WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Social Distancing Monitoring using YOLO Object Detection

Pradeep Mondal P, Rajath H R, Saloni Dhar, Yogendra Kumar V R, Prof. Asha M S, Dr. Saravana Kumar

Computer Science Engineering,Dayananda Sagar Academy of Technology and Management,Bangalore 560082,India

A B S T R A C T

WHO recommends social distancing as a preventive measure to minimize the spread of COVID-19 and other infectious diseases. All government and healthcare authorities recommend 2 meters as the optimal distance to practice social distancing in crowded areas like schools, markets, etc. The Project uses a Deep-Learning BasedNeural Network model on common CCTV footage to automatically detect people, track them and estimate inter-people distances. We use YOLO object detection framework to accurately track people, even in situations of occlusion, partial visibility and variations in lighting. The output can be used to enforce social distancing in crowded public places to prevent the spread of infectious diseases and can have further applications like identifying areas to prioritize disinfecting and cleaning..

Keywords:. —"Social Distancing; COVID-19; People Detection and Tracking; Deep Learning; Convolutional Neural Networks

I. INTRODUCTION

Social distancing refers to a precautionary action to prevent spread of infectious diseases, by minimizing the distance of people physical interactions in dense public places like schools, bus stations, etc. to reduce infection risk.

Artificial Intelligence (AI)has a major role in monitoring social distancing. Computer vision has become more important with various application like video surveillance, self-driving cars and more. It has beenused in solving various health related problemswhichhelps us to extract complicated features from the data. This information is used for understanding images by analysing the features.

As per WHO requirements, “the minimum distance between people must be at least 6 feet in order to effectively maintain social distancing”. This study helps in reducing the spread of the coronavirus by providing an automated solution for monitoring and detecting social distancingviolations among people using “Convolutional neural network (CNN) which is a neural network which has multiple convolutional layers which are used mainly for segmentation, image processing, classification”. The proposed model uses a YOLO based framework for bird’s eye transformation for accurate people detection and social distancing monitoring in various conditions.The model uses common CCTV security cameras for automated people detection, tracking, and inter-people distances estimation in the crowd.

II. LITERATURE SURVEY

“N. Dalal" [5] paper discusses methods used for detecting and tracking of people in a video footage by using the HOG featureshelps to train aSupport Vector Machine (SVM) classifier in a frame. Window of a particular size are cropped out from the image one by one in a Convolutional manner and Histogram of Oriented Gradients (HOG) features is used to train a Support Vector Machine (SVM) classifier for detecting human beings in any frame. HOG is used to train the SVM for human detection.

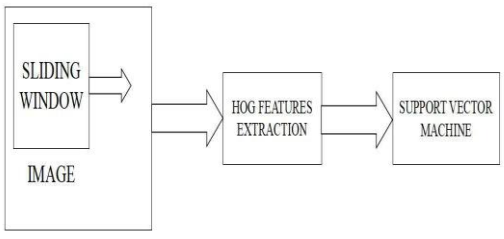

Fig 1. Block diagram for Human using HOG

Once the human in the frame are detected we need to track them. Initially we find the current position of each person in the frame and continuously compare them to the upcoming frames. k-means clustering is used to identify the location, the points are located for fixed iterations so that we can find a group which has least distance among them and path by connecting all points from previous frame to next frame.

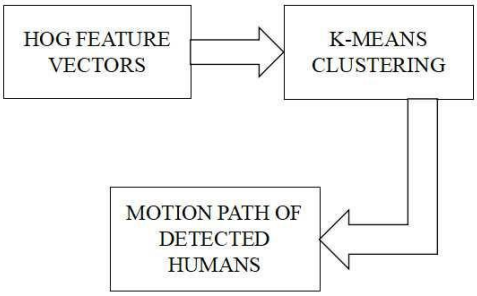

Fig 2. Block Diagram of Post Recording Tracking

"Ammar Abbas and Andrew Zisserman" [6] perform a bird eye view transformation which gives us a top view of a scene. We apply this method to an image in order to transform the perspective of the image taken to a top view of the image. To help in more accurately measuring the distance between people. This requires only four parameters to be specified (the vanishing point and vanishing line). We show that a CNN trained using our four scalar parameterisations to detect horizon and transform the image.
In any given applications that use a single camera to understand the 3D layout of the scene is always difficult.

Fig 3(a). Input Scene from single camera

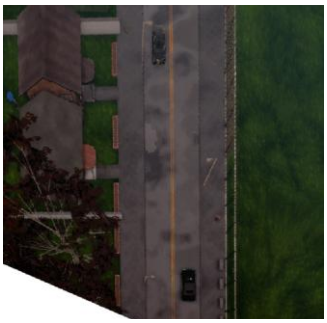

Fig 3(b).Output Image after Bird's eye transformation

Hence, we can use bird eye transformation which can be used for accurate distance prediction as the image is mapped to the top-view of the scene. This method is used in various computer vision applications like video surveillance for detection and tracking. First the image is tilted to get a horizontal view of the scene, then we identify the horizon and vanishing points.
The CNN is initially trained with pre-trained weights from ImageNet classification which contains different scene. Once the CNN is trained, we can use the model to convert the scene to top-view.

"Mahdi Rezaei, Mohsen Azarmi" [1] use a 3-stage model for social distancing which includes detection of people, tracking and distance estimation between the people. This can be used to monitor social distancing in any given place where any type of CCTV Camera is available with real time analysis. The model includes a DNN detector for extracting features and prediction and location of objects. The DNN model has to discover patterns in objects for faster prediction.The proposed method helps in reducing the spread of the coronavirus by providing an automated solution for monitoring and detecting social distancingviolations among individuals.

Fig 4. Examples of people detections

III. RESEARCH AREA

Artificial Intelligence (AI) has a major role in monitoring social distancing. Computer vision has become more important with various application like video surveillance, self-driving cars and more. It has beenused in solving various health related problems whichhelps us to extract complicated features from the data. This information is used for understanding images by analysing the features.

"Convolutional neural network (CNN) which is a neural network which has more than two convolutional layers which are used mainly for segmentation, image processing, classification**. "**

IV. METHODOLOGY

We used a three-stage model for social distancing which includes detection of people, tracking and distance estimation between the people. This can be used to monitor social distancing in any given place where any type of CCTV Camera is available with real time analysis. The model includes a DNN detector for extracting features and prediction and location of objects. The DNN model has to discover patterns in objects for faster prediction.

The detector is of two types: Two stage detector and One stage detector.

Two-Stage Detector initially extracts a set of objects by selective search and then feeds it to the CNN model which later classifies the object. Even though the Two stage detectors havehigh accuracy, systems with limited computational resources are not recommended. One stage detectoruses a single detection process called YOLO (You Only Look Once). Using regression analysis, it calculates the coordinates of boundary boxes and interprets their probabilities. It checks the probability of an object existence in each cell of the grids by mapping the image pixels to the enclosed grids, offering substantial improvements in speed and efficiency.

After detection, the next step is to track people and initialize ID for each individual. For tracking we used the combination of SORT tracking technique and Hungarian algorithm. The "Simple Online and Real-time or SORT" tracking technique uses a Kalman filter which predicts the position based on current and next frame and modelling of the human movement. Hungarian algorithm helps assign Unique ID by examining the person in the current frame and previous frame.

Stereo-vision technique is popular for distance estimation but it is not a feasible approach in our project as we want an efficient solution in real time for existing CCTV cameras. We therefore use a monocular solution, which has the drawback of giving unrealistic pixel-distances between the objects due to perspective effect from the projection of a 3D scene into a 2D perspective image plane.

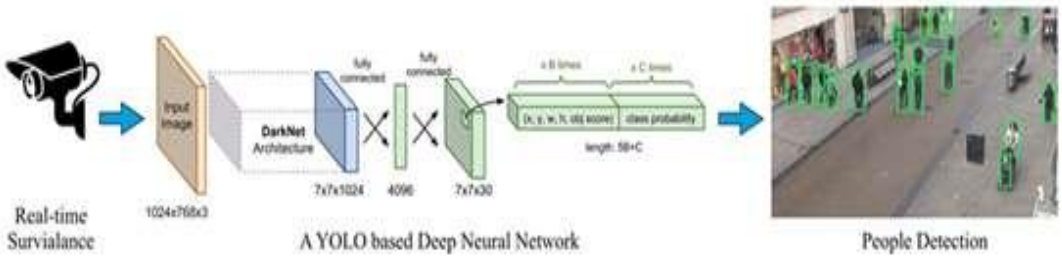

Fig 5. Model for People Detection

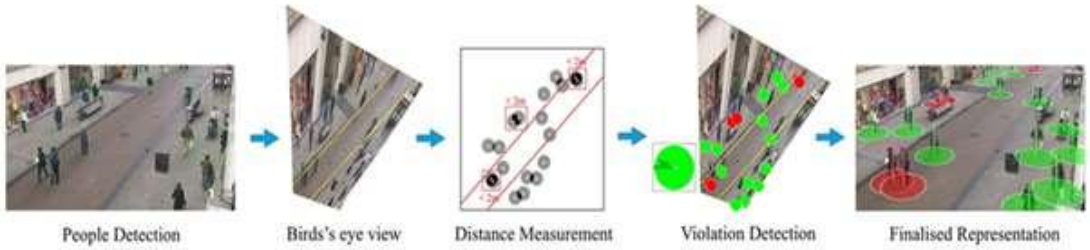

Fig 6.People Detection and Image Transformation

For each frame in CCTV video, we pass the image to the model for People Detection and convert the scene to bird eye view for position identification then we calculate distance between people, if the distance is less than threshold value then red boxes will be drawn around them marking that they are not social distancing, if people are following social distancing then a green box will be drawn around them. We can then combine all the modified frames into a video to playback the result.

Fig 7. Tracking people

The proposed method can be used in bright environment where there is sufficient light for the camera to captured and therefore, we have more information in the image for image processing which results in more accurate results.

In the night/low-light environment we can use thermal camera to capture the source of heat in the image. Generally, the temperature of human body is 37^{o}C as the human move, they leave a trail of heat which can be used to track people movement, if the person doesn't move for long time the heat map turns to redin that particularlocationsand then we can identify the places which are most risk areas.

Data is obtained from people movement and tracking;this would be more usefulfor analysis of the density and thelocation of the people where particularly the social distancing measures are being violated

Fig 8. Long-term 2D heat map

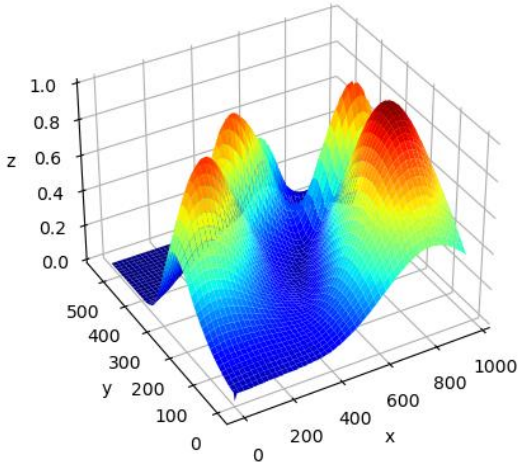

Fig 9.Long-term 3D heat map

This helps to identify zones based on risk, or bring more restrictions to certain areas and limited movement is allowed toparticular zones or re-design the scene of the environment to make it a saferplace. Based on the nature of environment and situation the restriction rules are varied. We can categorize people into various categories like safe, high-risk.

V. CONCLUSION AND FUTURE WORK

The proposed method helps in reducing the spread of the coronavirus by providing an automated solution for monitoring and detecting breaking rules of social distancing among individuals. The proposed people detector model is able to detect and track stationary and moving people in public places and monitor and maintain social distancing standards in COVID-19 era and later. This Method is applicable in various placewhich uses any type of CCTV cameras.

This helps to identify zones based on risk, or bring more restrictions to certain areas and limited movement is allowed toparticular zones or re-design the scene of the environment to make it a saferplace. Based on the nature of environment and situation the restriction rules are varied. We can categorize people into various categories like safe, high-risk.

VI. ACKNOWLEDGEMENT

- We extend our sincere thanks to Dr. Saravana Kumar, Dept of CSE, DSATM
- We also would like to thank our project guide Assistant Prof. Asha M S, Dept of CSE, DSATM

REFERENCES

[1] Mahdi Rezaei, Mohsen Azarmi, "DeepSOCIAL: Social Distancing Monitoring and Infection Risk Assessment in COVID-19 Pandemic"

[2] World Health Organisation. WHO Corona-viruses Disease Dashboard (August 2020). Available at https://covid19.who.int/table.

[3] WHO Director, Generals. Opening remarks at the media briefing on COVID-19 (2020). WHO generals and Directors speeches.

[4] Nida M. Zaitoun, Musbah J. Aqel, "Survey on Image Segmentation Techniques"

[5] N. Dalal, B. Triggs and C. Schmid, "Human Detection and Tracking for Video Surveillance"

[6] Ammar Abbas and Andrew Zisserman, "A Geometric Approach to obtain a bird's eye view from an image"