DETECTING AND TRACKING OBJECTS USING OPENCV: A SOCIAL DISTANCING ALERT SYSTEM

Mr.E.Ramesh Reddy, Associate professor, Department of CSE, Narayana Engineering College Gudur.
A.Niranjan, Department of CSE, Narayana Engineering College, Gudur.
A.S.Ramcharan, Department of CSE, Narayana Engineering College, Gudur.
M.Rohit, , Department of CSE, Narayana Engineering College, Gudur
G.Sai Vinay, Department of CSE, Narayana Engineering College, Gudur.

ABSTRACT:

A computer allows you to do some amazing things. One of those most amazing things is detecting & tracking objects using computer vision. Computer vision allows us to detect objects and tracking the object's movements and behavior. Let us take a use case; one of the biggest problems around the globe is the current ongoing pandemic situation which is COVID-19. The virus is rapidly spreading, we need to take necessary precautions, one way of limiting the spread of the virus is to practice social distancing. It is an action taken to minimize contact with other individuals. "It has been suggested that maintaining a distance of approximately '2m' from another individual result in a marked reduction in transmission of most flu virus strains, including COVID- 19". One of the major issues of practicing social distance is monitoring whether people are maintaining social distances in public places like classrooms, seminar halls, and others places. Hence, we have come up with a solution for one of these problems by using this detecting and tracking objects using open CV techniques.

Here the objects are students and their behavior is movements if any student violates social distancing practices, then the system produces an alert sound to alert the students.

Keywords: COVID-19, Social distancing, Computer vision, Open CV.

I.INTRODUCTION:

The people around the world we're living peacefully, greeting each other, and gathering in huge numbers publicly but all of a sudden, a pandemic situation broke out as people are not allowed together publicly due to the spread of a deadly virus (COVID 19). To be safe from that virus and reduce its spread around the world, the WHO (World Health Organisation) has declared few safety measures and the government of the respective countries has requested its people to socially distance themselves from one another. It is one of the measures to contaminate the rapid spread of that deadly virus[1][2][3]. The major problem in following these social distancing practices is monitoring whether the people are properly socially distancing themselves from each other. If they are not monitored properly the virus may spread rapidly and it may lead to death because of its severity and it can be transmitted from one person to another person. Hence, social distancing is important and must be monitored properly[4][5].

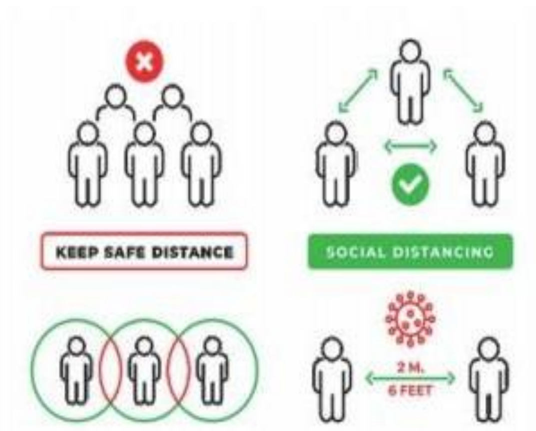


Fig 1. Social distancing demonstration

The above figure indicates how topractice social distancing with each other. To solve the problem that was mentioned previously in monitoring the social distancing practices we have taken a real- time example and proposed a solution to it[6][7]. The real-time example is monitoring social distances among students in classrooms, Auditoriums, etc… and the solution to it is 'A social distancing alert system' which is developed based on detecting and tracking object techniques using open-source computer vision (OpenCV[8][9][10]). The logic here is very simple "Students must socially distance themselves from other students by maintaining 6 feet or 2m distance among each other", if any student found to be violating this condition i.e. less than 2m/6 feet, then immediately an alarm sound is produced to alert the students so that they will come to know that they are getting closer and distance themselves from each other.

II. RELATED WORK:

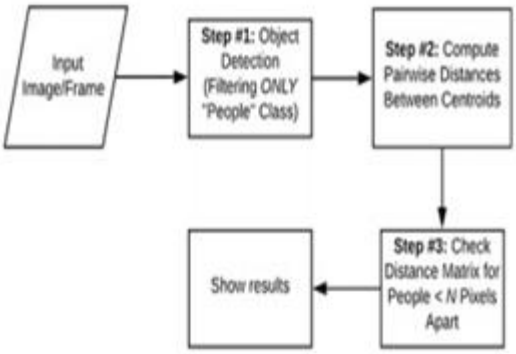
Ever since the outbreak of the novel COVID-19 which is also known as the coronavirus, all countries around the world have declared lockdown and other safety measures. It is one tough task for the government to keep monitoring the public all the time. Hence, various countries have taken various actions for it like creating apps, sending drones to monitor public gatherings, etc. To monitor public gatherings using drones, webcams, and other cameras, object recognition techniques are used. In this case, we have to detect the object's position first and then track down their movements to check how well students are practicing social distance measures, which is the real-time example we have taken here. The current object detecting approach uses CNN (Convolutional Neural Networks) and the CNN family of algorithms[10][11][12]. The convolutional neural network (CNN, or ConvNet) is a class of deep neural networks, most commonly applied to analyze visual imagery. They are also known as shift invariant or space invariant artificial neural networks (SIANN), based on the shared weight architecture of the convolution kernels or filters that slide along input features and they can also provide translation equivariant responses known as feature maps[13][14].

The CNNs are regularized versions of multilayer perceptrons. Multilayer perceptrons usually mean fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The "full connectivity" of these networks makes them prone to overfitting data. Typical ways of regularization, or preventing overfitting, include: penalizing parameters during training (such as weight decay) or trimming connectivity (skipped connections, dropout, etc.) CNN's take a different approach towards regularization: they take advantage of the hierarchical pattern in data and assemble patterns of increasing complexity using smaller and simpler patterns embossed in their filters. Therefore, on a scale of connectivity and complexity, CNNs are on the lower extremity. But after a clear study, it is being said that YOLO (You Look Only Once) approach and in that YOLO family of algorithms we use YOLOv3 algorithm. Hence all the CNN-based approaches are using grouping techniques, another approach YOLO is considered a relapse- based strategy to dimensionally isolate the bounding boxes and decipher their class probabilities. YOLO uses a different approach. YOLO is a clever convolutional neural network (CNN) for doing object detection in real-time. The algorithm applies a single neural network to the full image, and then divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted by the predicted probabilities[15].

III. METHODOLOGY:

Let us discuss the various methods and strategies used for tracking and detecting objects (students).

Fig 2.Workflow of social distancing alert system



1. Inputs :

Videos captured through the camera are given as input frames. Here we can classify video capturing into two types namely, online and offline video capturing. The live video that is being recorded is said to be online capturing and the previously recorded video is said to offline capturing. Online capturing is very much useful in monitoring and alerting the students in real- time.

2.Detecting and tracking objects :

Object detection is one of the classical problems in computer vision where you work to recognize *what* and *where* specifically what objects are inside a given image and also where they are in the image. The problem of object detection is more complex than classification, which also can recognize objects but doesn't indicate where the object is located in the image. In addition, the classification doesn't work on images containing more than one object.

YOLO is popular because it achieves high accuracy while also being able to run in real-time. The algorithm "only looks once" at the image in the sense that it requires only one forward propagation pass through the neural network to make predictions. After non-max suppression (which makes sure the object detection algorithm only detects each object once), it then outputs recognized objects together with the bounding boxes. With YOLO, a single CNN simultaneously predicts multiple bounding boxes and class probabilities for those boxes. YOLO trains on full images and directly optimizes detection performance. This model has several benefits over other object detection methods:

- YOLO is extremely fast
- YOLO sees the entire image during training and test time so it implicitly encodes contextual information about classes as well as their appearance.
- YOLO learns generalizable representations of objects.

To understand the YOLO algorithm, first, we need to understand what is being predicted. Ultimately, we aim to predict a class of an object and the bounding box specifying object location. Each bounding box can be described using four descriptors:

1. Center of the box
2. Width (BW)
3. Height (**b h**)
4. Value **c** corresponding to the class of an object

Along with that, we predict a real number **pc**, which is the probability that there is an object in the bounding box.YOLO doesn't search for interesting regions in the input image that could contain an object, instead, it splits the image into cells, typically a 19x19 grid. Each cell is then responsible for predicting K bounding boxes.

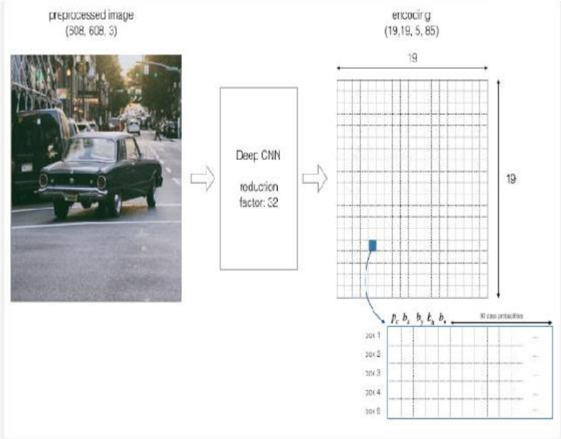


Fig 3. Here we take K=5 and predict the possibility for 80 classes

An Object is considered to lie in a specific cell only if the center coordinates of the anchor box lie in that cell. Due to this property, the center coordinates are always calculated relative to the cell whereas the height and width are calculated relative to the whole image size. During the one pass of forwards propagation, YOLO determines the probability that the cell contains a certain class. The equation for the same is :

$$score_{c,i} = p_c \times c_i$$

The probability that there is an object of certain class 'c

The class with the maximum probability is chosen and assigned to that particular grid cell. A similar process happens for all the grid cells present in the image. After computing the above class probabilities, the image may look like this :

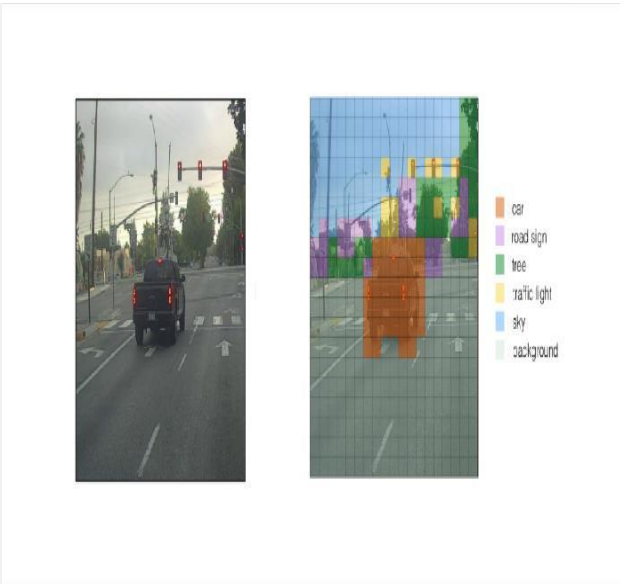


Fig 4. Comparability grid cells

This shows the before and after of predicting the class probabilities for each grid cell. After predicting the class probabilities, the next step is Non-max suppression, it helps the algorithm to get rid of the unnecessary anchor boxes, like you can see that in the figure below, there are numerous anchor boxes calculated based on the class probabilities.

Fig 5.Anchor boxes

To resolve this problem Non-max suppression eliminates the bounding boxes that are very close by performing the IoU (Intersection over Union) with the one having the highest-class probability among them.

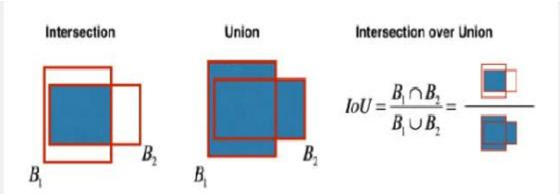


Fig 6.IoU operation

It calculates the value of IOU for all the bounding boxes respective to the one having the highest-class probability, it then rejects the bounding boxes whose value of IoU is greater than a threshold. It signifies that those two bounding boxes are covering the same object but the other one has a low probability for the same, thus it is eliminated. Once done, the algorithm finds the bounding box with the next highest- class probabilities and does the same process, it is done until we are left with all the different bounding boxes.

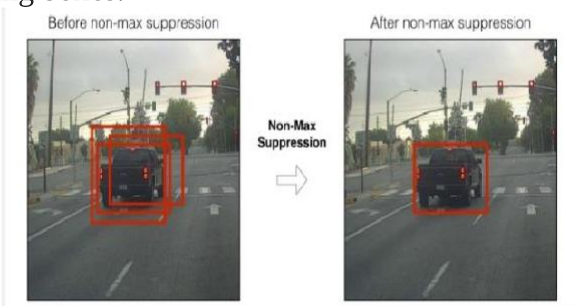


Fig 7.Before and After of Non-max suppression

After this, almost all of our work is done, the algorithm finally outputs the required vector showing the details of the bounding box of the respective class. The overall architecture of the algorithm can be viewed below :

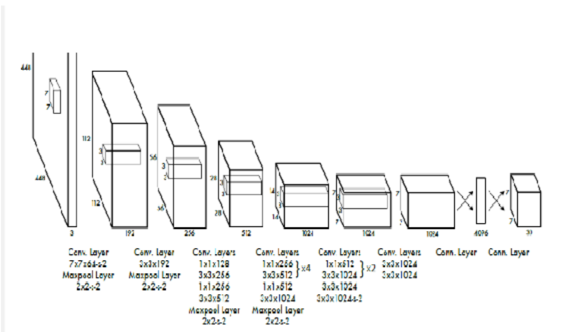
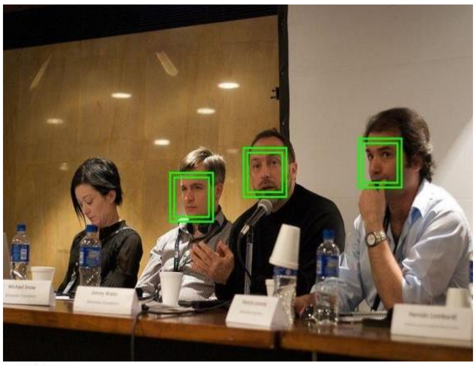


Fig 8.YOLO Architecture

IV. RESULTS AND ANALYSIS:

From the below picture which is given as the input dataset, we can see how our proposed theory on object detection and tracking approaches has satisfied to use to monitor people gathered in large numbers. The below figure represents the tracking of objects and detecting whether they are socially distanced from each other or not with the specification of the social distancing violation count in numbers also.

Fig 9. Social distancing violation count 3 Fig 10. Social distancing violation count 4



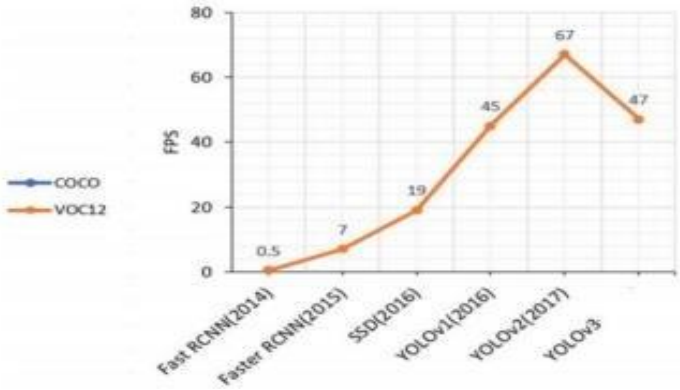


Fig 11. Checking social distancing through face to face detection Fig 12. mAP Performance on MS-COCO and PASCAL-VOC datasets.

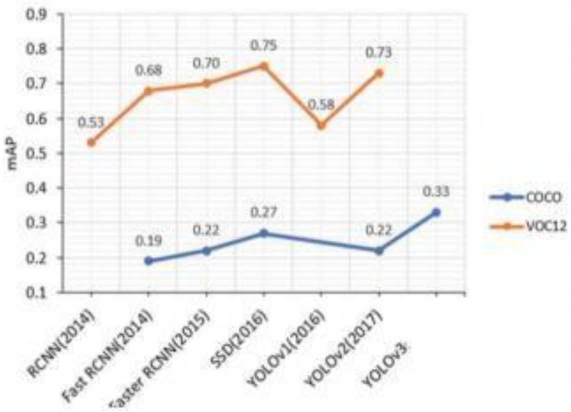


Fig 13. FPS Performance on MS-COCO and PASCAL-VOC datasets.

The Dataset: Yolo was checked against the COCO dataset which has 80 labels. And PASCAL-VOC dataset which has 20 classes. The train/Val data contains 11,530 images. Hence it has been considered YOLO works better than the others.

V.CONCLUSION:

The social distancing alert system will provide us a faster monitoring facility as this system is not only used to solve the image clarity issues but also speeding up the mathematical calculations by using the newly proposed method and also producing the alert sound. Since the social distancing alert system is developed with help of computer vision, it can be used in any sort of camera for monitoring purposes. Technology is known for development; it will improve and grow day by day so this social distancing alert system. For the time being, this proposed system is the need to monitor the social distancing practices in this pandemic situation. Some day this system may get an upgrade but as of now, this is the best possible solution for one of the crucial problems being handled during this pandemic situation.

VI. REFERENCES:

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

[2] YOLO — You Only Look Once, A State- of-the-Art Algorithm for Real-Time Object Detection System by Manishgupta, Machine Learning, Computer Vision Enthusiastic

[3] Overview of the YOLO Object Detection Algorithm by ODSC - Open Data Science

[4] K. Prem, Y. Liu, T. W. Russell, A. J. Kucharski, R. M. Eggo, N. Davies, S. Flasche, S. Clifford, C. A. Pearson, J. D. Munday, et al., "The effect of control strategies to reduce social mixing on outcomes of the covid19 epidemic in Wuhan, China: a modeling study," The Lancet Public Health, 2020.

[5] C. Adolph, K. Amano, B. Bang-Jensen,
N. Fullman, and J. Wilkerson, “Pandemic politics: Timing state-level social distancing responses to covid-19,” medRxiv, 2020.

[6] K. E. Ainslie, C. E. Walters, H. Fu, S. Bhatia, H. Wang, X. Xi, M. Baguelin, S. Bhatt, A. Boonyasiri, O. Boyd, et al., "Evidence of initial success for china exiting covid-19 social distancing policy after achieving containment," Wellcome Open Research, vol. 5, no. 81, p. 81, 2020

[7] J. Harvey, Adam. LaPlace. (2019) Megapixels.cc: Origins, ethics, and privacy implications of publicly available face recognition image datasets. [Online].
Available: https://megapixels.cc/

[8] N. Sulman, T. Sanocki, D. Goldgof, and

[9] R. Kasturi, “How effective is human video surveillance performance?” in 2008 19th International Conference on Pattern Recognition. IEEE, 2008, pp. 1–3.

[10]Venkateswara Rao, P., Ramamohan Reddy, A., Sucharita, V.An approach of detecting white spot syndrome of peaneid SHRIMP using improved FCM with hybrid back propagation neural network, International Journal of Pharmacy and Technology, 2016, 8(4), pp. 22351–22363

[11]Sucharita, V., Venkateswara Rao, P., Bhattacharyya, D., Kim, T.-H. Classification of penaeid prawn species using radial basis probabilistic neural networks and support vector machines International Journal of Bio-Science and Bio-Technology, 2016, 8(1), pp. 255–262

[12]Mandava Geetha Bhargava, Modugula TS Srinivasa Reddy, Shaik Shahbaz, P Venkateswara Rao, V Sucharita Potential of big data analytics in bio-medical and health care arena: An exploratory study, Global Journal of Computer Science and Technology 2017/8/5

[13]P. Venkateswara Rao , A. Ramamohan Reddy , V. Sucharita, Computer Aided Shrimp Disease Diagnosis in Aquaculture. International Journal for Research in Applied Science & Engineering Technology Volume 5 Issue II, February 2017 ISSN: 2321-9653

[14]S. Aslani and H. Mahdavi-Nasab, "Optical flow-based moving object detection and tracking for traffic surveillance," International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering, vol. 7, no. 9, pp. 1252–1256, 2013.

