

A hybrid deep learning based approach for the prediction of social distancing among individuals in public places during Covid19 pandemic

Santosh Kumar Sahoo*

Department of Electronics and Instrumentation Engineering, CVR College of Engineering, Hyderabad, Telangana, India

Abstract. Social distance is considered one of the most effective prevention techniques to prevent the spread of Covid19 disease. To date, there is no proper system available to monitor whether social distancing protocol is being followed by individuals or not in public places. This research has proposed a hybrid deep learning-based model for predicting whether individuals maintain social distancing in public places through video object detection. This research has implemented a customized deep learning model using Detectron2 and IOU for monitoring the process. The base model adapted is RCNN and the optimization algorithm used is Stochastic Gradient Descent algorithm. The model has been tested on real time images of people gathered in textile shops to demonstrate the real time application of the developed model. The performance evaluation of the proposed model reveals that the precision is 97.9% and the mAP value is 84.46, which makes it clear that the model developed is good in monitoring the adherence of social distancing by individuals.

Keywords: Covid19, social-distancing, deep learning, Detectron 2, Intersection over Union, video object detection

1. Introduction

The concept, ‘Social distancing’ has gained greater importance in the present day. Ever since the outbreak of the Covid19 pandemic, social distancing has been considered one of the most vital measures to break the chain of the spread of the virus. With the first and second waves of Covid19 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 to

break the chain of infections till a majority of the population are vaccinated. In the field of artificial intelligence, researchers are coming up with solutions that could be implemented in order to ensure that the Covid19 protocols are strictly followed by the people. However, there are cases in countries like India, France, Russia, and Italy where either people are heavily populated or do not adopt preventive measures like social distancing in crowded places [1]. Social distancing is a healthy practice or preventive technique which evidently provides protection from transmission of the Corona (Covid19) virus. It has been identified that when there is a maintenance of a minimum of 6 feet distance between two/more people [18] the spread of the virus could be avoided. Social distance doesn't have to be always people's

*Corresponding author. Santosh Kumar Sahoo, Department of Electronics and Instrumentation Engineering, CVR College of Engineering, Hyderabad, Telangana, India. E-mail: santosh.kr.sahoo@gmail.com.

preventive technique rather it could also be a practice that could be adhered to reduce physical contact towards transmitting diseases from a virus-affected person to another healthy being [9]. Deep learning like its role in every field has been found to create a significant impact in addressing this problem [4]. Manually monitoring, managing and maintaining distances between people in a crowded environment like colleges, schools, shopping marts and malls, universities, airports, hospitals and healthcare centres, parks, restaurants and more places would be impractical and therefore 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 faster detecting options. Faster RCNN is found to be more advantageous, faster convergence and also provides higher performance [3] even under a low light environment.

In the regions of America, Europe and South-East Asia due to poor social distancing and improper measures against Covid19 by the people, in the years 2020 and 2021 many fatalities were recorded and most of them which were because of the violation of the threshold in social distancing measures during Covid19 [10]. The concept of adherence to social distancing has been implemented by machine learning researchers in the past through techniques like face recognition, object (human) identification, people monitoring, object identification, image processing and so on [29, 32].

Deep learning and machine learning-based based Artificial Intelligence (AI) models surpass various criteria like time consumption, cost reduction and reduction of human intervention. Machine learning-based solutions help in achieving results with greater accuracy, minimal loss and errors in addition to reducing the prediction time from several years to a few days [16]. Majorly R-CNN (region-based CNN) and faster R-CNN are applied for accurate and faster predictions. To measure the predicted outcomes, researchers generally utilize the metric evaluation approaches/ techniques like regressors (Random-Forest-Regressor, Linear Regressor), IOU (intersection-over-union) and more. For object detection, the IOU approach is commonly adopted to evaluate performance.

The uses and scope of RCNN in object detection models are mainly to identify/detect the image/video and bounds the class targeted by researchers. The bounding box separates the class needed from other

classes and objects in an image. Thus to validate the bounding box, investigators can code them with colors (for instance *green* for no error, *red* for error) and also use RCNN to label the images under categorized files [7]. Similarly, the RCNN application is found to be majorly identified in the defense system by the militaries, automobiles using object detection like Tesla, surveillance cameras and securities in banking, malls, hospitals, jewelry shops, and more that are crowded and need an advanced security system [27]. Recently, due to advancements in technology and the advantages of surveillance, shops with larger-area to monitor have adopted surveillance cameras (with heat, objects, light, and other sensory detections) to secure their shops [11]. Thus RCNN in object detection has higher scope in identifying, classifying and labeling images/videos.

This research has developed a customized deep learning RCNN-based architecture. A hybrid approach combining Detectron2 and Intersection over Union method has been applied in order to predict whether the social distance is maintained by the individuals or not, especially in crowded places and public places.

Purpose: The study is mainly motivated towards examining and analyzing the social distancing among the individuals (i.e. person class), in public and crowded areas with the use of RCNN architecture where the IOU method is used for evaluating the performance of the developed model. Detectron2 has been used to identify 'person class' from other classes in an image to measure the social distance between the identified class, people. A similar research as a 'case study' approach would also be examined for adaptation of deep learning and object identification through image as input to measure social distancing between individuals.

2. Related works and research background

Social distancing post Covid19 was vigorously studied by many researchers where prediction models were developed to maintain and monitor social distancing. Authors [31] focused on the prediction of social distancing post Covid19 outbreak in India with DNN architecture. [19] utilized the TWILIO and the authors [31] utilized the YOLOv3 as their base architectures respectively for addressing the same problem. These studies utilized the faster R-CNN and found that the architecture is more flexible and faster than other existing approaches. The conclusions from

the studies revealed that efficient and effective object-detection outcome is attained with the YOLOv3 with faster R-CNN model. However authors [14] developed a YOLOv4-based ML model and tested the model with COCO datasets. Their model architecture was also based on DNN (Deep-Neural-network). The latter model was found to render 99.8% accuracy with a speed of 24.1 frames per second (fps) when tested on a real-time basis. Authors [26] had developed an AI-based social distancing model which measured the distances among two-or-more people with social-distancing norms. The authors adopted YOLOv2 with the Fast R-CNN for detecting the objects. The objects considered by them were 'Humans' through person class. The accuracy was 95.6%, precision was 95% and recall rate was 96%.

Authors [22] examined the different techniques of metrics in evaluating the prediction models to monitor the social distancing via 'Computer-vision' as an extensive-analysis with existing social distancing reviews. They concluded that mAP is found effective in evaluating the prediction models. The study also focused on examining the performances of security-threats identification and facial expressions-based models that had adopted deep learning and 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. Henceforth, from the study it's streamlined that, adopting two-stage object detectors is wiser, efficient, effective and accurate.

The study by [6] investigated the social-distancing concept, post Covid19, by developing the vision and critical-density based detection system. They developed the model based on two major criteria. First criterion was to identify violations in social-distancing via real-time vision-based monitoring and communicating the same through deep-learning-model. Second the precautionary measures were offered as audio-visual cues, through the model towards minimizing the violation threshold to 0.0, without manual supervision by 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%. Similarly, accuracy at 92.80% and recall score at 95.94% were also achieved.

The study [10, 11] focused on deep-learning architecture that utilizes the social-distancing concept as a base for their evaluation with people monitoring/ management as a secondary focus, post the

Covid19. The authors [10] have utilized YOLOv3 for identifying humans and faster R-CNN as the social-distancing algorithm. The model achieved 92% accuracy in tracking, without transfer-learning and 98% with transfer-learning; similarly, 95% of tracking accuracy. The study insisted that social-distancing detection through YOLO with tracking transfer-learning technique is highly effective than other deep-learning models. Authors [17] proposed a study to monitor Covid19 based social distance with person detection and tracking with fine-tuned YOLOv3 and Deep sort techniques. The authors developed the deep learning model for task automation, to supervise social distances between people, using surveillance videos as datasets. The model used YOLOv3 for object detection, towards segregating background and foreground people; similarly, deep sort methods were utilized towards tracking the identified people, with assigned identities and bounding boxes. The deep sort tracking method together with YOLOv3 showed good outcomes with balanced scores of fps (frames-per-second) and mAP (mean-Average-Precision) towards supervising real-time social-distancing, among people.

Similarly a study by [3] proposed an advanced social-distance monitoring model to surpass the drawback of existing models where low light surroundings are the major hindrance in the identification of people, as the foreground. Identifying and monitoring people, especially in crowded and low-light environments in deep-learning model prediction models, especially in circumstances like Covid19 and other pandemic circumstances has made researchers to develop new models and improvise existing models, where social distancing is a must. The developing disease caused by the SARS-covid-2 virus (Covid19) has acquired a worldwide crisis, with its deadly distribution, all-over the globe. Due to isolation and social distancing, people try to walk out of their homes, during the nighttime with their families to breathe in some fresh air. In such circumstances, it is essential to consider efficient steps towards supervising the criteria of safety distancing to avoid positive cases and to manage the death toll.

Thus, from the existing researches and studies it is observed that YOLO for object detection has been adopted by the investigators majorly, with RCNN and Faster RCNN in their architecture. Though YOLO detects the objects meaningfully, it lacks behind in detection speed and thus affects the tracking process. Whereas, according to [24] the Detectron2 model is found effective and rapid in tracking processes than

Table 1
Reviews of social distance prediction during covid-19

S. No.	Author & Year	Technology used	Benefits of Technology	Approach Used
1	[19]	Deep Neural Network	Helps developers to analyze the data publicly	TWILIO
2	[31]	Computer Vision and Deep Learning	Monitors social distancing at workplaces and public places	YOLOv3
3	[14]	Deep Neural Network	Helps to examine the social-distancing of people post Covid19 and infectious assessment	YOLOv4
4	[26]	Artificial Intelligence System	Classifies social distancing people using thermal images	YOLOv2 v
5	[22]	Deep Learning	Monitors whether people are maintaining social distance among each other, by examining real time streams of video, from camera in workplace and public places	YOLOv3
6	[6]	Artificial Intelligence	Measures social distancing using real time detectors of object	YOLOv4
7	[10]	Deep Learning	Tracks social distance using an overhead prospect	YOLOv3
8	[9]	Faster-RCNN	Detects human in images	YOLO
9	[17]	Deep sort approach	To track the people with the use of assigned identities and bounding boxes	YOLOv3
10	[3]	Deep Learning	Monitors the distance criteria of safety to avoid positive cases and to handle the toll of death	YOLOv4
11	[13]	Deep Learning	Parallel Deep Learning model that uses community detection technique in 16 real-world large CNs (complex-networks)	BP-PSO (Back-propagation-particle swarm optimization)

Source: Author.

YOLO. Thus, in this research, a deep learning-based method is proposed using the RCNN with Detetcron2 for measuring social distances among individuals and detection of real-time object-class (i.e. person class) with a single motionless ToF (time of flight) camera.

The below Table 1 shows the reviews of social distance prediction during covid-19:

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. However the usage of models with different neural networks, algorithm and evaluation metrics are observed to be lesser through the reviews undertaken.

Research gap: There have been several research works on solutions to overcome the impact of pandemic after the outbreak of Covid19, especially through social distancing as a norm. Numerous machine learning studies were attempted to predict the social distancing in public areas. Models in predicting the social distances by authors [21] focused on hospitals, ATMs and malls as crowded area with DNN model; [8] developed a stereo-vision based model for monitoring and predicting social distances; [28] developed a model that predicts social distancing and monitors face-masks in peo-

ple with object detection based mobilenetv2 model; [20] developed an IoT with Passive-Infrared sensors based social distancing prediction model; [25] developed a computer-vision based prediction model; [2] developed a social-distancing prediction model with image depth estimation; [15] developed a Infrared-array based sensor model that detects and monitors social distancing; authors [33] developed a UAV (unmanned-aerial vehicle) based PeelNet incorporated social distancing model; [12] developed a VSD model through automated estimation of distances based on inter-personal spaces between the people identified; [30] developed a MobileNet based model that predicts social distancing through sound-wave technique and [13] developed a parallel deep learning model that used BP-PSO (Back-propagation-particle swarm optimization) and community detection techniques.

The present study has developed a novel approach to detect whether social distancing has been maintained by individuals at public places. The current approach has combined “Detectron2” and “IOU” techniques and developed an R-CNN based customized machine learning model for detecting the social-distancing between people, especially in the crowded areas. This research has developing and tested a customized model for social distancing pre-

diction, through video object detection using COCO dataset.

Novelty of the work: Authors [10] have proposed and implemented an approach for prediction of social distance. His model adapted YOLOv3 as the base model for transfer learning. The model rendered an accuracy of 95%. Similarly the authors [23] have proposed an approach for social distance prediction and mask prediction using IoT using Faster RCNN and IoT. However, they had not developed any novel model. Authors [19] proposed a model with R-CNN architecture where TWILIO as communication-platform. The model achieved 96% accuracy with loss of 3%. Contrarily, authors [6, 26] each developed R-CNN model with YOLOv4 and YOLOv2 respectively, in predicting social distancing between people in crowded areas, aiming at achieving higher accuracy than existing R-CNN based YOLO models. Howsoever the both studies achieved 95% accuracy, which was not higher than existing models.

In this research, the author has developed a Detectron2 model with IOU metrics for predicting the social distancing between individuals in crowded public areas that has not been attempted by any authors, yet. This paves way in the research to adopt new evaluation techniques and prediction models than attempting to adopt existing models and comparing with other similar models.

3. Proposed system

The system proposed here is a social distancing technique for analyzing objects that utilizes the deep learning, computer vision along with python language, for detecting intervals (distance) among people for maintaining safety and monitoring crowd. RCNN was used to develop the model along with deep learning algorithms, computer vision and CNN (convolutional neural networks). To detect objects (people) from video, object detection method was used where ‘Detectron’ algorithm is applied. From CoCo datasets and obtained results with class-as-objects “People” as the class was alone filtered, by disregarding other classes. To map people from other classes in the video or frame, bounding boxes are mapped with different colors. Finally, distances between identified people are measured through obtained outcomes.

Proposed approach: The research focuses on social-distancing between people through video frames as dataset. The datasets are obtained from

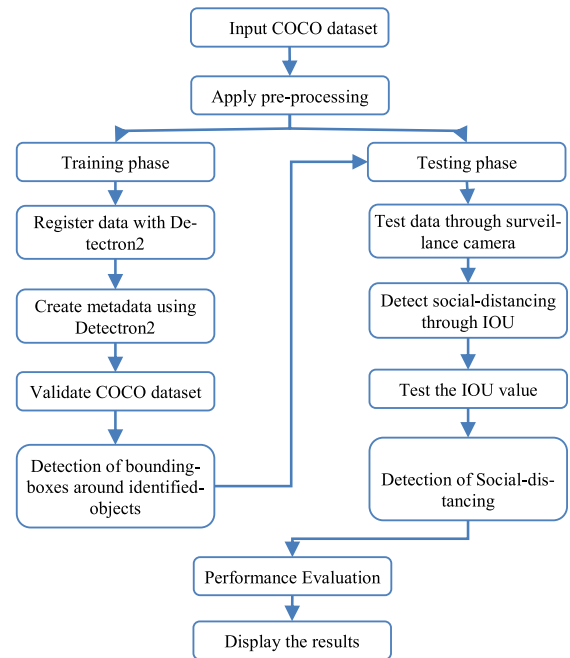


Fig. 1. Flow-chart of proposed research.

existing library and programming language used is python. Faster RCNN is used for developing the model with detectron2. Object detection algorithm is used to differentiate humans (people) from other classifications. The approach proposed is represented through flow-diagram in the following section.

Flow-diagram: The flow-diagram (refer to Fig. 1) represents the research approach in detailed manner:

Proposed Architecture: Among the R-CNN, faster R-CNN and fast R-CNN, faster R-CNN is viewed and identified as effective in terms of performance. Though there are numerous methods and approaches in object detection using machine language based deep learning, faster R-CNN with Detectron based algorithm has been proved to be effective than the previous models [18].

The architecture (refer to Fig. 2) is developed with faster R-CNN. The reason for adopting and utilizing faster R-CNN as the backbone in the developed model is for its features namely: accuracy, precision and speed. The developed model architecture includes 5 Convolutional layers with a configuration of 128, 256, 512 and 1024. The research includes deep learning with 8layers of ReLU activation function. 3*3Conv Kernel size has been adapted.

The image is fed as input in the first layer of proposed detectron model with image size of 224x224x3. In the middle layer of the model, 2d convolutional,

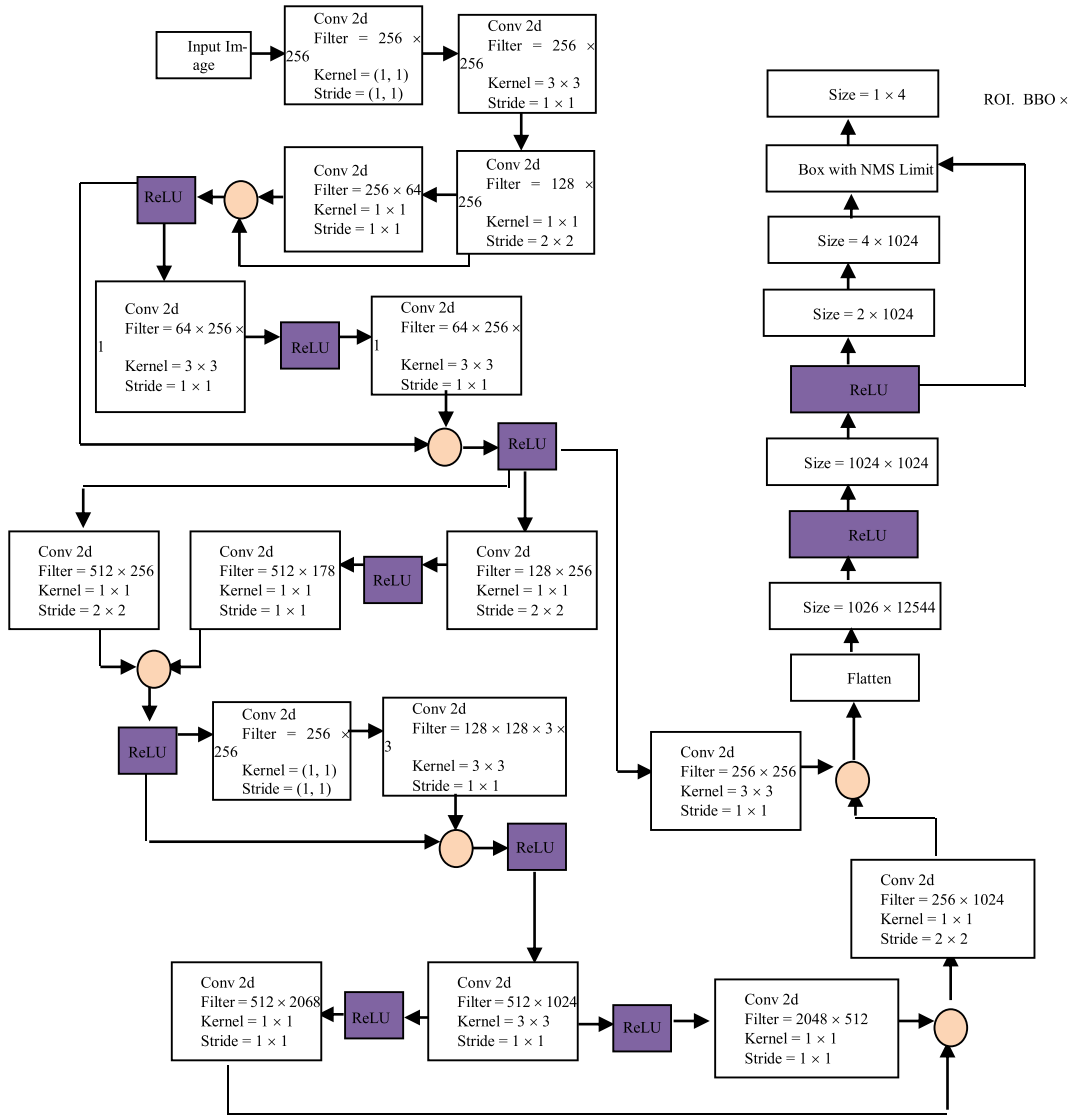


Fig. 2. Architecture of the customized detectron2 with RCNN model.

batch normalization ReLU and max-pooling are included. The convolutional layer here maps image features, where the filter is set-as 5x5. The width along with height of image in regions, is defined through filters. In regularizing model and eradicating the overfitting challenge/issue, the batch normalization is included in the model as middle layer. Later, the ReLU loss-function is used in the neural-network to introduce non-linearity. Finally, the max-pooling were included to obtain down-sampled images in pooling-region.

The stride-size is set at 2x2. Detectron2 ResNet50 is utilized for object detection in the model with similar layers.

Based on the architecture the study adopts the following schematic representation (refer to Fig. 3). The image is passed through the convolutional layers (Backbone network) where RPN transforms the image as feature maps. In the transformation RPN (Regional proposal network) identifies the regions with objects known as ‘proposals’. Identified proposals will be examined for feature extraction, with RPN as backbone network. If there are no proposals found, the ROI extracts features directly from RPN and passes on to classifier for object detection network layer. Thus the faster-RCNN in this research is developed for object detection along with Detectron.

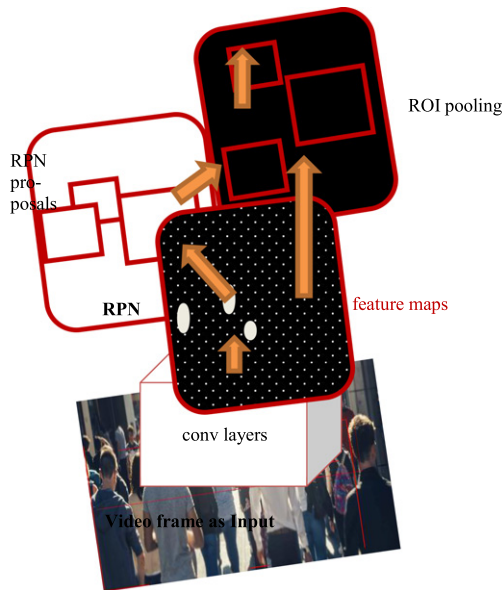


Fig. 3. RCNN schematic diagram.

Model developed in Classifying Social Distancing: In the developed model the following algorithms were used:

RCNN: 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) is 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_frcnn.py* as file in python keras;

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

Detection: The algorithm for the developed social-distancing evaluation through object detection for detedctron2 with faster R-CNN50 and 101 is designed as:

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

Step 2: Initially the images are passed through the ConvNet of Faster R-CNNs of 50 and 101;

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

Step 4: Choose the Learning-Rate with good outcome; here 300 is predicted to be enough for the obtained sample dataset;

Step 5: Next, focus on maintaining 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 and bounded through object detection;

Step 8: The same process is repeated in testing and if the identified person class satisfies the norms for social distancing criteria; where 'green bounding-box' is applied for the image with proper social-distancing and 'red bounding-box' is applied onto image with no social-distancing. Finally, the results obtained are stored in a file under the person-class folder.

Thus the social-distancing along with detecting people (person-class) upon video frames as images is done in this research. The steps involved are represented in the following illustration (refer to Fig. 4).

IOU: To predict and bound the class in identified object class (person) the IOU is used and the model's predicted results are compared with estimated outcome. IOU is also used for metric evaluation to evaluate the model's precision and accuracy. The IOU (Intersection-over-Union) is calculated through algorithm:



Fig. 4. Classification and social-distancing of person-class.

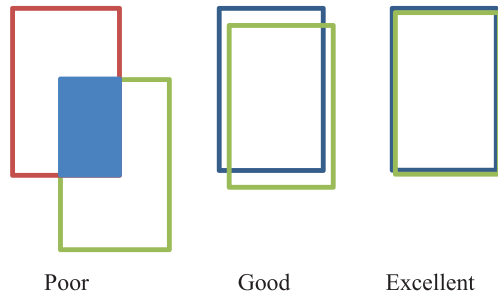


Fig. 5. IOU estimation in object detection.

Step 1: Initialize epsilon at 0.5 ($1e-5$) where the two boxes are pre-defined as lists;

Step 2: Represent the bounding box, with IOU values with 'X1,Y1' as box's upper-left corner and 'Y2,X2' as upper-right corner;

Step 3: Initially the value is set at 0 and after each-iteration the value is increased by 1;

Step 4: Later, estimate the area through overlapping of boxes where height ($Y2-Y1$) and width ($X2-X1$) intersects. If there are no overlapping (i.e. intersection of boxes) then the value is set at 0;

Step 5: Combined areas are then measured as two-sections, firstly, sum of a ($a[2]-a[0] \times (a[3]-a[1])$) and b ($b[2]-b[0] \times (b[3]-b[1])$), secondly, by estimating $(area_a)+(area_b) - (area_overlap)$; similarly combined area-ratio is measured with IOU estimation $(area_overlap) / (area_combined+epsilon)$;

Step 6: Boundingbox is estimated and applied on the input-image to classify the object (person-class) to estimate social-distancing;

Thus IOU is estimated through bounding-box intersection (refer to Fig. 5) as:

Generally, the more the bounding-boxes overlaps the excellent the score is and vice-versa. Hence in the research, 0.75 as threshold for bounding-box in object detection for person class is used in IOU metric evaluation.

4. Methodology

The research is done to measure the social distancing between people as "person class" through object detection method in deep learning technique. For the proposed aim, the developed model adopts neural networking (RCNN) and deep learning with python as programming-language. Initially the datasets are acquired, trained-and-tested and evaluated for its efficiency through IOU evaluation method. For

Table 2
Classes in COCO dataset

– 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.

detection of objects ('person class') 'detectron2' is utilized.

Datasets Used: The research uses the video streams and frames for detecting objects to identify the target-object (people as person-class) class from the other class-categories; henceforth, person class alone will be used for identifying and detecting objects for the developed model. The research is to examine the social distancing between people and thus other categories would be ignored.

The datasets were obtained from COCO dataset. The dataset has 95categories of classes (refer Table 2), where, researchers opt for required object detection classes alone.

Here 'person' as class from Humans category is alone opted and other 9categories are ignored for the prediction model. The data is obtained from open source through "Kaggle" [5].

Training: The model developed is trained with small sample datasets, to test the performance and accuracy of predictions made. If the outcomes are unsatisfactory the model is modified according to the Training:

Input: COCO 2017 dataset

Output: Bounding boxes around objects detected

1. Filter the COCO dataset and annotations specific for person category only;
2. Register the data using DataCatalog feature of Detectron2;
3. Create Metadata using MetadataCatalog feature of Detectron2 which is a key-value mapping

that contains information contained in the entire dataset;

4. Train the model on COCO dataset;
5. Use the model on validation COCO dataset and evaluate mAP scores.

Followed by the training phase, the research uses the SGD algorithm for gradient descent in developed model.

Overview of gradient descent (GD): The research utilizes the SGD algorithm for its implementation and efficient usage in neural networking. However, it lacks in two features. Firstly, it requires large iterative numbers and regularization parameters as hyper-parameters and secondly, its hyper-sensitive towards feature scaling.

Generally, the GD variants could be basically categorized as SGD (stochastic gradient descent (SGD)), BGD (batch gradient descent) and MBGD (mini-batch gradient descent). When compared against other gradient descents (Eq.1 and 2), the SGD (Eq.3) is found effective as loss function for object detection. The BGD also known among researchers as ‘Vanilla GD’ is formulated as:

$$\varphi = \varphi - \eta \cdot \nabla_{\varphi} K(\varphi) \tag{1}$$

This formula is mostly utilized for cost function gradient by researchers since this doesn’t allows the researcher to update gradients more than once and thus slower in operation processes.

The MBGD is formulated as:

$$\varphi = \varphi - \eta \cdot \nabla_{\varphi} K(\varphi; a^{(m:m+x)}; b^{(m:m+x)}) \tag{2}$$

Though the MBGD has stable convergence it allows only smaller batches (50in size) for evaluation and thus SGD is mostly preferred by researchers. The SGD formula utilized here is:

$$\varphi = \varphi - \eta \cdot \nabla_{\varphi} K(\varphi; a^{(m)}; b^{(m)}) \tag{3}$$

where φ is gradient descent optimization value, $a^{(m)}$ is the training example value and the $b^{(m)}$ is the label value.

Testing: The testing of model is carried-out through the following steps:

Input: Images from a surveillance camera

Output: Detecting social distancing amongst people

1. Testing the above model for test dataset;
2. Using Intersection Over Union (IOU) method, detect the social distancing between people using the overlapping (intersection) area;

3. If the value of IOU is non-zero, then it can be said that people are not at proper social distancing from each other;

Loss Function and other formulae used:

a) *ReLU:* The ReLU (Rectified Linear Units) is better and performs higher than Tanh or Sigmoid in Loss Function estimation. Here, the ReLU is estimated since it rectifies and evades vanishing-gradient issues. Similarly, it is less-expensive when computations are compared against Tanh and Sigmoid and thus researchers adopt ReLU loss function a lot; it also provides simpler, easier and faster mathematical operations.

However, it has certain drawbacks, like: it shouldn’t be utilized out of hidden-neural layers of network model, fragile gradient loss during training with dead neurons, activation blow-ups and ReLU might stop replying to error variations when gradient range is 0 (equation 4).

The formula for ReLU activation could be measured through:

$$\max(0, z) \tag{4}$$

b) *R-CNN: (Regional – Convolutional Neural Network):* The research makes use of bounding-box regression to improvise the performance of localization. The ground-truth prediction of bounding-box localization is estimated with its size and location, through the formula:

$$\widehat{GT}_a = P_v d_a(P) + P_a \tag{5}$$

$$\widehat{GT}_b = P_f d_b(P) + P_b \tag{6}$$

$$\widehat{GT}_v = P_v \exp(d_v(P)) \tag{7}$$

$$\widehat{GT}_f = P_f \exp(d_f(P)) \tag{8}$$

where the initial two-equations (5a and b) specifies the translation of scale-invariant of P’s centre (a and b) and the latter two-equations (5c and d) specifies transformation of log space with the height (f) and width (v).

c) *IOU (intersection-over-Union):* The IOU is measured through dividing the area of bounding boxes, where: overlapping (intersection of bounding-boxes) is divided by union (combined area of bounding-boxes). The IOU is estimated through [34]:

$$IoU = \frac{OverlappingArea}{UnionArea} \tag{9}$$

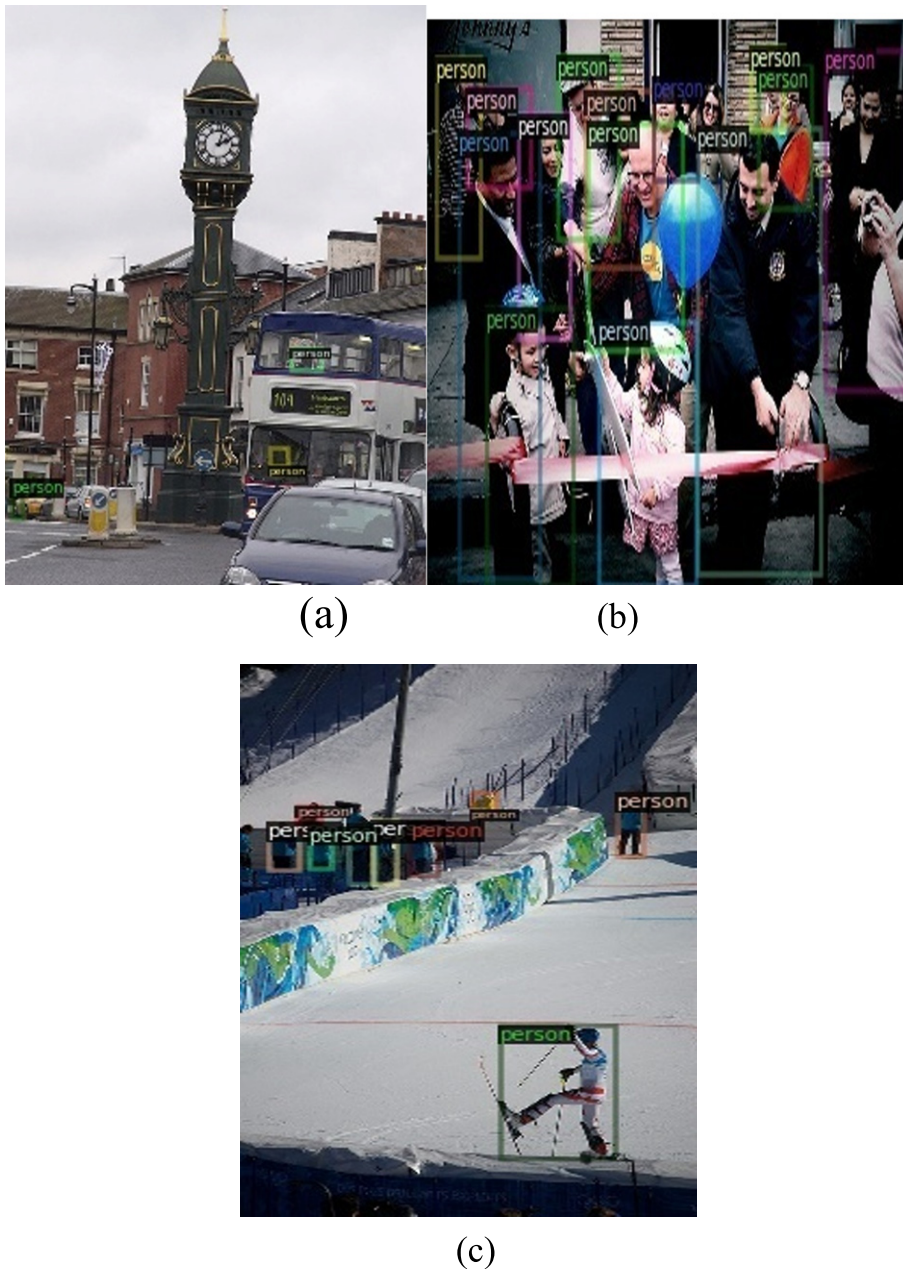


Fig. 6. Input dataset selection.

5. Results and performance evaluation

Random Sample Dataset: The datasets for training the model are selected randomly (refer to Fig. 6a, b and c) to predict its reliability and accuracy towards comparing the results obtained against estimated outcomes.

Ground truth versus predicted by model: The following outcomes (predictions) are the results

obtained from the trained-tested-model, where the ground truth is compared and weighed against its' predictions, based on the algorithm. The results from trained detectron2 with faster R-CNN model are as follows:

a) **Classification of Individual Person:** The figure (refer to Fig. 7a) represents the ground truth of an identified person in the image versus the Fig. 7b representing the prediction made by the model by

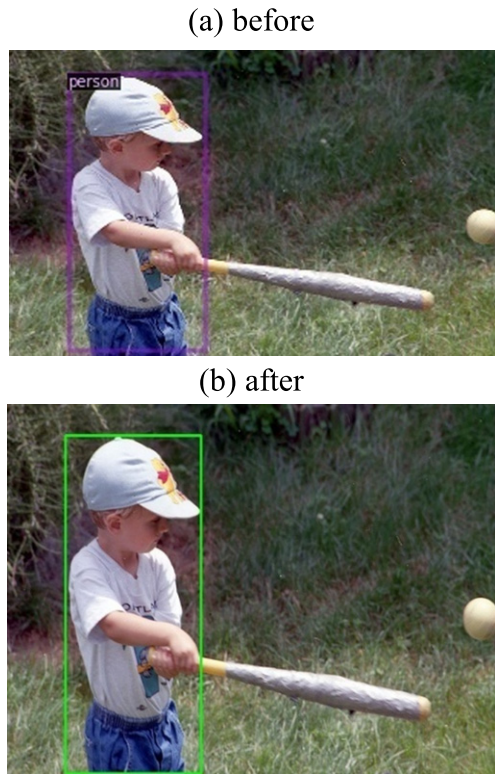


Fig. 7. Individual as person-class identification.



Fig. 8. Group as person-class identification.

identifying the person-class only. It could be inferred that obtained accuracy and precision are similar to the estimated outcomes.

Interpretation: However, the study is focused on social-distancing with green-bound-box as correct distance and red-bound-box as incorrect distance, between two-or-more people. Since there is no more than single person in the image, the algorithm identified the individual with green-bound-box, stating that, norms of social-distancing are satisfied by the 'person'.

b) Classification of Group of People versus Objects: The following figure (refer to Fig. 8a) represents the ground truth of identified person in the image versus Fig. 8b representing the prediction made by the model with identification of person-class only.

Interpretation: The image is of mass-people, where the ground truth identified 8 individual people in Fig. 8a with class-person. The predicted model provided outcomes (8b) with same head-counts of 8 people as 'person-class'. The model also identified

that there is no social-distancing between the people and thus the resulted image 8b is obtained with red-bounding-boxes.

c) Classification of People and Animals: The figure (refer to Fig. 9a) represents the ground truth of identified person in the picture versus the Fig. 9b representing the prediction made by the model with identifying person-class only.

Interpretation: The class 'person' is identified from other class-categories and evaluated for social distancing. According to the developed algorithm, only people-class is identified for social-distancing evaluation in 9a; thus the predicted outcome is found to be negative, where there is no social-distancing between 'identified persons' in the result 9b.

d) Classification of people from other categories:

Interpretation: From the figures (refer to Fig. 10a, b, f and g), it is interpreted that, there are people who maintained social-distancing and thus the green-bound-box; contrarily the remaining Fig. 10c, d and e has red-bounding-box indicating that there are no social-distancing between the identified person-class.

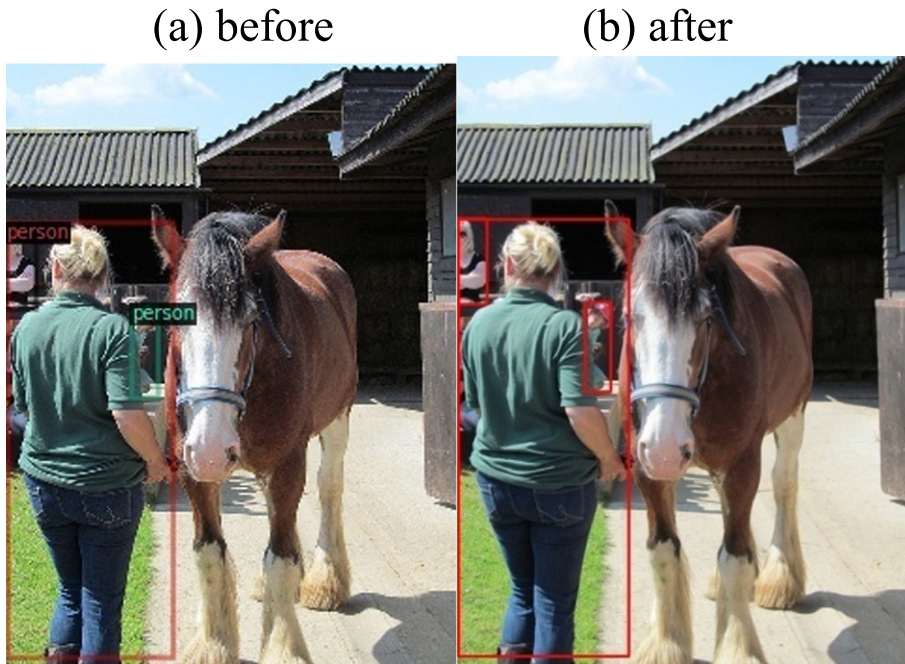


Fig. 9. Identified person-class social distancing.



Fig. 10. Prediction and social-distancing by developed model.



Fig. 11. Identification of individuals (with person class).



Fig. 12. Identification of people in crowded areas.

6. Case study with respect to Textile shops

Case study based on the developed model: The following example illustrates how the developed model can be used to check and measure the social distancing between people on a real time basis. The model is trained and tested with images of the people in textile shop in order to illustrate whether social distance has been followed by people in a crowded textile shop. Few images as samples of the people in the textile shop were tested and trained for social distancing. The results are:

Interpretation: The developed model can be explained as below. The green-bound-box represent

the proper distance among individuals (people) and the red-bound-box denotes there exists no proper distance between two-or-more people.

Since, there is just one person in the image (refer to Fig. 11-before), the model identified the person with green-bound-box (refer to Fig. 11-after), stating that the norms of social-distances between people is justified by the identified 'person'.

Interpretation: The model identified the mass/people (refer to Fig. 12-before), where 8 individuals under person-class has been identified. The model predicted 8 head-counts of the people as 'person-class' and bounded them with 'person' label; according to the social distancing algorithm it



Fig. 13. Measuring the social-distancing through the developed model.

resulted with 8 red-bounding boxes (refer to Fig. 12-after), where it indicates there is no social-distancing between them.

Interpretation: Figures 13(a–f) denotes that, there are only few people with proper social-distancing in the textile shops, thus the green-bounding-box. However, the red-bounding-box in the Fig. 13(a–f) represents that improper or no social-distancing among the individuals identified as person-class.

f) Loss function: Based on the estimated versus obtained outcomes, the graphical representation of loss function (the time, data time, total loss and eta-seconds) of the developed model had been evaluated in python and represented below in the graphical representation of the Fig. 14a–d:

Interpretation: It is inferred through the figures (refer to Fig. 14a–d) that no huge variations have been identified between the results obtained from the model. Similarly, the estimated outcome and obtained result of total loss is less at 0.1, thus proving that developed model is effective in predictions.

The figures (refer to Fig. 15a and b) denote the anchor values (negative-positive), whereas the Fig. 15c and d represents the Background samples and Foreground samples of the input images, for training the model.

Interpretation: From the above figures (refer to Fig. 15a–d) it's inferred that there are no huge variations in the rpn (region-proposal-networks)

numbers of negative and positive anchors; similarly, the roi_head-numbers (region-of-interest) for foreground and background samples in pooling have no huge-variations. The predicted outcomes and ground results are similar. Thus the developed model is a good-fit.

The Fig. 16a–c represents the loss_rpn with loc (loss-rpn-localization), class and regression of bounding box outcomes. The results evidently insist that the regions are overlapped. Thus NMS (non-maximum suppressions) is used towards minimizing the proposal numbers. The loss is minimized at 0.1, as total loss of the developed model.

The outcomes from Fig. 17a–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. Obtained results were higher than estimated outcomes with 75% (threshold value) and above stating the performance was effective.

g) Scheduling LR: The learning rate (as shown in Fig. 18) is set-to 4k intervals where the developed model attained successful learning rate at 1000k. It remained the same till 23k iterations stating that, there are no sudden dropping-down in the learning rate, but rather, the LR steadily decreased and remained constant from 1000k-70k iterations.

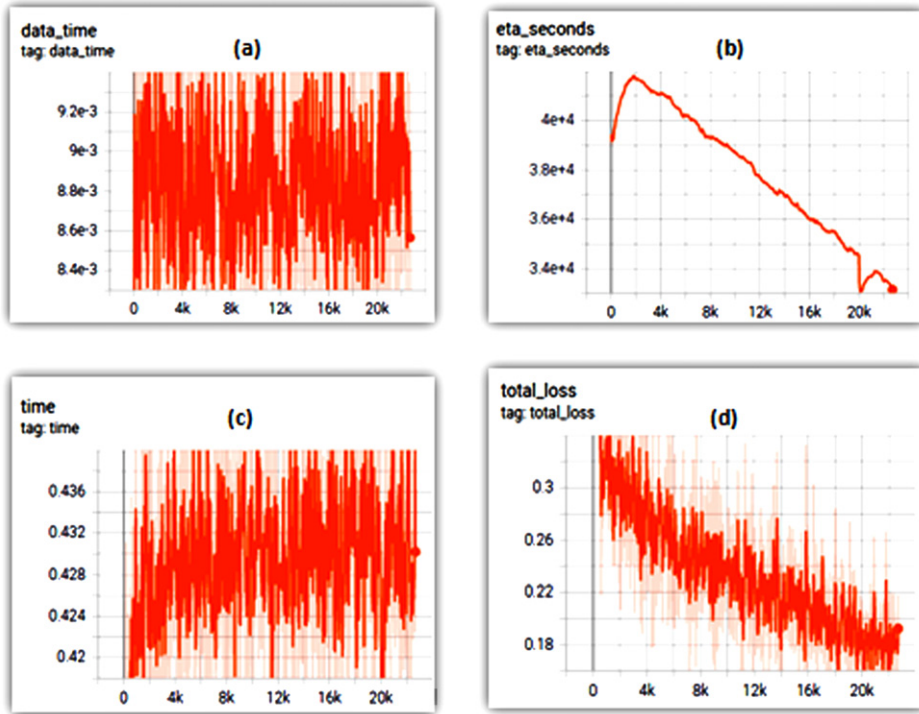


Fig. 14. (a) Time, (b) Data time, (c) ETA seconds and (d) Total loss.

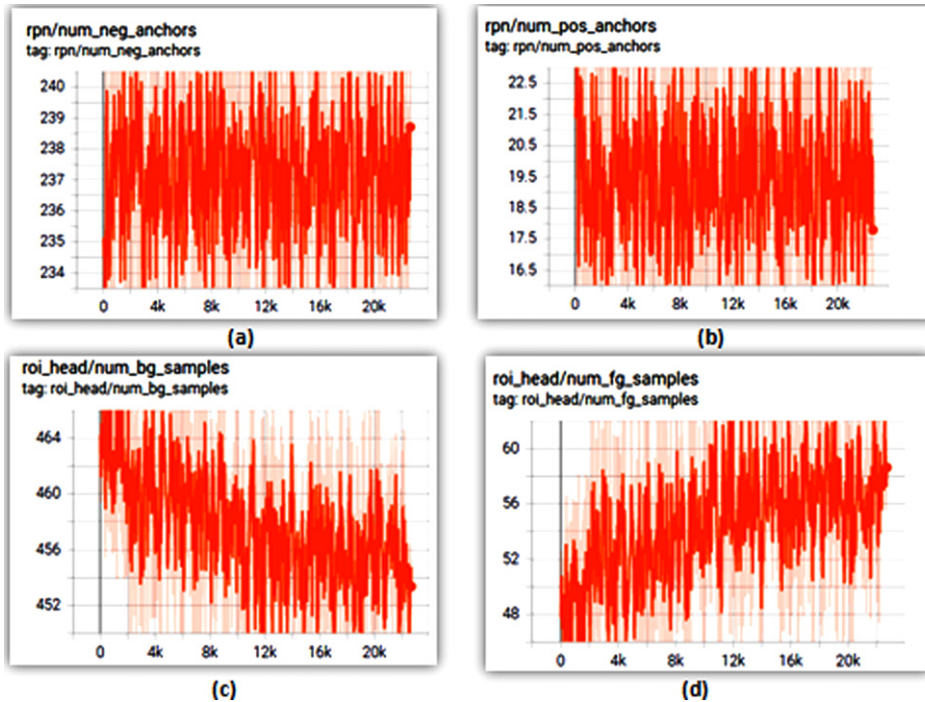


Fig. 15. RPN number (a) negative anchors and (b) positive anchors; ROI-head number (c) BG-samples and (d) FG samples.

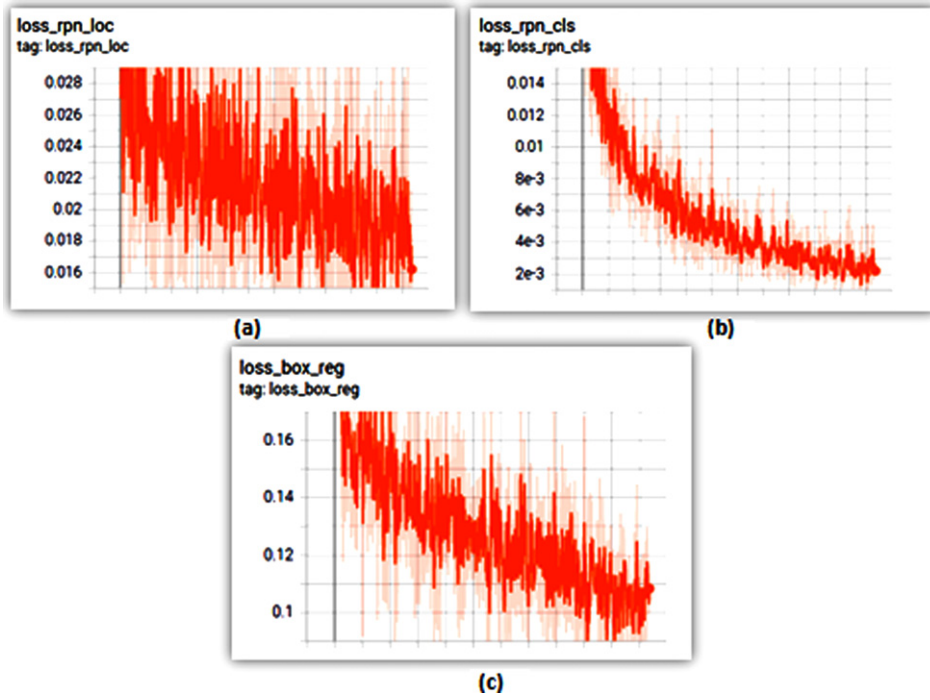


Fig. 16. Loss RPN (a) LOC, (b) Class and (c) Box regression.

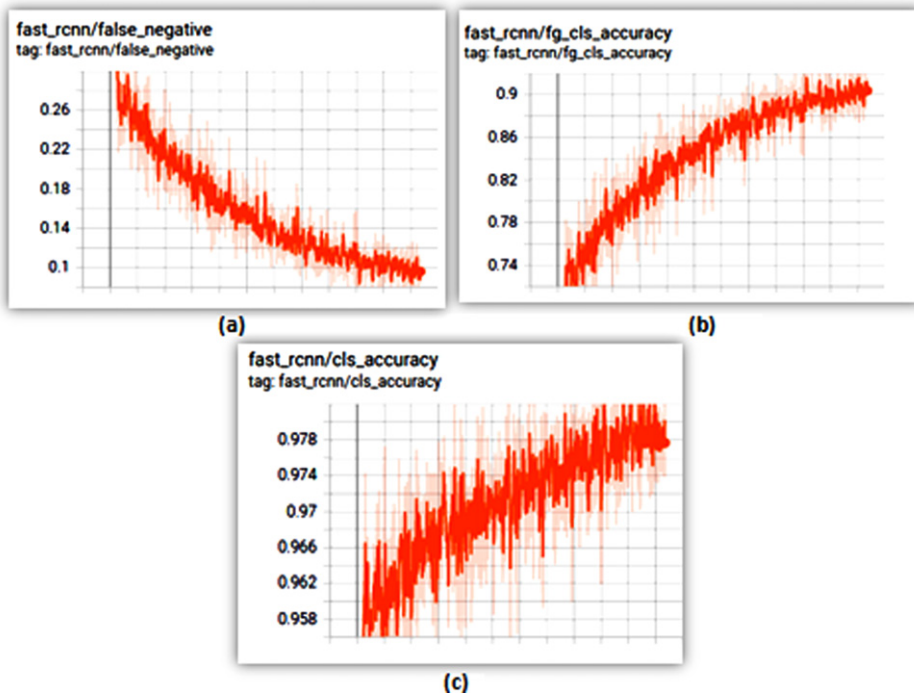


Fig. 17. Faster RCNN (a) False negative, (b) FG class accuracy and (c) Class accuracy.

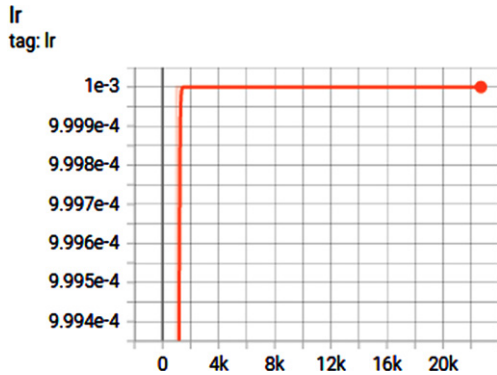


Fig. 18. Learning rate.

Table 3
Outcome of the trained model

mAP	Precision (IOU @ 0.5)	Recall	Total loss
84.466	97.9	87.0	0.1

Scores post testing and training the model: The scores for the developed model after testing and training the processed dataset have been obtained (Table 3) 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.

h) 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 (refer Table 4) lack precision in human detection towards social-distancing threshold violation measures (refer Fig. 19).

Inference: The Fig. 19 exemplifies that among the existing models the developed model with detectron2 with faster R-CNN architecture acquired higher

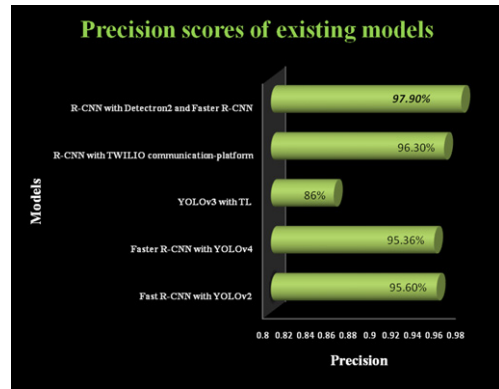


Fig. 19. Performance evaluation and comparison with existing approaches.

precision rate of 97.90% (98%) 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%;

Thus the researcher examined and evaluated the datasets with developed detectron2 model where it's evidently concluded that, the developed model is a success and good-fit for object detection based 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.

Table 4
Comparative analysis of models and architecture

S. No	Author	Year	Architecture	Performance
1	Saponara et al.,	2021	Fast R-CNN with YOLOv2	95.6%
2	Yang et al.,	2021	Faster R-CNN with YOLOv4	95.36%
3	Ahmed et al.,	2021	YOLOv3 with TL	86%
4	Pandian	2020	R-CNN with TWILIO communication-platform	96.3%
5	Proposed approach	2021	Detectron2 with IOU	97.9%

Source: Author.

7. Conclusion and future recommendation

Conclusion: The investigation particularly aimed at analyzing and developing a better model with object detection-based machine language adopted architecture where it could attain a higher precision rate than average metric scores attained by the existing models. The study adopted an IOU metric evaluation that uses mAP (mean-average of precision) which examines and evaluates the developed object detection model. The model developed is of detectron2 and faster R-CNN, where outcomes are estimated through ground-truth detection and prediction with an algorithm 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 an 87% recall score where IOU was at 0.5.

Limitations: The research has exclusively focused on identifying whether social distance has been maintained between individuals in crowded areas. This study does not involve other applications like detection of face mask, hand hygiene and so on.

Future work: In the future, the developed model could be converted into an android application and integrated with IoT-based hardware such that images of people in public places are captured on a dynamic basis and proper notification is given to the relevant authorities when an individual fails to adapt social distance at a crowded place.

References

- [1] A. Brodeur, A. Islam, D. Gray and S.J. Bhuiyan, A Literature Review of the Economics of COVID-19, *IZA- Institute of Labor Economics (Discussion paper series)* **13411** (2020), 1–61.
- [2] A. Mingozzi, A. Conti, F. Aleotti, M. Poggi and S. Mattocchia, Monitoring Social Distancing With Single Image Depth Estimation, in *IEEE Transactions on Emerging Topics in Computational Intelligence*, (Early Access: In Print), 1–12. doi.10.1109/TETCI.2022.3171769
- [3] A. Rahim, A. Maqbool and T. Rana, Monitoring social distancing under various low light conditions with deep learning and a single motionless time of flight camera, *PLoS ONE* **16**(2) (2021), e0247440, 1–19.
- [4] A.B. Payedimarrì, D. Concina, L. Portinale, M. Canonico, D. Seys, K. Vanhaecht and M. Panella, 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) (2021), 1–11.
- [5] C. Werner, Human dataset for object detection. <https://www.kaggle.com/constantinwerner/human-detection-dataset>
- [6] D. Yang, E. Yurtsever, V. Renganathan, K.A. Redmill and U. Ozguner, A Vision-Based Social Distancing and Critical Density Detection System for COVID-19, *Sensors* **21**(4608) (2021), 1–15.
- [7] D.D. Triphená and V. Karunakaran, Deep Learning based Object Detection using Mask RCNN, *Communication and Electronics Systems* **6** (2021), 1684–1690.
- [8] H. Ziran and N. Dahnoun, A Contactless Solution for Monitoring Social Distancing: A Stereo Vision Enabled Real-Time Human Distance Measuring System, *Embedded Computing – IEEE* **10** (2021), 1–6.
- [9] I. Ahmed, M. Ahmad and G. Jeon, Social distance monitoring framework using deep learning architecture to control infection transmission of COVID-19 pandemic, *Sustainable Cities and Society* **69**(102777) (2021b), 1–11.
- [10] I. Ahmed, M. Ahmad, J.J.P.C. Rodrigues, G. Jeon and S. Din, A deep learning-based social distance monitoring framework for COVID-19, *Sustainable Cities and Society* **65**(102571) (2021a), 1–12.
- [11] K. Khan, W. Albattah, R.U. Khan, A.M. Qamar and D. Nayab, Advances and Trends in Real Time Visual Crowd Analysis, *Sensors* **20**(5073) (2020), 1–28.
- [12] M. Cristani, A.D. Bue, V. Murino, F. Setti and A. Vinciarelli, The Visual Social Distancing Problem, *IEEE Access* **8** (2020), 126876–126886.
- [13] M.N. Al-Andoli, S.C. Tan and W.P. Cheah, Distributed parallel deep learning with a hybrid backpropagation-particle swarm optimization for community detection in large complex networks, *Information Sciences* **600** (2022), 94–117.
- [14] M. Rezaei and M. Azarmi, Deep Social: Social Distancing Monitoring and Infection Risk Assessment in COVID-19 Pandemic, *Appl Sci* **10**(7514) (2020), 1–29.
- [15] M. Rezzouki, S. Ouajih and G. Ferré, Monitoring Social Distancing in Queues using Infrared Array Sensor, in *IEEE Sensors Journal*, (Early Access: In Print), 1. doi.10.1109/JSEN.2021.3104669.
- [16] N. Ahmed, R.-I. Alam, S.N. Shefat and Md.T. Ahad, Keep me in Distance: An Internet of Things based Social Distance Monitoring System in Covid19, *Int J Advanced Networking and Applications* **13**(5), 5128–5133.
- [17] N. Iqbal and M. Islam, Machine learning for dengue outbreak prediction: A performance evaluation of different prominent classifiers, *Informatica* **43** (2019), 363–371.
- [18] N.S. Punn, S.K. Sonbhadra, S. Agarwal and G. Rai, Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLOv3 and Deepsort techniques, <https://arxiv.org/pdf/2005.01385.pdf>
- [19] P. Pandian, Social Distance Monitoring and Face Mask Detection Using Deep Neural Network, https://www.researchgate.net/publication/347439579_Social_Distance_Monitoring_and_Face_Mask_Detection_Using_Deep_Neural_Network
- [20] P.R. Raje, S.S. Ubhe, P. Tambe, N.N. Sakhare and Nurjahan, Social Distancing Monitoring System using Internet of Things, *Electronics and Renewable Systems* **2** (2022), 511–517.
- [21] P. Shukla, R. Kundu, A. Arivarasi, G. Alagiri and J. Shiney, A Social Distance Monitoring System to ensure Social Distancing in Public Areas, *Computational Intelligence and Knowledge Economy* **4** (2021), 96–101.
- [22] S. Arya, L. Patil, A. Wadegaonkar, N. Shinde and P. Gorsia, Study of Various Measure to Monitor Social Distancing

- using Computer Vision: A Review, *International Journal of Engineering Research & Technology (IJERT)* **10**(5) (2021), 329–326.
- [23] S. Meivel, N. Sindhwani, R. Anand, D. Pandey, A.A. Alnuaim, A.S. Altheneyan, M.Y. Jabarulla and M.E. Lelisho, Mask detection and social distance identification using internet of things and faster RCNN algorithm, *Hindawi-Computational Intelligence and Neuroscience* **2022** (2103975), 1–13.
- [24] S. Noor, M. Waqas, M.I. Saleem and H.N. Minhas, Automatic Object Tracking and Segmentation Using Unsupervised SiamMask, *IEEE Access* **9** (2021), 106550–106559.
- [25] S.P. Das, D. Majumdar and R.K. Gayen, Monitoring Social Distancing Through Person Detection and Tracking Using Computer Vision, *Electronics, Materials Engineering & Nano-Technology (MENTech)* **5** (2021), 1–5.
- [26] S. Saponara, A. Elhanashi and A. Gagliardi, Implementing a real-time, AI-based, people detection and social distancing measuring system for Covid-19, *Journal of Real-Time Image Processing - Springer*, (2021), 1–11.
- [27] S. Singh, U. Ahuja, M. Kumar, K. Kumar and M. Sachdeva, Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment, *Multimed Tools Appl* **80** (2021), 19753–19768.
- [28] S. Swetha, J. Vijayalakshmi and S. Gomathi, Social Distancing and Face Mask Monitoring System Using Deep Learning Based on COVID-19 Directive Measures, *Computing and Communications Technologies (CCT)* **4** (2021), 520–526.
- [29] S. Tuli, S. Tuli, R. Tuli and S.S. Gill, Predicting The Growth And Trend Of Covid-19 Pandemic Using Machine Learning And Cloud Computing, *MedRxiv, Pre-Print* (2020), 1–13.
- [30] V. Rupapara, M. Narra, N.K. Gunda, S. Gandhi and K.R. Thipparthi, Maintaining Social Distancing in Pandemic Using Smartphones With Acoustic Waves, *IEEE Transactions on Computational Social Systems* **9**(2) (2022), 605–611.
- [31] V. Vinitha and V. Velantina, Social Distancing Detection System With Artificial Intelligence Using Computer Vision And Deep Learning, *International Research Journal of Engineering and Technology (IRJET)* **7**(8) (2020), 4049–4053.
- [32] Y. Chen, C. Han, Y. Li, Z. Huang, Y. Jiang, N. Wang and Z. Zhang, SimpleDet: A Simple and Versatile Distributed Framework for Object Detection and Instance Recognition, *Journal of Machine Learning Research* **20** (2019), 1–8.
- [33] Z. Shao, G. Cheng, J. Ma, Z. Wang, J. Wang and D. Li, Real-Time and Accurate UAV Pedestrian Detection for Social Distancing Monitoring in COVID-19 Pandemic, *IEEE Transactions on Multimedia* **24** (2022), 2069–2083.
- [34] Shyam Prasad Devulapalli, Srinivasa Rao Chanamallu, Satya Prasad Kodati: A hybrid ICA Kalman predictor algorithm for ocular artifacts removal. *Int J Speech Technol* **23**(4) (2020), 727–735.