Human Tracking and Profiling for Risk Management

Ranjith K.H.V.S[1], Jayasekara A.S[2], Ratnasooriya K.A.L.L[3], J.L Thilini Randika[4], Dr. Lakmal Rupasinghe[5] and Ms. Chethana Liyanapathirana[6]

[1] Sri Lanka Institute of Information Technology (SLIIT)

Received: 30 September 2021 Accepted: 20 October 2021 Published: 2 November 2021

Abstract

Infectious viruses are conveyed via respiratory droplets produced by an infected person when they speak, sneeze, or cough. So, to combat virus transmission, the World Health Organization (WHO) has imposed severe regulations such as mandatory face mask use and social segregation in public spaces. The ?Human Tracking and Profiling for Risk Management System (HTPRM)? is an online application that identifies the risk associated with failing to follow proper health practices. This proposed approach, which is divided into four components, utilizes ?You Only Live Once YOLO (V3)? to detect facemask danger, which would be determined based on two factors: wearing the face mask properly and the type of mask (Surgical, k95, homemade, and bare). The second phase is to use Open CV and SSDMobilenet to evaluate the value of a one-meter space (Social Distance) between people. The system recognizes the maximum number of individuals that can be in the vicinity of the specific hall that uses YOLO(V3) and image processing as the third procedure. In the last processing, the system identifies each person?s behavior, classifies it as uncommon or not, and calculates the risk associated with each category. Finally, the system computes the overall risk and generates a warning alarm to notify the user that they are in a dangerous scenario.

Index terms— YOLO (V3), SSD (single shot detector), mobile- net, open-CV, image processing, open pose, tenser-flow.

1 Introduction

iruses have been a point of controversy for humans since even before our species evolved into its contemporary form. Vaccines and antiviral drugs have allowed to limit the widespread spread of certain viral diseases and have helped sick patients to recover.Some viruses, such as Marburg, Ebola, Rabies, HIV, Smallpox, SARS -Cov, and MERS -Cov, have been unable to be eradicated, leading to a rise in new cases. However, destroying viruses is a challenging task. Some viruses have migrated from animals in recent decades, triggering large epidemics and claiming thousands of human lives. From 2014 to 2016, the virus that caused the Ebola epidemic in West Africa killed approximately 90% of those infected, making this the most dangerous member of the Ebola family. But other viruses are just as deadly, and some are even more deadly. It is true that the death rate from some viruses, such as the new corona-virus, which is currently spreading around the world, is still high, but infections pose a serious threat to public health as human cannot yet combat germs. The goal of the HTPRM is to identify risk by investigating at four different virus transmission modes. Crowd Counting is one technique that determines the population of a scene. Public places are overcrowded possible to spreading the virus among peoples. To minimize the spreading viruses, must be limited the number of people in a public area. The density risk percentage calculate by comparing the number of people in an specific area at a given moment to the maximum number of people permitted in that region. This is one method to help prevent virus propagation in densely populated places, since it is a highly robust strategy in today's society.

The primary source of risk is that COVID-19 spreads by touch or being in close proximity to an infected individual. Social Distancing is the only way to prevent the spread of COVID-19. Maintaining a safe distance from one another is the most effective method of preventing the transmission of this disease at least until a vaccine is found. Social distancing can have a negative impact on social well-being and health by resulting in social isolation and physical activity limitations.

One of the another critical of risk is behavior control. If one person is infected with a virus, there is a high possibility of the infection spreading due to his behavior. To avoid or mitigate such a circumstance, a human behavior recognition function has been developped for risk management systems using Human Tracking and Profiling. This capability is mostly concerned with human behavior in a given area.

Another critical aspect of combating this epidemic is to use a face mask. Wearing a face mask has received general support as a means of delaying the transmission of viruses. Speed of virus transmission is dependent on the sort of face-mask used (Surgical, k95, homemade, and bare). Until the virus is completely destroyed, daily use of face masks is crucial for infection prevention and protection against airborne infectious germs. This approach not only determines whether or not someone is wearing a mask, but also determines whether or not a face mask is worn properly. The danger varies according to the type of face mask used, the average risk is determined as well. This research will conduct an in-depth examination of the usage of masks to prevent the transmission of the lethal coronavirus. This version introduces a novel multi -facemask detection in real-time. Finally, the Human Tracking and Profiling system calculates the average risk associated with four components and then calculates the total average risks involved with those four components.

2 II.

3 Literature Review

The main intention is to implement automated Human tracking and profiling for risk management application to avoid speed of spreading the virus infections in the world. Social Distance Risk, Face-mask risk, Density Risk and Human Actions and behaviors Risk are the main four components to minimize the percentage of deaths due to viruses. Lot of applications developed for analysis the risk and get the crowd/ visitors count in the frame but there is not single system to analysis the density risk in a particular area to minimize the spreading viruses among people. In the current pandemic situation in world, one Deep-CNN Crowd Counting Model for enforcing social distancing application is implemented in Saudi Arabia's public places for avoid spreading the viruses among peoples [1]. Actually, above proposed method is based on CNN model to count people who appear in video frames in public places [1]. Another, people counting system developed in post COVID-19,which is counting people through infrared detection and this system count and update based on people moving in/out through the area/premise [2]. There already exists a few work that pedestrian counting systems [[3], [4]]. so, in this proposed density risk solution go beyond above systems and in the first step, a video first frame user [system owner] must selects the area where he/she wants to measure the density risk using four mouse click points. Then according to area width and length system estimate the maximum people/visitors count allowed in that area. After that, system take the real time total people/visitors count inside the user selected area (where user wants to measure the density risk) in each video frames and system comparing the these both real time people count, and maximum people count allowed in this area and analysis get density risk. If the real people count is higher than the maximum people count allowed in the area, then that area is a high-risk place. This is the novelty of density risk analysis. When it comes to the face-mask risk, the majority of the research studies reviewed focused exclusively on identifying the face mask. Researchers used a variety of machine learning and deep learning algorithms to assess whether or not they were wearing a face mask. Using image processing, the device developed by a team led by S. Balaji detected the passengers' facemasks [9]. Additionally, the team, which includes Amit Chavda, has presented a method that uses a Convolutional Neural Network to detect individuals who use facemasks [7] and some of research papers used to detect facemask by utilizing Faster-RCNN [[6], [8], [10]]. Numerous devices have been proposed and implemented to detect facemasks using various methodologies, however analysers all have significant limitations. Numerous facemask types have been introduced to the market. Even if individuals are masked, it is impossible to demonstrate that they are passing hatred from one individual to another. This is because they must cover their nose and mouth and secure it beneath their chin, even if they are wearing a mask. Additionally, it clings snugly to their chin. The National Center for Immunization and Respiratory Diseases (NCIRD) has confirmed that the viral transmission rate varies between different types of masks. Certain masks are designed and tested to ensure consistent performance in preventing the transmission of COVID-19. These masks are labeled with the criteria they comply with. KN95 masks provide approximately 98.5% protection, whereas surgical masks provide 56.1 percent protection. Some folks make their homemade masks. This results in a 51.4% guard. We concentrated on that and identified the facemask as surgical, KN95, and homemade using the YOLO principle (V3). Thus, our proposed system analyzes the multi-person real-time face mask type and analyzes the risk of face masks and unmasking. According to the investigation, if it exceeds 75%, it is considered a risky zone. As soon as it becomes a risky area, the head of the location is warned through SMS. Another important aspect of human profiling and estimating for risk management systems is the estimation of human actions and behaviors. In the study publications, systems to detect an individual's activity were introduced, but systems to recognize the action of a group of individuals were not found. Zhe

Cao and colleagues focused on a critical component of acquiring a deep understanding of humans in photos and videos: human two-dimensional posture estimation-or the difficulty of localizing anatomical important points or ”parts.” Human estimating has always been primarily concerned with locating individuals’ body components [14]. Federico Angelini and his colleagues proposed Action Pose: a two-dimensional pose-based technique for human action recognition at the pose level [15]. They retrieved low-level and high-level features fo the Action XPose from the human body posture and fed them into a LSTM (Long Short-Term Memory Neural) Network and a 1D Convolutional Neural Network for classification. Action XPose, a 2D pose-based algorithm for posture-level Recogninizing Human Action, was introduced by Zeyu Fu and his team [16]. However, our suggested system recognizes the action of human and classifies actions such as leap, run, and walk. The risk is assessed based on the classification of human behavior .To accomplish this, we use a computation algorithm that we created ourselves. As with other systems, if someone behave incorrectly, it sends a notification to the location chief.

When analyzing the risk of social distance to get a better accuracy this paper used the SSD-mobilnet model. In each frame need to be more accurate. Therefore to get better detection for bad visibility areas used CLAHE prepossessing method to identify objects [11]. Most of the researchers only considered about detect social distance [13]. This paper is most related to analyze the risk of percentage. Based on these percentages, can easily detect whether the area is bad or good.

4 III. Methodology

Fig. **??**: Overall System diagram According to the system overview diagram Fig. **??**, initially system gets CCTV footage as a input and same CCTV footage goes through four sub risk analysis functions separately and estimate the risk status. For the estimating the overall risk percentage, divide the 100% equally among four sub functions and each function gets 25%. if one function is totally violating, then added each sub function 25% percentage to the total overall risk. finally if the total risk percentage is greater than 75% email will be send, informing that area will be a risky place. a) Density Risk Analysis Fig. **??**: An image of a density risk system overview

The proposed system analysis the density risk in a particular area at a particular time. According to Fig. **??** the system gets video frames as input and then in the first video frame user must select the area where they want to measure the density risk using four mouse click points. Then this area is a Polygon shape rectangle area. Then according to the width and length of the area, the system estimates the maximum people count allowed in that particular area according to Fig. 3 using predefined formula. After that, the system takes the realtime people/visitors count inside the area in each video frame. finally, the system comparing both the Maximum people count and real-time people count inside the area, and if the real-time people count higher than the Maximum people count that area is a High-Risk area, email will be send, informing that area will be a risky place.

Object detection and tracking are one of the main parts of this function.Yolov3 is used to detect an object in the frame. Yolov3 is a unique neural network that predicts bounding boxes and class probabilities directly from complete images in a single evaluation. The Yolov3 configure (cfg file) and weight file trained on the detect 80 classes objects. However, People/visitors type object detection is only needed for the density risk estimation. So then did some transfer learning (hyperparameter changing -max batches, filters, classes) for the Yolov3.cfg file (model architecture file) and re-trained using Google’s Open Images, then generated the new weight file and it used for the people/visitors object detection in the video frames and also shapely python libraries and Open CV techniques used for the estimate the length and width of the area in the video frame. This algorithm calculates distances between people and draw different colors of bounding boxes with fulfilling above steps. Used SSD-MobileNet model for object detection. For better detection for bad visibility areas used CLAHE prepossessing method.

5 c) Human Behaviour Risk Analysis

This proposed Human Tracking and Profiling for Risk Management another main part is Human behaviour recognition part. Briefly in this part, Estimate the human actions and then recognize what are the actions using previously estimated actions. In this scenario, mainly there are two main parts in the human action recognition.

6 i. Human action estimation

In this part there are also two parts, Which is, single person post estimation and multi person action estimation. Multi-person action estimation is more difficult than single person action estimation, Because, there are more than one object should be locked in the each frame. According to this proposed system we had to use multi person human action estimation method. There are lot of multi human action estimation method. for a example, Open Pose, Alpha Pose , Deep Cut and Mask RCNN. from them, Open Pose, and Deep sort algorithm methods are used to develop this function, Because, it gives more accuracy than other methods, and there are more capable facility to get real time human actions. And another advantage is, Open Pose follows Bottom-Up approach. In the bottom-Up approach, first initially detect the human joints and the connect each joint for each related person. Deep Sort algorithm is mainly used for track multi people.

7 ii. Recognize human Behaviours

This is the second part of this function. In this function recognize the what are the human behaviours using previously estimated human actions. To do that, we used a machine learning model that we created using more than 4100 image data as the data set.

8 d) Face Mask Risk Analysis

Facemask risk was monitored in real time using a deep learning approach for detecting face masks. This section identifies the type of facemask and calculates the risk by comparing it to the recommended risk values. Two distinct YOLO (V3) object detection models are used to determine whether a face mask is present or absent and to classify the type of face mask. YOLO is an ingenious convolutional neural network (CNN) for real-time object detection. The algorithm applies a single neural network to the entire image, then divides it into regions and predicts their bounding boxes and probabilities. Here the feature maps are obtained by 81,79 and 91 convolutional neural network layers in three detections. In this detection is accomplished by applying detection kernels to feature maps of three distinct sizes located in three distinct locations throughout the network.

Due to the difficulty of obtaining a sufficiently large dataset for training the two models, custom datasets were used. A dataset of 6000 images was used to classify four types of facemasks: Surgical, KN95, Homemade, and Bare. Additionally, to determine whether or not to use facemasks, we used a dataset of 4000 images. Where transfer learning was used to train YOLO (V3) models using custom data.

Then, using Python, author created an algorithm based on the risk value assigned by the (NCIRD), Division of Viral Diseases. If the area is dangerous, the head of location is notified via SMS. Twilio's Python library assists in creating a new instance of the Message resource by allowing you to specify the message's To, From, and Body parameters.

9 Results and Discussion

10 a) Density Risk Analysis

Yolov3 network was first trained with a single class dataset of 4000 people images. So the dataset only has one class, the maximum batch was set to 8000, the steps were set to 6400, 7200, and 18 filters in the three convolutional layers before the YOLO layers, and the number of classes in the YOLO layers was set to 1 and also set network size width 608 height 608 in Yolov3.cfg file. Map value test on 500 people images and got 85% map value for our yolov3 trained model [5]. Table **??** [5] compares our Yolov3 approach to a variety of different object detection methods in terms of mAP.Fig. 6 shows the results of the density risk estimation. if the real-time visitors count inside the area is higher than the maximum count that area is a high-risk area and an email is sent to the nearest police station. Authors tested the proposed model using a video stream and images. In each frames were also labelled as unsafe and safe accordingly. To bad visibility areas proposed using CLAHE prepossessing technique. It is vital to have individuals moving continuously while utilizing the webcam, or else the detection will be wrong.

11 Fig. 7: Final output of social distance detection c) Human Behaviour Risk Analysis

Another main part in this system is Human behaviour risk analysing. To estimate the human actions, we mainly used two pre-trained Open Pose models to estimate the human actions. The main part of this function is recognizing human actions using estimated human actions. To do that we used over 4000 image data to train a model. After the training, we were able to get a 98.3 percentage training accuracy and 95.9 percentage of test accuracy.

12 d) Face Mask Risk Analysis

The wearing of masks correctly and consistently is a vital step that everyone can take to avoid contracting and spreading COVID-19. Masks are most effective when everyone wears them, but not all masks offer the same level of protection. Consider how well a mask fits, how well it filters the air, and how many layers it has when purchasing one. For the purpose of this research, the data sets which have two classes (MASK and No Mask) and four classes (Surgical, KN95, homemade and bare) were obtained. For the facemask risk detection using facemask type, a YOLO(v3) model was pre-trained with Pytorch Geometric using custom dataset imported from YOLO v3 achieving a train mean average precision of 99.24% and test mean average precision of 73% with 6000 images in training and 2000 test images under 4 classes in validating the model. Figure **??**0 is shown it efficiently. With these findings, our model has also demonstrated success in detecting face masks in images beyond the our training and validation range. We initialized our learning rate at (LR=0.001), the number of training epochs at (EPOCHES = 45000), and the batch size at (BS = 64) for the testing phrase. Figure **??** depicts various scenarios for detecting different sorts of face masks in real time from a live-stream. Additionally, Table 1 discusses the importance of performance indicators in gaining a better knowledge of how suggested models behave throughout the testing process. The result analysis demonstrates that our suggested approach for face mask detection based on several types of masks performs really well despite the fact that testing data is limited. The authors were unable to obtain individual outputs for this system after considering the above situations (Social distance Risk, Social Density Risk, Human Behavior Risk, and face mask Risk). Because different health guidelines infractions can occur in the same public space As a result, overall risk must be estimated utilizing social distance risk, social density risk, human behavior risk, and face mask risk. The authors provide a new formula to calculate the overall risk using each functionality.

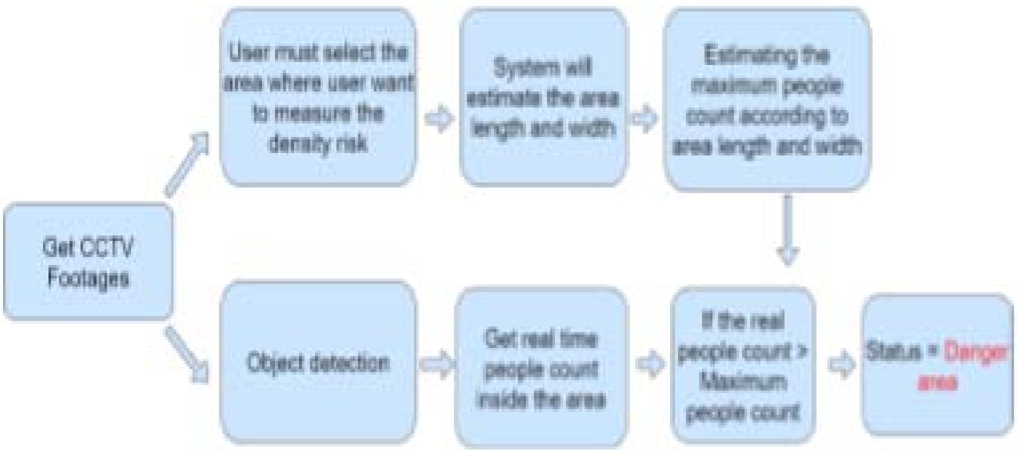
13 Fig. 11: Overall Risk Calculation Formula

Total risk categories must be determined in order to compute the overall risk. There are four risk categories, according to the system (Social distance Risk, Social Density Risk, Human Behavior Risk and face mask Risk). The total number of risk categories that have been breached should next be determined. Finally, using these variables, compute the overall risk. If the aggregate danger exceeds 75%, the area is considered high risk. If the entire risk is between 25% and 75%, the area is considered low risk. Finally, if the threat is less than 25%, the location is considered safe.

V.

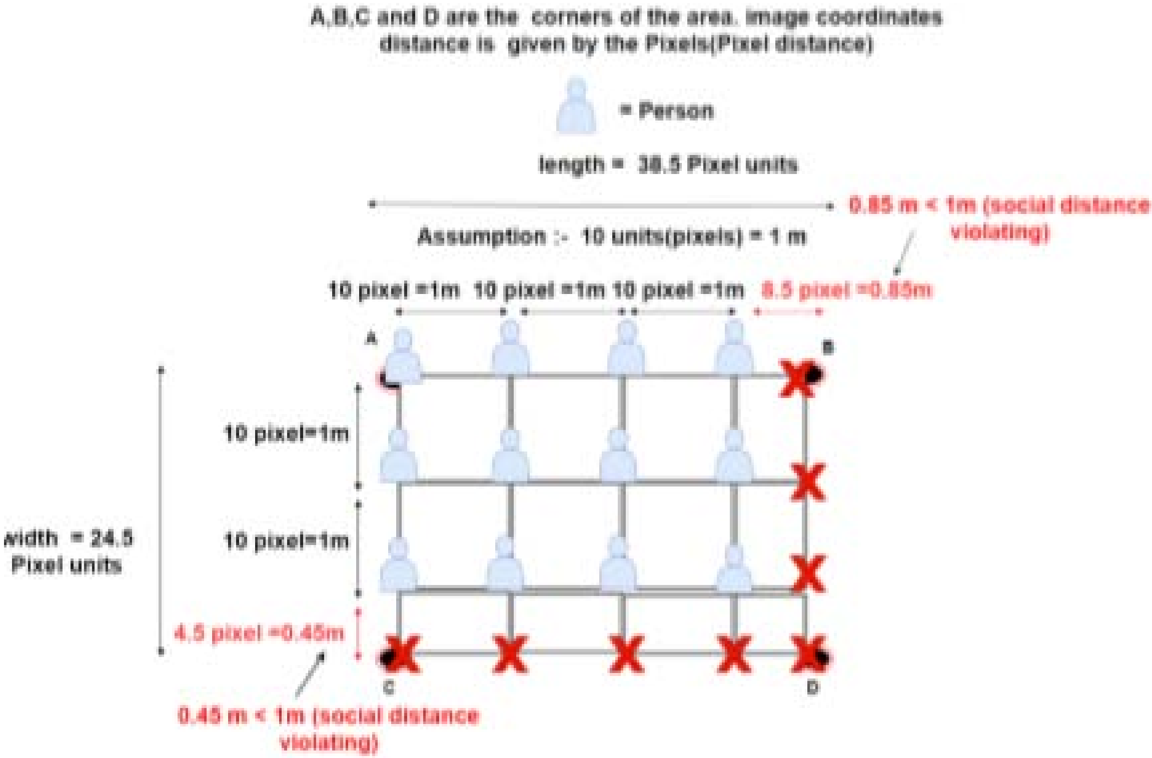
14 Conclusion

The use of machine learning becomes more common. By using the image processing and deeplearning techniques, i.e. YOLO, SSD, Open Pose, and Deep-Sort methods, we provide a comprehensive realtime person recognition system. Mainly covered four main scenarios. Those are density detection and analyze the risk, social distance detection and analysis of the risk, face-mask detection and analyze the risk, and human pose detection and analyze the risk. Test average precision (mAP) for detect humans and detect facemask with facemask type respectively 85.0 %, 73.0 %. To detect human behavior the system got 95.0% present of test accuracy. Our social distancing risk detection and estimating area length and width for density risk detection did not use correct camera calibration, which means that pixel distances to measurable real units were not (easily) mapped to (i.e., meters, feet, etc.). Therefore, the first step to improving our social distancing risk detection and estimating area length and width for density risk from the distance between our social systems is therefore to use a good camera calibration. That way, the results will be better and can calculate measurable units actually (rather than pixels).This work can be used as the basis for estimating the risk of each function. In the end, we are come up with the four individual average risks. Based on that we are calculating the total risk for a particular place. [1]



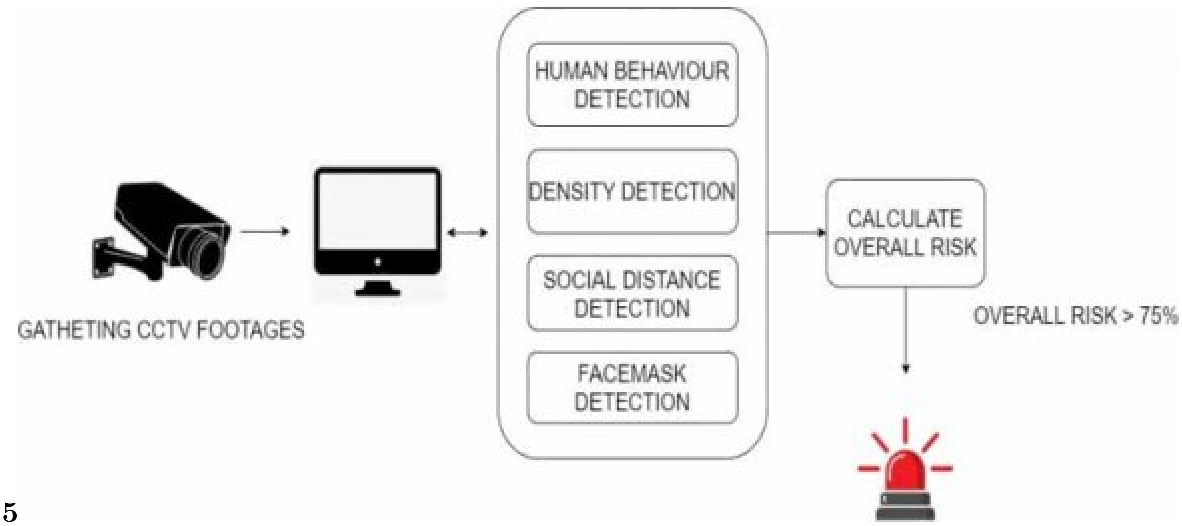
3

Figure 1: Fig. 3 :



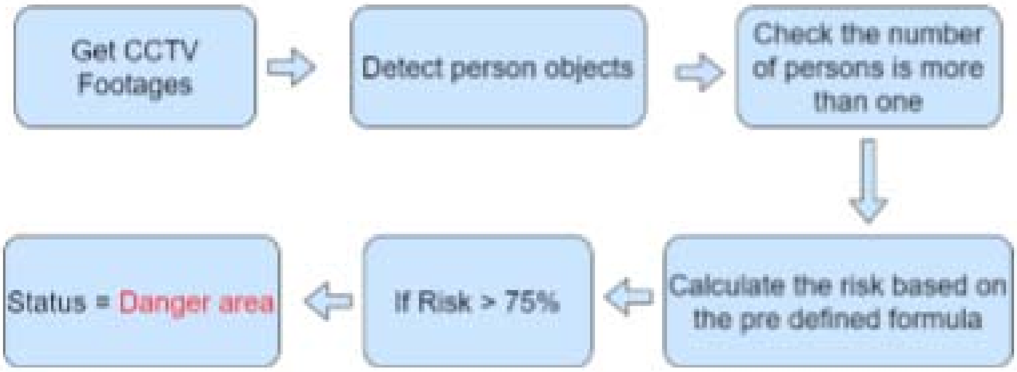
4

Figure 2: Fig. 4 :



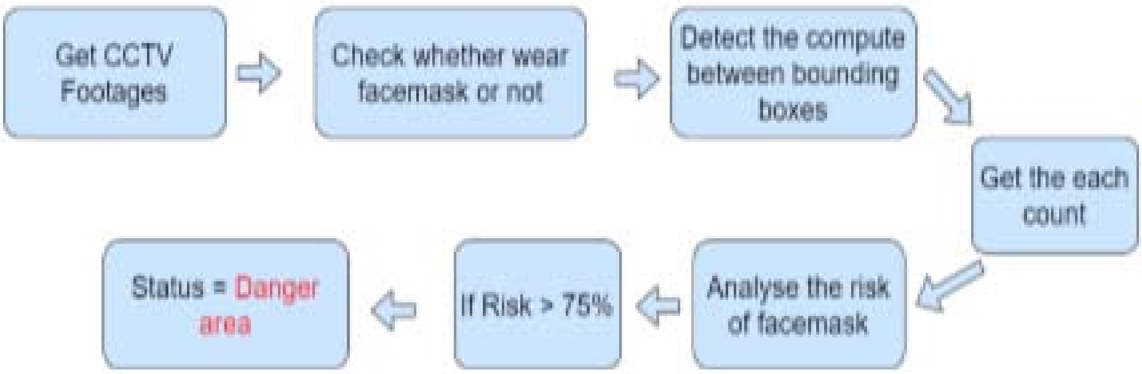
5

Figure 3: Fig. 5 :



6

Figure 4: Fig. 6 :



8

Figure 5: Fig. 8 :

910

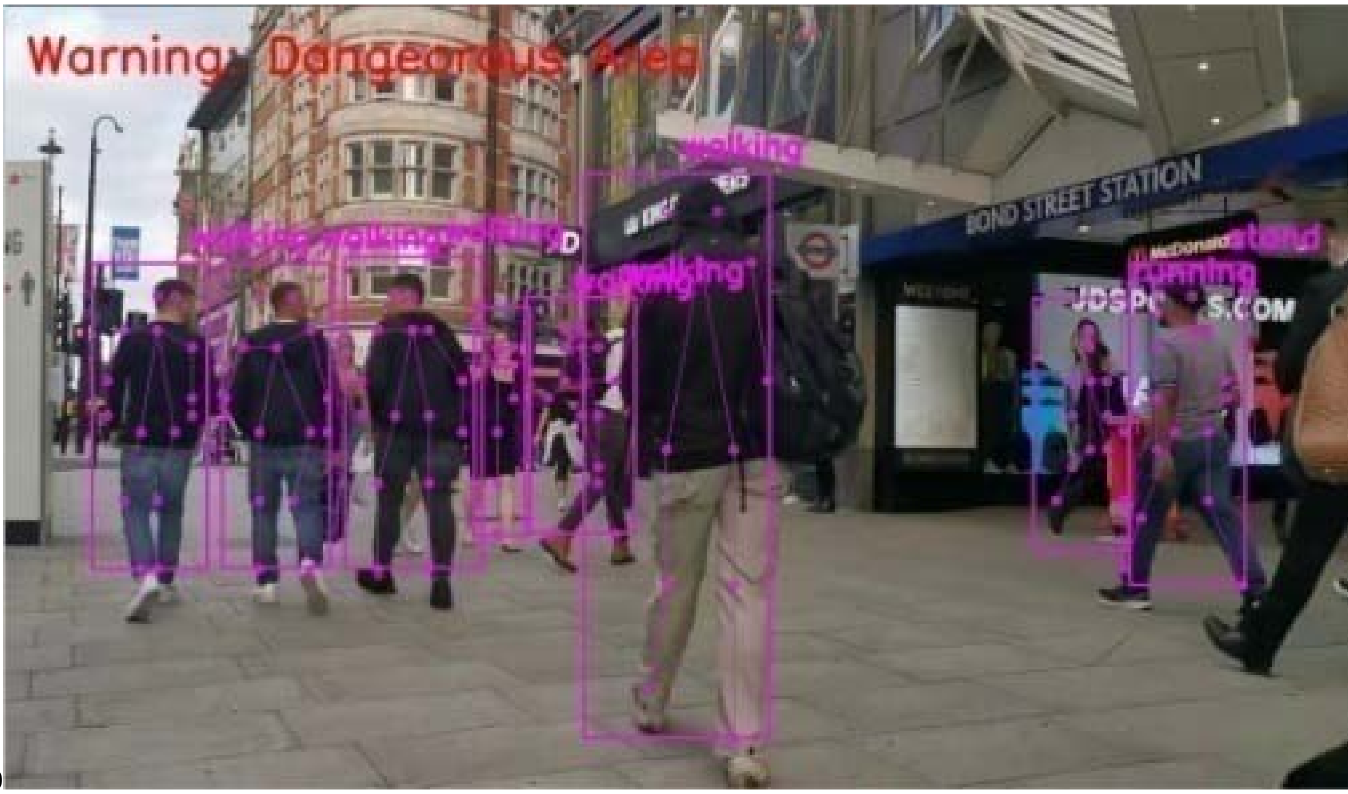


Figure 6: Fig. 9 :Fig. 10 :

1

Model	Dataset	MAP (%)
Our YOLOv3	Google Open Images	85.0%
Alexey AB YOLOv3	Pascal Voc	87.0%
R-CNN	Pascal Voc	53.2%

Figure 7: Table 1 :

[Balaji1 et al. ()] ‘A brief Survey on AI Based Face Mask Detection System for Public Places’. S Balaji1 , B Balamurugan2 , T Ananth Kumar3 , * , R Rajmohan4 , P Praveen , Kumar . *A brief Survey on AI Based Face Mask Detection System for Public Places*, 2021.

[Salma Kammoun Jarraya et al.] *A Deep-CNN Crowd Counting Model for Enforcing Social Distancing during COVID19 Pandemic: Application to Saudi Arabia’s Public Places*, Salma Kammoun Jarraya , Manar Maha Hamdan Alotibi , Salamah Ali1 . p. 2020.

[Angelini et al.] Federico Angelini , Student Member , Zeyu Ieee , Student Fu , Member , Yang Ieee , Senior Long , Member , Ling Ieee , Shao . *Senior Member, IEEE, and Syed Mohsen Naqvi, Senior Member, IEEE ”ActionXPose: A Novel 2D Multi-view Pose-based Algorithm for Real-time Human Action Recogniti on*,

[Angelini et al.] Federico Angelini , Student Member , Zeyu Ieee , Fu , Member , Yang Ieee , Senior Long , Member , Ling Ieee , Shao . *Senior Member, IEEE, and Syed Mohsen Naqvi, Senior Member, IEEE ”2D Pose-based Realtime Human Action Recognition with Occlusionhandling*,

[Yadav et al.] *Contrast limited adaptive histogram equalization based enhancement for real time video system*, Garima Yadav , Saurabh Maheshwari , Anjali Agarwal . p. 2020.

[Jiang and Learned-Miller ()] ‘Face Detection with the Faster R-CNN’. Huaizu Jiang , Erik Learned-Miller . *IEEE 12th International Conference*, 2017.

[Pethani and Vora] *FACE MASK DETECTION USING YOLO V3 ARCHITECTURE*, Nisarg Pethani , Harshal Vora . p. 2020.

[Akhter and Khadivizand ()] *IoT Enabled Intelligent Sensor Node for Smart City: Pedestrian Counting and Ambient Monitoring*, Fowzia Akhter , Sam Khadivizand . 2019. (Hasin Reza Siddiquei, Md Eshrat E. Alahi and Subhas Mukhopa dhyay)

[Melbourne-Pedestrian Counting System (2021)] *Melbourne-Pedestrian Counting System*, `https://www.melbourne.vic.gov.au/about-melbourne/research-and-statistics/city-population/Pages/pedestrian-counting-system.aspx` March 2021.

[Miller4 et al.] Harapan Miller4 , Abram L Harapanid5 ; Yihan Lu8 , Wagne . *Social distancing in response to the novel coronavirus (COVID-19) in the United States*, 6.

[Cao et al.] *OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields*, Gzhe Cao , Student Member , Gines Ieee , Student Hidalgo , Member , Tomas Ieee , Shih-En Simon , Yaser Wei , Sheikh . p. 2020.

[Nurul Iman Hassan et al.] *People Detection System Using YOLOv3 Algorithm*, Fadhlan Nurul Iman Hassan , Hafizhelmi Kamaru , Nooritawati Zaman , Md , Habibah Tahir , Hashim . p. 2020.

[Rahman1 et al.] *Post COVID-19 implementation of a bidirectional counter with reduced complexity for people counting application*, A S Rahman1 , S Yaakob1 , A R Razak1 , R Ramlee2 . p. 2021.

[Keniya and Mehendale] *Real-time social distancing detector using Social distancing Net19 deep learning network*, Rinkal Keniya , Ninad Mehendale .

[Zhang Jialin1 et al. ()] ‘Research on medical mask detection method based on fast Fourier transform and linear Gauss’. Zhang Jialin1 , Jiang Liu Yaofeng1 , Zhang Zhendong1 , Heng1 . *IEEE/ASME International Conference on Advanced Intelligent Mechatronics Hong Kong*, (China) July 8-12, 2019.

[Susanto et al. ()] ‘The Face Mask Detection For Preventing the Spread of COVID-19 at Politeknik Negeri Batam’. Febri Alwan Susanto , Riska Putra , Analia . *IEEE International Conference*, 2021. (Ika Kar lina Laila Nur)