Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Real-time social distance measurement and face mask detection in public transportation systems during the COVID-19 pandemic and post-pandemic era: Theoretical approach and case study in Italy

Marco Guerrieri a,⁎, Giuseppe Parla b

⁎ Corresponding author.
E-mail address: marco.guerrieri@unitn.it (M. Guerrieri).

Contents lists available at ScienceDirect
Transportation Research Interdisciplinary Perspectives
journal homepage: www.sciencedirect.com/journal/transportation-research-interdisciplinary-perspectives

ARTICLE INFO

Keywords:
Covid-19
Pandemic and post-pandemic era
Social distancing
Face mask detection
Deep learning
YOLOv3

ABSTRACT

Due to its remarkable learning ability and benefits in several areas of real-life, deep learning-based applications have recovered to be a research topic of great importance in the last few years. This article presents a method devoted to guaranteeing safety conditions in public transportation systems (PTS) during the COVID-19 pandemic and post-pandemic era. The paper describes a viable real-time model based on deep learning for monitoring social distance between users and detecting face masks in stop areas and inside vehicles of public transportation systems. Detections are made using the deep learning approach and YOLOv3 algorithm. The safety rule violations are represented by red bounding boxes and red circles in a bird’s eye view as output of the video surveillance analysis. The datasets used to train the neural network are the “Caltech Pedestrian Dataset” and the “COVID-19 Medical Face Mask Detection Dataset”. Metrics, such Loss Accuracy, and Precision, obtained in the testing process of the neural network were used to evaluate the performance of the model in detecting users and face masks. The proposed method was recently tested in the Public Transportation System of the Municipality of Piazza Armerina (Italy). The results show a significant reliability of the method in detecting real-time interactions between users of the PTS in terms of over-time variations in their mutual distancing, as well as in recognising cases of violation of the imposed social distancing and FFP2 face mask use.

Introduction

The World Health Organization (WHO) on 11 March 2020 claimed that the international outbreak of the new coronavirus SARS-CoV-2 infection was to be considered a pandemic. Coronavirus disease (COVID-19) can cause multiple symptoms with different levels of severity, ranging from asymptomatic disease to acute respiratory failure, even death. The coronavirus can be spread through the air by droplets (5 to 10 μm) and fine aerosols (smaller than 5 μm) exhaled from infected individuals when breathing, speaking, coughing, or sneezing (Prather et al., 2020). The virus is highly contagious, the incubation period varies from a minimum of two days to a maximum of fourteen days depending on the variant considered; the median incubation period was estimated to be 5.1 days (Lauer et al., 2019). The disease spread quickly around the world and globally had a shocking impact on the population’s health, welfare and economics. In accordance with the warnings of the WHO, physical distancing between people helps limiting COVID-19 infections from spreading widely and thus it has been taken as a key strategy for reducing the number of infected people since the beginning of the pandemic. Following the advice of the World Health Organization, the public authorities have proposed 1 to 2 m of interpersonal distancing. Another effective way to prevent COVID-19 infections is to wear face masks (Tang et al., 2009) and therefore, in several countries, wearing face masks has become compulsory in crowded areas (e.g. in public transportation systems). As the spread of the virus mainly depends on the interaction among people, social distancing and wearing face masks can reduce the spread of the disease. Reducing the rate of spread of the disease (e.g. in terms of the effective reproductive number R0) means less infection, a smaller number of deaths and even less undesirable effects on other domains of life. Therefore, health authorities around the world have suggested wearing face masks, maintaining adequate “social distancing” in indoor environments, (at least one meter) and avoiding crowded places even outdoors. These are the only non-pharmaceutical public health measures that, if effectively respected by the majority of
the population, provide indirect benefits like a lower probability of oversaturation of the health system of a given country and a significant reduction in the rate of infections and deaths (Fig. 1). The evolution of the spread of an infectious disease can be analysed through epidemiological models, the best known of which is the SIR model by Kermack and McKendrick (Kermack & McKendrick, 1927). Recently, Eskin et al. (2019) have developed an epidemiological model largely based on the SIR model but suitably modified to include the effect that compliance with an appropriate social distancing can lead to the evolution of the spread of the infection in a given examined population. Hence, with the use of adequate epidemiological models, it is possible to formulate a reliable estimate of the future evolution of the pandemic and then adequately program the types and intensity of health measures to contrast the phenomenon.

Public transportation systems (buses, trains, trams, trolleybuses, ferries, rapid transit, etc.) play a crucial role in the transmission of the virus (De Vos, 2020), given the densely crowded conditions normally found at station stops and even inside vehicles (Grant and Booth, 2009; Corriere, et al. 2013; Guerrieri, 2019; Calderin Peralvo, et al., 2022). Therefore, in public transportation systems, it may be useful to introduce suitable techniques for monitoring the social distance between users and detecting whether users wear face masks correctly along the journey (Park and Kim, 2021). Clearly enough, conventional monitoring techniques cannot be used for this type of control; instead, smart control measures can play a fundamental role in this field. Today artificial intelligence (AI) is applied to many fields of life and numerous are the applications in the sector of transportation engineering. The ability to “learn” useful information, to propose decision-making processes and to produce useful results are the most surprising characteristics of deep learning. These distinctive features of deep learning will lead to intelligent automation of transportation systems and reduction of operating costs. Computer vision, moreover, allows the computer to evaluate visual data. The recent evolution of Deep Learning (DL) algorithms makes it easy for a computer to analyse digital data and offers the possibility of detecting a huge number of objects of interest in an increasingly reliable way starting from the analysis of images or videos. Consequently, the DL can be used as a method for detecting whether users wear face masks correctly, for analysing the instantaneous distancing between users of local public transport systems and identifying cases in which there are violations of the minimum social distancing established by the government bodies (Ahmed, et al., 2021). This paper aims to giving a viable contribution to real-time techniques for evaluating interpersonal distancing and detecting whether users wear face masks correctly on public transportation systems both at stations and on-board vehicles. After having designed, calibrated and validated a deep learning-based model, several experiments have been conducted in the public transportation system of an Italian city using low-cost devices. The paper is organized as follows. Sect. 2 explains the related work on object detection and recognition systems based on the deep learning approach and YOLO v3 algorithm. Sec. 3 briefly explains the main characteristics of the datasets used for users’ detection, the neural network training process and the outputs. Sect. 4 explains the procedure applied for tracking users in the consecutive frames of the surveillance video by means of the Inverse Perspective Mapping (IPM) and the Kalman filter. Sect. 5 includes the description of the case study, the experimental activities and the main results. Finally, conclusions are given in Sect. 6.

Deep learning-based technique

Deep learning (DL) constitutes a subset of artificial intelligence (AI) and is now used in many real-time technical applications for human life facilitation, including the detection of people in an image or video sequences of interest. In general, transportation system user and face mask detection can be considered as object detection and recognition problems (Ottakath, et al., 2022). methods of the YOLO family are very effective in detecting people in other types of algorithms (Faster RCNN, SSD, etc.). The second generation of YOLO, called YOLOv2 increases the detection accuracy by utilizing batch normalization to convolution layer, anchor box, multi-scale training and fine-grained features (Ito, et al., 2022). Despite these advantages, the detection accuracy of YOLO v2 is very low for small objects.

Consequently, YOLOv2 is not appropriate for face mask detection. The third generation of YOLO, called YOLOv3, includes 53 convolutional levels and 23 residual levels. YOLOv3 is able to detect in real time the “objects” of interest in a given scene, even if the objects are very small (Guerrieri and Parla, 2021). YOLOv3 performs multi-scale...
classification employing various logistic classifiers. In YOLOv3 the small
feature maps give semantic information, instead large ones give finer-
grained information. In short, the YOLOv3 algorithm has the qualities
of multiscale prediction and multi-label classification. Due to these
peculiar characteristics, in the research presented here, the YOLOv3
algorithm, whose schematic architecture is represented in Fig. 3, was
used as a fast real-time algorithm for face masks and social distance
detection and measure. YOLOv3 utilizes Darknet-53 as the core inter-
connection interface, which acts as the attribute extractor for classifi-
cation. Darknet-53 is characterized by 53 fully convolution layers with
residual conjunctions (cf. Fig. 4).

In YOLOv3 the input image is subdivided into $S \times S$ grids. Each grid
cell predicts three bounding boxes. The method of predicting the
bounding box (cf. Fig. 5) is given by the relationships (1):

$$
\begin{align*}
    b_x &= \sigma(t_x) + c_x \\
    b_y &= \sigma(t_y) + c_y \\
    b_w &= p_w e^{t_w} \\
    b_h &= p_h e^{t_h}
\end{align*}
$$

(1)

In which $t_x$ and $t_y$ represent the relative position coordinates of
the centre of the bounding box, $t_w$ and $t_h$ are the width and height of
the bounding box. $c_x$ and $c_y$ represent the net, and $p_w$ and $p_h$ are the width
and height of the predicted front bounding box. Finally, $b_x$, $b_y$, $b_w$, and $b_h$
are the true coordinates of the centre of the bounding box and the true

---

**Fig. 2.** State-of-the-art detection methods for people and face mask detection (adapted from: Wang et al. 2021).

**Fig. 3.** YOLOv3 Network structure.

**Fig. 4.** Darknet-53 framework in YOLOv3.
width and height of the bounding box obtained after prediction.

The model predicts a confidence score ($t_i$) representing the probability that the grid will detect an object for each bounding box by logistic regression (Wan et al., 2021):

$$\sigma(t_i) = P_{r_{\text{object}}} \times \text{IOU}(b, \text{object}) \quad (2)$$

The probability is 1 when the prediction bounding box coincides with the ground truth location to the cut-off threshold or else the probability value is 0 (Wan et al., 2021).

In order to evaluate the performance of the model in detecting the users of the public transportation system and face masks, metrics such as Precision, Recall and Accuracy were used in this research. The metrics were determined with the following expressions:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

In which TP, TN, FN and FP variables identify the true positives, true negatives, false negatives and false positives respectively. The TP are referred for objects that were labelled true and predicted as true. The TN are for objects that were labelled true but predicted as false. FP are referred for objects that are labelled as false but predicted as false; FN are the images that are labelled false but predicted as true (Ottakath, et al. 2022).

One of the most important criteria for evaluating the performance of the detection model is the loss function. In Yolov3 the loss function is divided into three main components: classification loss, localization loss (errors between the predicted boundary box and the ground truth) and confidence loss (the objectness of the box). The loss function can be expressed as follows (Hui, 2018):

$$\lambda_{\text{coord}} \sum_{i=0}^{N} \sum_{j=0}^{B} l_{ij}^\text{box}[(x - \tilde{x}_i)^2 + (y - \tilde{y}_i)^2] + \lambda_{\text{coord}} \sum_{i=0}^{N} \sum_{j=0}^{B} l_{ij}^\text{loc}[(\sqrt{w_i} - \sqrt{w_{\text{gt}}})^2 + \sqrt{h_i} - \sqrt{h_{\text{gt}}})^2] +$$

$$\sum_{i=0}^{N} \sum_{j=0}^{B} l_{ij}^\text{obj}(C - \tilde{C}_i)^2 + \lambda_{\text{obj}} \sum_{i=0}^{N} \sum_{j=0}^{B} l_{ij}^\text{obj}(p(c) - \tilde{p}(c))^2$$

\(6\)
or without protective face masks.

The neural network training process was performed using a workstation with Intel(R) Core(TM) i7-4510 CPU @ 2.00 Hz 2.60 GHz – Memory RAM 6 GB, Windows 10 Home. In all, the process requested a Time Elapsed of 00:50:30 h and 60 Epoch (1 Epoch = 25 iterations); the following values were assumed: learning rate = 0.001; L2 regularization factor = 0.0005, penalty threshold = 0.5.

Starting from a surveillance video, for each processed image, the algorithm evaluates:

- the coordinates of the position of the Bounding Box related to each user of the public transportation system;
• the coordinates of the position of the Bounding Box related to each face mask;
• the statistical confidence level associated with each Bounding Box;
• the label that identifying the class name of the identified object (user of the transportation system and face mask).

Figs. 8–10 illustrate the Loss, Accuracy and Precision values obtained in the testing process; it is clear that when accuracy increases, loss decreases.

Tracking of users and social distancing evaluation

Starting from the analysis of a video recording, each user of a given public transportation system is detected with the procedure explained in the previous sections. Reliable user tracking in the different frames of a video recording can be achieved by using the Kalman filter. Consequently, for each instant of time the proposed technique identifies clusters of people who maintain a social distance (d) lower than the minimum social distance rule (d_{min}) and evaluates the exposure time. In this research, the estimated social distance is that one between the people centroids projected on the ground plane (i.e. road pavement surface for stop areas or vehicle floor) by the inverse perspective mapping (IPM) procedure (Dorj & Lee, 2016). With the use of IPM the position of the bounding box of each detected user in the perspective view is converted into a top-down view (Fig. 11). Denoting with:

- \{F_w\} = \{X_w, Y_w, Z_w\} the world frame centered at the camera optical center;
- \{F_c\} = \{X_c, Y_c, Z_c\} the camera frame;
- \{F_i\} = \{u, v\} an image frame (Fig. 19);

- h the height of the camera frame (Fig. 9) with respect to the ground plane,

under the hypothesis that the optical axis has no roll (i.e. the camera frame X axis is in the world frame X_wY_wZ_w plane), the projection point on the road plane of each point \( p = \{u, v, 1, 1\} \) on the image plane can be estimated by Eq. (7) (Aly, 2008):

\[
\hat{T} = h \begin{bmatrix}
\frac{1}{f_u} c_2 & \frac{1}{f_v} s_1 s_2 & \frac{1}{f_u} c_4 c_2 - \frac{1}{f_v} c_s s_1 s_2 - c_1 s_2 & 0 \\
\frac{1}{f_v} s_1 c_1 & \frac{1}{f_v} c_s s_2 - \frac{1}{f_u} c_1 c_2 - c_1 c_2 & 0 & 0 \\
0 & \frac{1}{f_v} c_1 c_1 + s_1 & 0 & 0 \\
0 & \frac{1}{h f_v} c_1 & \frac{1}{h f_u} c_1 - \frac{1}{h} s_1 & 0 
\end{bmatrix}
\]

where:

- \( f_u, f_v \) are the horizontal and vertical focal lengths respectively;
- \( (c_u, c_v) \) are the coordinates of the optical centre;
- \( c_1 = \cos \alpha, \ c_2 = \cos \beta, \ s_1 = \sin \alpha, \) and \( s_2 = \sin \beta \).

Similarly, given a point on the road plane \( w = \{x_w, y_w, -h, 1\} \), its subpixel coordinates can be calculated on the image frame by the equation \( \hat{p} = \hat{T} w \), i.e. by the inverse of the transform (Aly, 2008):

\[
\hat{T} = \begin{bmatrix}
f_u c_2 + c_1 & c_2 c_1 - s_2 f_u & -c_s s_1 & 0 \\
 c_1 c_2 - s_1 & f_s c_2 - f_u c_1 & c_1 c_2 & 0 \\
 s_z(s_1 c_1 - f_v s_1) & c_1 c_2 + c_2 f_s s_1 & -f_v c_1 - f_v c_s s_1 & 0 \\
c_1 s_2 & c_1 c_2 & -s_1 & 0
\end{bmatrix}
\]
A Bird’s-eye view of users’ centroids can be obtained from Eq. (7). Consequently, we can project the centroids of all users of the public transportation system from the input image onto the ground plane and then calculate the social distances. In fact, for each user identified in the scene of each frame of the video under analysis, the model creates a circumference of radius \( r = \frac{d_{\text{min}}}{2} \) (Fig. 12). Denoting with \((x_1, y_1)\) and with \((x_2, y_2)\) the coordinates of the centroids of two detected people (Id1 and Id2 respectively, cf. Fig. 12) the social distance violation occurs if \( d < d_{\text{min}} \).

The Euclidean distance \( d \) can be evaluated by Eqs. (9), (10), (11): \[
d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2} - 2r = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2} - \frac{d_{\text{min}}}{2}
\]

\[
d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2} - \frac{d_{\text{min}}}{2} \geq 0 \rightarrow \text{no social distance violation}
\]

\[
d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2} - \frac{d_{\text{min}}}{2} < 0 \rightarrow \text{social distance violation}
\]

In the final output the social distancing violations are automatically identified by red circles in a “bird’s eye view” of the scene and by red bounding boxes in the original images.

Since one or more users can be detected in successive frames of the analysed video, it is crucial to apply a tracking algorithm. The methodology used in this work is summarised in Fig. 13. In particular, when a user of the public transportation system comes into the scene, it is considered as a new tracking object, then the algorithm assigns the Id number and initializes the bounding box for that user. To decrease the noise introduced by inaccurate or imprecise detections the linear Kalman filter is employed. As it is well known, the Kalman filter (Kalman, 1960) is a recursive predictive filter that assesses the status of a dynamic system. The Kalman filter equation is (Welch and Bishop, 2006):

\[
x_{n+1} = A_n x_n + B_n u_n
\]

Considering the error covariance (Welch and Bishop, 2006; Niu, 2018):

\[
P_{n+1} = A_n P_n A_n^T + Q_n
\]

where \( x_n \) is the state value at step \( n \), \( A_n \) is the state transition matrix, \( u_n \) is...
the measurement and the input at step n. \( Q_n \) is the white noise covariance. Kalman gain value is given by the following relationship (Niu, 2018):

\[
K_n = P_n C^T (C P_n C^T + R_n)^{-1}
\]

(14)

where \( C \) is the measurement matrix and \( R \) is the measurement noise.

Actual measurement value at the updated time and error covariance is the relationship (Niu, 2018):

\[
P_n = (I - K_n H) P_n
\]

(15)

In which \( K_n \) is the measurement value and \( H \) is the mapping matrix from true state to observation.

An example of user trajectory calculated with the Kalman filter is given in Fig. 16a.

**Experimental activities and results**

The complete description of the experiment executed in this research is given in this section. The experimental activities were undertaken in January and February 2022 in the Public Transportation System (PTS) of the Municipality of Piazza Armerina (Italy), operated by the company Savic Sciclone Ltd. In particular, some video cameras were installed at stops and inside buses (Fig. 9). The recorded video streams were split up into frames of sequences for the accurate estimation of the distance values between each couple of the detected users.

The most important challenge of detecting social distancing from videos is the accuracy of the measurement of the actual distance between public transportation users (Zuo, et al., 2021). To solve this problem, the holography technique can be applied to morph the video frames from a perspective view into a top-bottom view. Considering prefixed objects with known dimensions as references, distances can be computed in the transformed frames (Szeliński, 2010). Therefore, the first phase of the experimental activities concerned the camera calibration so as to obtain the extrinsic parameters of the cameras themselves (Fig. 14). To this end, Zhang’s algorithm was used and applied to a set of twenty-two different images of a chessboard of known size. The chessboard is employed to scale the pixels in the image acquired by the camera to compute the focal length, the principal point etc (Gad, et al. 2020). The camera calibration parameters, including the height from the ground level in the scene, the yaw and the pitch, are used in the inverse perspective mapping (IPM) process to create a bird’s-eye view of the images of interest. In conclusion, by using the procedures set out in the previous sections, each user of the PTS is detected within a certain video recording. The original input frames are subject to a Gaussian smoothing filter to reduce noises in the images. Subsequently, an order number is assigned to each user and the position of his centroid in the coordinate system centred in the photographic objective, and the mutual distances are calculated between the users identified in the scene. The system then checks whether the minimum social distancing between single pairs of users, or between different pairs of users belonging to the same cluster, is respected or not. The overall procedure is summarised in Fig. 15.

This procedure is applicable for analysing safety condition external to the vehicles of the PTS (i.e. at stop areas as shown in Figs. 16 and 17 or for users entering and exiting vehicles) and those inside vehicles (Fig. 18). As a result of the analyses performed by the proposed algorithm, the violations of the minimum social distancing are highlighted with red bounding boxes and can also be correlated to the time of exposure to the risk of infection, that is the interval of time during which the minimum social distancing is violated. Finally, the proposed technique is able to identify whether users are wearing respiratory protective FFP2 face masks correctly or not (Fig. 19). As can be observed in Figs. 18 and 19, users maintaining the minimum permissible social distance are identified using green bounding boxes, whereas the users violating the threshold limit for social distancing are highlighted by red bounding boxes or circles. Fig. 17c shows an example of coordinates for user centroids obtained by a bird’s-eye view (cf. transformation matrix given in Eq. (7) and Eq. (8)) for a specific instant of time. After determining the coordinates (in pixels) of each user’s centroid, the pixels in a bird’s-eye view are converted into metres in order to compute the social distancing between a couple of users.

As part of the experiments only few cases of false positives and negatives were generated, but in general the results have made it possible to highlight a considerable reliability of the proposed method in detecting real time interactions between users of the PTS in terms of over-time variations in their mutual distancing, as well as in recognising cases of violation of the imposed rules. The number of false negatives and false positives decreases when the longitudinal, lateral and vertical acceleration of the camera decreases during the vehicle motion. An increase in camera oscillation can be observed near horizontal curves (Vaiana et al., 2018), road intersections (Gallelli, et al., 2019; Mauro and Guerrieri, 2016; Guerrieri, et al. 2018) or in case of road pavement damages. Therefore, in order to increase the precision of the proposed technique it is essential to use vibration reduction lenses and vibration sensor technologies. These technologies help reduce the amount of blur on a video caused by movement of the camera due to vehicle oscillations with consequent benefits in terms of accuracy in object detection. Experimental outcomes have demonstrated that the proposed technical approach is characterised by notable accuracy and detection speed and can effectively detect the users of the public transportation systems who maintain or breach the permitted minimum social distance as well as calculate the number of violations. In addition, the Loss, the Accuracy and the Precision values achieved in the training process of the neural network and the experimental results show that the algorithm proposed...
has high accuracy in face mask detection.

**Conclusions**

This article presents a model based on computer vision, deep learning and YOLOv3 as a contribution to preserve a safe environment by scanning the areas inside vehicles or around the stops of a public transportation system in order not to let COVID-19 spread out. The technique proposed aims not to prevent the social distance and the use of face mask violations but rather to decrease their excessive occurrence in public transportation systems. Such a technique, jointly with a user’s face recognition algorithm, can speed up the process of contact tracing that is a key strategy for breaking chains of transmission of SARS-CoV-2 and decreasing COVID-19-associated mortality. In addition, the procedure can be associated with visual and acoustic real-time signals that immediately alert users whenever they are about to breach the

Fig. 16. a) Example of a user’s trajectory obtained using Kalman filter. b) Example of minimum social distance violation (red circles); c) Example of a Bird’s Eye view. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

(a)  (b)   (c)

Fig. 17. a-b) Detection of users in a stop area of the PTS; c) social distance evaluation.

(a)  (b)   (c)

Fig. 18. Detection of users inside a vehicle of the PTS and social distance measurement.

Fig. 19. Detection of users’ face masks (FFP2 type) inside a vehicle of the PTS and social distance measurement.
mandatorily minimal distance from other people. In this research the neural network was trained by using a sample of images of the “Caltech Pedestrian Dataset” and the “COVID-19 Medical Face Mask Detection Dataset” were used. Although the proposed method still requires a vast experimental campaign for properly validating the algorithms, the first results achieved in the case study of the public transportation system (PTS) of the municipality of Piazza Armerina (Italy) showed that the technique is able to accurately measure the reciprocal distance between users and to detect the presence of face masks (type FFP2). In fact, face detectors perform very well even if their accuracy may be further improved. Such a technique can also track both transient and steady-state people in order to measure social distance at each instant of time both in stop areas and on-board vehicles. The outputs of the model, in terms of bounding boxes, help in identifying couples or groups of people that satisfy the required minimum social distance and the number of violations, also correlated to the time of exposure to the risk of infection. It is worth underlining that the methodological approach suggested in the study, although limited only to one PTS, is general and can be applied to other urban public transportation systems as well as railways, metro, light rail systems etc.

CRediT authorship contribution statement

Marco Guerrieri: Conceptualization, Methodology, Writing—original draft, Writing—review & editing. Giuseppe Parla: Investigation, Data curation, Visualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

This research was developed within the Project SAFEMOBILITY-C19 financed by the University of Trento, Internal Call 2020 “Covid 19”. The authors are grateful to Giuseppe Scichilone and Aldo Scichilone of the company Savit Scichilone srl. for the constant willingness during the development of this research.

References

Ahmed, I., Ahmad, M., Rodrigues, J.J.P.C., Jeon, G., Din, S. 2021. A deep learning-based social distance monitoring framework for COVID-19. Sustainable Cities and Society, 65, art. no. 102571.
Aly, M., 2008. Real time Detection of Lane Markers in Urban Streets. IEEE Intelligent Vehic. Symposium, Proc. 7-12.
Calderón Peralvo, F., Cazorla Vanegas, P., Avila-Ordóñez, E., 2020. A systematic review of COVID-19 transport policies and mitigation strategies around the globe. Transp. Res. Interdiscip. Perspect. 15, 100652.
Corriere, F., Di Vincenzo, D., Guerrieri, M. 2013. A logic fuzzy model for evaluation of the railway station’s practice capacity in safety operating conditions. Archives of Civil Engineering, 2013, 59(1), pp. 3–19.
De Vos, J., 2020. The effect of COVID-19 and subsequent social distancing on travel behavior. Transp. Res. Interdiscip. Perspect. 5, 100121.
Dewi, C., Chen, R.-C., Yu, H., 2020. Weight analysis for various prohibitory sign detection and recognition using deep learning. Multimedia Tools Appl. 79 (43–44), 32897–32915.
Dorji, B., Lee, D.J., 2016. A Precise Lane Detection Algorithm Based on Top View Image Transformation and Least-Square Approaches. J. Sens. art. no. 4058093.
Erick, C., Paarporn, F., Weitz, J.S., 2019. Systematic biases in disease forecasting: the role of behavior change. Epidemiology, 27, 96–105.
Farman, H., Khan, T., Khan, Z., Halib, S., Islam, M., Ammar, A., 2022. Real-Time Face Mask Detection to Ensure COVID-19 Precautionary Measures in the Developing Countries. Appl. Sci. (Switzerland) 12 (8).
Gad, A., Elbary, G., Alkhedher, M., Ghazal, M. Vision-based Approach for Automated Social Distance Violators Detection (2020). 2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies, 3ICT 2020, art. no. 9311969.
Gallelli, V., Guido, G., Vitale, A., Vaina, R., 2019. Effects of calibration process on the simulation of rear-end conflicts at roundabouts. J. Traffic Transp. Eng. (English Edition). 6 (2), 175–184.
Grant, M.J., Booth, A., 2009. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Inf. Libr. J. 26, 91–108. https://doi.org/10.1111/j.1471-8422.2009.00848.x.
Guerrieri, M., 2019. Catenary-Free Tramway Systems: Functional and Cost-Benefit Analysis for a Metropolitan Area. Urban Rail Transit 5 (4), 289–309.
Guerrieri, M., Mauro, R., Parla, G., Tollazzi, T. 2018. Analysis of kinematic parameters and driver behavior at turbo roundabouts. Journal of Transportation Engineering Part A: Systems, 2018, 144(6), 04018020.
Guerrieri, M., Parla, G., 2021. Deep learning and yolov3 systems for automatic traffic measurement. Displays, 73, art. no. 102235.
Hui, J. 2018. Real-time Object Detection with YOLO, YOLOv2 and now YOLOv3. Available: https://jonathanhui.medium.com/real-time-object-detection-with-yolo-yolov2-28b1b93e2088.
Kalman, R.E., 1960. A new approach to linear filtering and predictions problems. J. Basic Eng. 82 (D), 35–45.
Kernack, W. O., McKendrick, A. G. 1927. Contributions to the mathematical theory of epidemics-i.
Lauer, S.A., Grantz, K.H., Bis, Q., Jones, F.K., Zheng, Q., Meredith, H.R., Azman, A.S., Reich, N.G., Lessler, J., 2019. The incubation period of coronavirus disease (CoVID-19) from publicly reported confirmed cases: Estimation and application. Ann. Intern. Med. 172 (9), 577–582.
Lu, H., Chen, T., Shi, L., 2022. Research on Small Target Detection Method of Traffic Signs Improved Based on YOLOv3. 2022 2nd International Conference on Consumer Electronics and Computer Engineering, ICCECE.
Mauro, R., Guerrieri, M., 2016. Comparative life-cycle assessment of conventional (double lane) and non-conventional (turbio and flower) roundabout intersections. Transp. Res. Part D: Transp. Environ. 48, 96–111.
Niu, M. 2018. Object Detection and Tracking for Autonomous Driving by MATLAB toolbox. Ohio State University. Available: https://etd.osu.edu/apexprod/rws_etd/send_file/send?accession=osu152978995395614&disposition= inline (accessed on 10 August 2021).
Ottakath, N., et al. 2022. ViDMASK dataset for face mask detection with social distance measurement. Displays, 73, art. no. 102225.
Park, J., Kim, G., 2021. Risk of COVID-19 Infection in Public Transportation: The Development of a Model. Int. J. Environ. Res. Public Health. 18, 12790. doi:10.3390/ijerph18127990.
Prathir, K.A., Wang, C.C., Scholey, R.T., 2020. Reducing transmission of SARS-CoV-2. Science 368 (6498), 1422–1424. https://doi.org/10.1126/science.abe6197.
Szeliski, R., 2010. Computer Vision: Algorithms and Applications. Springer Science & Business Media.
Tang, J.W., Liebner, T.J., Craven, B.A., Settles, G.S., 2009. A sclerotherapeutic treatment of the human cough with and without wearing masks for aerosol infection control. J. R. Soc. Interface 6, S727–S736.
Tassif, S., Anirudh, G., Tombrhune, J., 2022. Object detection using YOLO: challenges, architectural successors, datasets and applications. Multimedia Tools Appl. https://doi.org/10.1007/s11042-022-13644-y.
Vaiana, R., Iauele, T., Gallelli, V., Rogano, U., 2018. Demanded versus assumed friction along horizontal curves: An on-the-road experimental investigation. J. Transp. Saf. Eng. 10 (4), 318–344.
Wang, B., Zheng, J., Philip Chen, C.L. 2021. A Survey on Masked Face Detection Methods and Datasets for Fighting Against COVID-19. Journal of IEEE transactions on artificial intelligence, 00(0), December 2021.
Welch, G., Bishop, G. 2006. An Introduction to the Kalman Filter. Department of Computer Science University of North Carolina at Chapel Hill Chapel Hill, NC 27599-3175. Available: https://www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf (accessed on 1 August 2021).
Zuo, P., Gao, J., Kurkcu, A., Yang, H., Ozay, K., Ma, Q., 2021. Reference-free video-to-real distance approximation-based urban social distancing analytics amid COVID-19 pandemic. Journal of Transport and Health, 21, art. no. 101032.