Proceedings of the Third International Conference on Intelligent Sustainable Systems [ICISS 2020]
IEEE Xplore Part Number: CFP20M19-ART; ISBN: 978-1-7281-7089-3

Monitoring Pandemic Precautionary Protocols using Real-time Surveillance and Artificial Intelligence

Prof. B. Sathyabama
Dept. of Information Technology
SRM Institute of Science and Technology
Chennai, India
sathyabb@srmist.edu.in

Ashutosh Devpura
Dept. of Information Technology
SRM Institute of Science and Technology
Chennai, India
ashutoshdevpura@gmail.com

Mayank Maroti
Dept. of Information Technology
SRM Institute of Science and Technology
Chennai, India
marotimayank@gmail.com

Rishabh Singh Rajput
Dept. of Information Technology
SRM Institute of Science and Technology
Chennai, India
rishabhsinghrajput081@gmail.com

2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) | 978-1-7281-7089-3/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICISS49785.2020.9315934

Abstract— **The worst possible situation faced by humanity, COVID-19, is proliferating across more than 180 countries and about 37,000,000 confirmed cases, along with 1,000,000 deaths worldwide as of October 2020. The absence of any medical and strategic expertise is a colossal problem, and lack of immunity against it increases the risk of being affected by the virus. Since the absence of a vaccine is an issue, social spacing and face covering are primary precautionary methods apt in this situation. This study proposes automation with a deep learning framework for monitoring social distancing using surveillance video footage and face mask detection in public and crowded places as a mandatory rule set for pandemic terms using computer vision. The paper proposes a framework is based on YOLO object detection model to define the background and human beings with bounding boxes and assigned Identifications. In the same framework, a trained module checks for any unmasked individual. The automation will give useful data and understanding for the pandemic's current evaluation; this data will help analyse the individuals who do not follow health protocol norms.**

Keywords— Face Mask, Social Distancing, Pandemic, Artificial Intelligence, YOLO, TensorFlow.

I. INTRODUCTION

A report from the World Health Organization (WHO) presented that coronavirus disease (COVID-19) has globally infected millions of people and caused deaths on a large scale. This virus was initially taken into account in December in Wuhan, and later it was found it had spread across 216 countries, declared this as a pandemic. Many healthcare facilities, medical experts, and scientists are in the process of creating medicines or a vaccine for this deadly virus. However, to date, the development processes are in the latter stages, but they are still not sure of the effect caused by this unknown virus. Making the countries stop practising their daily activities on which their livelihoods depend, leads to a severe economic crisis during the pandemic period. The higher official came up with a temporary solution to re-establish the economy by opening up the countries by practising health protocols prescribed by WHO. In this pandemic, people look for alternative precautionary practices to overcome this deadly virus. Social distancing and face cover are the best precautionary methods used in today's time. The rules set by the World Health Organisation (WHO) give a detailed review of the outdoor activities that one must strictly follow: WHO follow up rules:(1) Mask/Face cover should be worn before going out in public. (2) The person stepping out from his private space must ensure that social distancing is maintained during the time he spends in public areas [1]. Hence, the absence of any medical aid as a permanent solution is not available. A Few steps are being taken worldwide to use AI and statistics as a preventive or predictive measure against Covid-19, such as compartmental models in epidemiology. The idea of a tool to oversee social distancing and face mask covering at the workplace with the help of feed on the security cam that can analyse live and offline to see the social spacing and face mask covering in the official, public, or any social place will help to monitor health protocols [2]. The Practising health protocol saves lives and allows to control the outbreak.

II. BACKGROUND STUDY AND RELATED WORK

Social distancing and face masks both together are the most feasible solution to control the pandemic; recent studies exhibit that from the day the government-issued social distancing orders it significantly slowed the Covid-19 case count, leading to a reduction of more than 600,000 cases in the US within a few weeks of commencement of the model created from these findings and mathematical algorithms used show that state-wide social distancing implementation has reduced the number of reported Covid-19 cases by nearly 1,600 cases in one week after issuing orders. The use of face masks was a prevention method issued and stated by the World Health Organisation (WHO), by which the findings for the benefit of face mask in public areas are astounding, as it shows that a small preventive measure could help a load. The study says that about 4,000 lives could have been saved in the UK if the face-covering was practised strictly; this is a helpful study that gives a critical perspective to this idea. YOLO and OpenCV have been used previously to detect objects using computer vision [3]. Drowning detection using machine learning and Drowsiness detection using Deep learning are applications of object detection algorithms and Computer Vision in the Field of Artificial Intelligence. Chest X-ray radiographs are used for the detection of Corona Virus using three significant models of CNN, namely Inception, ResNetV2, InceptionV3 and ResNet50. Amongst the existing systems, the highest performance and classification accuracy is provided by the ResNet50 model [4]. A few of the applications of object detection algorithms are The Facial Recognition algorithm is used for face detection using the live video feed [5]. People counting algorithm is used for tracking humans and analysing crowd statistics [6]. Distance Estimation algorithms are used for Object detection, localisation, and Distance tracking system for smart mobility applications like traffic road and railway environment [7].

The framework is built on such ideas to detect social distance and face-mask covering in real-time with the help of computational statistics.

III. Object Detection and Tracking Model

A. *Object Detection*

It is a fundamental concept used in this framework. The base structure to the foundation is built upon this very idea. Object detection allows for the recognition, detection, and localisation of multiple objects within an image. Object detection helps us find instances of real work objects such as faces and Individuals from images or videos in real-time. Object detection generally uses learning algorithms to recognise instances of an object category. Figure 1 shows the framework for the object detection algorithm.

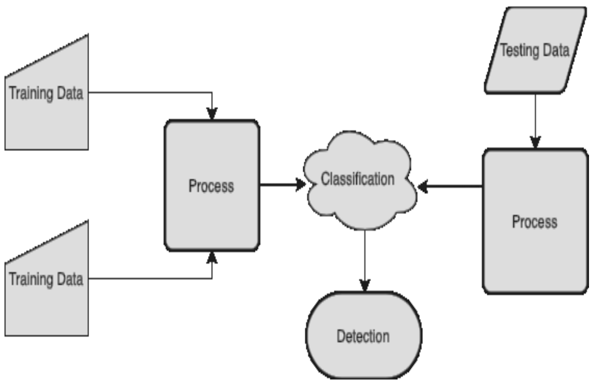

Fig. 1. Object Detection Algorithm Framework

1) Yolo V3

YOLO or 'You Only Look Once' can detect over 9000 object categories and specifications. YOLO is simple to understand as it takes an input image and then learns the class probabilities instantaneously. The same model is used in this framework for the application. The model is straightforward and comfortable; the input images are rendered through CNN, and then it is classified into a class. YOLO V3 uses a few improved tricks to clean detection and identification performance. It has a better classifier, and a strong backbone structure to support the multiple moving objects [8]. Many trained data supremely formulate the detection into the model Framework report.

a) Object Detection in YOLO V3

YOLO uses a unique neural network using the characteristics of the entire image to predict multiple boxes. Each image is divided into grids of equal size, and then a class and bounding box of objects present in each cell of the grid is located. The class of the cell with no object will be different from the class of the cell in which an object is present, fully or partially. In this way, all the cells in the grid are labelled. If the centre of the object falls into a grid cell, the cell is responsible for detecting the object. Each cell predicts a certain number of bounding boxes with a confidence score for each of the boxes. Each object in the bounding boxes will be given a class label and object is detected. Confidence score means how confident the model is that the object contains an object and how accurately the box has predicted the object boundaries [8].

2) TensorFlow

TensorFlow is an open-source AI library. Data flow library graphs to build a network structure; this network structure helps create large scale neural projects with many networks; it helps to build the model easily with an intuitive easy-to-use set of APIs that make it simple to implement machine learning, deep learning and scientific computing. In TensorFlow, computation is approached as a data flow graph. TensorFlow can run on multiple CPUs and GPUs, and its flexible architecture allows for easy deployment of computation across various platforms [9].

a) Object Detection in TensorFlow

Input data, which is a set of images is provided and using TensorFlow the model is trained. The model is trained using deep learning, and the main objective of the model is to extract the features. These features are visual and based on edge detection, facial recognition, people counting and many more. These features are extracted, and the model is created. Test data is provided to test the model, and output is generated [9].

3) R-CNN

It uses a search selective method to find and detect objects after it goes from a convolutional network. It uses 2k areas that are processed to an SVM classifier, and then they are computed[10]. It is highly used as the image is passed to the pre-trained CNN only once. The input images are used to obtain and assemble detected objects.

4) Resnet50

The Resnet50 is a deep neural network that can also be used for object detection. It is fast and efficient. It has ImageAI, which gives a more effective analysis of the input image after it is passed through the network. Very useful and can be trained for many other image skills used other than detection[11].

IV. Proposed Approach

Deep learning has changed and bought unparalleled results and advancements in science and technology, not in a particular field, serving in various kinds such as medical diagnosis, machine translation, speech recognition, and many more. Other than that, the second instalment to the framework has also developed into a very innovative technique and technology that is face recognition with masks. These tasks include discovery, frame recognition, and classification. The convolution neural network (CNN) has shown many advancements and updates in its efficiency, which gives equality between mAP and FPS on standard benchmark datasets, deep learning and face recognition based framework. This idea usher the tasks of monitoring face masks and distance between the people but needs accuracy and speed due to the live tracking feature included in the framework.For Social Distancing detection, previous studies show that CPU based framework obtains less FPS and slow processing for the real-time environment. This study shows the adoption of utilizing an NVIDIA CUDA-capable GPU and installation of OpenCV's "dnn" module with NVIDIA GPU support to increase the FPS and processing speed for the real-time environment. For face mask detection, previous studies show the use of CNN for masked /no masked data set and obtained a classifier that was approximately 86% accurate. This study shows the adoptation of MobileNetV2 on this framework for masked/no masked datasets, which gets the accuracy of approximately 98%.

The YOLO V3 is used with TensorFlow, by this, each objected detected is monitored within a box created by the framework, and then they are compared with each other to check for distancing between them; if the distancing is green, then they are at a safe distance, but if it is in red, then the subject distance is significantly less. This is the basic rundown of the detection process; now, for the recognition process, the software is trained with 3800 images of the non-masked and masked people, and its accuracy is above 96%. There are two folders with the mask and without the mask. There are 1900 images with mask and same without a mask from Kaggle, google images and open-source image libraries. Data set is present into a local directory. Images from the data set will be converted into arrays to create a deep learning model. After running the model, the model computed/visualised accuracy with the help of matplotlib. The data is partitioned into training and testing using 80% of the data for training and the remaining 20% for testing. This model will detect anyone not wearing a mask in red box. This shows the norms have not been followed in the public section of the feed passed into the model to analyse. The workflow for the given study is:

- The object detection model discover and classify the object in the feed.
- The feed is processed through the object model and given a tag for reference.
- Each individual is associated with three-dimensional feature space (x, y, d), where (x, y) corresponds to the centroid coordinates of the bounding box and d defines the individual's depth as observed from the camera.

$$d = ((2 * 3.14 * 180) / (w + h * 360) * 1000 + 3) \qquad (1)$$

where w is the width of the bounding box, and h is the height of the bounding box [12]. Figure 2 shows the framework calculating centroid coordinates of objects.

Fig. 2. Calculating Centroid Coordinates of Object

- Anyone mixing in with a fellow mate is indicated with a different colour then the previous one to formulate a group for easy tracking and to find the group if infected by any chance.
- By this, any violation will be traced on the go.
- The feed also has been passed through a face recognition model to check for unasked objects within the parameter.
- The model has been trained with ample training data available online and creates a box structure to the face while tracking the object in its path.
- As the analysis is done, which takes about 3 seconds for a moving object and a sec for a still object within this frame reference, the analysis is completed.
- The red box appearing on a person/object will represent an unmasked object, and it is easily traceable through the id given to the object.
- This is the full control measure and equipped application for following norms set for is protection during this pandemic.

V. IMPLEMENTATION

A. Social Distancing Detection

1) Detection of Individuals using Computer Vision

For monitoring social distancing, the framework uses YOLO object detection algorithm to detect people using surveillance cameras that are preinstalled in public places. Using YOLOv3 with OpenCV2 requires more output processing than other object detection methods (such as R-CNN), so any YOLO object detection logic is encapsulated. NumPy, along with a computer vision library, is used to compute and operate high-level mathematical function to operate multi-dimensional arrays. A function is then defined to accept parameters such as the frame from surveillance cameras, the pre-trained object detection algorithm, i.e. YOLOv3 and the YOLOv3 convolutional neural network output layer's name. Once the function parameters are accepted, result list is initialised, and it consists prediction probability of the number of individuals in the frame, probability of the bounding boxes surrounding the Individual and centroid of the individual that is an object.

2) Implementation of Social Distancing using Computer vision and Deep Learning

The Euclidean distance formula is used to identify the distance between the centroids of the objects with OpenCV [13]. COCO labels are loaded, and YOLO path is defined to load the model into memory. Path to surveillance cameras is initialised, and then frame by frame is processed to determine if social distancing among individuals is followed or not. Figure 3 shows the graphical user interface for the social distancing framework.

B. Face mask Detection

1) Training Face mask Detection Dataset

To monitor if a person is wearing face cover or not, the framework will use the surveillance cameras that are preinstalled in public places. Face mask detection dataset is bundled from memory. A model is trained on this data set using Keras/TensorFlow and then published the trained data back to memory. First, a database is created with photos without the mask and then add a mask on them with the help of computer vision by building a data set of wearing a face mask. Computer vision is used to bound the box around the face and extracts that region. Using facial landmarks are identified, the area of the faces such as nose, lips and eyes. A transparent image is added to different types of mask to different images in the dataset. A function is then defined to accept parameters from surveillance cameras [14]. Scikit-learn will print a classification report, segregate dataset and classify class labels. It will segregate data into training data and testing data. MobileNetV2 is fine-tuned to establish the baseline for the model, and data is trained using deep learning[15]. Figure 3 includes the sample of the unmasked data set and Figure 4 includes the sample of the masked data, which is bundled in the memory.

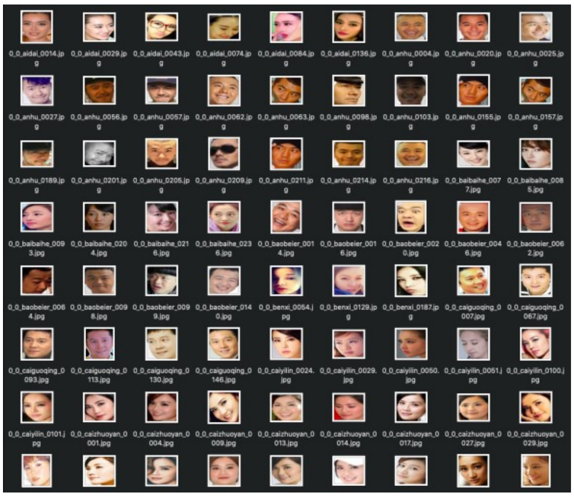

Fig. 3. Unmasked Data Set

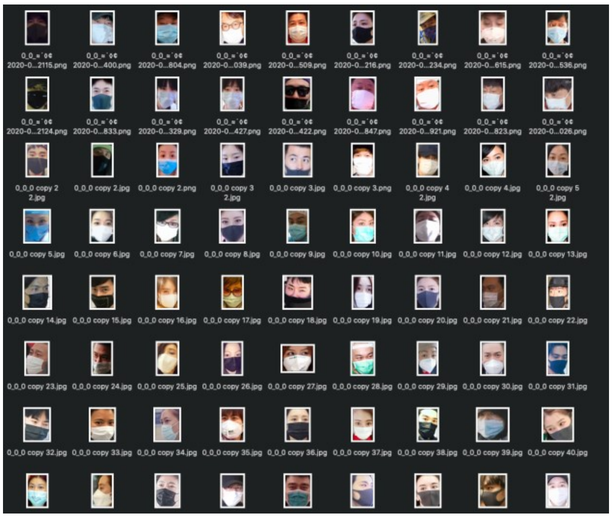

Fig. 4. Masked Data Set

2) Deployment of Face Mask Detection Framework

Once the dataset is trained, the data is analysed to perform mask detection, face detection and then allocate each face as masked or not masked. Computer vision library is used for displaying image and image manipulation. TensorFlow/Keras is used to pre-process the image that is bundled and also to load the Market model. Face detection model and face mask classifier model, both are now loaded. As both deep learning models are already loaded in memory, input image/video is processed, and the model detects the faces. The framework now puts bounding boxes around the face using NumPy slicing and perform mask detection.

VI. OUTPUT AND FINDINGS

The outputs for the project are very conclusive to its kind; they are performed in real-time to test the framework of the application. The framework is based upon both real-time and stored feeds, as per the use conditions. During this pandemic situation, few precautionary measures are strictly implemented in society. The protocols to follow during the pandemic are mainly Social Distancing and Face Mask detection. The information obtained by training the data with variable and unfavourable conditions has shown impeccable results which can help to oversee the measures being implemented around the socially active places. The feed goes through the rendering in the confined framework, which then makes a box structure to represent the human object, the object then coming in contact shows a different colour which makes both of them of the same group. These colour variants then give a proper means to track and find the individual as seen in Figure 5. The study of monitoring social distancing provides the working capabilities for this application. The 21 seconds stream was processed in 54 seconds on NVIDIA CUDA-capable GPU. The problem with the current configuration is speed. The social distancing detector is obtaining approximately 4.2 FPS. The framework also gives a high accuracy result of face mask detection by training its data in the TensorFlow framework and recognising the potential masked and unmasked people within the crowd. Figure 6 shows the detection of an unmasked individual. Figure 7 shows the detection of a masked individual. The video stream received by the trained module in the tensor flow, which gives immediate feedback of mask/unmasked object. Figure 8 shows Facemask detection module training's accuracy/loss curve with high accuracy and low data overfitting signs, which implicates that training for the detection has a great response and has acceptable reliability. There are little signs of overfitting, with the validation loss lower than the training loss. The study is optimistic about its outcome in the real-time environment. The framework can be used at the prominent public places like shopping complexes, entrances of schools and offices, and many other enclosed areas. The social distancing frameworks can be deployed at crowded public places to identify individuals not following the social distancing parameters set by the government. The other side of the study is face mask detection that is simultaneously used to identify individuals not pursuing face-covering in public places.

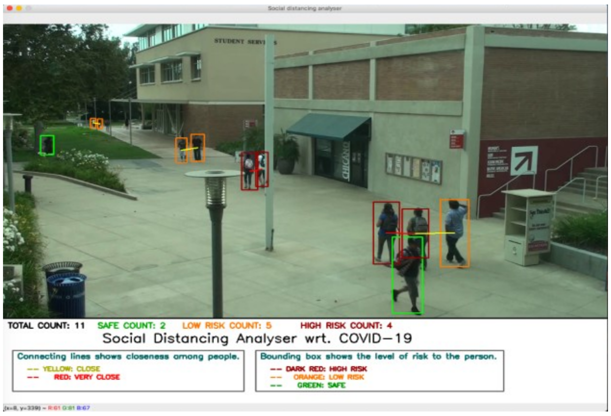

Fig. 5. UI of Social Distancing Framework

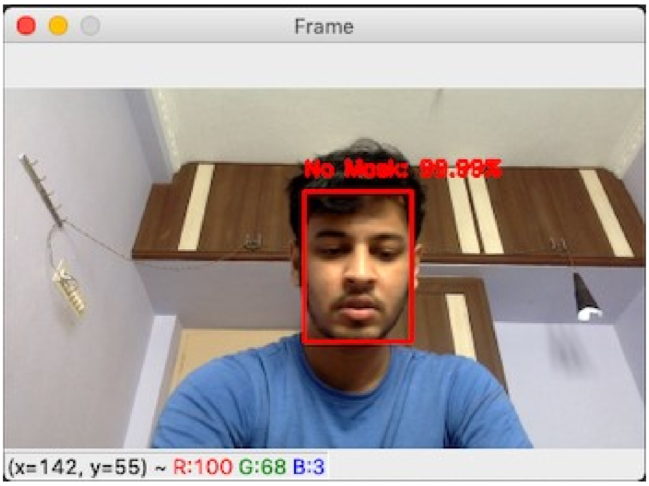

Fig. 6. Detection of an Unmasked Individual

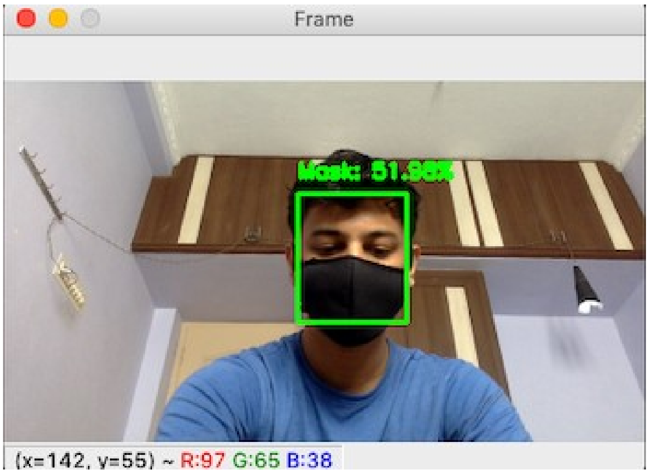

Fig. 7. Detection of a Masked Individual

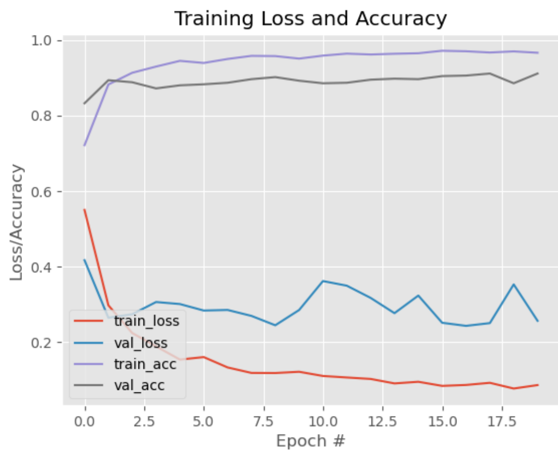

Fig. 8. Training Loss and Accuracy of Face Mask Detection

VII. FUTURE SCOPE AND CHALLENGES

It needs further accuracy, and data frame enhancements would give more accurate precision in monitoring. With this in place, one might see this as a hindrance to private spaces, which can be solved with consent forms in a closed promise. For social distancing framework, fast-moving individuals can be occasionally missed by the system. The processing at lower GPU is slow and can be improved in future enhancements. For face mask detecting framework, individuals can try to cheat the system by just covering their face with fabric and not wearing a face mask. There is a approximately 5 per cent probability that the system may not detect this.

VIII. CONCLUSION

The system works great, but it required some effort to overcome some challenges. For Social distancing framework, the calibration of the model by simulating a 3D depth factor based on the camera position and orientation gives better analysis. For Face mask detection small faces were hard to detect and even harder to track hence using pose estimation, which is using the body to estimate the head bounding box and using facial key points instead if they are visible. Installation of this system on surveillance cameras for monitoring health protocols by the government will help to control the outbreak. The proposed framework can detect social spacing and face-covering correctly. Experimental results show that YOLO weights and the face data set had achieved promising performance.

REFERENCES

[1] WHO (2020). "Advice for the public on COVID-19". Retrieved August 2, 2020, from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public.

[2] Pahuja, R. (2020, June 26). "How AI is helping maintain social distance at Indian factories". Economic Times. Retrieved 2020, from https://cio.economictimes.indiatimes.com/news/strategy-and-management/how-ai-is-helping-maintain-social-distance-at-indian-factories/76636929.

[3] 3. Suma, V. "Computer Vision for Human-Machine Interaction-Review." Journal of trends in Computer Science and Smart technology (TCSST) 1, no. 02 (2019): 131-139.

[4] Dhaya, R. "Deep Net Model for Detection of Covid-19 using Radiographs based on ROC Analysis." Journal of Innovative Image Processing (JIIP) 2, no. 03 (2020): 135-140.

[5] J. Dhamija, T. Choudhury, P. Kumar and Y. S. Rathore, "An Advancement towards Efficient Face Recognition Using Live Video Feed: "For the Future"," 2017 3rd International Conference on Computational Intelligence and Networks (CINE), Odisha, 2017, pp. 53-56, doi: 10.1109/CINE.2017.21.

[6] L. Koraqi and F. Idrizi, "Detection, identification and tracking of objects during the motion", 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 2019, pp. 1-6, doi: 10.1109/ISMSIT.2019.8932833.

[7] Z. Chen, R. Khemmar, B. Decoux, A. Atahouet and J. Ertaud, "Real Time Object Detection, Tracking, and Distance and Motion Estimation based on Deep Learning: Application to Smart Mobility," 2019 Eighth International Conference on Emerging Security Technologies (EST), Colchester, United Kingdom, 2019, pp. 1-6, doi: 10.1109/EST.2019.8806222.

[8] Redmon, J., & Farhadi, A. (2018). "YOLOv3: An Incremental Improvement". ArXiv. Retrieved 2020, from https://pjreddie.com/darknet/yolo.

[9] Abadi, M., Agarwal, A., Barham, P., et al. (2015, November 9). "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems". Retrieved 2020, from http://download.tensorflow.org/paper/whitepaper2015.pdf.

[10] Gandhi, R. (2018, July 09). R-CNN, Fast R-CNN, Faster R-CNN, YOLO - Object Detection Algorithms. Retrieved from https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e.

[11] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. doi:10.1109/cvpr.2016.90

[12] Pias, P., "Object detection and distance measurement," https://github.com/paul-pias/Object-Detection-and-Distance-Measurement, 2020, [Online; accessed 01-August-2020].

[13] Rosebrock, A. (2016, April 4). "Measuring distance beten objects in an image with OpenCV". Retrieved 2020, from https://www.pyimagesearch.com/2016/04/04/measuring-distance-beten-objects-in-an-image-with-opencv/.

[14] G. Deore, R. Bodhula, V. Udpikar and V. More, "Study of masked face detection approach in video analytics," 2016 Conference on Advances in Signal Processing (CASP), Pune, 2016, pp. 196-200, doi: 10.1109/CASP.2016.7746164.

[15] Singhal, M. (2020, July 07). Object Detection using SSD MobilenetV2 using Tensorflow API : Can detect any single class from... Retrieved from https://medium.com/@techmayank2000/object-detection-using-ssd-mobilenetv2-using-tensorflow-api-can-detect-any-single-class-from-31a31bbd0691.