agencies to tackle the pandemic effectively.

2. PROBLEM DEFINITION

According to the World Health Organization (WHO), COVID-19 pandemic is causing the major global health crisis and so the effective protection methods is wearing a face mask and maintaining social distance in public areas. Governments across the world are forced to impose lockdowns to prevent virus transmissions. Our goal is to automate the identification of whether the person on image/video stream is wearing a face mask or not and monitor the social distancing with the help of computer vision and deep learning.

3. PROPOSED SYSTEM

Our main aim is to propose an efficient AI and computer vision based approach focused on the real-time automated monitoring of people that can detect faces in real-world videos and identify if the detected faces are wearing masks or not monitor if proper social distancing measures is maintained .The proposed model ensures the automatic monitoring of the public whether they wore masks and follow social distancing norms.

The process starts from

i. Capturing of video streams through webcam/recorder
ii. Divides image/videos into frames
iii. Gives as an input to the trained model
iv. Detects the risk factors in violating both the condition or any of the condition (thus recording violations).

3.1. METHODOLOGY

In our project, the proposed model ensures the automatic monitoring of public whether they wore masks and follow social distancing norms.

The main contribution of the proposed system is three components:

i. Person detection
ii. Face mask detection
iii. Safe distance measurement between detected persons

Real-time person detection is done with the help, transfer learning is used on the top of high performing pre-trained SSD model for face detection with MobileNetV2 architecture as the backbones to create a lightweight model that is accurate and computationally efficient which will be easier to deploy the model. Even if Single Shot Detector could detect multiple objects in a frame, it is limited to the detection of a single person in this system.

The custom face crop dataset used comprises of about 3165 images annotated in mask and no mask. Annotated images are used to train a deep learning binary classification model which classifies the input image into the mask and no mask categories using the output class confidence. The result of SSD model extracts a person's mask and displays a bounding box. The proposed system monitors public places continuously and when a person without mask is detected it records the violations.

In social distance detector, we will input the video/ image and apply object detection (using YOLO) which filters to detect only people.YOLOv4 is extremely fast, easy to train, robust, stable and gives promising results even for tiny objects, hence, we selected it as our object detector of choice. This effectively means that the same model is used for both person detection to track social distancing and for masked-face detection for facemask monitoring. This significantly boosts overall efficiency and simplicity significantly. Thus finding the total people and coordinates to check the distance between their centroids using Euclidean distance formula. Hence will check whether they are apart from each other (about 100px) or close to each other (distance greater than 100px). If the people are far away (lesser than 100px) from each other then will categories as "following" and otherwise will categories as "not following". Thus, we could see the final results.

3.2. FLOW DIAGRAM

3.2.1. FACE MASK DETECTOR

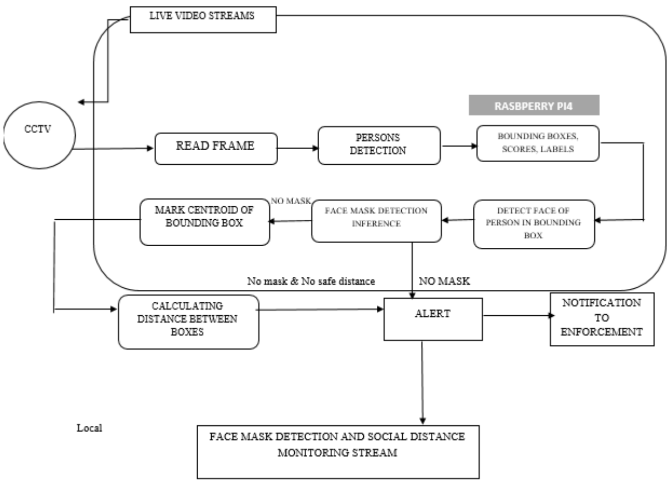


Chart-1: Face mask and Social distance detection flow chart

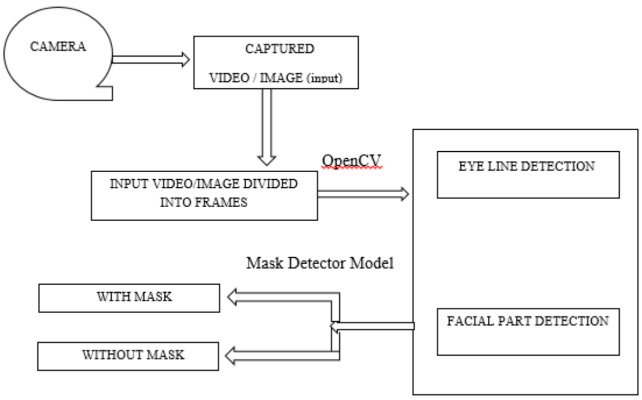


Chart -1.1.a: Face mask detection flow chart

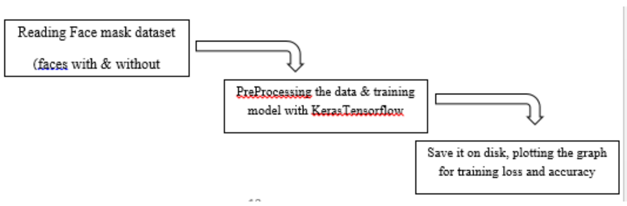


Chart -1.1.b: Training Face mask detector

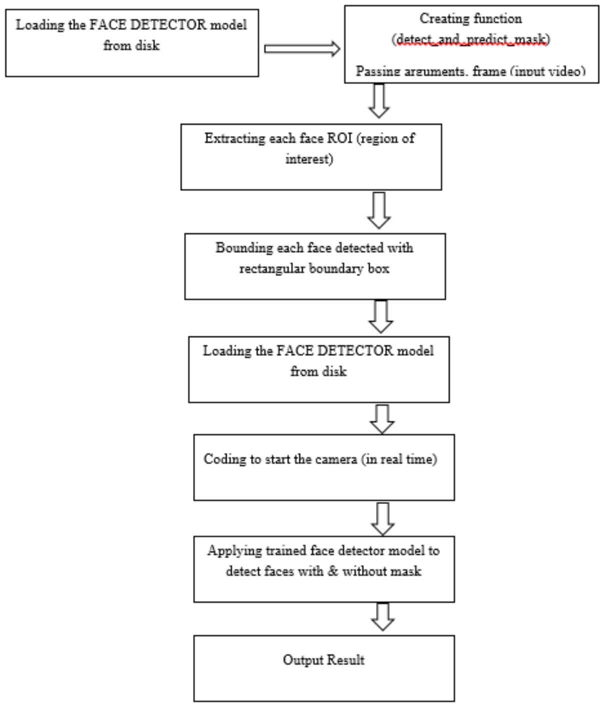


Chart -1.1.c: Applying Face mask detector model

3.2.2. SOCIAL DISTANCE DETECTOR

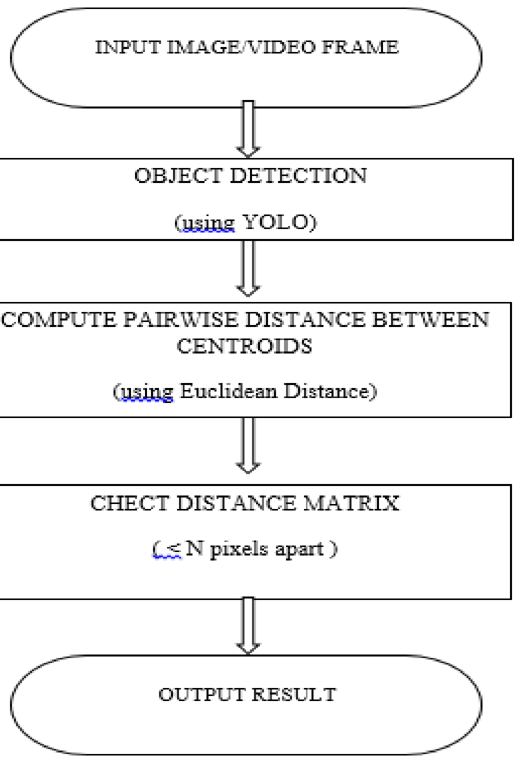


Chart-2: Social distance detection flow chart

3.3. IMPLEMENTATION

3.3.1. INSTALLING THE REQUIREMENTS

a) *TensorFlow* - Python library for fast numerical computing created and released by Google. It is a foundation library that can be used to create Deep Learning models directly or by using wrapper libraries that simplify the process built on top of TensorFlow.

b) *Keras* - Python library for developing and evaluating deep learning models. It wraps the efficient numerical computation library TensorFlow and allows you to define and train neural network models in just a few lines of code

c) *Imutils* - a series of convenience functions to make basic image processing functions such as translation, rotation, resizing, skeletonization, and displaying Matplotlib images easier with OpenCV

d) *NumPy* - python library used for working with arrays. It also has functions for working in domain of linear algebra, fourier transform, and matrices.

e) *OpenCV* – open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products.

f) *MatplotLib* - Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy.

g) *SciPy* - SciPy uses NumPy arrays as the basic data structure, and comes with modules for various commonly used tasks in scientific programming, including linear algebra, integration (calculus), ordinary differential equation solving, and signal processing.

3.3.2. DATA PREPROCESSING

The dataset consists of 3165 images to train our proposed face mask detector. Before the custom face mask image dataset is labelled and divided into the training set and testing set. The training dataset should consist of 80% images inorder to train the algorithm effectively. Testing dataset should consist of 20% images to test the prediction accuracy of the algorithm. The images in the training data collections are classified into two categories: with mask and without mask.

In face mask detection, the following process takes place

i. Resizing Images – processing images to fit in the same size
ii. Converting Images to Array
iii. Pre-Processing Images for MobileNetV2
iv. One-Hot Encoding using Label Binarizer
v. Using NumPy library to process them in the form of Arrays
vi. Splitting datasets into Train and Test

In social distance detection, the following process takes place.

i. Filtering the person class from detection
ii. Get Bounding box centroid for each person detection
iii. Calculate Euclidean distance between centroids
iv. Checking person boundary box close to each other
v. Creating Green bounding boxes and Red bounding boxes
vi. Risk analysis and recording the risk factors

3.3.3. DATA TRAINING

80% of our total custom dataset is used to train the model with a single shot detector which takes only one shot to detect multiple objects present in an image using a multi-box. The custom dataset is loaded into the project directory and the algorithm is trained based on the labelled images.

i. Constructing the training image generator for data augmentation
ii. Constructing head of the model that will be placed on top of the base model
iii. Placing head FC model on top of the base model (training model)
iv. Loop over all layers in the base model and freeze them
v. Compile and train the head of the model

In pre-processing steps, the image is resized to 224*224 pixels, converted to NumPy array format and the corresponding labels are added to images in the dataset before using SSD model as input to build our custom model with MobileNetV2 as the backbone to train our model using TensorFlow Object Detection API.

Before model training begins, TensorFlow helps in Data augmentation and downloads pre-trained ImageNet weights to make the algorithm's prediction accurate. After downloading the pre-trained weights and creating new fully-connected head (FC), the SSD algorithm is trained with both pre-trained ImageNet weights and annotated images in the custom dataset by tuning the head layer weights without updating weights of base layers. We trained our model for 1000 steps using the Adam optimization algorithm, the learning decay rate for updating network weights and the binary cross-entropy for mask type classification. We used webcam video for social distance monitoring and when a person has been identified (object detection using YOLO), start with bounding box coordinates and computing the midpoint between top-left coordinates and bottom-left along with the top-right and bottom-right points. We calculate the Euclidean distance between the centroid points to determine the distance between the people in future. We will check whether they are apart from each other or close to each other using social distance threshold value.

3.3.4. MODEL PREDICTION

The proposed model detects all persons in the range of cameras and shows a green bounding box around each person who is far from each other and red bounding box around people who are close to each other simultaneously. Face mask detection is also achieved by showing bounding boxes on the identified person's faces with mask or without mask labelled with confidence scores. If face mask is not visible or if social distance is not preserved, then the system analyze the risks and records the risk indicators. This system detects social distancing with a precision score of 91.7%, with confidence score of 0.7, precision value 0.91 and the recall value 0.91 with FPS = 28.07.

4. MODEL TESTING

Our work on face mask and social distance detection comprises of data collection to tackle the variance in the kinds of face masks worn by the workers. Face mask detection model is a combination of face detection models to identify the existing faces from camera feeds and then running those faces through a mask detection model Thus, social distance detection records the risks indicating by human violations from camera feeds or webcam video and then running the results.

5. IMPLEMENTATION RESULTS
5.1 FACE MASK DETECTION

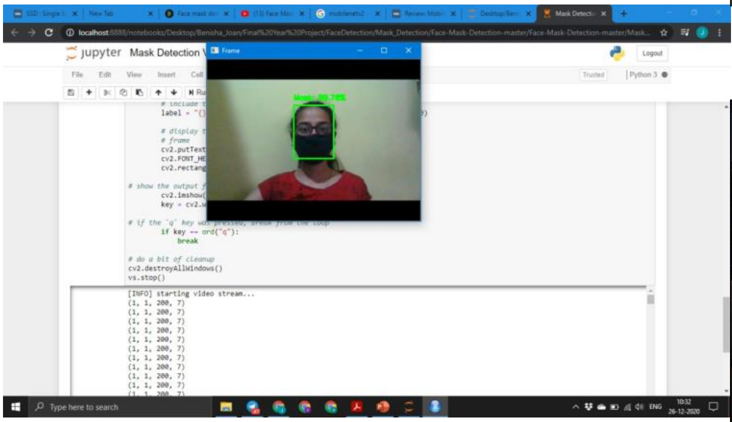


Fig-1: Face Mask Detection Output (with mask)

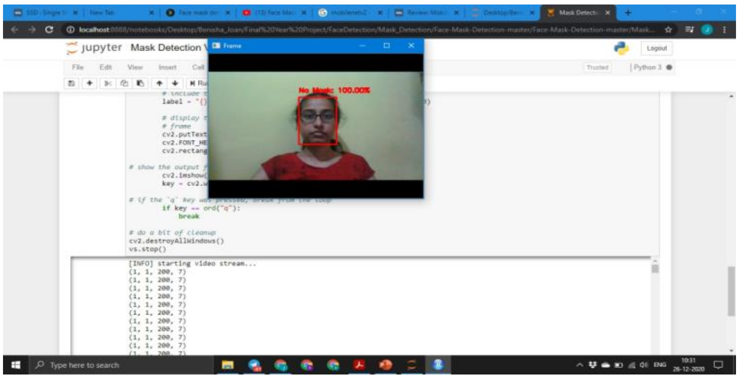


Fig-2: Face Mask Detection Output (without mask)

5.2 SOCIAL DISTANCE DETECTION

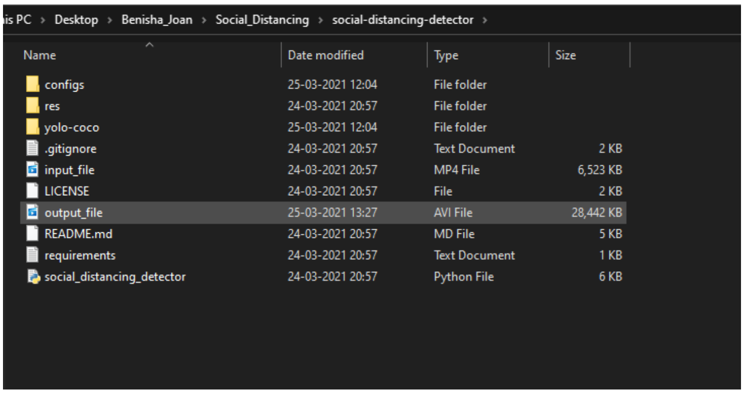
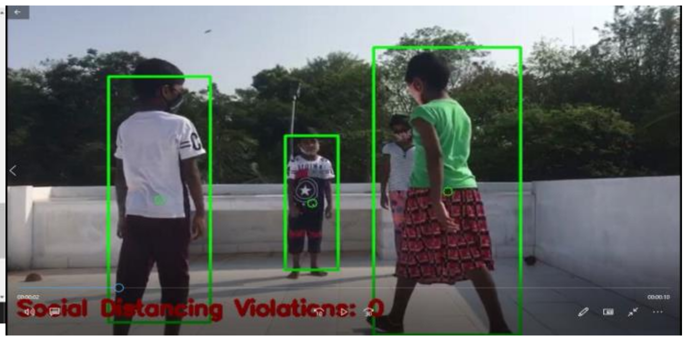


Fig-3: Social Distance Detection showing the violations (webcam videos)



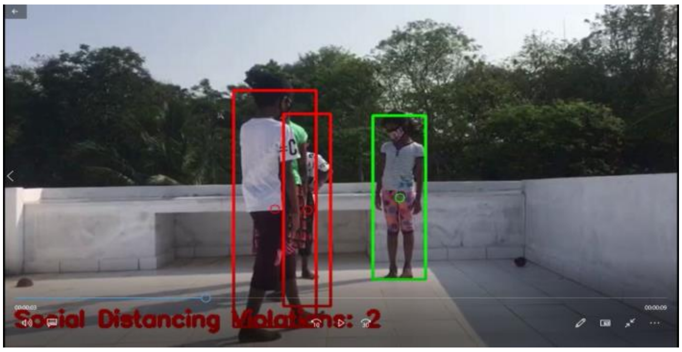
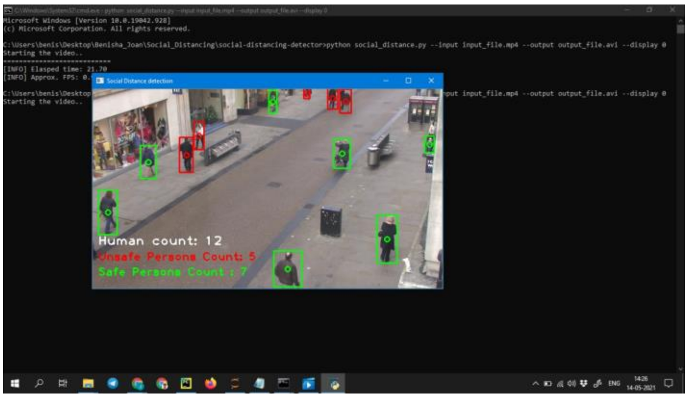


Fig-4 & 5: Social Distance Detection Output (mobile camera videos)



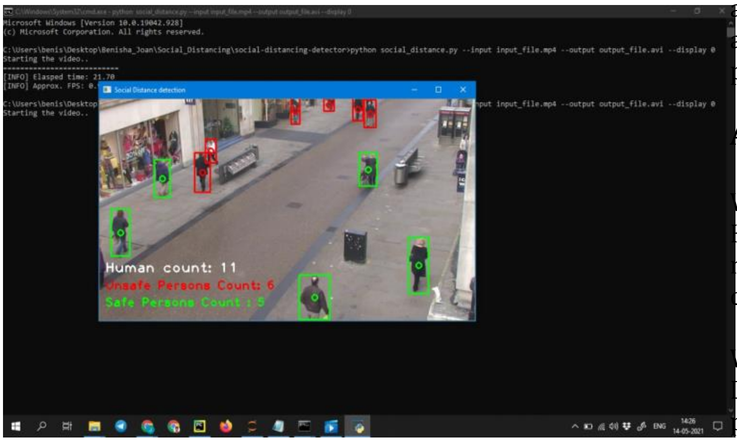


Fig-6 & 7: Social Distance Detection output with Human counts, Safe and Unsafe Human Counts

6. CONCLUSION

In this paper, we proposed an approach that uses computer vision and YOLO to help maintain a secure environment and ensure individuals' protection by automatically monitoring public places to avoid the spread of the COVID-19 virus. On the available datasets, some face detectors have achieved extremely high performances and it seems to be somehow difficult to further improve them. However, the current scenarios are much more challenging than expected for containing faces captured at unexpected resolution, illumination and occlusion.

In particular, the detection of masked faces is a critical undertaking. In this developed a deep learning model for face mask detector, we have implemented mask detection using Python, Keras, and OpenCV. We have trained the model. Training the model was the first part of this project and testing using webcam was the second part. We proposed a Deep Neural Network-Based social distance detector model to detect and track static and dynamic people in public places in order to monitor social distancing metrics in COVID-19 era and beyond. In this system each individual is identified with the help of bounding boxes. The generated bounding boxes aid in identifying groups of people satisfying the closeness property computed with the help of pairwise approach. The number of violations are getting confirmed.

The YOLO technology was evaluated using large and comprehensive datasets and proved a major development in terms of accuracy and speed compared to three state-of-the-art techniques. The extensive trials were conducted with popular object detection model YOLO v4 which illustrated the efficient performance. This applications can be used to analyse for mask detection and social distancing in a public area and perform important moves to higher address the pandemic. Automating the task will lead in effective moves taken in short time hence equipping us better to address the situation. Applications are applicable in various environments using CCTV surveillance cameras. This two applications are very useful in many areas like Hot-spot areas, in offices, colleges, hospitals, airports, railway stations, public places like banks, ATM, Government offices, etc.

ACKNOWLEDGEMENT

We express our profound gratitude to the management of Kumaraguru College of Technology for providing us with the required infrastructure that enabled us to successfully complete the project.

We express our profound gratefulness to our Principal, Dr.J.Srinivasan, for providing us necessary facilities to pursue the project.

We would like to acknowledge Dr.P.Devaki, Professor and Head of the Department, Computer Science and Engineering, for her support and encouragement throughout this project.

We thank out project coordinator Dr.Latha, Professor, Department of Computer Science and Engineering and guide Mr.V.Senthil Kumar, Assistant Professor, Department of Computer Science and Engineering, for their constant and continuous effort, guidance and valuable time.

Our sincere hearty thanks to staff members of Computer Science and Engineering of Kumaraguru College of Technology for their good wishes, timely help and support rendered to us during our project.

REFERENCES

[1] Glass RJ, Glass LM, Beyeler WE, Min HJ. Targeted social distancing architecture for pandemic influenza. Emerging Infectious Diseases. 2006; 12:1671–1681

[2]] S. Ge, J. Li, Q. Ye and Z. Luo, "Detection of Masked Faces in the Wild with LLE-CNNs," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 426-434.

[3] S. S. Mohamed, N. M. Tahir and R. Adnan, "Background modelling and background subtraction efficiency for object detection," 2010 6th International Colloquium on Signal Processing & its Applications, Mallaca City, 2010, pp. 1-6, doi: 10.1109/CSPA.2010.5545291.

[4] S. S. Farfade, M. J. Saberian, and L. Li. Multi-view face recognition using deep convolutional neural networks. In ACM ICMR, pages 643– 650, 2015

[5] C.Fu, W.Liu, A.Ranga, A. Tyagi, A. Berg, "DSSD: deconvolutional single shot detector model," arXiv preprint arXiv:1701.06659, (2017)

[6] Lin, Tsung-Yi, PiotrDollár, Ross Girshick, Kaiming He, BharathHariharan, and Serge Belongie, "Type Pyramid Networks for Object Detection," IEEE Conference Proceedings on Computer Vision and Pattern Recognition, pp. 2117-2125. 2017.

[7] Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H. Wu, P. Yi, K. Jiang, N. Wang, Y. Peiet al., "Masked face recognition dataset and application," arXiv preprint arXiv:2003.09093, 2020.

[8] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review,"IEEE transactions on neural networks and learning systems, vol. 30, no. 11, pp. 3212–3232, 2019.

[9] A. Kumar, A. Kaur, and M. Kumar, "Face detection techniques: a review," Artificial Intelligence Review, vol. 52, no. 2, pp. 927–948, 2019.D.-H. Lee, K.-L. Chen, K.-H. Liou, C.-L. Liu, and J.-L. Liu, "Deep learning and control algorithms.