ISSN No:-2456-2165

Covid-cop (Social Distancing Violation Detector)

Asst. Prof. Rumana Shaikh (*Author*)
Computer Science and Engineering
Parul Institute of Technology
Vadodara, India

Bharat Singh (*Author*)
Ramesh Singh Rajpurohit (*Author*)
Bharat Singh Rajpurohit (*Author*)
Gaurav Chouhan (*Author*)
Computer Science and Engineering
Parul Institute of Technology
Vadodara, India

Abstract:- COVID-19 is currently threatening human life, health, and productivity. Population vulnerability grows as a result of a lack of effective remedial agents and a scarcity of vaccines against the virus. As a result, social distancing is considered an adequate precaution (norm) against the pandemic virus's spread. The danger of viral propagation can be decreased by avoiding physical contact between people. This paper gives overview about how to automate the task of monitoring social distancing using input video and images, which is motivated by this idea. To distinguish humans from the background, this methodology employs the YOLO v4 model for detecting objects. With respect to mean average precision (mAP), frames per second (FPS), and loss values given by object classification and localization in input frame, the YOLO v4 model's results are compared to those of other popular models, such as FRCNN and SSD. The pairwise vectorized L2 norm is computed using the 3D feature space obtained by using the bounding box's centroid coordinates and dimensions. Random forest regression is used to predict violations using data collected from the Yolo v4 model. According to the results of the experimental analysis, YOLO v4 and random forest regression produced the best results with balanced mAP and FPS score to monitor and predict future possibility of infection using generated data.

Keywords:- *m YOLO-v4, Object Detection, OpenCV, Random Forest regression.*

I. INTRODUCTION

Since the COVID-19 pandemic swept the globe, governments around the world have taken tough but necessary measures to stem its spread. Normal day-to-day activities came to a complete halt as a result of this. When the curve flattens in several countries after months of lockdown, the community becomes agitated. Relevant authorities, such as the World Health Organization (WHO), have issued guidelines to limit people's exposure to the virus. Wearing masks and keeping a distance of 3 feet (1 meter) from another person are two safety precautions that people are encouraged to take.

Following social distancing and wearing a mask is a regulation measure to reduce SARSCoV-2 spread, according to a new report, because individuals with mild or no symptoms may accidentally transmit crowd infection and spread the virus to others. Consider data-based and numerical-based models, both of which are constantly useful decisions. Social distancing has been shown to be an effective and productive measure in preventing the spread of the corona-virus. As the name implies, it advocated for people to maintain physical distance from one another, reducing close contact and thus limiting virus spread. Existing works will be used to improve the proposed current system. By providing a technical analyzer tool to monitor schools, colleges, public areas, hospitals, and workplaces, the system to be developed aims to encourage people to follow social distancing. to detect social distance violations and generate real-time warnings, as well as collect data and forecast future events.

We used a computer vision and deep learning model, as well as a convolution neural network, for this project. The most effective solutions for measuring social distancing among humans across variable moving video frames are computer vision with image processing techniques, machine learning, and deep learning techniques. Computer vision extracts important information from given input images and videos in order to have a correct understanding input and predict visual input, similar to how the human brain can predict images in real life. Objects are detected in real-time using YOLO (You Only Look Once), which supports convolutional neural networks for determining and detecting the distancing between the human in the frames using clusters of people during a neighborhood by taking the input from a video input and along with this Random Forest regression a give possible event that may occur in the future, so we can be prepared.

II. WORKING

A. Data Collection Using YOLO-v4

The Yolo algorithm (You Only Look Once) is a state-of-the-art algorithm that uses a real-time system based on deep learning to solve various Object Detection and Object Tracking problems. We'll take input frames from public surveillance cameras, videos, and photos, and run them through Covid-cop, which runs on YOLO-v4 and OpenCV. It detects objects in the frame, classifies them as human, and then locates those human objects in the frame, as shown in figure 5. We collect data on social distancing violations in a csv file based on the distance between human objects.

ISSN No:-2456-2165

B. Analysis of Data and Make Predictions

The Random Forest Regression algorithm uses the data in the csv file. Random forest is a supervised based learning algorithm. It creates a "forest" out of an ensemble of various decision trees, which use "bagging" method for training. It state that combination of different models improves the final result of the algorithm. Random forest is a multilateral, straightforward and simple to use machine learning algorithm which mostly produces superb results even without hyper-parameter tuning. Random Forest Regression algorithm is simple to use, that's why it is also known as one of the widely used algorithms (which can be use in regression and classification tasks both). We will use the random forest regression algorithm in this paper for prediction.

III. RELATED WORK

C. RCNN

As Seen[2] To get rid out of problem of selecting a huge number of regions in image, Ross Girshick et al. given a method in which they used selective search to generate regions in small number from an input image. It's result is, rather than trying to classify a large number of regions, we will able to work on only smaller number regions. This small number of region proposals is generated.

We propose an improved algorithm for small object detection based on a FRCNN (Faster R-CNN) [13]. Using the two-stage detection concept, the improved loss function is introduced for bounding box regression based on IoU(intersection over Union), and bilinear interpolation used for improving the RoI (regions of interest) pooling operation to resolved the issue of positioning deviation. In the stage of recognition, we use multi-resolution convolution feature fusion for making feature map contain more data, and used improved non maximum suppression to solve the problem of recognition deviation.

D. Histograms of Oriented Gradients for Human Detection

In 2005, Navneet Dalal and Bill Triggs [1] introduced the HOG (Histograms of Oriented Gradients) feature. The HOG is a feature descriptor that is used in image processing to detect objects. A feature descriptor is a represent an image or an image patch that extracts useful information from the image to simplify it.

The oriented gradients descriptor's histogram [11] is based on the notion that the distribution of intensity gradients or edge directions can represent the look and form of local objects in an image. The x and y derivatives of a picture (Gradients) are helpful because gradients have a significant magnitude along edges and corners owing to abrupt changes in intensity, and we know that edges and corners pack in a lot more information about object shape than flat regions. As a Result features in this descriptor uses directions of gradients histogram.

E. Single Shot Object Detection

The goal of SSD [15] with FFESSD(Feature Enhancement and Fusion) is to improve and exploit the shallow and deep features in the SSD algorithm's feature pyramid structure. They did this by introducing the Feature Fusion Module and two Feature Enhancement Modules, which they integrated into the SSD's traditional structure. FFESSD achieved 79.1 percent mean average precision (mAP) at a speed of 54.3 frames per second [5] (FPS) with the input size 300 300, while FFESSD with 512 512 sized inputs achieved 81.8 percent mAP at 30.2 FPS, according to experimental results on the PASCAL VOC 2007 dataset.

F. You Look Only Once

Object detection is a computer vision phenomenon [7] that entails detecting various objects in digital images or videos. People, cars, chairs, stones, buildings, and animals are among the objects detected.

YOLO algorithm is important because of Speed, High Accuracy, Learning capability. We can see in figure-1.

The following three techniques are used by the YOLO algorithm [20]:

- Bounding box regression
- Residual blocks
- Intersection Over Union (IOU)

Yolo has several versions, including YOLO, YOLO9000, YOLOv3 [20], YOLOv4 [4], and YOLO can detect the 80 COCO object classes, such as person, using the COCO dataset. Car, bicycle, airplane, bus, train, motorcycle, boat, and truck are all examples of transportation. a stop sign, a traffic light, a fire hydrant, a parking meter, and a bench.

In this paper we are going to use YOLOv4. It is a 4^{th} edition to the family of YOLO object detector models
YOLOv4 is made up of:

- Neck: SPP [25], PAN [49]
- Backbone: CSPDarknet53 [81]
- Head: YOLOv3

YOLOv4 has the following advantages over its predecessor, YOLOv3:

- It's a rapid and precise object recognition model that can be used to train a super-quick and accurate object detector by anybody with a 1080 Ti or 2080 Ti GPU.
- The effect of state-of-the-art "Bag-of-Freebies" and "Bag-of-Specials" item identification approaches was proven during detector training..

Method	Advantage	Disadvantage
R-CNN	The network transform the object detection problem into the classification problem and greatly improv the accuracy.	It generate partially overlapping candidate areas from each detection target.
SPP-Net	It introduces the spatial pyramid pooling layer after the last convolution layer, thus repetitive processing is eliminated.	Training is a multi-stage process with long training time.
Fast R-CNN	Its raining and testing are significantly faster than SPP-net. The input image can be any size.	The network still depend on candidate region selection algorithm.
Faster R-CNN	This network is faster than Fast R-CNN and no longer depend on region selection algorithm	The training process is complex, and there is still much room for optimization in the calculation process.
SSD	The multi-scale feature map is adopted and the processing speed is fast.	The robustness of this network to small object detection is not high.
YOLO	The network can meet the real-time requirements with using the full image as Context information.	It is relatively sensitive to the scale of the object, and the effect of small target detection is not good.

Fig 1: comparison of different-different algorithms

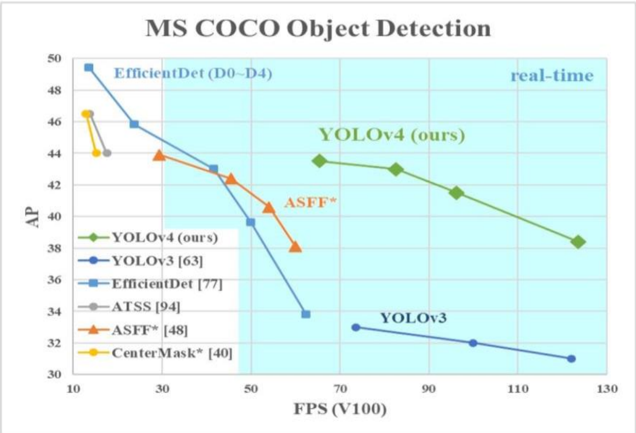

Fig 2 : YOLO v3 vs YOLO v4 comparison

G. Random Forest Regression

The random forest method [22], sometimes known as random choice forests, is an ensemble learning strategy for classification, regression, and other problems that necessitates the training of a large number of decision trees. For classification tasks, the random forest's output is the class chosen by the majority of trees. The mean or average forecast of the individual trees is returned for regression tasks.

It constructs a "forest" out of a collection of decision trees, which are often trained using the "bagging" approach. The main idea of the bagging approach is that mixing many learning models enhances the final output.

Ensemble Learning Comes in a Variety of Forms:

- Bootstrap Aggregation (Bagging)
- Boosting.

1. Bootstrap Aggregation (Bagging)

Random sampling with replacement is the bootstrap approach. Using Bootstrap, we can better grasp the dataset's bias and variance. The Bootstrap technique selects a small slice of data from the dataset at random.

It's a generic approach for lowering the variance of high-variance algorithms, such as decision trees. Bagging permits every model to run separately before combining the final result, with no bias for one model over the other.

2. Boosting

Boosting is a series of algorithms that employ weighted averages to transform poor students into strong ones. It's all about "Teamwork" when it comes to boosting. The features that will be prioritized in the next model are determined by each model that runs.

Boosting, as the name indicates, comprises learning from others, which improves learning.

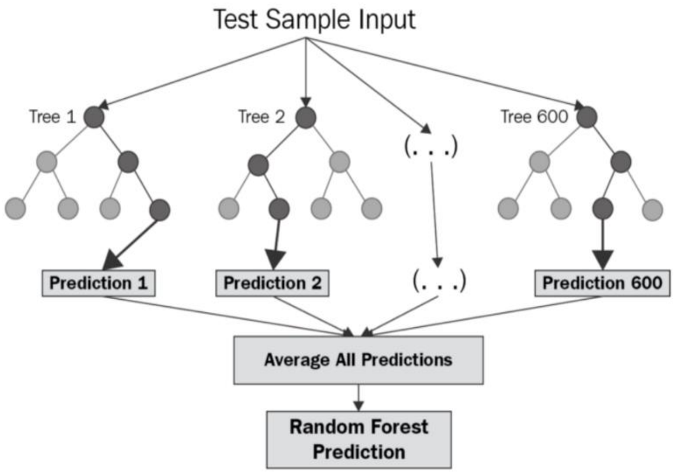

Fig 3: Random Forest Tree

Fig 4 : Inputted video clip

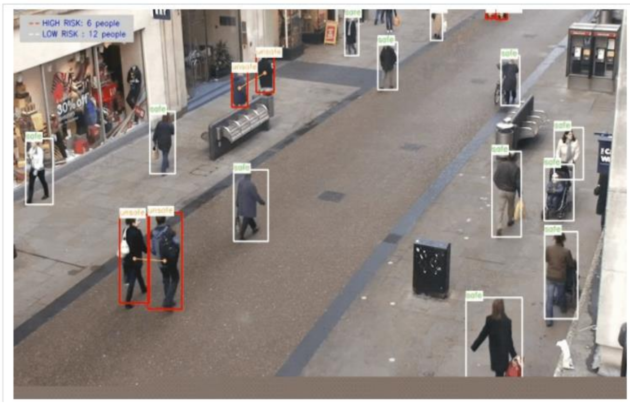

Fig 5: Outputted video clip

IV. RESULT

There are some of the snapshots of the output of working program model of Covid-cop. It is categorized into two stages, stage one is where the pedestrians are detected from the inputted videos as shown in Fig.-4 and the total count I displayed on the left upper corner as shown in fig.-5. In the stage two we are going to predict using random forest tree

V. CONCLUSION

The purpose of this research was to see whether there was a way to identify social distance for the Covid-19 incidents. People detection, which was done using Yolo, was used to detect objects for social distancing. Yolo v4 was implemented by Darknet for object identification, while Yolo v4 was implemented by Deep Sort for object tracking.

Calculating mAP was used to determine how well the model performed for object predictions, and it indicated that the average precision for a threshold of 0.25 percent was around 90% and higher, and for a threshold of 0.50 percent, the average precision was about 88 percent and higher. Similarly, the results for social distance were tested on a variety of film datasets, with crowded conditions factored in to increase the difficulty of detection.

REFERENCES

[1]. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, pp. 886-893 vol. 1, doi: 10.1109/CVPR.2005.177 [refer]

[2]. Girshick, Ross & Donahue, Jeff & Darrell, Trevor & Malik, Jitendra. (2013). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 10.1109/CVPR.2014.81. [refer]

[3]. Huang, Zhanchao & Wang, Jianlin & Fu, Xuesong & Yu, Tao & Guo, Yongqi & Wang, Rutong. (2020). DC-SPP-YOLO: Dense connection and spatial pyramid pooling-based YOLO for object detection. Information Sciences. 522. 10.1016/j.ins.2020.02.067. [refer]

[4]. Bochkovskiy, Alexey & Wang, Chien-Yao & Liao, Hong-yuan. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. [refer]

[5]. Liu, Wei & Anguelov, Dragomir & Erhan, Dumitru & Szegedy, Christian & Reed, Scott & Fu, Cheng-Yang & Berg, Alexander. (2016). SSD: Single Shot MultiBox Detector. 9905. 21-37. 10.1007/978-3-319-46448-0_2. .[refer]

[6]. Jiao, Licheng & Zhang, Fan & Liu, Fang & Yang, Shuyuan & Li, Lingling & Feng, Zhixi & Qu, Rong. (2019). A Survey of Deep Learning-Based Object Detection. IEEE Access. PP. 1-1. 10.1109/ACCESS.2019.2939201. . [refer]

[7]. Redmon, Joseph & Divvala, Santosh & Girshick, Ross & Farhadi, Ali. (2016). You Only Look Once: Unified, Real-Time Object Detection. 779-788. 10.1109/CVPR.2016.91. . [refer]

[8]. Hammedi, Wided & Ramirez, Metzli & Brunet, Philippe & Senouci, Sidi-Mohamed & Messous, Mohamed-Ayoub. (2019). Deep Learning-Based Real-Time Object Detection in Inland Navigation. 1-6. 10.1109/GLOBECOM38437.2019.9013931. [refer]

[9]. Community, Research. (2020). Comparative Analysis on YOLO Object Detection with OpenCV. [refer]

[10]. Pooranam, N & Sushma, Priya & Sruthi, S & Sri, Dhanya. (2021). A Safety Measuring Tool to Maintain Social Distancing On COVID-19 Using Deep Learning Approach. Journal of Physics: Conference Series. 1916. 012122. 10.1088/1742-6596/1916/1/012122. [refer]

[11]. Alvi, Nazia & Singh, Kshetra & Chandel, Garima & Vardhan, Yash. (2020). Convolution Neural Network Based Real Time Moving Object Detection. 29. 8134-8143. [refer]

[12]. Lee, Kelvin & Choo, Che & See, Hui & Tan, Zhuan & Lee, Yunli. (2010). Human detection using Histogram of oriented gradients and Human body ratio estimation. 18 - 22. 10.1109/ICCSIT.2010.5564984. . [refer]

[13]. Cao, Changqing & Wang, Bo & Zhang, Wenrui & Zeng, Xiaodong & Yan, Xu & Feng, Zhejun & Liu, Yutao & Wu, Zengyan. (2019). An Improved Faster R-CNN for Small Object Detection. IEEE Access. 7. 1-1. 10.1109/ACCESS.2019.2932731. [refer]

[14]. Szegedy, Christian & Toshev, Alexander & Erhan, Dumitru. (2013). Deep Neural Networks for Object Detection. 1-9. [refer]

[15]. Shi, Wenxu & Bao, Shengli & Tan, Dailun. (2019). FFESSD: An Accurate and Efficient Single-Shot Detector for Target Detection. Applied Sciences. 9. 4276. 10.3390/app9204276. [refer]

[16]. He, & Huang, Chang-Wei & Wei, & Li, & Anfu, Guo. (2019). TF-YOLO: An Improved Incremental Network for Real-Time Object Detection. Applied Sciences. 9. 3225. 10.3390/app9163225. [refer]

[17]. Thakkar, Heet & Tambe, Noopur & Thamke, Sanjana & Gaidhane, Vaishali. (2020). Object Tracking by Detection using YOLO and SORT. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. 224-229. 10.32628/CSEIT206256. [refer]

[18]. Magalhães, Rafael & Peixoto, Helton. (2019). Object Recognition Using Convolutional Neural Networks. 10.5772/intechopen.89726. [refer]

[19]. Ding, Sheng & Zhao, Kun. (2018). Research on Daily Objects Detection Based on Deep Neural Network. IOP Conference Series: Materials Science and Engineering. 322. 062024. 10.1088/1757-899X/322/6/062024. [refer]

[20]. Redmon, Joseph & Farhadi, Ali. (2018). YOLOv3: An Incremental Improvement. [refer]

[21]. Youzi, Xiao & Tian, Zhiqiang & Yu, Jiachen & Zhang, Yinshu & Liu, Shuai & Du, Shaoyi & Lan, Xuguang. (2020). A review of object detection based on deep learning. Multimedia Tools and Applications. 79. 10.1007/s11042-020-08976-6. [refer]

[22]. Liaw, Andy & Wiener, Matthew. (2001). Classification and Regression by RandomForest. Forest. 23 [refer]