A Seminar Report
on

SOCIAL DISTANCING DETECTION USING COMPUTER VISION

By

GAURI KAKULTE
Roll Number: C33342

TE (Computer Engineering)

Under the guidance of
PROF. P. A. DEOLE

Department of Computer Engineering
Smt. Kashibai Navale
College of Engineering,

Vadgaon (Bk.), Pune-411041

Accredited by NAAC

Affiliated to Savitribai Phule Pune University, Pune

2022-23

CERTIFICATE

This is certified that the Seminar Report entitled

SOCIAL DISTANCING DETECTION USING COMPUTER VISION

Submitted by

Gauri Kakulte

Has successfully completed his Seminar Report and Presentation under the supervision of Prof. P. S. Deole for the partial fulfillment of Third Year of Bachelor of Engineering, Computer Engineering of Savitribai Phule Pune University. This work has not been submitted earlier elsewhere for any degree.

Prof. Prgati A. Deole
Guide

Dr. R. H. Borhade
Head, Computer Engineering Department

External Examiner

Dr. A. V Deshpande,
Principal,
Sinhgad College of Engineering, Pune

Abstract

Covid-19 is a highly contagious disease that is caused by a new coronavirus called the SARS-CoV-2, the pandemic has affected the entire world causing more than a million deaths. To control the spread of this disease it is imperative to maintain distance between people because it can't be possible to know at all times if a person is infected or not. Social distancing while maintaining a minimum one meter distance is recommended by WHO. The approach aims at using Computer Vision enforcing this social distancing in public places by constantly monitoring the distance between people by a video feed and alerting the responsible person so as the required actions can be taken. This video feed can be very easily collected by pre-existing infrastructure across the public places such as CCTV Cameras. This would allow us to constantly check the distance between any two individuals in the public place. The scalability of this solution is very high, as cameras are installed at almost all public places.

Computer vision is one of the technologies that aim at digitally perceiving the real world at a higher level through digital images and videos. Object detection, a subset to computer vision is one of the prominent techniques in this area of research. Object detection is basically an algorithm based on either machine learning or deep learning approaches employed for classification of elements in diverse classes and localization in the image. R-CNN, Fast R-CNN, YOLO are the techniques in the trend which facilitates the developer in accomplishing the task of detecting an object in the image. These techniques train and compute the parameters of the model in reduced hence increase performance as compared to the traditional object detection techniques.

The model presents an approach for social distancing detection using deep learning to evaluate the distance between people to mitigate the impact of this coronavirus pandemic. The detection tool was developed to alert people to maintain a safe distance with each other by evaluating a video feed. The video frame from the camera was used as input, and the open-source object detection pre-trained model based on the YOLO algorithm was employed for detection. YOLO stands for You Only Look Once, this algorithm is used for Object Detection as well as Object Tracking, it is extensively used in field of object detection computer technology which is related to computer vision and image signal processing. It is used to detect instances of class objects like humans, buildings, cats, dogs, etc. It is used in locations like face detection, pedestrian detection, video surveillance,

image retrieval etc. The distance between people can be estimated and any noncompliant pair of people in the display will be indicated with a red frame. The developed technique can be further developed as a detection tool in real time application.

Keywords:

Social Distancing, Computer Vision, Object Detection, Distance Detection, YOLO

Acknowledgement

I take this opportunity to express our profound sense of gratitude and respect to all those who directly or indirectly helped me through the duration of this work. It gives me great pleasure to present a seminar on **"Social Distancing Detection using Computer Vision"**. I am extremely obliged to our seminar guide **Prof. P. A. Deole** in Department of Computer Engineering for helping and providing me with proper guidance. Timely suggestions made it possible for me to complete this seminar successfully.

I will fail in my duty if I don't acknowledge a great sense of gratitude to the head of computer engineering department **Dr. R. H. Borhade** and entire staff member of our department for their cooperation.

Gauri Kakulte(C33342)

List Of Figures

List Of Tables

TABLE OF CONTENTS

Chapter-1 Introduction

Covid-19 is caused by the SARS-CoV-2 virus. It can cause different ranges of problems ranging from mild fever to sever respiratory disease, which can even lead to death. The corona or covid-19 pandemic has created so much problem all across world and even after countries succeeding in vaccinating 50% of their population, still covid-19 spread is not stopping. People are being advised to wear masks and maintain social distance among them but they seem too careless for the same. Governments across the globe have made it mandatory to follow a minimum distance of 1 meter among them and to make sure that it is being followed this paper can provide a robust and fast solution. Especially we have to monitor those places in which there is predicted large number of people like hospitals, public gathering places, religious places, fairs etc. In these places there is less chances that people will follow social distancing rules hence we need to implement this model in these areas. With the help of CCTV cameras or files we can detect different people distances and henceforth can inform the designated authority about the same. This way people can be fined and punished accordingly.

1.1 PROBLEM STATEMENT

The pandemic has risen as a common enemy for the whole of humanity, killing people and destroying economies, yet it's not clear when the virus will be cured. Economies are being opened again, and everyday life activities are resumed, but this may eventually lead to another wave of the virus if the people don't abide by social distancing and safety regulations. To control the spread of diseases which are infectious, we use many measures but one major is social distancing which is a nonpharmaceutical way of avoiding this disease. Thus we aim to propose a model in which we perform real-time analysis and detection of people to see if social distancing is being followed or not.

1.2 OBJECTIVES AND OUTCOMES

The approach of developing a model for social distancing is focusing on surveillance of public places and detecting whether the people are maintaining social distancing or not.

There are different types of algorithms available for Object Detection as well as Object Tracking, YOLO (You Only Look Once) stands out from all the other present currently. It is extensively used in field of Object Detection Computer Technology which is related to Computer Vision and Image Processing. This approach uses YOLO for tracking the people in the frame, by using a video stream as input and YOLO algorithm, we can detect violation if people are not following social distancing.

Chapter -2 Literature Review

2.1 INTRODUCTION

Monitoring Social Distancing from surveillance video involves solving the pre-existing task of people detection in video/image data which is an extensive field of research. Over the years this task has been attempted to solve by using object detection techniques. The below section explains the history, approaches, and various significant model architecture used for human detection or in general object detection. Object detection is an underlying problem of computer vision that forms the base of many other tasks, such as autonomous driving, object tracking, human pose estimation, optical character recognition, etc. The simple task of detecting objects for humans is extremely hard for machines because of the large number of complications ranging from intraclass variations, objects under different viewpoints, illuminations, object rotation, scale changes, detection in crowded environments, dealing with occlusions, etc.

2.2. LITERATURE SURVEY

The simple-looking task of detection comes with lots of difficulties and challenges such as small size pedestrians in the frame, hard negatives, dense and occluded images, and the overhead complexity and computation cost for real-time detection of the pedestrian.

The traditional method for human or pedestrian detection was the Histogram of Oriented Gradients (HOG). It was an enhancement of the scale-invariant feature transform and shape contexts methods. To overcome the challenges described above for object detection the HOG descriptor performs dense matrix computation and uses overlapping local contrast normalization for improving robustness and accuracy of the model. Even though HOG can be used for general object detection its main motive was to achieve pedestrian detection.

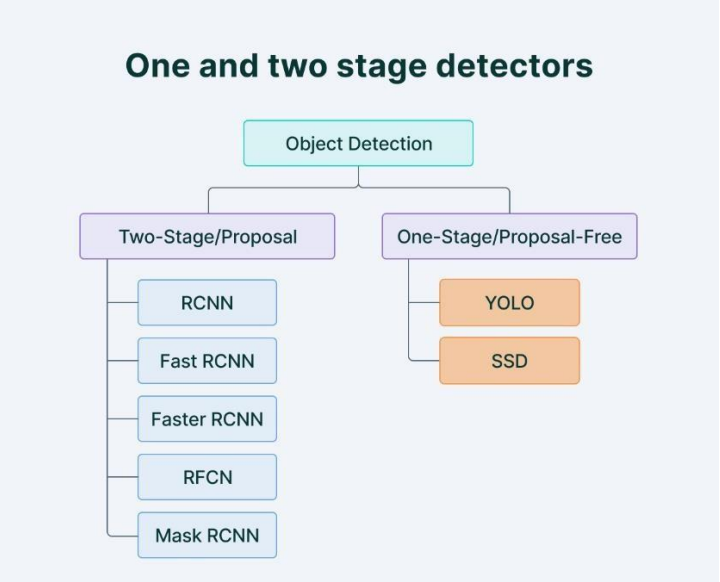

Fig. 2.1 Classification of Object Detection Techniques

High dimensional features of images were grasped or understood by the models with the introduction of deep Convolutional Neural Network in the year 2012. Two stage detectors - RCCN family Region-based Convolutional Network or RCNN marked the inception of object detection techniques by proposing region proposals by the Selective Search algorithm. Single stage detectors - YOLO (You Only Look Once) is the first one-stage detection model for object detection which unified different stages of object detection tasks into one single network.

Region with CNN (RCNN)

The algorithm proposed by Girshick et al employed selective search to extract just 2000 regions from the entire input image. These regions are referred to as region proposals. The subsets of these regions are identified in the image and are employed for classification of the objects in the regions. Therefore, instead of classifying big number of regions, just 2000 regions can be worked with. All these regions are generated by employing the selective search algorithm. The last layer of the CNN consists of the features which are extracted from the entire image and these features are further fed to a classification algorithm like support vector machine. This algorithm classifies the objects lying within the region proposal network.

Fast R-CNN

R-CNN model took a huge amount of time to train the network. Girshick et al built another faster object detection algorithm known as Fast R-CNN to circumvent this problem. The input image is fed to the CNN algorithm which further generates convolutional feature map. The model takes the entire image as an input with a set of different object proposals. The network firstly processes the entire input image using more than one convolutional and max pooling layers so as to produce the convolution feature map for the entire image. The region proposals are generated using an algorithm such as Edge Boxes. Further, for each of the object proposals, a region of interest (RoI) pooling layer converts the object proposals into a feature vector of a particular size.

YOLO

In contrast to the previous object detection algorithms, instead of complete image, the network looks at the portions of the image possessing immense probabilities that the object is present. You only look once or YOLO is an algorithm for object detection, which has features much different from the previous algorithms. This algorithm deploys bounding boxes and classifies the image hence predicting possibilities for these boxes using a unit convolutional network. YOLO works on input image by splitting it into an S x S grid, in each of which “m” bounding boxes are marked. In these bounding boxes, the network predicts class probability and offset values. The bounding boxes exceeding the threshold value for class probability are selected and are instrumental in locating the object inside the image. YOLO processes the image as fast as 45 FPS than the rest algorithms for object detection.

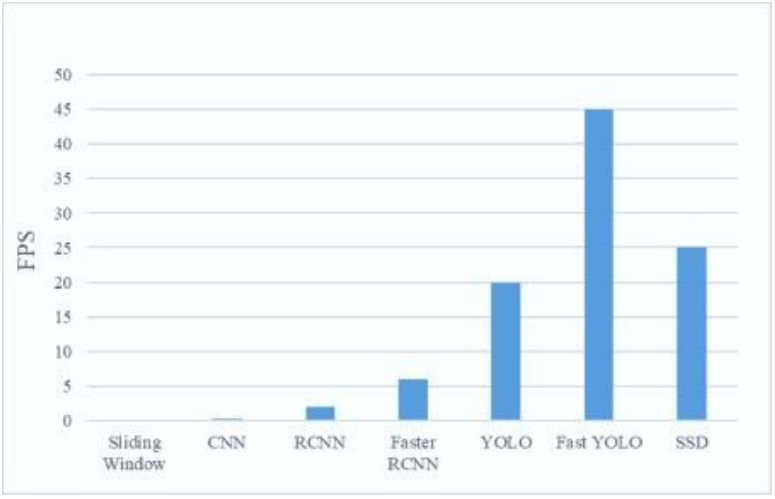

Fig. 2.2 FPS Chart for Object Detection Techniques

Unlike R-CNN, it is able to predict with a unit network evaluation instead of thousands for one image. This results in the algorithm being 1000x and 100x faster than R-CNN and Fast R-CNN respectively.

R-CNN, Fast RCNN and YOLO use different classification methods. YOLO can run in real time and is fast but the technique struggles to find small objects in a group like detecting a bird in the flock of birds. RCNN and Fast RCNN are slower than YOLO but can detect small objects. YOLO is good at regression than classification. YOLO has difficulty in classifying small objects. Both RCNN and Fast RCNN fails to perform real time detection but YOLO can perform real time classification with good speed.

Table 2.1 Summary of R-CNN, Fast RCNN and YOLO

R-CNN	Fast RCNN	YOLO
Based on classification	Based on classification	Based on regression
Uses 2000 conv Nets for each region	Uses single deep convent	Uses single convolution network for entire image
Uses selective search algorithm	Uses selective search algorithm	---
Needs 49 seconds to test one image	Needs 2.3 seconds to test one image	Needs less than 2 second to test on one image
Slow, cannot be implemented real time	Faster than RCNN	Can run real time
Uses SVM for classification	Uses softmax for classification	Uses regression for classification
Produces bounding boxes	Produces bounding Box regression head and classification head	Produces bounding box, prediction contextual concurrently
Can find small objects	Can find small objects	Struggles to find small objects that appear in groups

Chapter-3 Methodology

3.1 INTRODUCTION

This social distancing detection tool was developed to detect the safety distance between people in public spaces. The Computer Vision techniques are employed in this work. Initially, an open-source object detection network based on the YOLO algorithm was used to detect the pedestrian in the video frame. From the detection result, only pedestrian class was used and other object classes are ignored in this application. Hence, the bounding box best fits for each detected pedestrian can be drawn in the image, and these data of detected pedestrians will be used for the distance measurement.

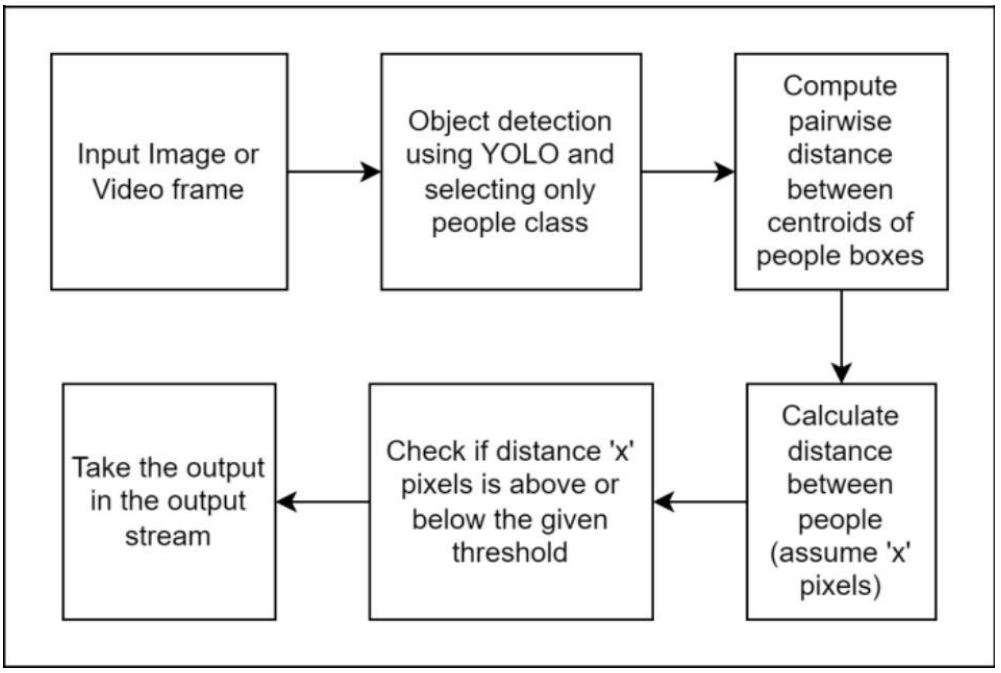

Fig. 3.1 Process Diagram for Social Distancing Detection between people

The distance between pedestrians can be measured and scaled. Depending on the preset minimum distance, any distance less than the acceptable distance between any two individuals will be indicated with red lines that serve as precautionary warnings.

3.2 SYSTEM REQUIREMENT ANALYSIS

3.2.1 COMPUTER VISION

Computer Vision is an interdisciplinary scientific field which is evolving with time that deals with how computers and other related machines can gain high-level understanding from digital images or videos. It includes different methods which includes concepts like acquiring, processing and analyzing digital images and videos and it is also used for extraction of high-dimensional data from real world scenarios in order to produce numerical or symbolic information which can be used by computers to process upon.

3.2.2 OPENCV

OpenCV is a special library mainly aimed at real-time computer vision. It is a cross-platform and free to use as it is open source. OpenCV also has special features like GPU acceleration for real-time operations in machines like computers. OpenCV CPU version is 9x faster. If we use DNN module implementation of OpenCV then it is faster. To judge speed of OpenCV, OpenMP takes 2 seconds when used with Dark net but OpenCV implementation takes mere 0.22 seconds! It has various use cases like object detection, landmark detection, face detection, tracking eye movements, extracting 3D models of objects, tracking modern camera movements etc.

3.3.3 YOLO FOR DETECTION

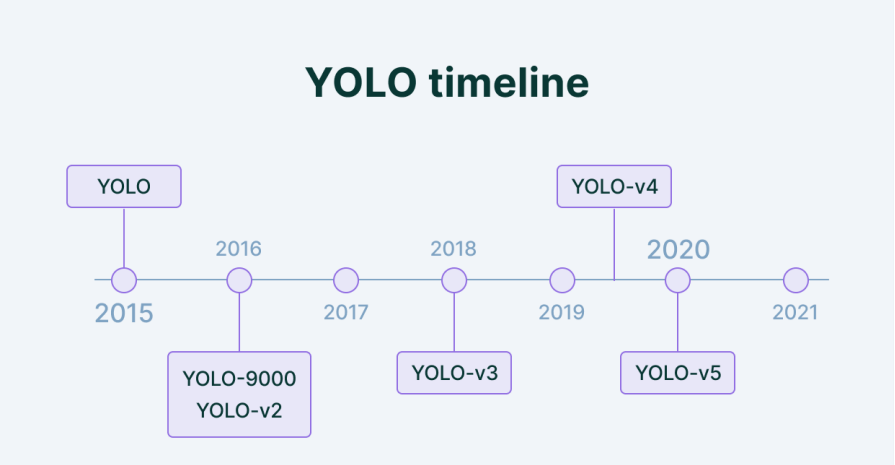

Fig. 3.2 YOLO Timeline

YOLO stands for You Only Look Once. It is extensively used in field of object detection computer technology which is related to computer vision and image signal processing. It is used to detect instances of class objects like humans, buildings, cats, dogs, etc. It is used in locations like face detection, pedestrian detection, video surveillance, image retrieval etc.

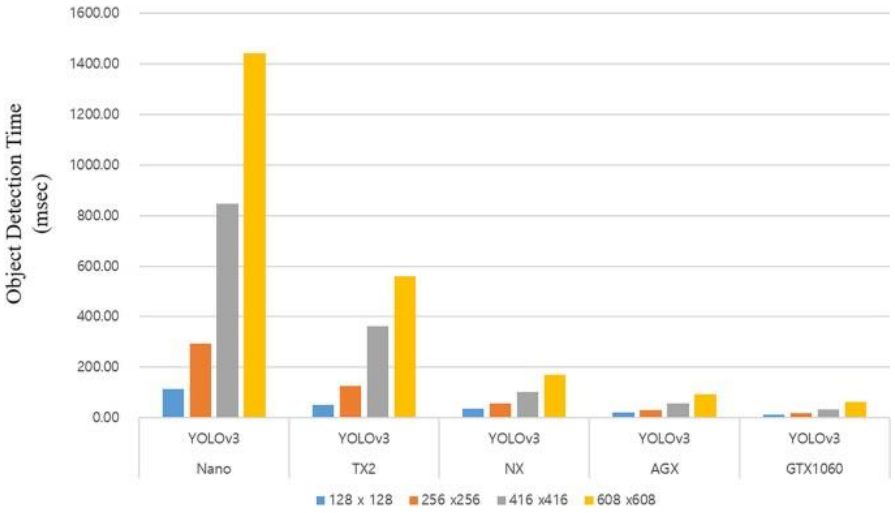

Fig. 3.3 YOLO Versions Object Detection Time

3.3.4 PROPOSED SYSTEM

Our proposed system is a four-level module that focuses on things like human identification, their tracking and then after their distance detection (mainly humans using YOLO object detection algorithm) and calculation using computer vision and deep learning by using OpenCV.

We will follow the following criteria given below: -

i. Pass the input which will be image or video and we will pass it frame by frame.

ii. Detect the human object in the image or video frame using YOLO object detection.

iii. Compute the pairwise distances between centroids of boxes drawn on the people.

iv. Check if the distance between any two persons is ‘x’ pixels and check whether it is above or below the given threshold.

v. Generate the output in the output stream.

3.3 SYSTEM DESIGN WORKFLOW

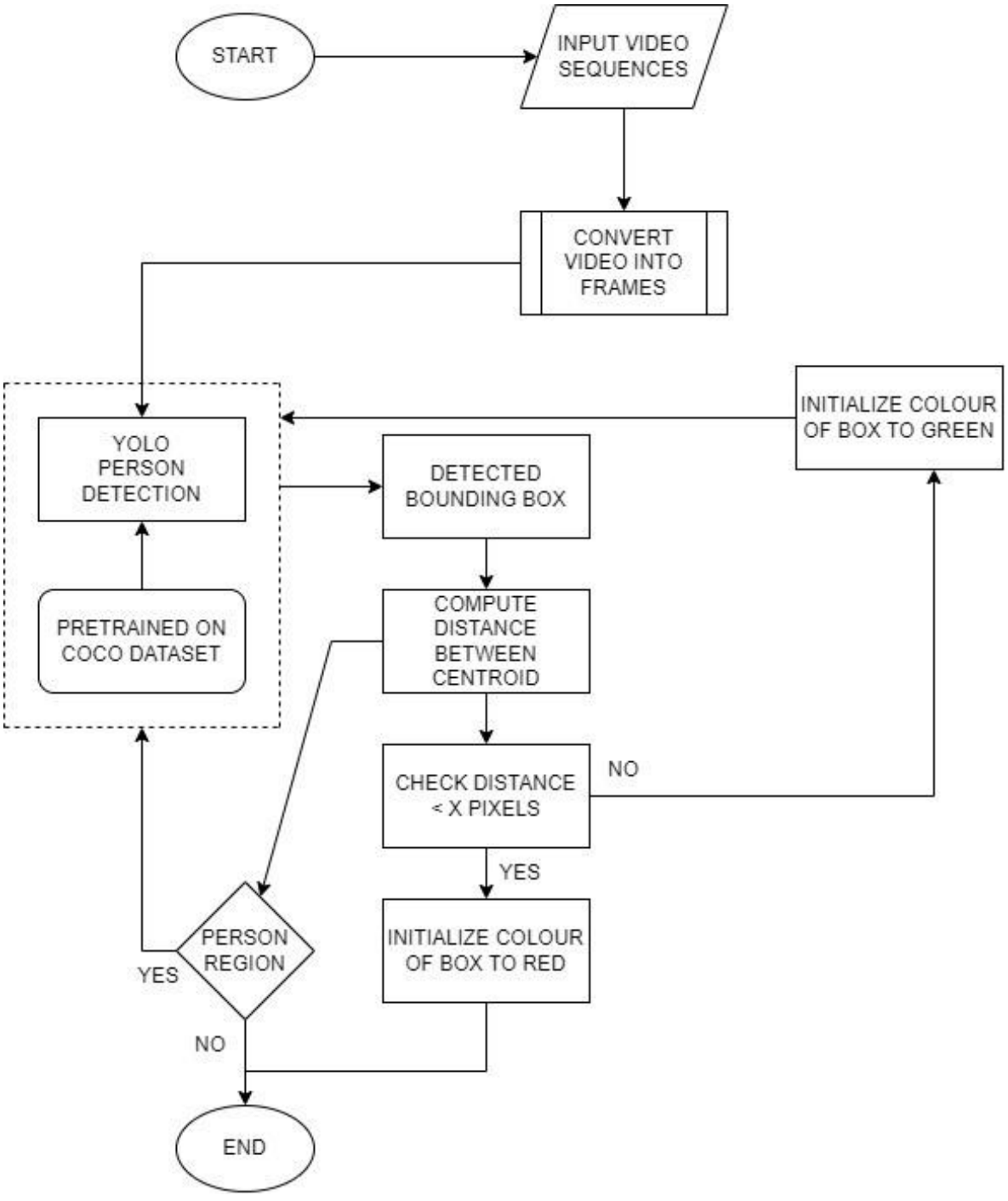

Fig. 3.4 Flow Chart for Social Distancing Detection using YOLO pretrained on COCO

3.3.1 IMAGES / VIDEO INPUT

In this section, capturing and supply of Digital Images in form of Images or Video Stream is done as an input to the model. It can either use Realtime Camera to capture frame from the scene or extract frames from Video Scenes.

When we will use videos then no. of frames per seconds will be much more hence it will be difficult to process hence, it would be beneficial to use high performance GPUs as the processing of the Digital Images is done at faster rates.

3.3.2 OBJECT DETECTION

When it comes to Object Detection and Object Tracking, the YOLO model is considered one of the state-of-the art object detectors which can be demonstrated to provide significant speed advantages will suitable for real-time application.

Now we are going to detect people object using YOLO. The People Class filter is used to filter only People out of many classes available. The YOLO trained on the COCO dataset which consists of 80 labels including People Class. In this model, the only box coordinates, object confidence and People Object Class from detection result in the YOLO model are used for detection.

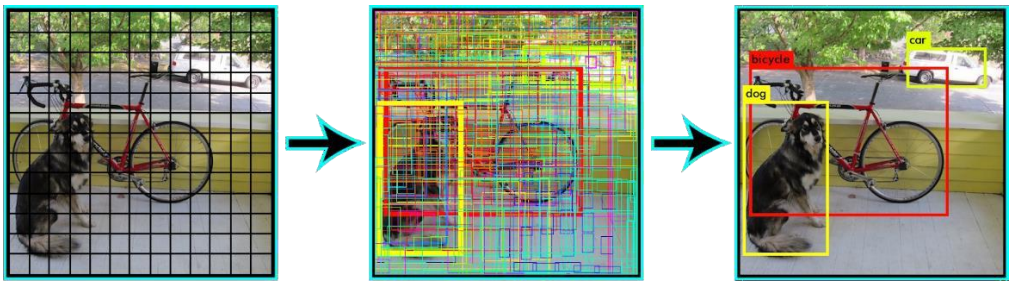

Fig. 3.5 YOLO Object Class Detection

3.3.3 DISTANCE MEASUREMENT

After detecting people available in the frame provided as input, the model is going to detect the person and draw different bounding boxes on the person. So, this can create many boxes around the person and this can be avoided by using an algorithm called as Non-maximum suppression (NMS). NMS is going to take into consideration the box which will have highest

probability of covering that person and hence only one box will be drawn around the person. After this step we are going to use this box for further calculations :

(I) After the bounding box is drawn, we are going to find centroid of bounding box.

(II) After finding centroid, we are going to find distance between two these two centroids which will find the distance between the bounding boxes.

(III) After finding distance we are going to compare this distance with 'x' pixels distance already predefined and check whether it is above or below the mentioned limit.

(IV) Add the total combinations of boxes which are not following social distance norms and put it in some storage class. We will use this afterwards to display real time number of people disobeying social distancing rules.

(V) After storing the number of comparisons falling below threshold value, we are going to add these up and display it under violation label. The output stream can be anything like computer screen, monitors, web displays, TVs, etc.

3.3.4 ALERT SYSTEM DETECTION

After storing the number of comparisons falling below standard set value, these are added up and displayed under Violation label. The output stream can be anything like computer screen, monitors, web displays, TVs, etc.

In this part, the system gives alert based on distance is greater than set value.

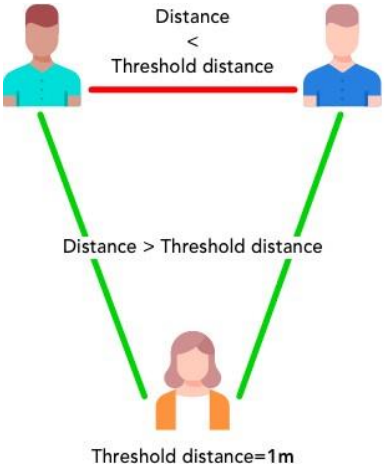

Fig. 3.6 Alert based on distance

Chapter-4 Future Scope

The algorithm detects the persons whose full body is included in the frame, which degrade the accuracy in crowded scenes, which could be improved in the future. It can't detect duplicates when someone re-enters into the frame the person has counted again as a new person. The algorithm could be extended to overcome this problem. The training could be extended to detect persons using the head for better handling of crowded inputs. Slightest improvements in these algorithms can improve accuracy.

We can make a dashboard to display various different parameters along with number of people disobeying social distancing norms. We can display data like frequency of having maximum population in specific time intervals and getting the time in which there is maximum violation of social distancing. Also, we can improve the camera calibration or take that images or videos as input in which people are clearly visible in frame to get more accurate results. We used boxes to calculate centroids and then calculated distances between them but if we want more accuracy, we can take topdown approach instead of boxy approach. This way we can calculate circle centers and find distances between them so that we can obtain better distance approximations among different combinations of people.

Chapter-5 Conclusion

In these times, social distancing combined with other primary sanitation measures is very crucial for controlling the spread of the Covid-19. Though, social distancing helps to keep the spread of the virus in check, individual precautions are equally vital to reduce the spiking rise in the number of fatalities. We have devised a methodology to control the spread of novel coronavirus by utilizing the modern technology of Computer Vision.

RCNN, Fast RCNN and YOLO are the common techniques employed for object detection. RCNN and Fast RCNN are slower than YOLO but can detect small objects. YOLO is good at regression than classification. YOLO has difficulty in classifying small objects. Both RCNN and Fast RCNN fails to perform real time detection but YOLO can perform real time classification with good speed. The choice for the type of object classification algorithm employed depends on type of data set, type of images, amount of training/testing time and the application requiring detection of object and the type of object.

Chapter-6 References

[1] Priyanka Malhotra and Ekansh Garg, “Object Detection Techniques: A Comparison”, IEEE 7th International Conference on Smart Structures and Systems ICSSS, 2020

[2] Sneha Madane and Dnyanoba Chitre, “Social Distancing Detection and Analysis through Computer Vision”, 6th International Conference for Convergence in Technology (I2CT), 2021

[3] Juhi Shah, Mahavir Chandaliya, Harsh Bhuta, Pratik Kanani, “Social Distancing Detection Using Computer Vision”, Fifth International Conference on Computing Methodologies and Communication (ICCMC), 2021

[4] Yew Cheong Hou , Mohd Zafri Baharuddin, Salman Yussof, Sumayyah Dzulkifly, “Social Distancing Detection with Deep Learning Model”, 8th International Conference on Information Technology and Multimedia (ICIMU), 2020

[5] Abdalla Gad, Gasm ElBary, Mohammad Alkhedher, Mohammed Ghazal, “Vision-based Approach for Automated Social Distance Violators Detection”, International Conference on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT), 2020

[6] Xinrui Zou, “A review of Object Detection Techniques”, International Conference on Smart Grid and Electrical Automation (ICSGEA), 2019

[7] Myint Myint Sein, Khaing Suu Htet, Ken T. Murata and Somnuk Phon-Amnuaisuk, “Object Detection, Classification and Counting for Analysis of Visual Events”, IEEE 9th Global Conference on Consum er Electronics (GCCE), 2020

[8] Wen Boyuan and Wu Muqin, “Study on Pedestrian Detection”, IEEE 6th Global Conference on Computer and Communications, 2020

[9] Suyash Paradkar, Shreya Tekade, Komal Ganatra, Saloni Sarwade and Prof. Rahul Bhole, “Social Distance Detector Using Computer Vision, OpenCV and YOLO Deep Learning Algorithm”, International Research Journal of Engineering and Technology (IRJET), 2022

[10] Sheshang Degadwala, Dhairya Vyas, Harsh Dave and Arpana Mahajan, “Visual Social Distance Alert System Using Computer Vision & Deep Learning”, Fourth International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020

[11] Marco Cristani ,Alessio Del Bue, Vittorio Murino, Francesco Setti, Alessandro Vinciarelli, “The Visual Social Distancing Problem”, 2020

[12] Chandan G, Ayush Jain, Harsh Jain, Mohana, “Real Time Object Detection and Tracking Using Deep Learning and OpenCV”, Proceedings of the International Conference on Inventive Research in Computing Applications (ICIRCA 2018)

[13] Emre DANDIL, Kerim Kürşat ÇEVİK, “Computer Vision Based Distance Measurement System using Stereo Camera View”, 2020

[14] Chengji Liu1 ,Yufan Tao1 ,Jiawei Liang1 ,Kai Li1 ,Yihang Chen1 ,“Object Detection Based on YOLO Network”, 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC 2018)

[15] Md. Bahar Ullah, “CPU Based YOLO: A Real Time Object Detection Algorithm”, 2020 IEEE Region 10 Symposium (TENSYMP), 5-7 June 2020