



An Optimized Deep Learning Approach for the Prediction of Social Distance Among Individuals in Public Places During Pandemic

Santosh Kumar Sahoo¹ · G. Palai² · Baraa Riyadh Altahan³ ·
Sk Hasane Ahammad⁴ · P. Poorna Priya⁵ · Md.Amzad Hossain^{6,7} ·
Ahmed Nabih Zaki Rashed⁸ 

Received: 17 October 2022 / Accepted: 27 December 2022 / Published online: 2 January 2023

© The Author(s), under exclusive licence to The Japanese Society for Artificial Intelligence and Springer Nature Japan KK, part of Springer Nature 2023, corrected publication 2023

Abstract

Social distancing is considered as the most effective prevention techniques for combating pandemic like Covid-19. It is observed in several places where these norms and conditions have been violated by most of the public though the same has been notified by the local government. Hence, till date, there has been no proper structure for monitoring the loyalty of the social-distancing norms by individuals. This research has proposed an optimized deep learning-based model for predicting social distancing at public places. The proposed research has implemented a customized model using detectron2 and intersection over union (IOU) on the input video objects and predicted the proper social-distancing norms continued by individuals. The extensive trials were conducted with popular state-of-the-art object detection model: regions with convolutional neural networks (RCNN) with detectron2 and fast RCNN, RCNN with TWILIO communication platform, YOLOv3 with TL, fast RCNN with YOLO v4, and fast RCNN with YOLO v2. Among all, the proposed (RCNN with detectron2 and fast RCNN) delivers the efficient performance with precision, mean average precision (mAP), total loss (TL) and training time (TT). The outcomes of the proposed model focused on faster R-CNN for social-distancing norms and detectron2 for identifying the human ‘person class’ towards estimating and evaluating the violation-threat criteria where the threshold (i.e., 0.75) is calculated. The model attained precision at 98% approximately (97.9%) with 87% recall score where intersection over union (IOU) was at 0.5.

Keywords Social distancing · Deep learning · Detectron2 · Intersection over union · Object detection

1 Introduction

First and second waves of Covid-19 hitting the globe and affecting nearly 180 countries across the globe, governments and health fraternities have realized that social distancing is the only way to prevent the spread of the disease and break the chain of infections. However, there are cases in countries such as India, France, Russia, and Italy where either people are heavily populated or do not adhere to follow preventive techniques like social distancing at crowded places [1]. Social distancing is a healthy practice or preventive technique which evidently provides protection from transmission of Corona (Covid-19) virus when there is a maintenance of minimum of 6 feet distance between two/more people [2]. Social distance does not have to be always people's preventive technique rather it could also be a practice that could be adhered to reduce physical contact towards transmitting diseases from the virus-affected person [3, 4]. Deep learning like its role in every field has found to be a significant technology in addressing this problem [5]. Manually monitoring, managing and maintaining distances between people in a crowded environment such as colleges, schools, shopping marts and malls, universities, airports, hospitals and healthcare center, parks, restaurants and more places would be evidently impractical and henceforth adopting the machine language, AI and deep learning techniques in automatic detection for social distancing is essential. Object detection with RCNN in machine language-based models had been adopted by researchers since it offers the investigators with faster detecting options where the faster RCNN is found to be more advantageous, supports faster convergence and also provides higher performance [6] even under low-light environment.

In regions of America, Europe and South-East Asia due to poor social distancing and improper measures against Covid-19 by the people, it has been confirmed that in 2020, many cases were legally recorded where people died due to violation of threshold in social-distancing measures during Covid-19 [4]. The social distancing has been adopted as a concept by the researchers so far to examine varied factors, namely with/without face mask, face recognition, object (human) identification, people monitoring and diseases monitoring during Covid-19 [7].

Deep learning and machine learning-based AI models surpass various criteria such as time consumption and labor intensive where human interventions are must in research and also to obtain results/outcomes that have minimal errors and loss with alternate predictions by shortening the period of investigations from years, months and days into several days [8]. Majorly R-CNN (region-based CNN) and faster R-CNN are utilized for accurate and faster predictions. To measure the predicted outcomes, the examiners generally utilize the metric-evaluation approaches/techniques such as regressors (Random-Forest-Regressor, Linear Regressor), IOU (intersection over union) and more. For object detection, the IOU is commonly adopted.

The research has developed a customized deep learning model with detectron2 and intersection over union that would predict whether social distance is maintained or not by the individuals, especially in crowded places and public places.

The studies by the authors Vinita and Velantina [9] have focused on the prediction of social distancing post-Covid-19 where the deep learning and the machine language have been adopted for their model development. The studies utilized the python for designing the model algorithm and networking. Pandian [10] utilized the TWILIO and the authors Vinita and Velantina [9] utilized the YOLOv3 for their architecture. The studies used the faster R-CNN and found that the architecture is flexible and faster than other approaches, and the conclusions revealed that efficient and effective object detection is attained with the YOLOv3 and faster R-CNN model.

Rezaei and Azarmi [11] developed the YOLOv4-based ML model with COCO datasets where the DNN (deep neural network) is utilized for the architecture. The model is the most identified and recognized model that had attained 99.8% accuracy with speed of 24.1 fps in real-time analysis. The model is developed to examine the social distancing of people post-Covid-19 and infectious assessment. The study is still identified as popular model-based investigation where the accuracy is higher than the existing model.

Saponara et al. [12] had developed AI-based social distancing and people detecting model post-Covid-19 which measures the distances among two or more people. The authors adopted YOLOv2 with the Fast R-CNN for detecting the objects where here it was humans. The study's algorithm and the architecture were solely developed to identify the social distances and the outcome where the accuracy was 95.6% with precision of 95% and recall rate at 96%.

Arya et al. [13] examined different measures towards monitoring the social distancing via computer vision as an extensive analysis of existing literature-based review. The study focused on security threats' identification and facial expressions based analyses and models that adopted deep learning, computer vision through real-time datasets with video-streams. The authors found YOLO to be effective in the detection models among other AI-based models and evidently concluded that two-staged detectors are far better and efficient than single detector which provides effective outcomes and reliable results that remain valid and constant in similar surroundings. Henceforth, adopting the two-stage object detectors is wiser, efficient, effective and accurate.

Yang et al. [14] investigated the social-distancing concept post-Covid-19 by developing the vision-based along with critical-density-based detection system. They developed the model based on two major criteria, first to identify the violations in social distancing via real-time vision-based monitoring and communicating the same to deep learning model, second precautionary measures are offered as audio-visual cues through the model, to minimize the violation threshold to 0.0 without manual supervision and thus reducing the threats and increasing the social distancing. The study adopted YOLOv4 and faster R-CNN where the precision average (mAP) was achieved higher in BB-bottom method with 95.36% with accuracy of 92.80% and recall score at 95.94%. Though the study provided better outcomes, the targets were initially small and occlusions occurred due to heavy-density accumulation. Finally, the authors were able to train and develop the model with huge mass resulting with 2% error in critical-density stating that it could be refined in future research with better understanding of the targeted

people and density accumulation algorithm. Thus, conclusion revealed that maintaining social-distancing practices within family members in crowded areas is very essential which had caused the major issue in the research which could be avoided in future research.

Ahmed et al. [3] and Ahmed et al. [4] developed deep learning-based architecture that utilizes the social-distancing concept as base for their evaluation and people monitoring/management post-Covid-19. The authors Ahmed et al. [3] utilized YOLOv3 for identifying humans and faster R-CNN as the social-distancing algorithm where they achieved 92% accuracy in tracking without transfer learning and 98% with transfer learning; similarly, the model obtained 95% as tracking accuracy stating social-distancing detection with YOLO with tracking transfer learning technique being effective. Pun et al. [2] proposed a study on monitoring Covid-19 social distance with person detection and tracking through fine-tuned YOLOv3 and deep sort techniques. The deep learning model is used for task automation of supervising social distance using surveillance video. The proposed structure uses the model of YOLOv3 object detection to segregate background people and uses deep sort methods to track the identified people with the use of assigned identities and bounding-boxes. The deep sort tracking method with YOLO v3 scheme shows good outcomes with balances score of FPS and mAP to supervise real-time social distance among people.

The main aim of Rahim et al. [6] study is to offer an efficient monitoring solution for social distance in low-light surroundings in pandemic circumstances. The developing disease of Covid-19 caused by the SARS-Covid-2 virus has acquired a worldwide crisis with its deadly distribution all over the globe. People find ways to come out of their homes with their families during nighttime to take fresh air. In such circumstances, it is essential to consider efficient steps to supervise the criteria of safety distance to avoid positive cases and to manage the toll of death. In this research, a deep learning-based method is proposed. The proposed structure uses the YOLOv4 structure for measuring social distance and detection of real-time object with a single motionless ToF (time of flight) camera.

Through the reviews, it is evidently understandable that YOLO with CNN, R-CNN and faster R-CNN are majorly utilized in the recent years towards social distancing-based object detection models.

The objective of this proposed research work is to:

- Implementing “detectron2” and IOU as machine language-based model with faster RCNN towards examining and detecting the social distancing between people, especially in the crowded areas.
- Developing a customized model for social distance prediction through video object detection.

The overall organization of the paper is presented as follows:

Section 1 describes the introduction and research objectives, Sect. 2 explains the research methodology and algorithm implementation in the proposed work. The experimental evaluation and analysis is presented in Sect. 3 followed by conclusion and future recommendation in Sect. 4.

2 Research Methodology

Following the input image of COCO dataset, the next step is to apply pre-processing and train the dataset by filtering and annotations specific for person category only. Then, validate the model and evaluate mAP scores for checking the performance as follows:

$$mAP = 1/|classes| \sum |TP_c/|FP_c|| + |TP_c| \tag{1}$$

For testing the dataset using intersection over union (IOU) method, detect the social distancing between people using the overlapping (intersection) area. If the value of IOU is non-zero, then it can be said that people are not at proper social distancing from each other. Hence, social distancing among people can be detected. Figure 1 represents the graphical form of the proposed research steps.

95 categories from MS-COCO datasets library have been created and only one class “person” is assigned as “head” class to identify people and categorize them through bound-box through people detection.

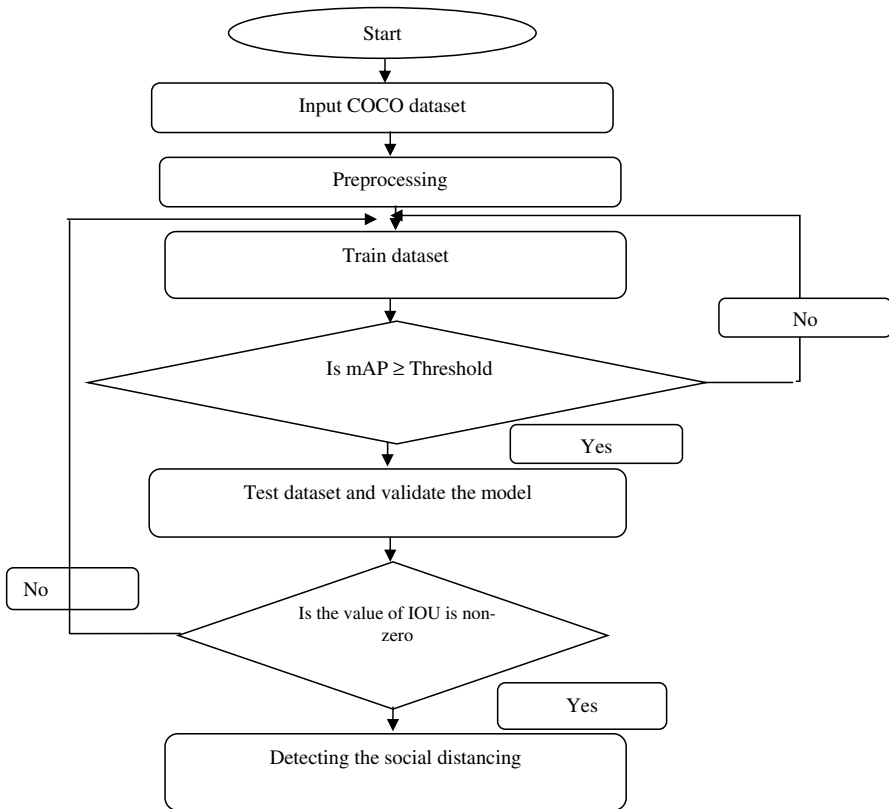


Fig. 1 Graphical form of the proposed research steps

Though detectron and detectron2 have no huge gaps in its function and ability in evaluating and processing datasets and training datasets, the more advanced and modular design-based detectron2 is identified to be extensible, highly flexible and also provides rapid training upon both multiple/single GPU servers. Henceforth, for object detection and training datasets, researchers majorly have been depending on faster R-CNN network with YOLO and similarly, MS-COCO along with PASCAL-VOC datasets, than other applications [2]; however, Wen et al. [15] had argued that, comparative to YOLOv4 models, the detectron models and datasets are not satisfying and considered as not a best configuring aspect for huge datasets but may assist the researchers for custom datasets and object detection.

The studies existing and most commonly utilized datasets and applications were thoroughly studied and examined and thus in this research, “detectron2” with MS-COCO datasets was adopted since the datasets are customizable and the model is layered with ‘Faster-R-CNN-ResNet50 and 101’ based CNN which provides, higher accuracy, precision, speed and minimal loss. The detectron2 as the machine language is an advanced machine learning-based software adopted by researchers towards detecting objects with more than 90 labels in the library with pre-defined utilities that could be installed which currently works with GPU only.

Once the detectron has been installed and hard-coded onto the computer into “database catalogue”, users could train his/her model, modify code inferences based on dataset configurations to evaluate scripts in-order to integrate the same with the final end-product. Currently, PyTorch allows the users to implement detectron2 which was initially identified as a “maskrcnn-benchmark” based detectron ground-up rewrite.

The researcher intends to adopt the IOU metric-evaluation to evaluate the precision rate with the models recall rate. In addition, the model is developed with detectron2 algorithm which makes customization of data more flexible.

In this research, the 95-COCO categories under the detectron2 library opted are given in Table 1 [16]. In this, different class and associated objects/items are presented for training the model. Though the existing models are of YOLO, RCNN, RFCN and SSD, the lack of detectron2 algorithm has motivated the researcher to develop the model with faster RCNN where the precision and recall rate could be predicted and evaluated for accuracy of object detection.

2.1 Intersection over Union

- (i) Initialize epsilon at $0.5 (1e-5)$ two boxes are pre-defined as lists;
- (ii) Define $X1, Y1$ to represent upper left corner box and $Y2, X2$ representing the right corner of box returning the IOU values;
- (iii) Initial value is set to 0 in returns of IOU where each iteration is increased by 1;
- (iv) The area is estimated with overlapping of boxes, i.e., intersection with width $(x2-x1)$ and height $(y2-y1)$ and if there is no intersection then the return is set to 0;

Table 1 Different classes of input dataset [16]

Class type	Items
Humans	Person
Vehicles	Bicycle, motorcycle, car, bus, airplane, train boat, truck
Outdoor objects	Bench, traffic light, stop sign, fire hydrant, parking meter
Animals and birds	Dog, cat, bird, horse, cow, sheep, elephant, bear giraffe, zebra
Accessories	Tie, backpack, suitcase, handbag, umbrella
Sports	Frisbee, skis, sports ball, tennis racket, kite, baseball glove, baseball bat, skateboard, snowboard, surfboard
Tableware	Bowl, bottle, glass, wine, cup, knife, fork, spoon
Eatables	Apple, banana, broccoli, sandwich, orange, carrot, hot dog, donut, pizza, cake
Furniture and plants	Dining table, potted plant, chair, bed, couch
Electrical equipment and others	TV, toilet, laptop, keyboard, mouse, remote, cell phone, toaster, microwave oven, sink, book, refrigerator, clock, scissors, vase, teddy bear, toothbrush, hair drier

- (v) Calculate the combined area as two sections: a representing sum of $(a[2]-a[0]) \times (a[3]-a[1])$ and b representing sum of $(b[2]-b[0]) \times (b[3]-b[1])$ where the combined area is calculated as $(\text{area}_a) + (\text{area}_b) - (\text{area_overlap})$; combined area ratio is calculated through IOU estimation where $(\text{area_overlap}) / (\text{area_combined} + \text{epsilon})$;
- (vi) Bounding-box (anchor box) is thus estimated and applied upon the person-class dataset to identify the person towards estimating the social distancing at threshold of 0.75.

Once the data are satisfying, the approach is applied on remaining datasets in runtime towards predicting the valid social distancing (green box) and invalid social distancing (red box) between two or more persons.

3 Proposed Model Architecture

Among the R-CNN, faster R-CNN and fast R-CNN, the CNN of faster R-CNN is viewed and identified as faster, but, however, all three ResNets are stated as similar approaches which identifies, detects, computes and classifies. Though there are numerous methods and approaches in object detection using machine language-based deep learning, faster R-CNN with detectron-based algorithm is proved to be effective than the previous models [2]. The flow of the developed model is represented in Fig. 1 where the layers of faster R-CNN are explained in detail.

3.1 Algorithm Used for Social Distancing

In the research, the focus is upon social distancing through image segmenting and object detection, and thus, the researcher was motivated to adopt the “detectron2”

as the object detection technique unlike the common techniques by existing research where YOLO is commonly adopted and compared with RCNN and FCN path-based models. Henceforth, the faster R-CNN with detectron2 as the base for the model is implemented through the following algorithms:

3.1.1 Faster R-CNN Algorithm

The faster R-CNN for the social distancing and object detection is carried through the following algorithm:

Step 1. Initially, cloning the repository for the faster R-CNN implementation is carried out;

Step 2. The folders (training datasets and testing datasets) along with the training file (.csv) are loaded to the cloned repository;

Step 3. Next, .CSV file is converted into .txt format/file with new set of data-frame and the model is then trained using train_frnn.py as file in python keras;

Step 4. Finally, the outcomes are the predicted images with detected objects as per the norms in the codes and the results are saved in separate folder as text images with bounding-box.

The architecture in this model (Fig. 2) is developed with faster R-CNN as the model for transfer learning. The reason for adopting and using faster R-CNN as the backbone in the developed model is for its features, namely: accuracy, precision and speed. In the developed model, the architecture includes 5 Convolutional layers of 128, 256, 512, 1024, and 2048. The research includes deep learning of 8 layers of ReLU activation. 3*3Conv Kernel size is utilized.

3.1.2 Detectron Algorithm for Object Detection and Social Distancing

The algorithm for the developed social-distancing evaluation through object detection for detectron2 with faster R-CNN-50 and 101 is designed as

Step 1. First, the images from COCO datasets are loaded and accessed as inputs for the developed model;

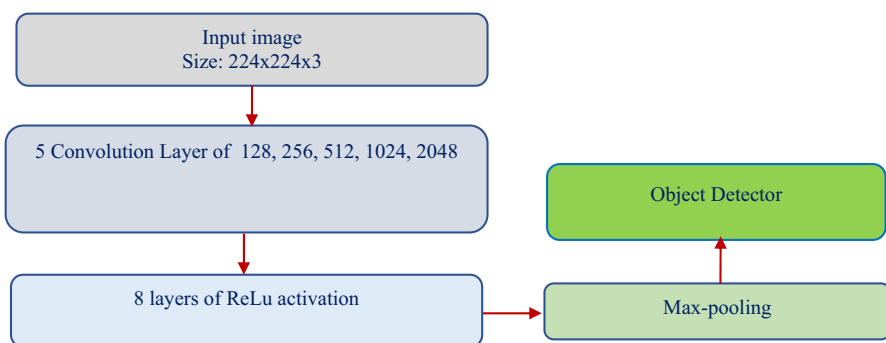


Fig. 2 Architecture of the customized detectron2 with RCNN model

Step 2. Initially, the images are passed through the ConvNet of faster R-CNNs of 50 and 101;

Step 3. Merge the model-zoo files from faster R-CNN-50 and train the sample dataset and initialize the training;

Step 4. Choose the LR with good outcome rate, where here 300 is predicted to be enough for the sample dataset;

Step 5. Next, focus on maintain the learning rate by preventing the decaying of learning rate;

Step 6. Set the ROI-head size as 512 for training datasets and label 'person class' is alone selected and the metric is obtained;

Step 7. Once the datasets are trained, the model is tested with remaining images with bounding-box identifying 'valid person' under class;

Step 8. The same processes for sample evaluation estimation are repeated in testing and if the identified person-class satisfies the norms for social-distancing criteria 'green bounding-box' is applied for the image, else 'red bounding-box' is applied onto image and the result obtained and stored in a file under the person-class.

3.1.3 Training the Model with Accumulated Datasets

The detectron2 model training is done by adhering to threshold value and norms as the following steps:

- Initially, the LR (learning rate) is set as 0.001 for the iterations of 30 k and then later the LR is gradually decreased up to 0.0001 for following 70 k iterations;
- Next, for about 40 k iterations, the LR is set at 0.00001 and then for the last round of 20 k iterations, the LR is set at 0.000001;
- The optimizer utilized here is the SGD (Stochastic-gradient-descent).

3.1.4 Procedure for Testing the Accumulated Datasets for Social Distancing

Initially, the bounding-boxes predicted through the model post-training are considered and examined with the set threshold (0.75) and if the scores are lesser, they are disregarded;

Each combined bounding-boxes, i.e., 2 boxes as a pair is evaluated for the metric calculation in the IOU at 0.5;

According to the IOU scores, if the $IOU > 0$, there is no social distancing between two people, and thus they are bounded with red box and if the score of $IOU < 0$, then the people are bounded with green box.

Thus, the algorithm, model and the norms of bound-box for social distancing are estimated/predicted and compared with the outcomes through the model developed.

4 Results and Performance Evaluation

4.1 Random Sample Dataset

The random sample dataset has been considered for implementation.

The datasets for training the model are selected randomly (refer Fig. 3a, b) to predict its reliability and accuracy and to compare the outcomes obtained with estimated outcomes.

4.2 Ground Truth Versus Predicted by Model

The following outcomes (predictions) are the results obtained from the tested model where the ground truth is compared and weighed against the models' predictions based on the developed algorithm and architecture. The results from trained detector2 with faster R-CNN model are:

4.2.1 Classification of Individual Person

The classification of individual person is explained as follows.

Figure 4a represents the ground truth of identified person in the picture versus Fig. 4b representing the prediction made by the model by identifying the person-class only. It could be inferred that accuracy and precision is similar to the outcomes. However, the study is focused on social distancing with green bound-box as correct distance and red bound-box as incorrect distancing between people. Since there is no other person involved in the picture, the algorithm identified the individual with green bound-box stating that, norms of social distancing is satisfied by the 'person'.

4.2.2 Classification of Group of People Versus Objects

Figure 5a represents the ground truth of identified person in the picture versus Fig. 5b representing the prediction made by the model by identifying the person-class only.



Fig. 3 a, b Input dataset selection

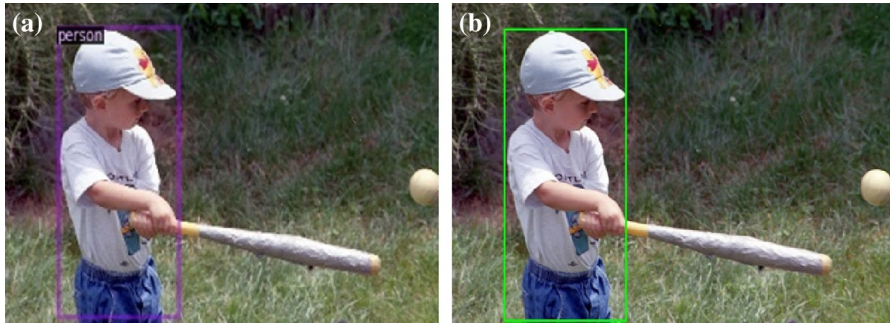


Fig. 4 a Before and b after



Fig. 5 a Before and b after

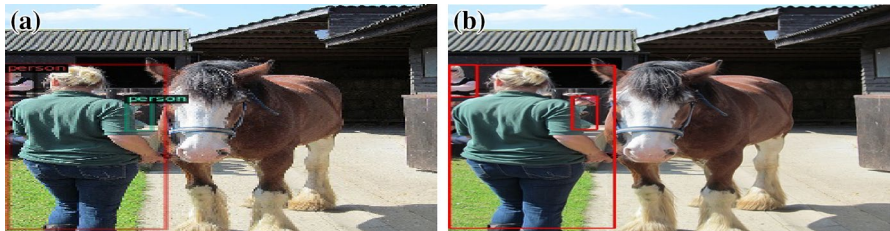


Fig. 6 a Before and b after

The picture is of mass-people where the ground truth identified 8 individual people in Fig. 5a with class-person and the predicted model provided outcomes with same head-counts of 8 people and also identified that there is no social distancing between the people and thus resulted Fig. 5b is obtained with red bounding-boxes.

4.2.3 Classification of People and Animals

Figure 6a represents the ground truth of identified person in the picture versus Fig. 6b representing the prediction made by the model by identifying the person-class only. The class ‘person’ is identified from the rest of categories and evaluated for social distancing. According to the developed algorithm, only people are

identified for social-distancing evaluation and the predicted outcome is found to be negative where there is no social distancing between people the identified in the input image .

4.2.4 Classification of People from Other Categories

See Fig. 7)

4.2.5 Graphical Outcomes

Based on the trained and tested model and estimated and obtained outcomes, the graphical representation of the time, data time, total loss and eta-seconds of the developed model had been evaluated in python and represented in the following graphical representation of Fig. 8. The losses have been calculated by considering the expected outcomes of the proposed model with the actual outcomes during the testing phase. Considering number of iteration and percentage of losses, the graph has been plotted. It has observed that the percentage of losses decreased from 0.75 to 0.3 with variation of numbers of iteration, which indicates the effectiveness of the proposed model. It is inferred through Fig. 9a–c that no huge variations are identified between the results obtained from the model and similarly the estimated outcome and obtained result of total loss is less at 0.15.

Figure 9a represents the Loss_RPN (region-proposal-networks)-LOC outcome response. There are no huge variations in loss for te class in pooling. At the 1.8th K iterations, loss minimized from 0.2 to 0.05%. Concluding that, the model is accurate and precise in detecting objects and perfectly measures the social distancing.



Fig. 7 a–g Prediction by model

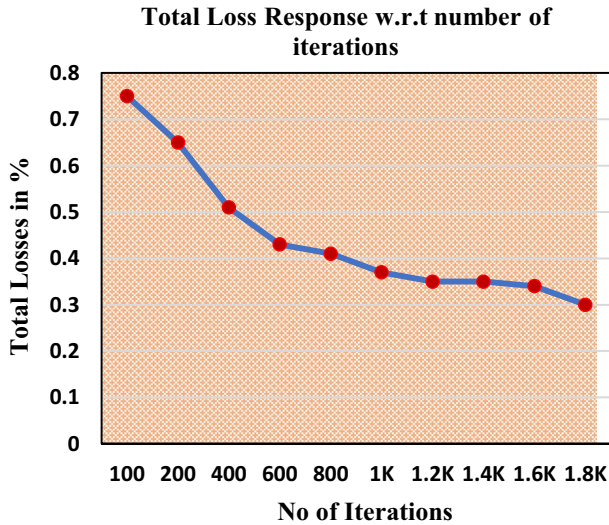


Fig. 8 Total loss response

Similarly in Loss_RPN_Class (Fig. 9b), the loss parameters are exponentially decaying from 0.03 to 0.01 approximately as the number of iterations increases.

This signifies the best outcomes of the proposed model. Again, Fig. 9c represents the loss from regression of bounding-box outcomes. The results evidently insist that regions are overlapped, and thus, NMS (non-maximum suppressions) is used towards minimizing the proposal numbers. Therefore, the loss is minimized at 0.15 from 0.34 as total loss of the developed model.

The outcomes from Fig. 10a–c denote that faster R-CNN outcome of accuracy is achieved at 98% where the foreground accuracy is achieved at 10% and false negative at 0.1% concluding that the model is accurate and precise in detecting objects and measuring social distancing with higher outcome than estimated/predicted outcome of 75% (threshold value) and above. Figure 10a interprets that there are no such remarkable variations of exact iteration value and moving average value in actual outcomes. From 600 to 1.8 k iteration, the value remains steady at 0.2 loss which is a good indication of systems performance.

With the help of fast RCNN model, the foreground (FG) class accuracy has been computed and the response is presented in Fig. 10b. The conclusion from the response is that as the order of iteration increases, the FG class accuracy also increases and maintained steady response from 1.6th K iterations.

The foreground class accuracy (Fig. 10c) was calculated in fast RCNN model. Initially, the class accuracy dropped from 0.4 to 0 at 200th iteration and again the accuracy increased with increasing in iterations. The steady response has spotted from 1.4th K iterations which is approximately equal to 0.8 or 80%

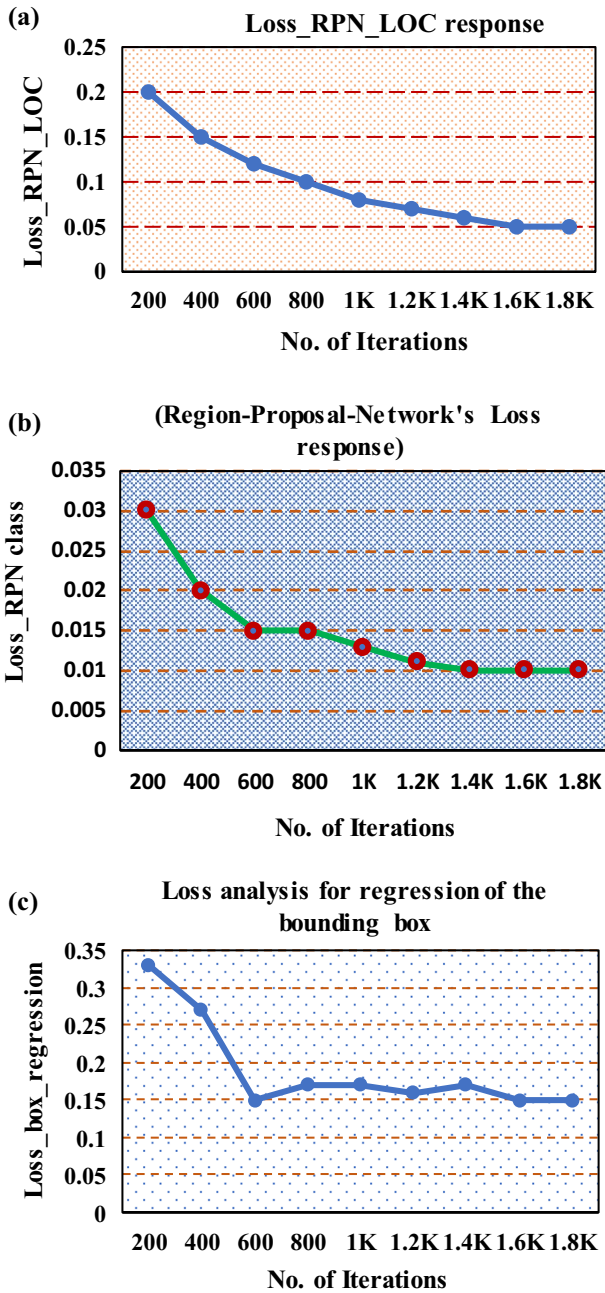


Fig. 9 a Loss_RPN_LOC response. b RPN loss response. c Loss analysis for box regression

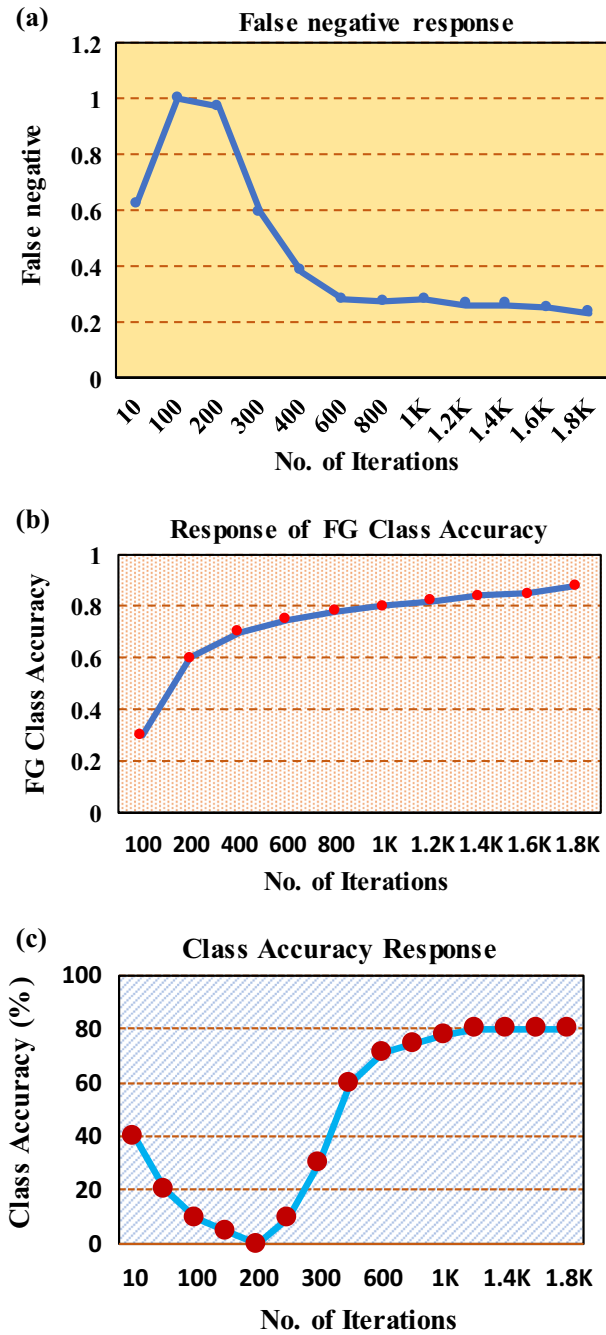


Fig. 10 a Faster RCNN false-negative response. b FG class accuracy response. c Faster RCNN class accuracy

4.2.6 Scheduling LR

The learning rate is set to 4 k intervals where the developed model attained successful learning rate at 1000 k and remained the same till 23 k iterations stating that there is no sudden dropping down in the learning rate but rather up to 70 k iterations, the LR steadily decreased and remained constant from 1000 k.

4.3 Scores Post-testing and Training the Model

The scores for the developed model after testing and training the processed dataset have been obtained, and through outcome values, the study concludes that the detectron2 model with faster RCNN as architecture where IOU metric is adopted to evaluate the model. The outcome of the trained model conforms a mAP of 84.5%, precision of 97.9%, 87% recall and with total loss of 0.1%.

4.4 Performance Metrics

The object detection (human) in the model developed has acquired recall rate of 87% and precision of 97.9% which is more than average 75% stating that the model is a success with effective precision with minimal total loss of 0.1 and mAP at 84.5%, where existing models has lack precision in human detection towards social-distancing threshold violation measures (refer Fig. 11). Thus, the researcher examined and evaluated the datasets with developed detectron2 model where it has evidently concluded that the developed model is a success and good fit for object detection-based

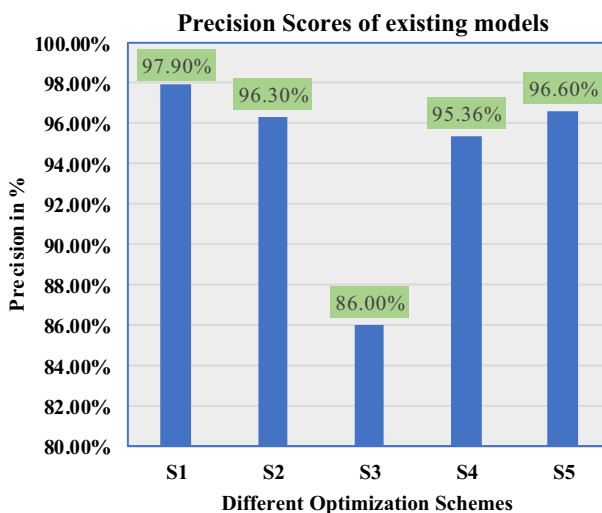


Fig. 11 Performance evaluation and comparison with existing approaches. Where S1: RCNN with detectron2 and fast RCNN (Proposed method). S2: RCNN with TWILIO communication platform. S3: YOLOv3 with TL. S4: fast RCNN with YOLO v4. S5: fast RCNN with YOLO v2

analysis models and for violation threshold-based applications in object detection and monitoring. Majorly for evaluating the social distance criterion, the model is reliable, accurate, precise and also has a fine recall score (87%) with better mAP (84.5%) that exceeds the average score of existing models.

Figure 11 exemplifies that among the existing models, the developed model with detectron2 with faster R-CNN architecture acquired higher precision rate of 97.90% (98% approximately) than other models, where

- R-CNN with TWILIO communication platform architecture attained 96.30%;
- Fast R-CNN with YOLOv2 architecture attained 95.60%;
- Faster R-CNN with YOLOv4 architecture attained 95.36%;
- TL with YOLOv3 architecture attained 86.0%.

The investigation particularly aimed at analyzing and developing a better model with object detection-based machine language-adopted architecture where it could attain higher precision rate than average metric scores attained by the existing models. The study adopted IOU metric-evaluation that uses mAP (mean average of precision) which examines and evaluates the developed object detection model. Table 2 summarizes the performance of both proposed model and other existing model in terms of training time (TT), number of iterations (NI) mAP and total loss (TL) during the simulation of training phase. It is observed that the developed model (RCNN with detectron2 and fast RCNN) architecture acquired higher precision rate of 97.90% (98% approximately) than other models.

5 Conclusion and Future Recommendation

The study mainly aimed to developing an Optimized deep learning approach for the prediction of social distance among individuals in public places in real time, which mainly focus with object detection-based machine language-adopted architecture where it could attain higher precision rate than average metric scores attained by the existing models. The extensive trials were conducted with popular state-of-the-art object detection model: RCNN with detectron2 and fast RCNN, RCNN with TWILIO communication platform, YOLOv3 with TL, fast RCNN with YOLO v4, and fast RCNN with YOLO v2. Among all, the proposed (RCNN with detectron2 and fast RCNN) delivers the efficient performance with precision,

Table 2 Performance parameters' evaluation

Model	TT in s	NI	mAP	TL	Precision in %
S1 (proposed)	5659	7500	0.845	0.01	97.90
S2	6552	1200	0.965	0.22	96.30
S3	8732	1500	0.985	0.35	86.00
S4	7754	3200	0.905	0.42	95.36
S5	6352	1143	0.958	0.35	96.60

mAP, TL and TT. The outcomes of the proposed model focused on faster R-CNN for social-distancing norms and detectron2 for identifying the human ‘person class’ towards estimating and evaluating the violation-threat criteria where the threshold (i.e., 0.75) is calculated. The model attained precision at 98% approximately (97.9%) with 87% recall score where IOU was at 0.5. In future, the study may be applied by researcher on similar context through two-stage detectron instead of single with fast R-CNN or DNN as the architecture to find variations and relationships between the same datasets under different techniques.

Acknowledgements The authors also extend their appreciation and gratitude to Al-Mustaqbal University College in Iraq for funding this project.

Authors Contributions The authors contributed equally in this work.

Funding This work is funded by Al-Mustaqbal University College in Iraq.

Availability of Data and Materials Simulation software.

Declarations

Conflict of Interest No competing interests.

References

1. Brodeur, A., Islam, A., Gray, D., Bhuiyan, S. J.: “A Literature Review of the Economics of COVID-19”, Discussion paper series, IZA- Institute of Labor Economics, DP No. 13411, pp. 1–61 (2020).
2. Punn, N.S., Sonbhadra, S.K., Agarwal, S., Rai, G.: Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLOv3 and Deepsort techniques, pp: 1–10. Retrieved on 19th July from <https://arxiv.org/pdf/2005.01385.pdf> (2021).
3. Ahmed, I., Ahmad, M., Jeon, G.: Social distance monitoring framework using deep learning architecture to control infection transmission of COVID-19 pandemic. *Sustain. Cities Soc.* **69**(102777), 1–11 (2021)
4. Ahmed, I., Ahmad, M., Rodrigues, J.J.P.C., et al.: A deep learning-based social distance monitoring framework for COVID-19. *Sustain. Cities Soc.* **65**(102571), 1–12 (2021)
5. Payedimarri, A.B., Concina, D., Portinale, L., Canonico, M., et al.: Prediction models for public health containment measures on COVID-19 using artificial intelligence and machine learning: a systematic review. *Int. J. Environ. Res. Public Health* **18**(4499), 1–11 (2021)
6. Rahim, A., Maqbool, A., Rana, T.: Monitoring social distancing under various low light conditions with deep learning and a single motionless time of flight camera. *PLoS ONE* **16**(2), e0247440 (2021)
7. Tuli, S., Tuli, S., Tuli, R., Gill, S.S.: Predicting the growth and trend of covid-19 pandemic using machine learning and cloud computing. *Internet Things* **11**, 100222 (2020)
8. Pooranam, N., Priya, S.P.N., Sruthi, S., Dhanya, S.K.: A safety measuring tool to maintain social distancing on COVID-19 using deep learning approach. *ICCEBS J. Phys. Conf. Series* **1916**(012122), 1–8 (2021)
9. Vinitha, V., Velantina, V.: Social distancing detection system with artificial intelligence using computer vision and deep learning. *Int. Res. J. Eng. Technol. (IRJET)* **7**(8), 4049–4053 (2020)
10. Pandian, K: Next-generation network - global market trajectory & analytics. *J. Next Gener. Comput. Syst.* **33**(2), 1–20 (2020)
11. Rezaei, M., Azarmi, M.: DeepSOCIAL: social distancing monitoring and infection risk assessment in COVID-19 pandemic. *Appl. Sci.* **10**(7514), 1–29 (2020)

12. Saponara, S., Elhanashi, A., Gagliardi, A.: Implementing a real-time, AI based, people detection and social distancing measuring system for Covid-19. *J. Real-Time Image Process.* **18**(6), 1937–1947 (2021)
13. Arya, S., Patil, L., Wadeganokar, A., et al.: Study of various measure to monitor social distancing using computer vision: a review. *Int. J. Eng. Res. Technol. (IJERT)* **10**(5), 329–326 (2021)
14. Yang, D., Yurtsever, E., Renganathan, V., et al.: A vision-based social distancing and critical density detection system for COVID-19. *Sensors* **21**(4608), 1–15 (2021)
15. Wen, H., Huang, C., Guo, S.: The application of convolutional neural networks (CNNs) to recognize defects in 3D-printed parts. *Materials* **14**(2575), 1–14 (2021)
16. <https://www.kaggle.com/datasets/awsaf49/coco-2017-dataset>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Santosh Kumar Sahoo¹ · **G. Palai**² · **Baraa Riyadh Altahan**³ ·
Sk Hasane Ahammad⁴ · **P. Poorna Priya**⁵ · **Md.Amzad Hossain**^{6,7} ·
Ahmed Nabih Zaki Rashed⁸ 

✉ Md.Amzad Hossain
mahossain.eee@gmail.com

✉ Ahmed Nabih Zaki Rashed
ahmed_733@yahoo.com

Santosh Kumar Sahoo
santosh.kr.sahoo@gmail.com

G. Palai
gpalai28@gmail.com

Baraa Riyadh Altahan
Baraa.riyadh@mustaqbal-College.edu.iq

Sk Hasane Ahammad
ahammadklu@gmail.com

P. Poorna Priya
hodece@diet.edu.in

¹ CVR College of Engineering, Hyderabad, Telangana 501510, India

² Department of Electronics and Communication Engineering, Gandhi Institute for Technological Advancement, Bhubaneswar, Odisha, India

³ Medical Instrumentation Techniques Engineering Department, Al-Mustaqbal University College, Hilla, Babil 51001, Iraq

⁴ Department of ECE, Koneru Lakshmaiah Education Foundation, Andhra Pradesh, Vaddeswaram 522302, India

⁵ Department of ECE, Dadi Institute of Engineering and Technology, Anakapalle, Visakhapatnam, India

- ⁶ Institute of Theoretical Electrical Engineering, Faculty of Electrical Engineering and Information Technology, Ruhr University Bochum, 44801 Bochum, Germany
- ⁷ Department of Electrical and Electronic Engineering, Jashore University of Science and Technology, Jashore 7408, Bangladesh
- ⁸ Electronics and Electrical Communications Engineering Department, Faculty of Electronic Engineering, Menoufia University, Menouf 32951, Egypt