Social Distancing Monitoring and Infection Risk -- Assessment in COVID-19 Pandemic

ABSTRACT

Social distancing is a recommended solution by the World Health Organization (WHO) to minimize the spread of COVID-19 in public places. The majority of governments and national health authorities have set the 2-meter physical distancing as a mandatory safety measure in shopping centres, schools and other covered areas. In this research, we develop a generic Deep Neural Network-Based model for automated people detection, tracking, and inter-people distances estimation in the crowd, using common CCTV security cameras. The proposed model includes a YOLOv4-based framework and inverse perspective mapping for accurate people detection and social distancing monitoring in challenging conditions, including people occlusion, partial visibility, and lighting variations. We also provide an online risk assessment scheme by statistical analysis of the Spatio-temporal data from the moving trajectories and the rate of social distancing violations. We identify high-risk zones with the highest possibility of virus spread and infections. This may help authorities to redesign the layout of a public place or to take precaution actions to mitigate high-risk zones. The efficiency of the proposed methodology is evaluated on the Oxford Town Centre dataset, with superior performance in terms of accuracy and speed compared to three state-of-the-art methods.

Keywords – Social Distancing; COVID-19; People Detection and Tracking; Inverse Perspective Mapping; Deep Learning; Convolutional Neural Networks.

Introduction

The novel generation of the coronavirus disease (COVID-19) were reported in late December 2019 in Wuhan, China. After only a few months, the virus was hit by the global outbreak in 2020. On May 2020 The World Health Organization (WHO) announced the situation as the pandemic. The statistics by WHO on 26 August 2020 confirms 23.8 million infected people in 200 countries. The mortality rate of the infectious virus also shows a scary number of 815,000 people. With the growing trend of patients, there is still no effective cure or available treatment for the virus. While scientists, healthcare organisations, and researchers are continuously working to produce appropriate medications or vaccines for the deadly virus, no definite success has been reported at the time of this research, and there is no certain treatments or rec-ommendation to prevent or cure this new disease.

Social distancing, refers to pre- caution actions to prevent the proliferation of the disease, by minimizing the proximity of human physical contacts in covered or crowded public places (e.g. schools, workplaces, gyms, lecture theatres, etc.) to stop the widespread accumulation of the infection risk .

According to the defined requirements by the WHO, the minimum distance between individuals must be at least 6 feet (1.8 meters) in order to observe an adequate social distancing among the people.

Recent researches have confirmed that people with mild or no symptoms may also be carriers of the novel Coronavirus infections[4]. Therefore, it is important all individuals maintain controlled behaviours and observe social distancing.

Figure a – Social Distancing Monitoring

Figure b – Accumulated Infection Risk

Many research works such as have proved social- distancing as an effective non-pharmacological approach and an important inhibitor for limiting the transmission of contagious diseases such as H1N1, SARS, and COVID-19.

It demonstrates the effect of following appropriate social distancing guideline to reduce the rate of infection transmission among individuals. A wider Gaussian curve with a shorter spike within the range of the health system service capacity makes it easier for patients to fight the virus by receiving continuous and timely support from the health care organizations. Any unexpected sharp spike and rapid infection rate (such as the red curve), will lead to the service failure, and consequently, exponential growth in the number of fatalities.

For several months, the World Health Organization believed that COVID-19 is only transmittable via droplets emitted when people sneeze or cough and the virus does not linger in the air. However, on 8 July 2020, the WHO announced:

" There has emerging evidence that COVID-19 is an air- borne disease that can be spread by tiny particles suspended in the air after people talk or breathe, especially in crowded, closed environments or poorly ventilated settings. "

Therefore, social distancing now claims to be even more important than thought before, and as one of the best ways to stop the spread of the disease in addition to wearing facial masks.During the COVID-19 pandemic, governments have tried to implement a variety of social distancing practices, such as restricting travels, controlling borders, closing pubs and bars, and alerting the society to maintain a distance of 1.6 to 2 meters from each other.

However, monitoring the amount of widespread and efficiency of the constraints is not an easy task. People require to go out for essential needs such as food, health care and other necessary tasks and jobs.

In such situations, Artificial Intelligence can play an important role in facilitating social distancing monitoring. Computer vision, as a subset of Artificial Intelligence, has been very successful in solving various complex health care problems and has shown its potential in Chest CT-Scan or X-Ray based COVID-19 recognition. Besides, deep neural net- works enable us to extract complex features from the data so that we can provide a more accurate understanding of the images by analyzing and classifying these features. Examples include diagnosis, clinical management and treatment, as well as the prevention and control of COVID-19.

The main contribution of this research can be highlighted as follows:

(i)This study aims to support the reduction of the corona virus spread and its economic costs by providing an AI-based solution to automatically monitor and detect violations of social distancing among individuals.

(ii) We develop one of the most (if not the most) accurate deep neural network (DNN) models for people detection, tracking, and distance estimation called ***Deep SOCIAL***.

(iii)We perform a live and dynamic risk assessment, by statistical analysis of spatio-temporal data from the people movements at the scene.

(iv)The developed model is a generic human detection and tracker, not limited to social-distancing monitoring, and can be applied for various real-world applications such as pedestrian detection in autonomous vehicles, human action recognition, anomaly detection, and security systems.

More details and further information will be provided in the following sections. In Section we discuss the related work, existing challenges, and research gaps in the field. The proposed mythology including the model architecture and our object detection techniques, tracking, and red-zone prediction algorithm will be proposed . In experimental results and performance of the system is investigated against the state-of-the-art, followed by discussions and concluding remarks.

Related Works

Many researchers have worked in the medical and pharmaceutical fields aiming at treatment of COVID-19 infectious disease; however, no definite solution has yet been found. On the other hand, controlling the spread of such an unknown respiratory infectious disease is another issue.

Variety of studies with different implementation strategies have proven that controlling the prevalence is a contributing factor, and social distancing is an effective way to reduce the transmission and prevent the spread of the virus in society. Effectiveness of social distancing practices can be evalu- ated based on several standard approaches. One of the main criteria is based on the reproduction ratio, R_o, which indicates the average number of people who may be infected from an infectious person during the entire period of the infection. Any $R_o > 1$ indicates an increasing rate of infection within the society and $R_o < 1$ indicates that every case will infect less than 1 person, hence, the disease rate is considered to be declining in the target population.

Since the R_o value indicates the disease outspread, it is one of the most important indicators for selecting social distancing criteria. In the current COVID-19 pandemic, the World Health Organisation estimated the R_o rate would be in the range of 2-2.5, which is significantly higher than other similar diseases such as seasonal flu with $R_o = 1.4$. In, a clear conclusion is drawn about the importance of applying social distancing for the cases with a high amount of R_o.

In another research based on the game theory on the classic SIR model, an assessment of the benefits and economic costs of social distancing has been examined[21]. The results also show that in the case of Ro < 1, social distancing would cause unnecessary costs, while R_o 2 im plies that social distancing measures have the highest economic benefits. In a similar research, Kylie et al. investigated the relationship between the stringency of social distancing and the region's economic status. This study suggests although preventing the widespread outbreak of the virus is necessary, a moderate level of social activities could be allowed.

Use location-specific contact patterns to investigate the effect of social distancing measures on the prevalence of COVID-19 pandemic in order to remove the persistent path of disease outbreak using susceptible-exposed-infected-removed models (SEIR).

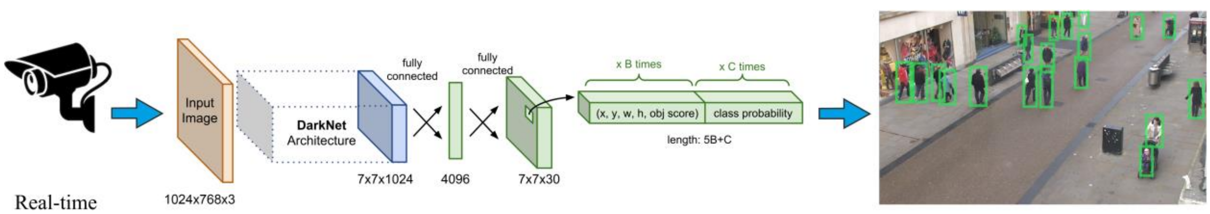
Since the onset of coronavirus pandemic, many countries have used technology-based solutions, to inhibit the spread of the disease. For example, some of developed countries, such as South Korea and India, use GPS data to monitor the movements of infected or suspected individuals to find any possible exposure among the healthy people. The India government uses the *Aarogya Setu* program to find the presence of COVID-19 patients in the adjacent region, with the help of GPS and Bluetooth. This may also help other people to maintain a safe distance from the infected person. Some law enforcement agencies use drones and surveillance cameras to detect large-scale rallies and have carried out regulatory mea-sures to disperse the population.

The utilization of Artificial Intelligence, Computer Vision, and Machine Learning, enables us to discover the correlation of high-level features. For example, it enables us to understand and predict pedestrian behaviours in traffic scenes, sports actions and activities, medical imaging, anomaly detection, etc. by analyzing Spatio-temporal visual information and statistical data analysis of the images sequences.

In health-related fields, it would be also feasible to predict the sickness trend of specific areas, to predict the density of people in public places, or determine the distance of individuals from the popular swarms, using a combination of visual and Geo-location cellular information. However, such research works suffer from challenges such as skilled labour or the cost of designing and implementing the infrastructures. On the other hand, recent advances in AI, Computer Vision, Deep Learning, and Pattern Recognition, enables the computers to understand and interpret the visual data from digital images or videos.

Such capabilities can play an important role in empowering, encouraging, and performing social distancing surveillance and measurements as well. For example, computer vision could turn CCTV cam-eras in the current infrastructure capacity into “smart” cameras that not only monitor people but can also determine whether people follow the social distancing guidelines or not. Such systems require a very precise human detection algorithms.

People detection in image sequences is one of the most important sub-branches in the field of object detection and computer vision. Although many research works have been done in the field of human detection and human action recognition, majority of them are either limited to indoor applications or suffer from accuracy issues under outdoor challenging lighting conditions. A range of other research rely on manual tuning methodologies to identify people activities, however, limited functionality has always been an issue.



Real-time Survialance   A YOLO based Deep Neural Network   People Detection

Similar to the other research, no statistical analysis is per- formed on the results of the distance measurement . The authors have made a comparison between two common types of DNN models (YOLO and Faster R-CNN); however, the system has been only tested on a basic dataset.

Since the topic is very recent, there has not been much research regarding the accuracy of detections, no experimental on challenging datasets has been performed, no standard comparison has been conducted on common datasets, and no analytical studies or post-processing have been considered after the detection-phase.

Methodology

We propose a 3-stage model including people detection, track- ing, inter-distance estimation as a total solution for social distancing monitoring and zone-based infection risk analy- sis. The system can be integrated and applied on all type of CCTV surveillance cameras with any resolution from VGA to Full-HD, with real-time performance.

People Detection

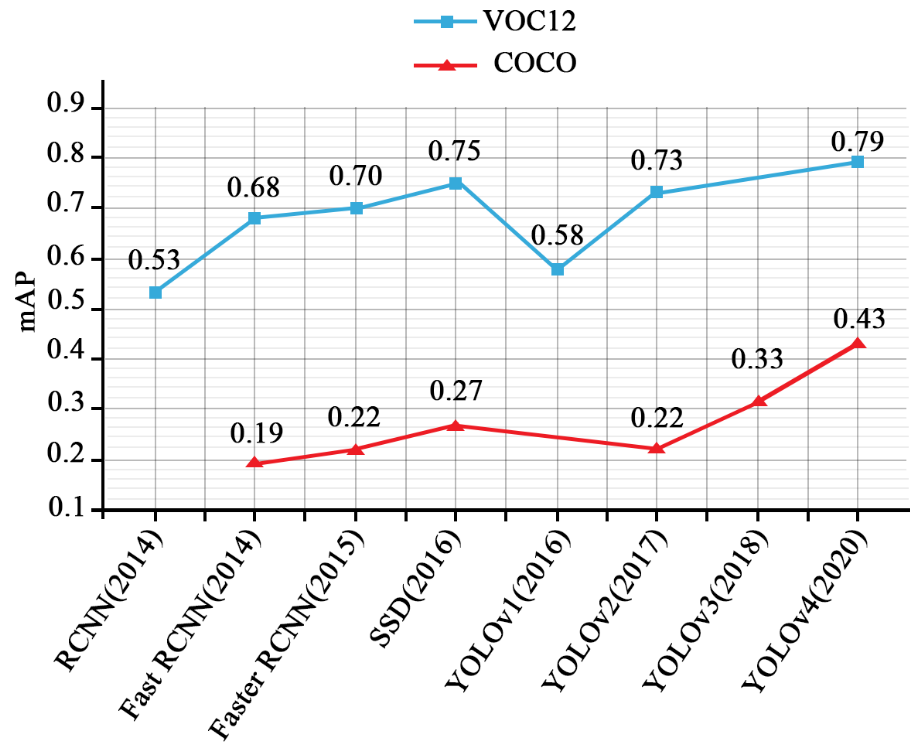
The objective is to develop a model to detect humans (people) with various types of challenges such as variations in clothes, postures, at far and close distances, with/without occlusion, and under different lighting conditions.

To gain this we inspire from the strength of cutting-edge re- search; however, we develop our own unique human classifier and train our model based on a set of comprehensive and multifaceted datasets. Before diving into further technical details, we overview the most advanced object detection techniques.

Modern DNN-based detectors consist of two sections: A *backbone* for extracting features and a *head* for predicting classes and location of objects. The feature extractor tends to encode model inputs by representing specific features that help it to learn and discover the relevant patterns related to the query object(s). Examples of feature extraction architectures can be seen in VGG16, ResNet-50, CSPResNeXt-50, CSPDarknet53, and EfficientNet-B0/B7. The head of a DNN is responsible for the classifying the objects (e.g. people, bicycles, chairs, etc.) as well as calculating the size of the objects and the coordinates of the correspondent bounding boxes.

There are usually two types of head sections: one-stage and two-stage. The two-stage detectors use the region proposal before applying the classification. First, the detector extracts a set of object proposals (candidate bounding boxes) by a selective search. Then it resizes them to a fixed size before feeding them to the CNN model. This is similar to R-CNN based detectors. There are usually two types of head sections: one-stage and two-stage. The two-stage detectors use the region proposal before applying the classification. First, the detector extracts a set of object proposals (candidate bounding boxes) by a selective search. Then it resizes them to a fixed size before feeding them to the CNN mo del. This is similar to R-CNN based detectors. In spite of the accuracy of two-stage detectors, such methods are not suitable for the systems with restricted computational resources.

On the other hand, the one-stage detectors perform a single detection process such as the work done by Liu et al. (known as SSD), or other works done by Redmon et al. and Bochkovski et al, known as "You Only Looks Ones" or YOLO detectors. Such detectors use regression analysis to calculate the dimensions of bounding boxes and interpret their class probabilities. It maps the image pixels to the enclosed grids and checks the probability of the existence of an object in each cell of the grids. This approach offers excellent improvements in terms of speed and efficiency.



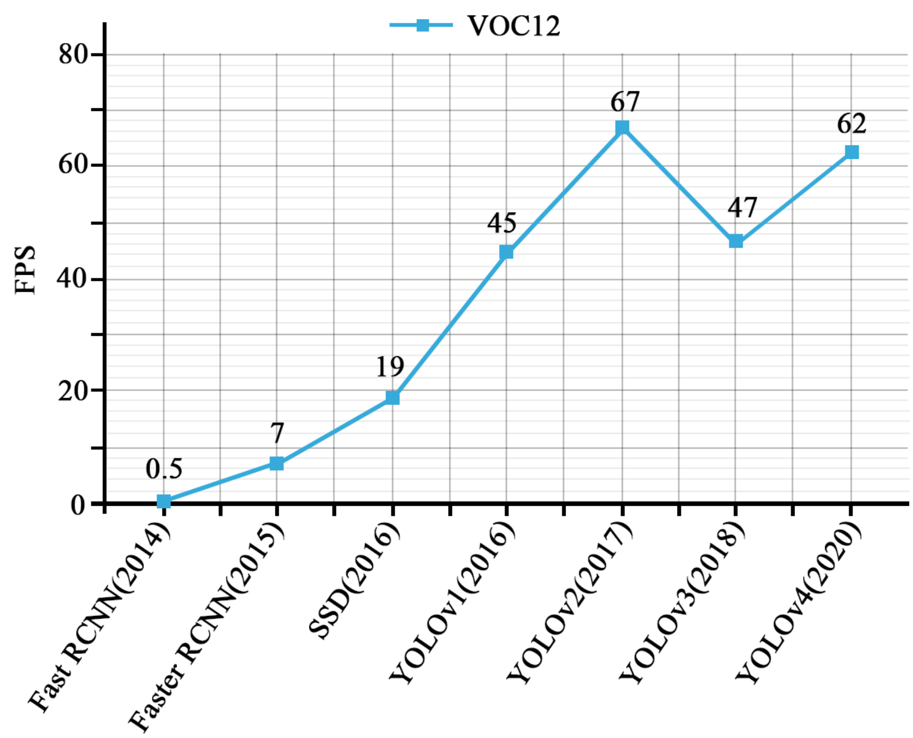


Table 1. Comparisons of three backbone models in terms of number of parameters and speed (*fps*) using an RTX 2070 GPU.

As illustrated in Figure 4, YOLOv4 offers the best trade-off for the speed and the accuracy; however, since YOLOv4 is an aggregation of various techniques, we undertook an in-depth study of each sub-techniques to achieve the best results for our people detection model and to outperform the state-of-the-art.

We considered two approaches to improve the accuracy of the backbone: A basic way to improve the accuracy of CNN-based detectors is to expand the receptive field and enhance the complexity of the model using more layers; however, using this technique makes it harder to train the model. We sug- gest using a skip-connections technique for ease of training, instead.

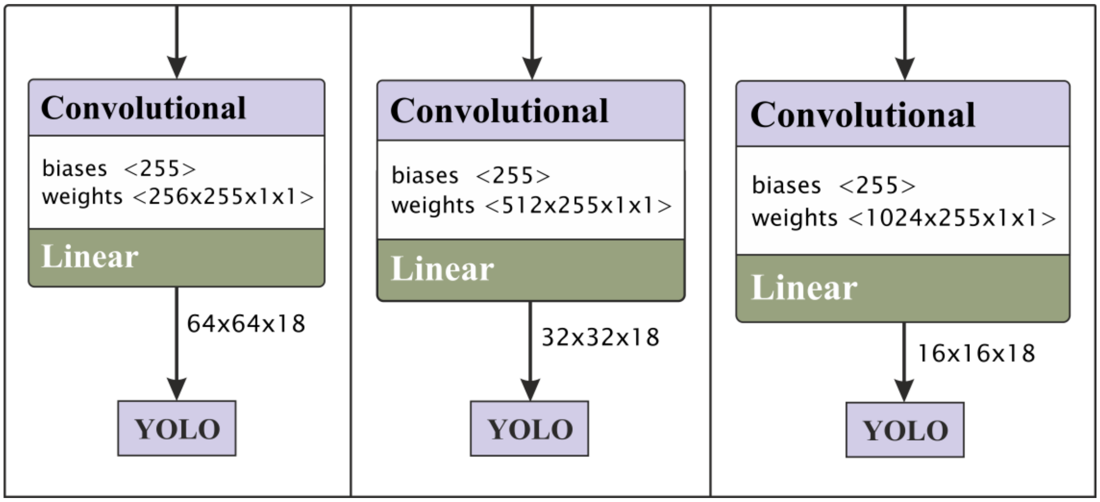
Various models use a similar policy to make connec- tions between layers, such as Cross-Stage-Partial-connections (CSP) or Dense Blocks (consisting of Batch Normalisation, ReLU, Convolution, etc.) in DenseNet. Such models are iments by us we concluded that CSPDarknet53 is the most optimal backbone model for our application, in spite of higher complexity (due to more number of parameters). Here, the higher number of parameters leads to the increased capability of the model in detecting multiple objects while at the same time we can maintain real-time performance.

Recently, some of the modern proposed models have placed some extra layers between the backbone and the head, called the *neck*, which is considered for feature collection from dif- ferent stages of the backbone network.

The neck section consists of several top-down and bottom- up paths to collect and combine parameters of the network in different layers, in order to provide a more accurate image features for the head section. Some of the techniques used in the neck section are the Feature Pyramid Network (FPN), Path Aggregation Network (PAN) and Spatial Attention Module (SAM).

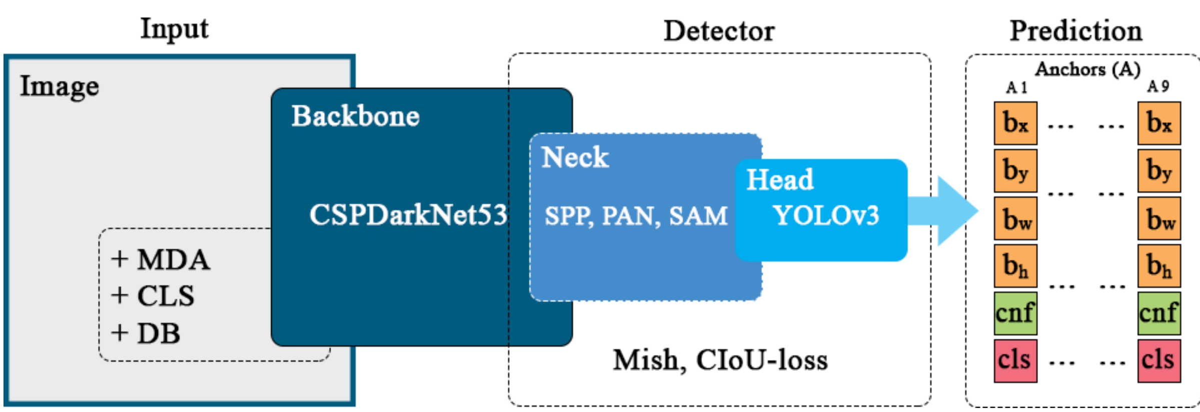
In YOLO-v4, the authors have dealt with two categories of training options for different parts of the network: *"Bag of Freebies"*, which includes a set of methods to alter the model's training strategy with the aim of increasing the generalisation; and *"Bag of Specials"* which includes a set of modules that can significantly improve the object detection accuracy in exchange for a small increase in training costs.

Many CNN-based models use Fully-Connected layers for classification part, and consequently, they can only accept fixed dimensions of images as input. This can lead to two types of issues: firstly, we cannot deal with low resolution images and secondly, the detection of the small objects would be difficult. These are in contradiction with our objectives where we aim to have our model applicable in any surveillance cameras with any input image sizes and resolutions.



The heads that we applied at different scales of the network for object detection at different sizes. YOLOv3 uses FPN to extract features of different scales from the backbone, however in YOLOv4 based approach we use a modified version of the Spatial Pyramid Pooling layer (SPP), instead of FPN, to address the issue of various spatial dimensions as well as dealing with multi-scale detection in head section.

In SPP module we see the input image as a “Bag of Words" feature-map which are divided into m x m bins, where m can be 1, 2, or 4. A max-pooling is then applied to each bin for each channel to gets best features. In section 4 we will examine the efficiency of this approach in improving the accuracy of our model.



.

For the training phase of the model, we considered the *person* category of the Microsoft COCO Dataset as well as the Open Images Datasets V6+ which includes 16 Million ground truth bounding boxes in 600 categories. The dataset is a collection of 19,957 classes and the major part of the dataset is suitable for human detection and identification. The dataset is annotated using the bounding-box labels on each image along with the corresponding coordinates of each label.

We also considered the category of human body parts such as the legs as we believe this allows the detector to learn a more general concept of a human being, particularly in occluded situations or in case of partial visibility e.g. at the borders of the input image were the full-body of the individuals cannot perceived.

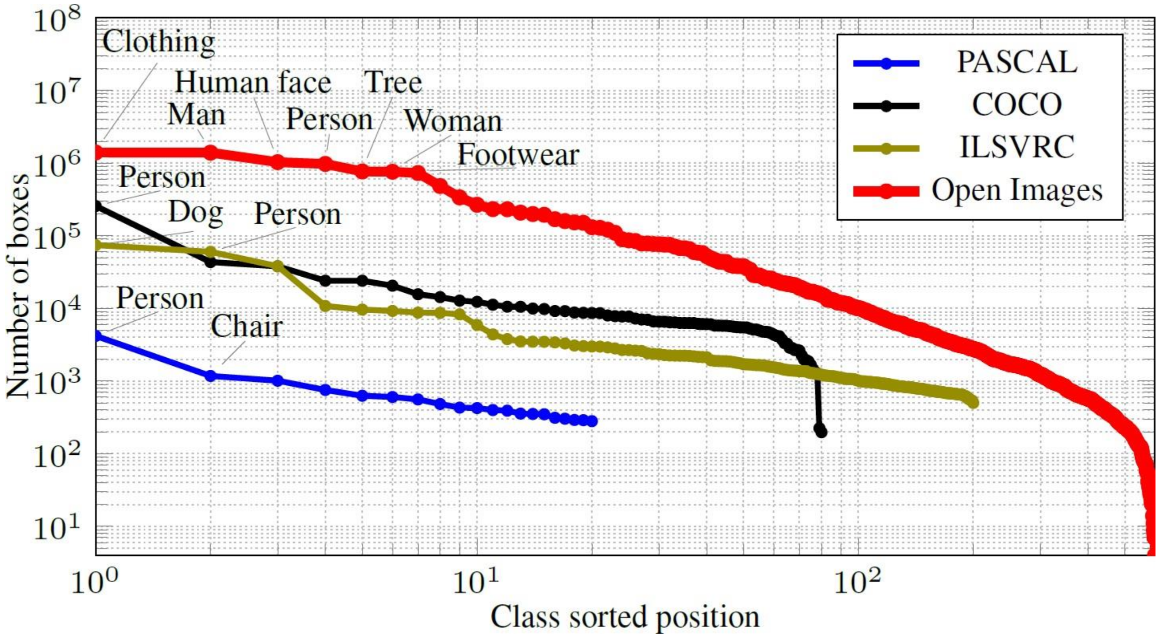
People Tracking

Stereo-vision is a popular technique for distance estimation such as in; however, this is not a feasible approach in our research when we aim at integration of an efficient solution, applicable in all public places using only a basic CCTV cam- era. Therefore we adhere to a monocular solution.

On the other hand, by using a single camera, the projection of a 3-D world scene into a 2-D perspective image plane leads to unrealistic pixel-distances between the objects. This is called perspective effect, in which we cannot perceive uniform distribution of distances in the entire image.

In three-dimensional space, the centre or the reference point of each bonding box is associated with three parameters (x, y, x), while in the image received from the camera, the original 3D space is reduced to two-dimensions of (x, y), and the depth parameter (x) is not available. In such a lowered-dimensional space, the direct use of the Euclidean distance criterion to measure inter-people distance estimation would be erroneous.

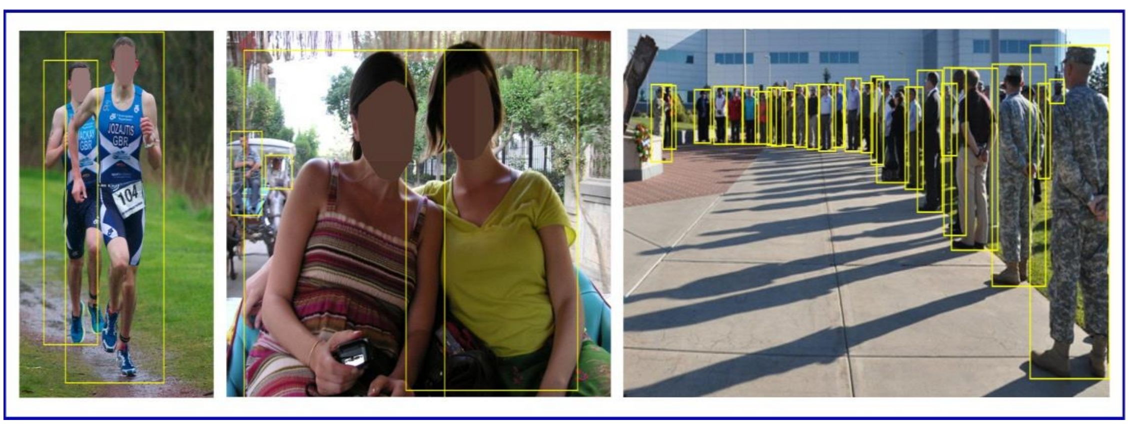
In order to apply a calibrated IPM transition, we first need to have a camera calibration by setting $x = 0$ to eliminate the perspective effect. We also need to know the camera location, its height, angle of view, as well as the optics specifications (i.e. the camera intrinsic parameters).



Experimental Results

In order to train the developed model, we considered a transfer learning approach by using pre-trained models on Microsoft COCO dataset followed by fine-tuning and optimisation of our YOLO-based model.

Four common multi-object annotated datasets including Pascal VOC, COCO, Image Net ILSVRC, and Google Open Images, were investigated in terms of the number of bounding boxes for human or person.



In order to test the performance of the propose model, we used the Oxford Town Centre (OTC) dataset[29] as a previously unseen and challenging dataset with very frequent cases of occlusions, overlaps, and crowded zones. The dataset also contains a good diversity of human specimens in terms of cloths and appearance in a real-world public place.

In order to provide a similar conditions for performance analysis of YOLO based models, we fine-tuned each model on human categories of the Google Open Images (GOI) data set. This has been done by removing the last layer of each model and placing a new layer (with random values of the uniform probability distribution) corresponding to a binary classification (presence or absence of a human). Furthermore, in order to provide an equal condition for the speed and gen-eralizability, we also tested each of the trained models against the OTC dataset.

We evaluated and compared our developed models against three common metrics of object detection in computer vision, including Precision Rate, Recall Rate, and FPS against three state-of-the art human/object detection methods.

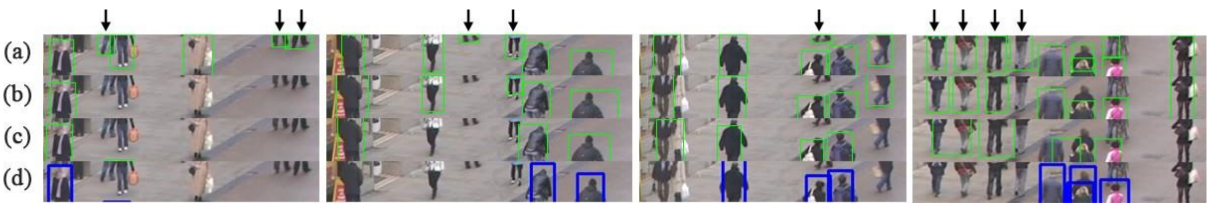
All of the benchmarking tests and comparisons were con- ducted on the same hardware and software: a Windows 10- based platform with an Intel Core i5-3570K Processor and an NVIDIA RTX 2080 GPU with CUDA version 10.1.

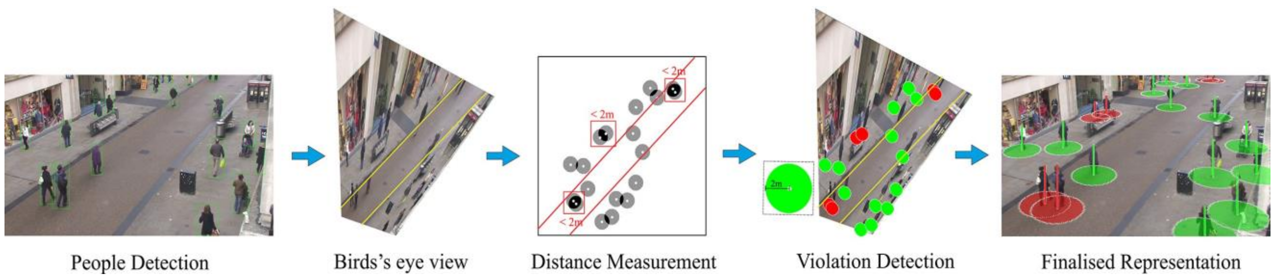
The development of loss function in training and validations phases for four versions of our Deep-SOCIAL model with different backbone structures. The graphs confirm a fast yet smooth and stable transition for minimising the loss function in DS version after 1090 epochs where we reached to an optimal trade-off point for both the training and validation loss. Interestingly, the Faster-RCNN model shows good generalizability; however, its low speed is an issue which seems to be due to the computational cost of the “region proposal” technique. Since the system requires a real-time performance, any model with the speeds lower than 10 *fps* and/or a low level of accuracy may not be a suitable option for Social Distancing monitoring. Therefore, SSD and Faster-RCNN fail in this benchmarking assessment, despite their popularity in other applications. YOLOv3 and v4 provide relatively better results comparing the other models, and finally, the proposed Deep SOCIAL-DS model outperforms in terms of both speed and accuracy.

It provides sample footage of the challenging scenarios when the people either enter or exit the scene, and only part of their body (e.g. their feet) are visible. The figure clearly indicates the strength of Deep SOCIAL in Row (a), comparing to the state-of-the-art. The bottom row (d) with blue bounding boxes shows the ground truth were some of the existing people with partial visibility are not annotated even in original ground truth dataset. Row (c), YOLOv3 shows a cou- ple of more detections; however, th e IOU of the suggested bounding boxes are low and some of them can be counted as False Positives. Row (b), the standard YOLOv4-based detector, shows a significant improvement comparing to row and is considered as the second-best. Row (a), the Deep- SOCIAL, shows 10 more true positive detections (highlighted by vertical arrows) comparing to the seconds best approach.

Social Distancing Evaluations

We considered the midpoint of the bottom edge of the de- tected bounding boxes as our reference points (i.e. shoes' location). After the IPM, we would expect to have the location of each person, in the homogeneous space of BEV with alinear distance representation.

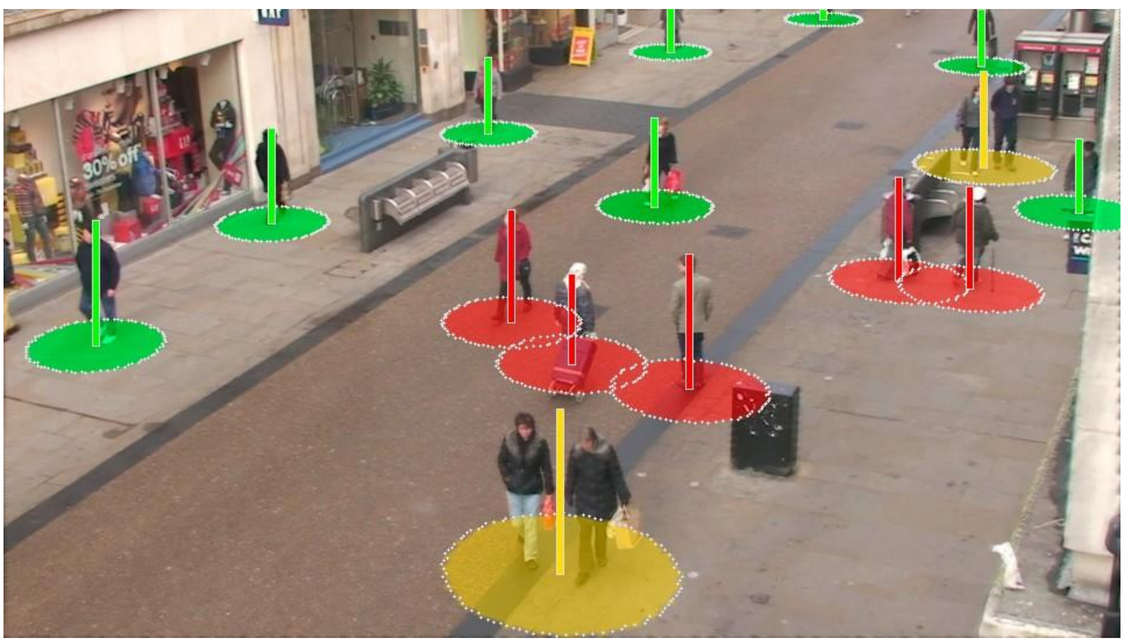




Any two people P_i, P_j with the Euclidean distance ofsmaller than r (i.e. the set restriction) in the BEV spaceare considered as contributors in social distancing violation: Depending on the type of overlapping and the violationassessment criteria, we define a violation detection function V with the input parameters of a pixel metrics ξ, the set safedistance of r (e.g 6ft or 2m), the position of the query human H_q, and the closest surrounding person P_o.

Regarding the Oxford Town Centre (OTC) dataset, every 10 pixels in the BEV space is equivalent to 98cm in the real world. Therefore, r 2 ξ and equal to 20 pixels. The inter-people distance measurement was measures based on the Euclidean L2 norm distance.

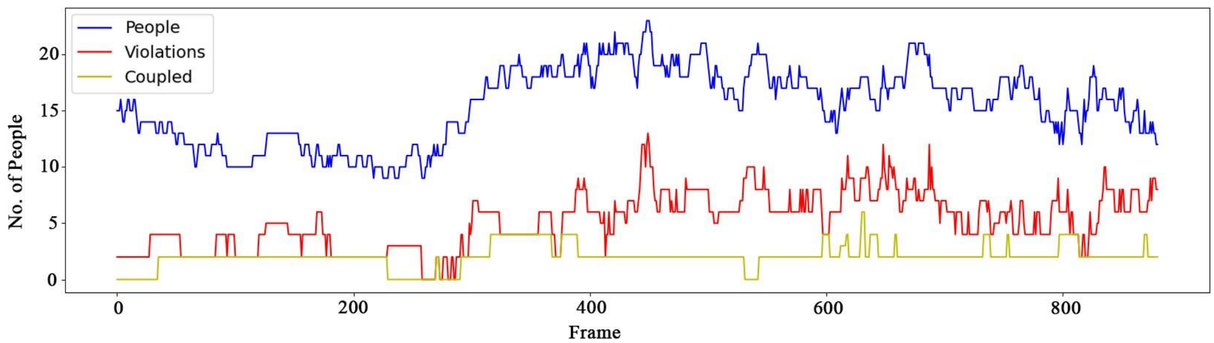
One of the controversial opinions that we received from health authorities were the way of dealing with family mem-bers and couples in social distancing monitoring. Some re- searchers believed social distancing should apply on every single individuals without any exceptions and others were advising the couples and family members can walk in a close proximity without being counted as breach of social distanc-ing. In some countries such as in the UK and EU region the guideline allows two family members or a couple walk to-gether without considering it as the breach of social distancing. We also considered a solution to activate the couple detection. This will be helpful when we aim at recognising risky zones based on the statistical analysis of overall movements and social distancing violations over a mid or long period (e.g. from few hours to few days).

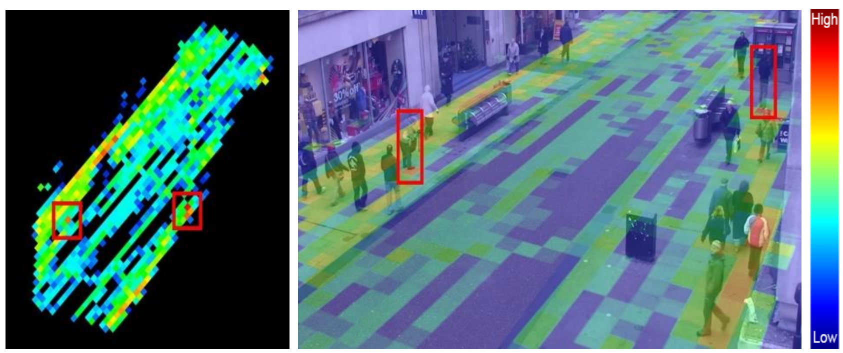


Zone-based Risk Assessment

We also tested the effectiveness of our model in assessing the long-term behaviour of the people. This can be valuable for health sector policy makers and governors to make timely decisions to save lives and reduce the consequent costs. Our experiments provided very interesting results that can be crucial to control the infection rates before it raises uncontrolled and unexpectedly. In addition to people inter-distance measurement, we considered a long-term Spatio-temporal zone-based statistical analysis by tracking and logging the movement trajectory of people, density of each zone, total number of people who violated the social-distancing measures, the total time of the violations for each person and as the whole, identifying high-risk zones, and ultimately, creating an informative risk heat-map.

In order to perform the analysis, a 2-D grid matrix $G_t \in R^{w \times h}$ (initially filled by zero) is created to keep the latestlocation of individuals using the input image sequences. $G_t$represents the status of the matrix at time t and w and h arethe width and height of the input image I, respectively.





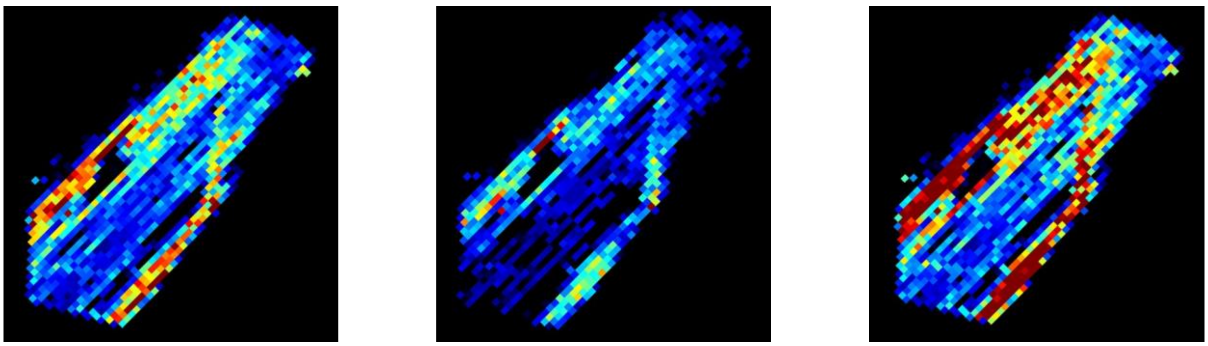
(a) BEV Heat Map **(b)** 2D Mosaic Heat Map

Since COVID-19 is an airborne virus, breathing by any s tatic person in a fixed location can increase the density of the contamination on that point (assuming the person may carrythe COVID-19), particularly in covered places with minimal ventilation. Therefore we can assign a more contamination weight to the steady state people.

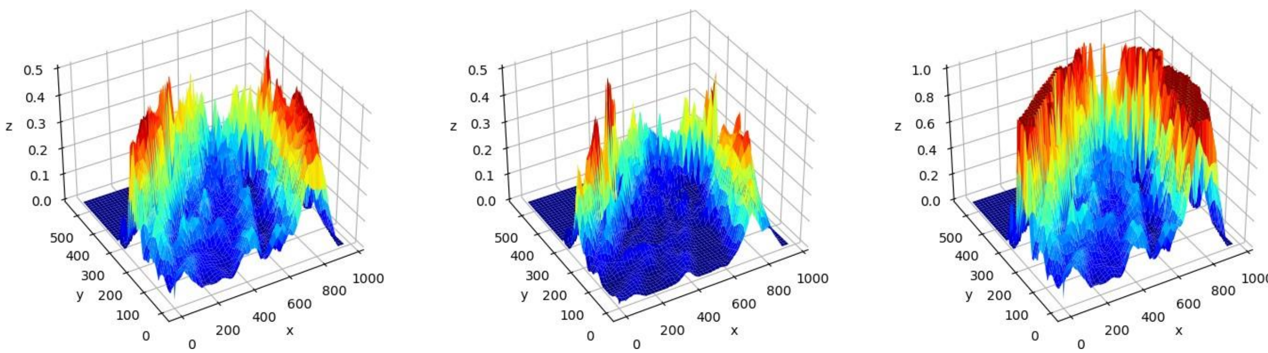
It shows two instances of cases wheretwo people have been steady in two particular location of thegrid for a long period; hence, the heat map is turning to red for those locations. Both sidewalks also show stronger heat- map than the middle of the street due to the higher traffic of people movements. In general, the more red grids potentiallyindicates more risky spots.

In addition to people raw movement and tracking data, thatwould be more beneficial analyse the people who particularlyviolated the social distancing measures.

(a) Moving trajectory and tracking map **(b)** Long-term 2D mosaic heat map of Violations **(c)** 2D mosaic heat map of Tracking + Violations



(d) Long-term BEV Tracking heat map **(e)** Long-term BEV Violation heat map **(f)** BEV Tracking + Violation heat map

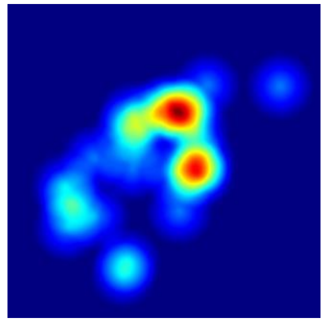


(g) 3D Tracking heat map **(h)** 3D Violation heat map **(i)** 3D Tracking + Violation heat map

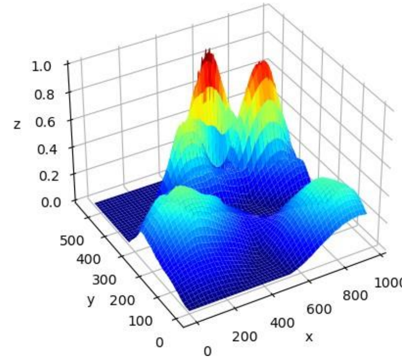
The above configuration was an accumulating approach were all the violations and risky behaviours were added together in order to highlight potentially risky zones in a covered area with poor ventilation.

We also thought about cases when there exist a good chance of ventilation were the spread of virus would not be necessarily accumulative. In such cases we considered both increasing and decreasing counts depending on the overall time spent by each individual in each grid cell of the image, as well as the total absence time of individuals which potentially allows bring down the level of contamination.

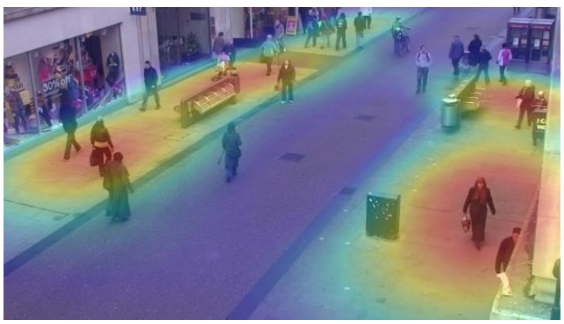
(a) Single-frame 2D crowd map



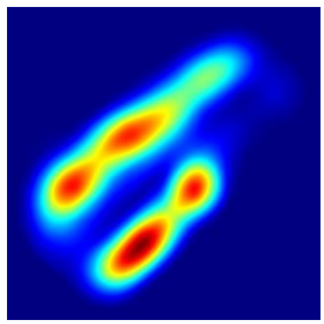
(b) Single-frame BEV crowd map



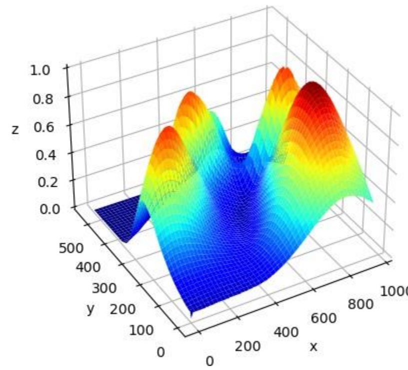
(c) Single-frame 3D crowd map



(d) Long-term crowd map



(e) Long-term BEV crowd map



(f) Long-term 3D crowd map

The 2D and 3D representation of violations and risk heat maps where we applied both increasing and decreasing contamination trends. The first row of the image represents a single frame analysis with two peak zones. As can be seen, those zones belong to two crowded zones were two large groups of people are walking together and breach the social distancing rule. However, in other parts of the street a minimal level or risk is identified. This is due to large inter people distances and a consequent time gaps which allows breathing and therefore a decreasing rate of contamination.

The second row of a long-term crow map which does not necessarily depend on the current frame. This can be a weighted averaging over all of the previous single- frame crow maps.

One of the extra research questions inis how to define appropriate averaging weights and coefficients (α, β, γ) in Equation and how to normalise the maps over the time. This is out of the scope of this research and needs further studies. However, here we aimed at showing the feasibility of considering a diversity of cases using the proposed method with a high level of confidence and accuracy in social distancing monitoring and risk assessment with the help of AI and Computer Vision.

Conclusion

We proposed a Deep Neural Network-Based human detector model called Deep SOCIAL to detect and track static anddynamic people in public places in order to monitor social distancing metrics in COVID-19 era and beyond. We utilised a CSPDarkNet53 backbone along with an SPP/PAN and SAM neck, Mish activation function, and Complete IOU loss function and developed an efficient and accurate human detector,applicable in various environments using any type of CCTV surveillance cameras. The system was able to perform in a variety of challenges including, occlusion, lighting variations, shades, and partial visibility. The proposed method was eval-uated using large and comprehensive datasets and proved a major development in terms of accuracy and speed comparedto three state-of-the-art techniques. The system performed real-time using a basic hardware and GPU platform.

References

1. World Health Organisation. WHO Corona-viruses Dis-ease Dashboard (August 2020). Available at https://covid19.who.int/table.
2. WHO Director, Generals. Opening remarks at the me- dia briefing on COVID-19 (2020). WHO generals and directors speeches.
3. Olsen, S. J. *et al.* Transmission of the severe acute respi-ratory syndrome on aircraft. *New Engl. J. Medicine* **349**,2416–2422, DOI: 10.1056/NEJMoa031349 (2003).
4. Adlhoch, C. *et al.* Considerations relating to social dis- tancing measures in response to the COVID-19 epidemic(2020). European Centre for Disease Prevention and Control, Technical report.
5. Ferguson, N. M. *et al.* Strategies for mitigating an in- fluenza pandemic. *Nature* **442**, 448–452, DOI: https://doi.org/10.1038/nature04795 (2006).
6. Thu, T. P. B., Ngoc, P. N. H., Hai, N. M. *et al.* Ef- fect of the social distancing measures on the spread of COVID-19 in 10 highly infected countries. *Sci. Total. Environ.* 140430, DOI: https://doi.org/10.1016/j.scitotenv. 2020.140430 (2020).
7. Morato, M. M., Bastos, S. B., Cajueiro, D. O. & Normey-Rico, J. E. An optimal predictive control strategy for COVID-19 (SARS-CoV-2) social distancing policies in Brazil. *Elsevier Annu. Rev. Control.* DOI: https://doi.org/10.1016/j.arcontrol.2020.07.001 (2020).
8. Fong, M. W. *et al.* Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings—social dis- tancing measures. *Emerg. infectious diseases* **26**, 976, DOI: 10.3201/eid2605.190995 (2020).
9. F, A., N, Z. & A, U. Effectiveness of workplace social distancing measures in reducing influenza transmission: a systematic review. *BMC public health* 1–13, DOI: 10. 1186/s12889-018-5446-1 (2018).
10. Australian Government Department of Health. Deputy chief medical officer report on COVID-19. *Dep. Heal. Soc. distancing for coronavirus* DOI: https://doi.org/10. 1136/bmj.m1845 (2020).
11. Rezaei, M. & Shahidi, M. Zero-shot learning and its applications from autonomous vehicles to COVID-19 di-agnosis: A review. *SSRN Mach. Learn. J.* **3**, 1–27, DOI: 10.2139/ssrn.3624379 (2020).
12. Toğaçar, M., Ergen, B. & Cömert, Z. COVID-19 de- tection using deep learning models to exploit social mimic optimization and structured chest X-ray images using fuzzy color and stacking approaches. *Comput. Biol. Medicine* 103805, DOI: https://doi.org/10.1016/j. compbiomed.2020.103805 (2020).
13. Ulhaq, A., Khan, A., Gomes, D. & Paul, M. Computer vision for COVID-19 control: A survey. *Image Video Process.* DOI: 10.31224/osf.io/yt9sx (2020). Nguyen, T. T. Artificial intelligence in the battle againstcoronavirus (COVID-19): a survey and future research di- rections. *arXiv Prepr.* **10**, DOI: 10.13140/RG.2.2.36491.23846/1 (2020).
14. Choi, W. & Shim, E. Optimal strategies for vaccination and social distancing in a game-theoretic epidemiological model. *J. Theor. Biol.* 110422, DOI: https://doi.org/10. 1016/j.jtbi.2020.110422 (2020).

15. C, E., K, P. & S, W. J. Systematic biases in disease forecasting–the role of behavior change. *J. Epidemics* 96—105, DOI: 10.1016/j.epidem.2019.02.004 (2019).

16. O, K. W. & G, M. A. A contributions to the mathematical theory of epidemics–i. *The Royal Soc. publishing* DOI: https://doi.org/10.1098/rspa.1927.0118 (1991).

17. Heffernan, J. M., Smith, R. J. & Wahl, L. M. Perspectives on the basic reproductive ratio. *J. Royal Soc. Interface* **2**,281–293, DOI: 10.1098/rsif.2005.0042 (2005).

18. Gupta, R., Pandey, G., Chaudhary, P. & Pal, S. K. Machine learning models for government to predict COVID-19 outbreak. *Int. J. Digit. Gov. Res. Pract.* **1**, DOI: 10.1145/3411761 (2020).

19. Nguyen, C. T. *et al.* Enabling and emerging technologies for social distancing: A comprehensive survey. *arXiv preprint* DOI: https://arxiv.org/abs/2005.02816 (2020).

20. C, R. T. Game theory of social distancing in responseto an epidemic. *PLoS computational biology* 1–9, DOI: 10.1371/journal.pcbi.1000793 (2010).

21. Ainslie, K. E. *et al.* Evidence of initial success for chinaexiting COVID-19 social distancing policy after achiev-ing containment. *Wellcome Open Res.* **5**, DOI: https://doi.org/10.12688/wellcomeopenres.15843.1 (2020).

22. Vidal-Alaball, J. *et al.* Telemedicine in the face of the COVID-19 pandemic. *Atencion primaria* **52**, 418—422,DOI: https://doi.org/10.1016/j.aprim.2020.04.003 (2020).

23. Sonbhadra, S. K., Agarwal, S. & Nagabhushan, P. Targetspecific mining of COVID-19 scholarly articles using one-class approach. *J. Chaos, Solitons & Fractals* **140**, DOI: 10.1016/j.chaos.2020.110155 (2020).

24. Punn, N. S. & Agarwal, S. Automated diagnosis of COVID-19 with limited posteroanterior chest X-ray im-ages using fine-tuned deep neural networks. *Image Video Process.* DOI: arXiv:2004.11676 (2020).

25. Punn, N. S., Sonbhadra, S. K. & Agarwal, S. COVID-19epidemic analysis using machine learning and deep learning algorithms. *medRxiv* DOI: https://doi.org/10.1101/ 2020.04.08.20057679 (2020).

26. Jhunjhunwala, A. Role of telecom network to manage COVID-19 in india: Aarogya Setu. *Transactions IndianNatl. Acad. Eng.* 1–5, DOI: 10.1007/s41403-020-00109-7 (2020).