

Social Distancing Detection Using Transfer Learning Algorithm

Deepak Yadav , M.Murali
Department of Computer Science and Engineering
SRM Institute of Science and Technology
Chennai, India

Abstract— COVID-19, declared as pandemic is a fast spreading disease that challenged the health services of the world. Social distancing has been recommended as one of the best practice that helps to restrain the curve of COVID-19 virus. Further, the practice of Social Distancing has definitely helped to decrease the transmission rate of the infectious COVID-19 worldwide. Furthermore, the lack of temporal understanding among the people may cause unintentional breach of the social distancing norms. Hence, it is necessary to bring in a vision based concurrent flow that will spot the social distancing violations. Social Distancing limits the physical contact among the people and by doing such, the danger of spreading COVID-19 can be decreased. The main purpose of the system proposed in this paper is to create a deep-learning system to detect social distancing in order to detect people in video sequences. The proposed system will employ YOLOv3 object recognition algorithm. The significance of this model is improvised through the transfer learning process. The pre trained algorithm is coupled with the trained layer which uses an additional data that will help in the detection process. In order to compute the pairwise distances of objects from the identified bounding box centroid, Euclidean Distance is used, while the bounding box information helps to identify the objects. A social distancing violation threshold will be set to examine whether the distance value among the people exceeds minimal barrier that has been set for social distance. A critical social distance value is defined in this work. Also, the system will show that the pedestrian density is held under the value defined. Thus, the chance of a Social Distancing violation could be prevented.

Keywords—*Social Distancing Detection, Euclidean Distance, Computer Vision, YOLO Algorithm, Object Detection, Object Tracking*

I. INTRODUCTION

COVID-19 (Coronavirus disease) started from Wuhan, China was declared a pandemic, has spread to over 180 nations, resulting in total cases 517,860,190 and total death of 6,278,347 worldwide as of May 10, 2022. The pandemic of COVID-19 has put a severe strain on the healthcare sector. It would take time for all the people to get fully vaccinated after few of the vaccines were clinically approved for use, clinical management now focuses on precaution, identification, and supportive care for the infected cases. Also, there has been immense research on the detection of variants so as to control the spread of new variants and its mutants, social distancing is the major effective measure. The population's vulnerability is exacerbated due to insufficient medicines, vaccines and immunity to COVID-19 infected people. Also, vaccines not being available in all the countries, social distance is one of the most efficient strategies for combating the epidemic.

Coronavirus, which was first reported in Wuhan, China is dangerous disease that affects mostly human respiratory system. WHO (World Health Organization) named it as Coronavirus disease 2019, shortly named COVID-19 on February 12, 2020. In comparison to the already known SARS (Severe Acute Respiratory Syndrome) and MERS (Middle East Respiratory Syndrome), the COVID-19 is exceedingly infectious disease. Moreover, it can spread through droplets and be easily transmitted between human beings, even with the minimal contact, and spread through asymptomatic virus carriers. Approximately 27 crore people have been infected by COVID-19 pandemic all around the world. Cough, fever, and lung inflammation are all symptoms of the virus, which potentially leads to acute respiratory distress syndrome. The common symptoms of coronavirus are fever, cough and dyspnea. The condition has the potential to induce deadly consequences in those who are susceptible, particularly the elderly with comorbidities. Early detection of the disease can enable individual patients receive treatment more quickly and allow for more appropriate isolation, which will manage the sick person and prevent the virus spread quickly and prevent infection.

The effective measure of Social distancing has helped to decrease the transmission rate of the infectious COVID-19 worldwide. Social distancing is usually referred to practice of minimum physical contact among people in public places, that can be up to 2 meters, also avoiding public gatherings. When it is implemented, social distance can help to decline the transmission of the COVID-19 and prevent the pandemic disease from reaching its peak. Ultimately, this measure will bring down the death percentage by making sure that the total no. of active cases will not surpass the community health care system's capabilities.



Social Distancing is quite contradictory to the long time followed normal behavior of humans. Hence, the outcome of automated and non-intrusive solutions that helps in social distancing practice can be handy for this period of time. Computer vision, deep learning, and machine learning techniques have showed promise in a variety of real-world situations during the last few decades. Deep learning has recently improved, making object detection more effective. The solution based on Computer vision (CV) are apt for automated controlling of social distancing compliance. Hence, the justification to benefit the implementation of social distancing compliance, we have proposed this technological solution in this paper.

The distance between people will be determined using clustering and distance-based techniques. Distance calculation between the people with an overhead perspective will result in a superior distance estimate and bigger coverage of the scene if a top-down method is used, i.e. an overhead perspective view approach.

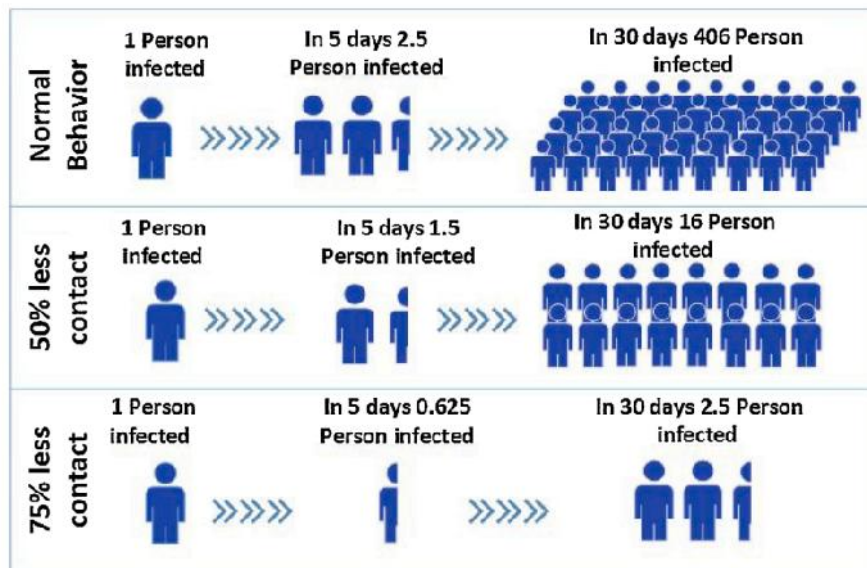


Fig. 1: Social Distancing Importance

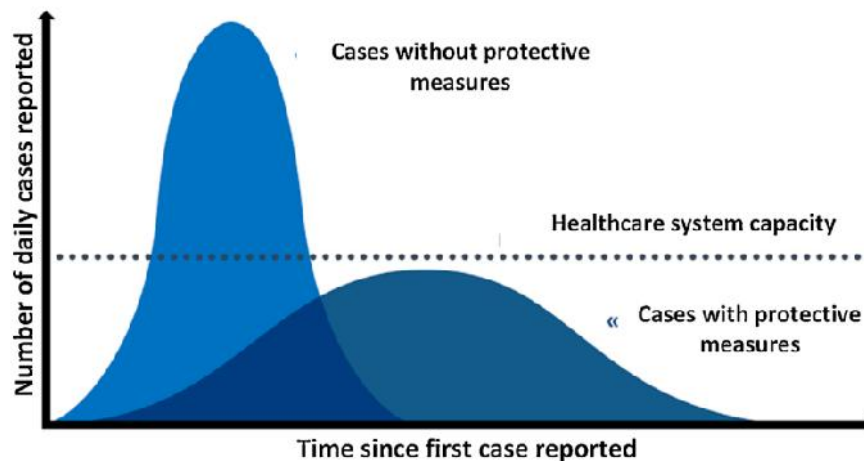


Figure 2 - Social Distancing Outcomes: the curve of COVID-19 cases is declining and being compatible with available health care capacity

Here, an overhead view will be used to establish a successful framework that will monitor the social distance among peoples. The above perspective has a wider area of view, and removes occlusion issues, thriving to make it suitable for monitoring social distancing and calculating the distances among people. The purpose is to develop framework that will help to monitor social distancing that comprises of deep-learning method. A deep learning method named You Only Look Once - YOLOv3 is utilized to recognize people in the crowd. Further the testing of the present model, overhead dataset is required that

has is either pre-trained on normal or frontal view data set. Transfer learning algorithm will help in increasing the accuracy of the detection model.

The detection model will detect the humans and provide the information related to bounding box. Using the centroid information of the bounding box that has been detected, the Euclidean distance among every centroid pair is determined following the human being detection. Using pixel to distance assumption, a preset minimum social distance violation criterion will be created. Then, the estimated data will be compared to the violation threshold so that the computed distance is inside the threshold/density defined for the social distancing violation or not. The bounding box's color will be originally set to green. Further, the bounding box's color will change to red if it is in the violation set. To track if someone has crossed the social distance threshold, the centroid tracking approach will be used.

The objectives of the proposed system are given below:

- To provide a social distance monitoring system based on deep learning.
- To use a pre-trained version of YOLO v3 to recognize humans and compute the bounding box centroid information. Additionally, apply transfer learning algorithm in order to increase the model's efficiency.
- An overhead data set is required to the additional training, and the layer that has been trained newly will be added to the method that has previously been trained.
- To track the social distance among the people, the distance among every pair of the centroid of the identified bounding box is approximated using the Euclidean distance. Also, using pixel to distance calculation, social distance breach threshold will be set.
- To keep track of an individual, a centroid tracking approach will be employed who crosses the social distance requirement.
- Determine how the pre trained YOLO v3 on overhead data set performs. The result from the detection framework would be evaluated in both approach: with and without transfer learning.

This has paved the way for deploying an active monitoring system like this necessitates careful ethical considerations as well as intelligent system design. In this regard, a fully convolutional machine learning system, for example, a deep neural network which is not having any of the feature-based input space is far more equitable, including a single discretion: have equitable training data distribution.

II. RELATED WORK

Few related deep learning based human detection research has been discussed in this section. The process of localization and categorization of its shape in the video footage, the detection of human being is known as object detection process in Computer Vision field. "Nguyen et al. has provided an in-depth study of the state of the art in current developments and the problems which occurs in the detection of human being [14]." Machine learning techniques, Human descriptors, real-time detection and occlusion are all covered in the survey. On a variety of image recognition benchmark approaches based on deep convolutional neural networks (CNN) have been proven to outperform others [15]. The CNN model, which is most suited in feature learning approaches and robust in detecting the objects in many situations, was one-off the categories in deep learning for the detection objects in images. Deep learning have neural network structure that helps to self-construct object descriptors helps in increasing its effectiveness in object recognition and learn high-level properties that aren't directly presented in the dataset.

Even in the near future, it is quite difficult to completely eliminate COVID-19, however an automated system for tracking the infected person and assessing social distancing measures would be extremely beneficial to us. Pedestrian Detection: The pedestrian detection is generally considered in two ways: (i) first being sub-task of normal object detection (ii) another as a separate task that will be dedicated solely to detect the pedestrians. Here, we will find a thorough examination of 2D object detectors, the datasets, the metrics and basics that go with them. YOLO, SSD, and EfficientDet are some of the most popular models. The detectors can be divided into two categories: anchor-based and anchor-free techniques.

The existing state-of-the-art object detectors having deep learning model have its own benefits and dis benefits in aspect of speed and accuracy. Within an image, the item may have various spatial locations as well as the aspect ratios. Hence as an outcome, the real time object detection algorithm and methods based on the Convolutional Neural Network model, such as YOLO and R-CNN, has been generated to acknowledge multi-classes in distinct regions of images. With aspect to both speed and accuracy, YOLO is most popular deep CNN based object identification algorithm.

The distance between people will be determined using clustering and distance-based techniques. A deep learning approach named YOLO v3 (You Only Look Once v3) will be used to recognize humans. The detection model will detect the humans



and provide us the bounding box information. For the detection of Euclidean distance among all the recognized centroid pair, the detected bounding box with its centroid information will be used after the detection of humans.

III. PROPOSED SYSTEM

The Proposed System of Social Distancing Detection is based upon three different algorithms:

- i. Object Detection
- ii. Object Tracking
- iii. Distance measurement between detected objects

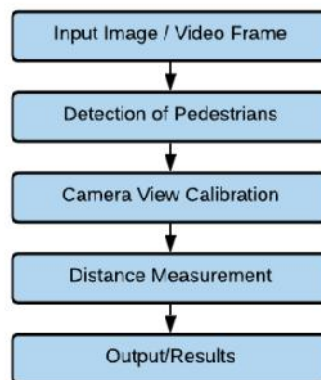


Fig. 3: Workflow of proposed system

A. Object Detection

For the Object Detection, we are using YOLO algorithm. YOLO is an algorithm that employs CNN (Convolutional neural network) to detect the objects in a real time environment. This algorithm divides the video frames and images into N grids, each having an equal dimension region of $S \times S$. Each of these grids is responsible for the detection and localization of the object it contains. YOLO algorithm can detect more than 9000 classes. In this system, we are using COCO dataset which is trained on 80 levels but we only use the partitioned classes. The process of localization and categorization of its shape in the video footage, the detection of human being is known as object detection process in the Computer Vision field. "Nguyen et al. has provided an in-depth study of the state of the art in current developments and the problems that occurs in the detection of human being [14]." Machine learning techniques, Human descriptors, real-time detection and occlusion are all covered in the survey. On a variety of image recognition benchmark approaches based on deep convolutional neural networks (CNN) have been proven to outperform others [15]. The CNN model, which is most suited in feature learning approaches and robust in detecting the objects in many situations, was one-off the categories in deep learning method to detect object in images. Deep learning have neural network structure that helps to self-construct object descriptors helps in increasing its effectiveness in object recognition and learn high-level properties that aren't directly presented in the dataset..

B. Object Tracking

After the detection of the partitioned classes, we need to track them i.e. the objects needs to be tracked once they are detected using the YOLO algorithm. We are assigning a new id to every detected partitions and drawing a box over them and measuring the centroid of the box.

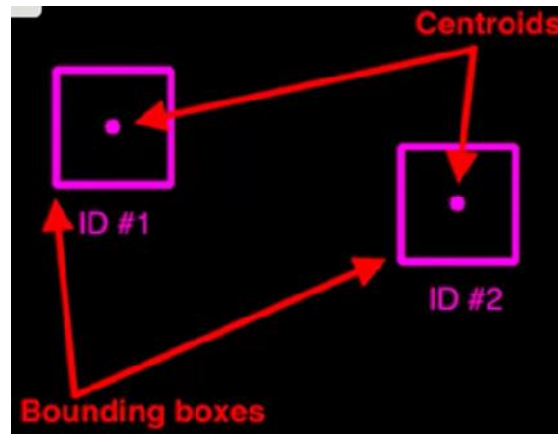


Fig. 4: Video Frame 1 - Detected partition classes

As seen in the figure 4, ID 1 and ID 2 has been assigned to the detected partitioned classes respectively and a bounding box has also been drawn over them. The centroid between each of the bounding boxes needs to be measured.

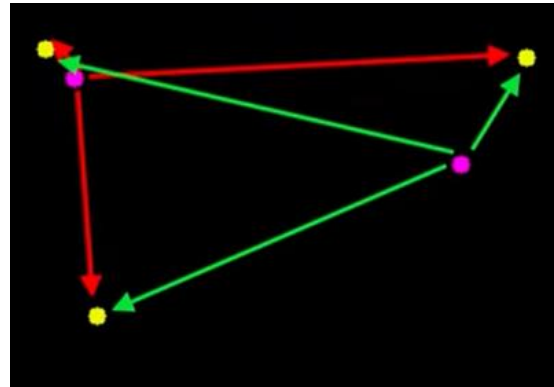


Fig. 5: Video Frame 2 – Movement of person from one point to another point

When the next frame of the video is captured, it is important for us to know that the person has moved from one place to another and also know that a particular person is the same person even after his/her movement. We are taking purple point as an old centroid and yellow point as a new centroid. For assurance, we are measuring the Euclidean Distance from every point i.e. from every old centroid to the new centroids and the close pairs will be detected as the same person. So, the yellow point is going to be the new centroid for the person and we are assigning a new ID to it.

C. Distance measurement between detected objects

After the tracking of the partition has been done, we are measuring the distances between them. To measure the distance between partition to partition in the photo frame, we are using the formula:

$$F = (P * D) / W$$

$$D = (W * F) / P$$

- F - Focal Length
- D – Distance from the object to the camera
- P – How much pixel the object is covering in the photo
- W – Object Weight

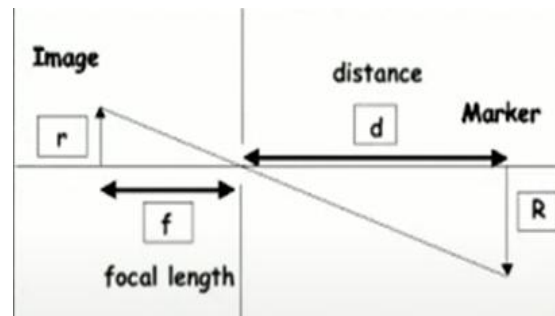


Fig. 6: Distance Measurement Illustration

The accuracy of the detection model is increased by making the use of transfer learning algorithm. The proposed model detects the humans and provide us the bounding box information. Furthermore, by making the use of centroid information of the bounding box, the Euclidean distance among every centroid pair is determined. A preset minimum social distance breach criterion would be created using the pixel to distance assumptions. The estimated data will then be compared to the violation threshold that has been set to determine if the computed distance is inside the critical density defined for violation or not. The color of the bounding box is originally set to green; furthermore, the color of the bounding box will change to red if it is in the violation set. The centroid tracking approach will also used to track someone who has crossed the social distance threshold.

IV. IMPLEMENTATION WORK

The implementation of this system is divided into three parts. They are as follows:

- i. Setting up the variable values
- ii. Creating the people detection function
- iii. Grab frames from the videos and make prediction measuring the distances of detected people

A. Setting up the Variable Values

The first phase of the implementation of this system is to set up the variable values to initialise minimal probability to get rid of the weak detections along with the threshold when applying non-maxima suppression. The minimum configuration value is set as 0.3, below which the detections will be as weak detection and we are filtering those weak detections. The non-maximal suppression (NMS) threshold is set as 0.3. It is basically for drawing the boxes over the detections that have been detected. 50 pixels has been defined as the minimal safe distance that two people can be from each other.

B. Creating the people detection function

While creating the people detection function, we are declaring a function whose parameters are going to be frame, net, In, personId. The frame which we are grabbing from the video, we are going to extract its coordinates and putting it in H and W and declare an empty list of result i.e. grabbing the dimensions of the frame and initializing the list of results.

We are using blobFromImage function so that a blob can be constructed from the input frame and then perform a forward of the grabbed frames. To get the bounding boxes and the associated probabilities, the data is forwarded into the YOLO object detector. The detected bounding boxes, its centroids and confidences needs to be done respectively and once the initialization has been done, the loop is run over all of the layered outputs and detections that extracts the class ID and confidences (i.e. probability) of the current object detection. YOLO algorithm returns the boxes' width and height, after the center (x, y) coordinates of the bounding box has been returned. Keeping this thing in mind, we scale the bounding box coordinates back corresponding to the size of the image. Also, to derive the top and left corner of the bounding box, the center (x,y) coordinates are used. By making the use of this the list of bounding box coordinates, centroids, and confidences are updated.

During the detection, we can have a problem like, a people is detected and it is detecting 4 or 5 boxes near its places. What is the perfect box? – By non-maxima suppression, we can get the perfect box. Applying the non-maxima suppression so that our weak detection or overlapping bounding boxes can be filtered. Basically, non-maxima suppression is a class of algorithm to select one entity (e.g. bounding box) out of many overlapping entities. If at least one person is detected, we are updating our box coordinates and appending it with the result.

C. Grabbing the frames from videos and make predictions measuring the distances of detected people

The first thing we have to do during this phase of the implementation is constructing the argument parse and parsing the constructed argument. The arguments are: input, output and display.



The data where the YOLO model are trained, the COCO class labels are loaded onto the same and then the paths to the YOLO weight and the model configuration are derived. The video streams and the pointer to the output video file are initialized so that the loop can be run over the frames captured from the video stream.

For accessing video stream, we are using cv2.VideoCapture to capture the frame from the video. We now need to resize the frames and detect the people in it. For this, we use people detecting function and store it in the results variables. We have to initialise the set of indexes which violates the minimal social distance. We also have to ensure that there must be at least two people detection in the frame which is necessary in order to compute the pairwise distance maps, and extract all centroid from result and compute Euclidean Distance between all pairs of centroids.



Fig. 7: Creation of Bounding Box

We have to run the loop (for i in range(0, D.shape[0]): and for j in range(i + 1, D.shape[1]):) over the upper triangular of the distance matrix to check to see if the distance between any two centroid pairs is less than the configured number of pixels. We also have to run the loop over the results to extract the bounding box and centroid coordinates and then initialize the color of the annotation. If the index pair exists within the violation set, then the color is updated.

As shown in the figure 7, a bounding box around the person and the centroid coordinates of the person is drawn. If the distance is violated, we are adding their centroid coordinates. From the output frame, we can get the total number of social distancing violations.

V. RESULTS AND DISCUSSION

The details of the different experiments that are carried out using the proposed model and approach to detect the social distance monitoring is mentioned here. To monitor the social distance, a dataset that contains the video frames are obtained. No restriction are put on the movement of the individuals all around the scene. Humans in the video frames can move freely; however their visual might get effected by the radial distance and camera's position. The appearance of the individual are not identical as the height, color, pose and the scales of the human keeps on differing in the data set. Opencv is used for the implementation of the proposed system. Then, the overall result from the experiment is segmented into two parts: firstly the pre trained model testing result is analyzed and then result of the detection model is discussed after the application of the transfer learning algorithm on the dataset in the second section.

A. Social Distancing Monitoring Result using Pre Trained Model

The testing result of social distancing framework is visualized using a pre trained model and using the different video sequences, the results of testing is evaluated. Human in the video can be seen having free movement in the scenes. The visual appearance of individual is not going to be uniform to the side view or the frontal view. The proposed model in this paper only consider human class; hence, only those objects which have appearance like human being are identified by the pre-trained model in the system. The pre trained model provides the principled result and detects different sizes individual bounding boxes. Many individual are entering in the scene and people can move freely in the video. To check if any individual in the video breaches the social distance or not, the distance among every identified bounding box is calculated after the detection of individual.



Fig.8: Tracking accuracy with pre-trained and trained YOLOv3 detection model

B. Social Distancing Monitoring Result using Transfer Learning Algorithm

Transfer Learning Algorithm helps to increase the accuracy and efficiency, hence we are using the transfer learning algorithm to improve the efficiency of detection model. Overall result of the experiment shows that transfer learning improves detection outcomes significantly. The model detects individuals at different locations of the scene. Individuals having different characteristics are identified effectively and the distance among the people are also computed. All the people entering and walking in the video frames are detected and monitored using the proposed system. Whenever, social distance among the people is violated, the framework detects the violating effectively and the bounding box is marked as red if the individuals are close to each other in the video frame. From the output frame, we can get the total number of social distancing violations.

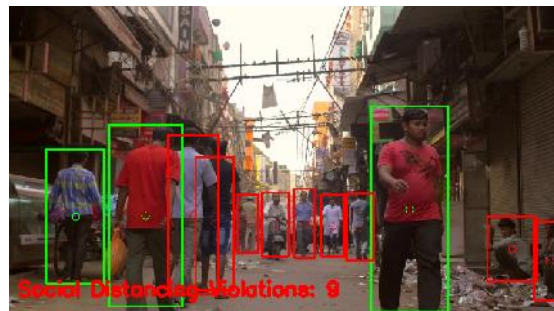


Fig. 9: Output – Total no. of Social Distancing Violation

VI. CONCLUSION

A method for analysis of social distance between the people is put forward in this paper. The system uses the deep learning and computer vision technique. Computer Vision helps to calculate the distance between each person and it helps in monitoring the social distancing. YOLO algorithm can detect more than 9000 classes. We are using COCO data set trained on 80 levels. The pre-trained YOLO v3 paradigm is used for the identification of human. An individual's size, shape, visibility, appearance, scale, and posture varies substantially. Therefore, the transfer learning process is applied so that it enhances the accuracy of the pre trained model. After training on pre trained data set and the detection of the partitioned classes, the objects are tracked.

A deep learning based detection paradigm uses the transfer learning algorithm to monitor the social distance. Bounding box information, including centroid coordinates, is provided by the detection model. For the computation of centroid distance in pair, the Euclidean distance is used among the identified bounding boxes. A red bounding box is displayed if any group of people is deemed to be violating the minimum acceptable threshold value. Already captured video stream of people having movement on a busy road is used in the system.



REFERENCES

- [1] Yang, D.; Yurtsever, E.; Renganathan, V.; Redmill, K.A.; Özgüner, Ü. "A Vision-Based Social Distancing and Critical Density Detection System for COVID-19". *Sensors* 2021, 21, 4608. <https://doi.org/10.3390/s21134608>
- [2] Narinder Singh Punn, Sanjay Kumar Sonbhadra and Sonali Agarwal, "Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques"
- [3] Sreetama Das, Anirban Nag, Dhruva Adhikary, Ramswaroop Jeevan Ram, Aravind BR, Sujit Kumar Ojha, Guruprasad M Hegde, "Computer Vision-based Social Distancing Surveillance Solution with Optional Automated Camera Calibration for Large Scale Deployment", Corresponding author: sreetama.d2005@gmail.com ORCID: 0000-0002-4523-8915, * equal contribution
- [4] Cong T. Nguyen, Yuris Mulya Saputra, Nguyen Van Huynh, Ngoc-Tan Nguyen, Tran Viet Khoa, Bui Minh Tuan, Diep N. Nguyen, Dinh Thai Hoang, Thang X. Vu, Eryk Dutkiewicz, Symeon Chatzinotas, and Bjorn Ottersten, "A Comprehensive Survey of Enabling and Emerging Technologies for Social Distancing Part I: Fundamentals and Enabling Technologies" DOI: 10.1109/ACCESS.2020.3018140 and "A Comprehensive Survey of Enabling and Emerging Technologies for Social Distancing Part I: Fundamentals and Enabling Technologies" DOI: 10.1109/ACCESS.2020.3018140
- [5] G V Shalini, M Kavitha Margret, M J Sufiya Niraimathi, S Subashree, "Social Distancing Analyzer Using Computer Vision and Deep Learning", To cite this article: G V Shalini et al 2021 J. Phys.: Conf. Ser. 1916 012039
- [6] C. Courtemanche, J. Garuccio, A. Le, J. Pinkston, and A. Yelowitz, "Strong social distancing measures in the united states reduced the covid-19 growth rate: Study evaluates the impact of social distancing measures on the growth rate of confirmed covid-19 cases across the united states." *Health Affairs*, pp. 10–1377, 2020.
- [7] D.T. Nguyen, W. Li, P.O. Ogunbona, Human detection from images and videos: A survey, *Pattern Recognition*, 51:148-75
- [8] Safa Teboulbi, Seifeddine Messaoud, Mohamed Ali Hajjaji, and Abdellatif Mtibaa, "Real-Time Implementation of AI-Based Face Mask Detection and Social Distancing Measuring System for COVID-19 Prevention" Article ID 8340779, <https://doi.org/10.1155/2021/8340779>
- [9] Mahdi Rezaei, Mohsen Azarmi, "DeepSOCIAL: Social Distancing Monitoring and Infection Risk Assessment in COVID-19 Pandemic" doi: <https://doi.org/10.1101/2020.08.27.20183277>
- [10] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection" <https://www.researchgate.net/publication/340883401>
- [11] Sonali Agarwal, Narinder Singh Punn, Sanjay Kumar Sonbhadra, P. Nagabhusahan, Praveer Saxena, "Unleashing the power of disruptive and emerging technologies amid COVID 2019: A detailed review", <https://www.researchgate.net/publication/341639677>
- [12] Yew Cheong Hou, Mohd Zafri Baharuddin, Salman Yussof, Sumayyah Dzulkifly, "Social Distancing Detection with Deep Learning Model"
- [13] Yilun Wang, Yuqing Zhang, Kim Bennell, Daniel Kenta White, Jie Wei, Ziyang Wu, Hongyi He, Shaohui Liu, Xianghang Luo, "Physical Distancing Measures and Walking Activity in Middle-aged and Older Residents in Changsha, China, During the COVID-19 Epidemic Period: Longitudinal Observational Study" <https://www.jmir.org/2020/10/e21632/>
- [14] D.T. Nguyen, W. Li, P.O. Ogunbona, "Human detection from images and videos: A survey", *Pattern Recognition*, 51:148-75, 2016.
- [15] A. Krizhevsky, I. Sutskever, G.E. Hinton, "Imagenet classification with deep convolutional neural networks", In *Advances in neural information processing systems*, pp. 1097-1105, 2012.
- [16] S.Amudha, M.Murali, 2020, 'Deep Learning based energy efficient novel scheduling algorithms for body-fog-cloud in smart hospital', *Journal of Ambient Intelligent Humanized Computing*, <https://doi.org/10.1007/s12652-020-02421-0>
- [17] M.Murali, 2015, 'Principal Component Analysis based Feature Vector Extraction', *Indian Journal of Science and Technology*, 8(35), 1-4
- [18] M.Murali, Shreya Sharma, Neel Nagansure, 2020, 'Reader and Object Detector for blind', ICCSP, July 28-30, INDIA
- [19] M.Murali, Simran Kour Awal, Kabir Singh Sethi, 2020, 'Detection and Classification of white blood cell cancers', *International Journal of Advanced Science and Technology*, 29(7), 937-94