Accident Analysis and Prevention 156 (2021) 106127

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Evaluation of pedestrian critical gap and crossing speed at midblock crossing using image processing☆

Y. Alver [a,*], P. Onelcin [a], A. Cicekli [a], M. Abdel-Aty [b]

[a] Ege University, Civil Engineering Department, Izmir, Turkey
[b] University of Central Florida, Department of Civil & Environmental and Construction Engineering, Orlando, FL, 32816, USA

ARTICLE INFO

Keywords:
Midblock crossing
Crossing speed
Critical gap
Image processing

ABSTRACT

Pedestrians confront risky situations at midblock sections due to the unyielding behavior of drivers. Thus, pedestrians have to wait for an appropriate gap to cross. This research investigates pedestrians' gap acceptance and crossing speed for midblock crossings by image processing methods in Izmir, Turkey. A total of 498 pedestrians have been tracked at two midblock crossings. The data were collected for one hour at each midblock crossing during the evening peak hour between 5.00–6.00 p.m. Three synchronized cameras were used to record pedestrian crossings. Then, by using image processing, vehicle and pedestrian trajectories have been obtained. Two cameras were mounted on telescopic tripods reaching up to 9 m, and the third camera was used to identify pedestrians' gender better. The parameters extracted from the recordings are; pedestrians' gender, group size, whether they carried items or not, and their accepted/rejected gaps. Pedestrian and item detection has been performed by YOLOv3 and YOLACT models. The accepted and rejected time gaps were extracted for pedestrians, excluding the pedestrians who crossed between stopped vehicles and crossed when an approaching vehicle did not exist within 100 m from the midblock crossing. Raff's method was used to estimate the critical gap using accepted/rejected gaps. The critical gaps ranged between 4.1 s and 6.2 s. The 15th percentile crossing speeds were found to be similar, ranging between 0.78 m/s and 0.80 m/s.

1. Introduction

Object recognition methods using deep learning have been the most progressive fields of study in recent years. With the development of graphics-based processor units (GPU), big data processing has been considerably reduced. Thanks to this technology, video analysis techniques have become much more useful. Object recognition methods decreased the margin of error in the analyses by making a high amount of transactions. Analyses made with human observations might include errors in terms of objectivity due to the human factor. Studies in transportation security have achieved more objective and accurate analysis results using computer-based analysis instead of human observations.

Pedestrians are the most vulnerable group in traffic, especially when they interact with vehicles. Accidents involving pedestrians can occur in rural or urban areas, and most of these accidents occur when pedestrians are crossing the road. An analysis of pedestrian accidents showed that the probability of fatality is higher at midblock locations compared with intersections. These crosswalks are mostly present at locations with little or no vehicle/vehicle conflicts that allow the drivers to elevate their speeds (Siddiqui et al., 2006). Pedestrians are at risk, particularly at uncontrolled intersections and at midblock sections. Pedestrians waiting for drivers to yield become impatient and choose smaller gaps to cross (WHO, 2013). WHO (2013) stated that low and middle-income countries had higher road traffic fatality rates per 100,000 population (24.1 % and 18.4 %, respectively). Road traffic deaths involving vulnerable road users (motorcyclists, pedestrians, and cyclists) constitute half of all road traffic deaths (WHO, 2013). More than one-third of accidents that result in fatalities are related to pedestrians and cyclists (Alkahtani et al., 2019). In Turkey, pedestrian fatalities constituted 22.3 % of road accident fatalities (GDH, 2018). The year 2019 is declared a "traffic year with priority for pedestrians" by the Turkish Interior Ministry. Drivers who do not yield the right of way to pedestrians at midblock crossings have to pay an administrative fine of 100 US dollars. Although pedestrians are protected by law, the drivers do not yield the right of way to pedestrians due to the lack of control. When pedestrians step into the

☆ This paper has been handled by associate editor Chennai Guest Editor.

* Corresponding author.
E-mail address: yalcin.alver@ege.edu.tr (Y. Alver).

https://doi.org/10.1016/j.aap.2021.106127
Received 14 September 2020; Received in revised form 2 April 2021; Accepted 5 April 2021
Available online 14 April 2021

midblock, the driver's expected behavior is to stop; however, this is unlikely observed in Turkey. Thus, severe interactions between pedestrians who have the right of way and vehicles can occur at midblock crossings.

This research investigates pedestrians' gap acceptance and crossing speed for midblock crossings by image processing methods in Izmir, Turkey. Although there are studies that focused on pedestrians, few have adopted image processing to analyze pedestrian behavior. Convolutional neural networks (CNNs) are widely used in image processing. CNN based methods give promising results in pedestrian detection with high accuracy. This study adopts a detection method based on CNN models and tracking by detection method to track pedestrians.

2. Literature review

Vision-based analyses are widely used in transportation safety, especially for urban traffic analysis (Atev et al., 2005; Saunier and Sayed, 2007; Ismail et al., 2010; St-Aubi et al., 2015; Alhajyaseen and Iryo-Asano, 2017; Essa and Sayed, 2018; Guo et al., 2019; Zhang et al., 2020; Essa and Sayed, 2018; Guo et al., 2019; Zhang et al., 2020). Most previous studies adopting image processing have focused on vehicle-vehicle interactions (Atev et al., 2005; Saunier and Sayed, 2007; St-Aubin et al., 2015; Essa and Sayed, 2018; Guo et al., 2019). On the other hand, image processing methods provide ease of computation in pedestrian-related researches as well. Ismail et al. (2010) studied pedestrian safety in a before-after evaluation using automated video data analysis. Zaki et al. (2013) detected traffic violation at a major signalized intersection using computer-vision techniques. Alhajyaseen and Iryo-Asano (2017) extracted pedestrian maneuvers and speed change events at signalized intersections using the image processing system TrafficAnalyzer. Zhang et al. (2020) analyzed pedestrians' red-light crossing intentions at signalized intersections using detection and tracking techniques. Fu et al. (2016) investigated pedestrian-vehicle interactions at non-signalized intersections using the open-source computer vision-based Traffic Intelligence project. According to the previous studies' successive attempts, vision-based tracking techniques are efficient in obtaining highly-accurate trajectory data and extracting safety indicators automatically from the data. Therefore, vision-based trajectory data analysis is promising in investigating pedestrian safety issues at non-signalized intersections, such as the issue associated with secondary interactions (Fu et al., 2019).

Technological developments have made machine learning applications widely used in vision-based methods. This approach makes vision-based methods faster, more accurate and enables the use of big data. Cai et al. (2019) investigated different strategies to improve the crash prediction accuracy, and they suggested that using a deep learning approach could significantly improve the results. Rahman et al. (2019) applied machine learning methods to analyze bicycle and pedestrian crashes. Brunetti et al. (2018) reviewed deep learning techniques used in pedestrian detection and tracking. In Turkey, although pedestrian safety is a critical issue in road safety studies, vision-based tracking techniques are not widely used yet.

In the early phase of image processing, classification of the objects was the primary concern, as the progress has continued detection and localization, semantic segmentation. Finally, instance segmentation can be made accurately (Garcia-Garcia et al., 2017). Object detectors are classified as one-stage and two-stage classifiers. Two-stage detectors give relatively better results compared with one-stage detectors. On the other hand, one-stage detectors are faster for real-time detections (Wu et al., 2020). You Only Look Once (YOLO) is a famous one-stage detector developed by Redmon et al. (2016). The input image is divided into an S x S grid. The object center's presence in the grid means that the specific grid cell is responsible for detecting that object. YOLO was not good at detecting small objects and objects at multiple aspect ratios. Thus, an improved model called YOLOv2 has been proposed, which used anchor boxes to predict bounding boxes (Redmon and Farhadi, 2017). YOLOv3 was another enhanced model, which used independent logistic classifiers instead of softmax, leading to better modeling of the data and performing better with small objects (Redmon and Farhadi, 2018). YOLACT is a fully-convolutional model for real-time instance segmentation. In order to overcome the difficulties of implementing instance segmentation into a one-stage detector, two parallel tasks are held; (a) a dictionary of non-local prototype masks over the entire image is generated, and (b) a set of linear combination coefficients per instance are predicted. For each instance, using the corresponding predicted coefficients, the prototypes are linearly combined, then cropped with a predicted bounding box (Bolya et al., 2019).

In literature, the effects of pedestrian characteristics such as gender

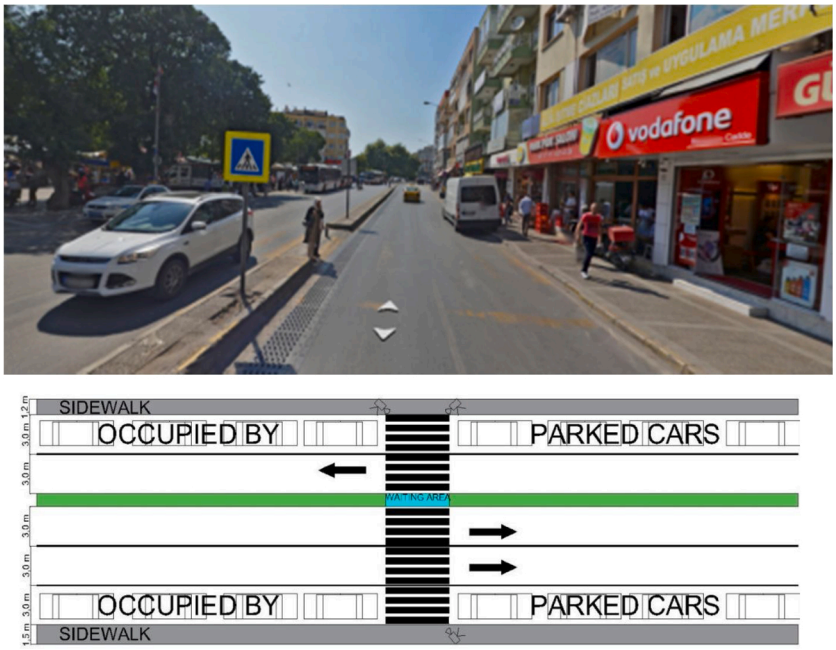


Fig. 1. Midblock crossing in Sirinyer.

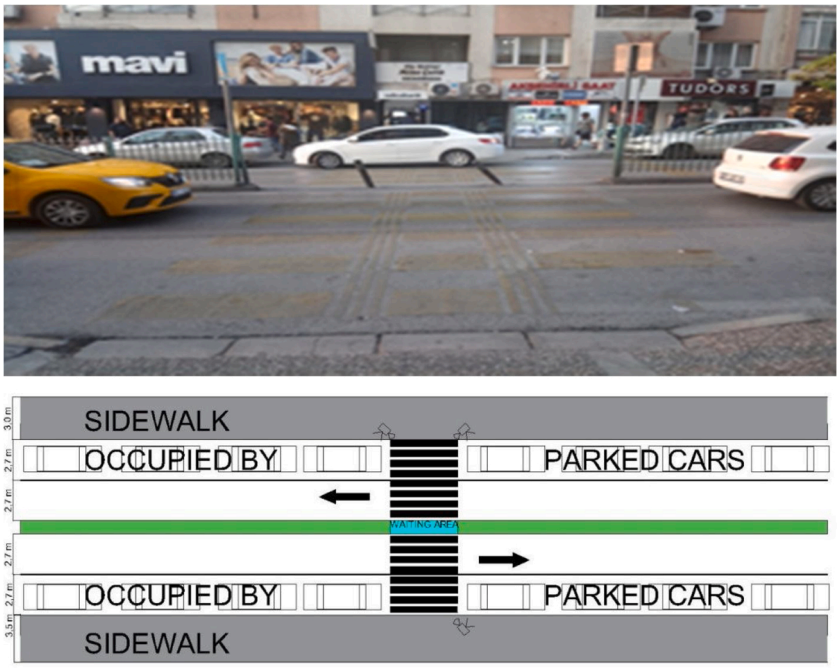


Fig. 2. Midblock crossing in Bornova.

(Rosenbloom, 2009; Guo et al., 2011; Wang et al., 2011; Yannis et al., 2013; Alver and Onelcin, 2018; Ravishankar and Nair, 2018); group size (Rosenbloom, 2009; Zhuang and Wu, 2012; Kadali et al., 2015; Ravishankar and Nair, 2018; Alver and Onelcin, 2018; Avinash et al., 2019;) on crossing speed and gap acceptance have been studied widely. Men generally are faster than women (Ravishankar and Nair, 2018). Yannis et al. (2013) found that men choose larger gaps, particularly when pedestrians accompany them. Pedestrians who crossed in groups showed less rolling gap behavior than individuals (Ravishankar and Nair, 2018). Pedestrians crossing in groups accept smaller vehicular gaps (Avinash et al., 2019), and as the pedestrian platoon increases, the decrease in the gap acceptance is remarkable (Kadali et al., 2015). Zhuang and Wu (2012) indicated that pedestrians in groups had lower crossing speed and vehicles yielded more often to pedestrians when they crossed in bigger groups.

Pedestrians who cross on multi-lane roads check gap availability in all lanes; however, as the pedestrians cross, oncoming vehicles on adjacent lanes lead to the use of rolling gaps (Shaaban and Abdel-Warith, 2017). Zhang et al. (2018) studied crossing strategies and lane-based gaps at midblock crossings and concluded that rolling gap crossing is the predominant crossing strategy among pedestrians. Pawar and Patil (2016) used different methods to determine spatial and temporal gap values at midblock crossings. They found that the actual gaps are lower than the adequate gaps.

Vision-based techniques have been used to identify conflicts generally. In this study, the authors aim to combine pedestrian behavior analyses with the image processing techniques and define the factors affecting the accepted gap by multiple linear regression (MLR) model and find the critical gap by Raff's method at midblock crossings. Previous studies have investigated pedestrian characteristics such as gender, group size, age, vehicle type, etc. However, these studies mainly focused on signalized intersections. This study investigates the effect of gender, group size, items carrying, and vehicle type at midblock crossings where pedestrians' crossing behavior is a critical issue. Another contribution of the study is the validation of image processing algorithms with the manually extracted data and proving that computer-vision techniques can replace human observers and save computational time.

Table 1
Traffic volume during the observation period (veh/hr).

Location	Vehicle type			
	Car-Taxi (veh/h)	Minibus (veh/h)	Heavy vehicles (veh/h)	Motorcycle (veh/h)
Sirinyer	503	87	33	28
Bornova	668	142	60	60

3. Study area

Two midblock crossings have been investigated to analyze pedestrians' crossing speed and gap acceptance in Izmir, Turkey. Izmir is the third biggest city in Turkey. There are 4,320,519 inhabitants in the city. The number of vehicles is 1,425,302 by 2019, of which 781,191 are private vehicles (TSI, 2019).

The major part of the pedestrian crossings is signalized in Izmir. A limited number of crosswalks have remained unsignalized. The two midblock crossings are located in Sirinyer and Bornova districts, where pedestrian density is high. There are shopping centers, banks, schools around the midblock crossings. Figs. 1 and 2 show photographs and measurements of the observed pedestrian crossings sites.

In these districts, the minibusses constitute approximately 15 % of the traffic composition. Minibusses pick up passengers anywhere along the road; hence a significant number of stop and goes are observed in the traffic, which causes delays and a decrease in other vehicles' speed. In Bornova, the road section has four lanes, and in Sirinyer, it has five lanes. Both road sections are divided by a pedestrian refuge. In Sirinyer, the lane width is 3.00 m, whereas, in Bornova, it is 2.70 m. In Bornova, the road section is separated by a median with a physical barrier built to avoid illegal crossings. At both observation sites, analyses were conducted in one direction. Although there were two lanes for each direction in Bornova, the nearest lane to the sidewalk was blocked by parking cars. Hence, only one lane could allow traffic flow. Pedestrians using the lane near the median were used to extract pedestrians' crossing speed and critical gap. In Sirinyer, there were three lanes in the observed direction. Here also, the nearest lane to the sidewalk was used as a parking lane. Hence, the middle lane and the far-side lane were considered. In

Fig. 3. YOLACT detections.

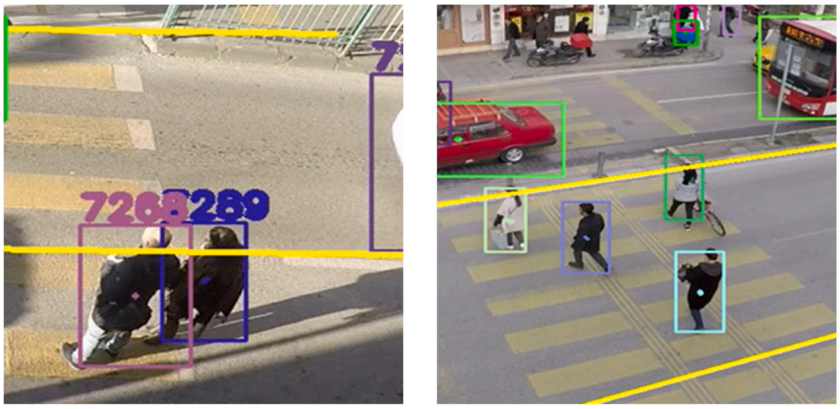


Fig. 4. Pedestrians crossing a) in groups b) individually.

this paper, the first lane is referred to as the lane that the pedestrian encounters first, either from the sidewalk or the median.

The observed traffic volume by vehicle type is presented in Table 1. In Bornova, the number of minibusses and heavy vehicles is comparatively higher than in Sirinyer.

4. Methodology

4.1. Data collection

The data were collected for a one-hour duration at each midblock crossing during the evening peak hour between 5.00–6.00 p.m. Three synchronized cameras were used to record pedestrian crossings. Two cameras were mounted on telescopic tripods reaching up to 9 m. The third camera was used to identify pedestrians' gender better, observing the midblock crossing only. One camera allowed making observations for approximately 100 m, which covered both the midblock crossing and the approaching vehicles. The second camera covered pedestrian crossing and the vehicles within the 25 m. GoPro Black Session 5 cameras were used with a wide-angle lens filming at 30 fps at a resolution of 1280 × 720.

4.2. Pedestrian detection

Two detectors were used to detect pedestrians, namely, YOLACT (Bolya et al., 2019) and YOLOv3 (Redmon and Farhadi, 2018). Both are one-stage detectors; however, YOLACT is a real-time instance segmentation framework.

The camera that covered the pedestrian crossing allowed for pedestrians' detection, and it could also be detected whether they carried items or not. Both YOLOv3 and YOLACT were used to detect pedestrians and items; however, YOLACT performed better in items (handbag, backpack, etc.) detection. Pedestrian counts were made using YOLOv3, and item detection was made using YOLACT. Meanwhile, gender was recorded on manual observations. Fig. 3 shows YOLACT detections for both observation sites. Pedestrians crossing individually and crossing in groups were determined by the distance of the centroids of the bounding boxes. Fig. 4 shows an example of this distinction. In Fig. 4b, none of the bounding boxes intersect; on the other hand, in Fig. 4a, the two bounding boxes intersect. Pedestrians are assumed to be crossing in a group if the distance between the centroids is 60 cm in the real-world. Its pixel correspondence is 60 pixels for Sirinyer midblock, 40 pixels for Bornova midblock. If the distance between the two centroids is lower than this distance for five consecutive frames, pedestrians were assumed to be crossing together. Familiarity with each other was not a necessary condition to count pedestrians as a group.

4.3. Pedestrian crossing speed

The crossing speed of pedestrians was computed as suggested by Noh et al. (2020) in Eq. 1. The pedestrians' velocity is found by dividing the pedestrian's distance moving between frames by the time interval between the frames. Pixel distance was transformed into the real-world distance in meters using the pixel-per-meter (P) constant.

$$Velocity = \frac{object\ distance}{F*P} \tag{1}$$

The actual length of the midblock crossings was measured on-site, and their pixel equivalents were obtained from the videos. Distance covered by the pedestrian was calculated in pixel coordinates. The pixel distance was calculated for each pedestrian that started crossing the lane of interest until he finished crossing the lane. As the camera's fps is

known (30 fps), the pedestrians' speed could be found. Then using P, the pixel coordinates were converted to real-world coordinates.

4.4. Pedestrian trajectory

In this study, the tracking by detection method was used to track pedestrians and vehicles. Tracking was performed by Simple Online and Real-time Tracking (SORT) algorithm. SORT is a multiple object tracking algorithm proposed by Bewley et al. (2016). SORT is an efficient implementation of a tracking-by-detection framework. Detections from the previous and current frames are presented to the tracker. Although the original work was proposed to track pedestrians, convolutional neural networks based detectors can detect various objects, such as vehicles. This algorithm combines the Kalman filter and the Hungarian algorithm.

Kalman filter was presented as a recursive solution to the discrete data filtering problem (Bishop and Welch, 2001). Kalman filter has been widely used in many fields, including object tracking. The next position of the object is estimated based on the object's previous position and current observation (Farahi and Yazdi, 2020). The state vector is represented by x_k, which is to be estimated from the measurement z_k. In Eqs. 2 and 3, **A** is the transition matrix, k is the discrete-time, **H** is the measurement matrix, w_{k-1} is the Gaussian process noise, and v_k is the Gaussian measurement noise. Q and R, are covariance matrices of w_{k-1} and v_k, respectively.

$$x_k = \boldsymbol{A}x_{k-1} + w_{k-1} \tag{2}$$

$$z_k = \boldsymbol{H}x_k + v_k \tag{3}$$

$$p(w) \sim N(0,\ Q) \tag{4}$$

$$p(v) \sim N(0,\ R) \tag{5}$$

The *a priori* estimate of error covariance (P_k^-) is given by Eq. 6.

$$P_k^- = AP_{k-1}A^T + Q \tag{6}$$

The *a priori* estimate of state ($\widehat{x}_k^-$) is given by Eq. 7.

$$\widehat{x}_k^- = \boldsymbol{A}\widehat{x}_{k-1} + w_k \tag{7}$$

Finally, as a linear combination of a priori estimate $\widehat{x}_k^-$, and a weighted difference between the measurement z_k and a measurement prediction $H\,\widehat{x}_k^-$, the a posteriori state estimate $\widehat{x}_k$ is computed. K_k is the Kalman gain. The higher values of the Kalman gain corresponds to the higher measurement accuracy. Otherwise, the Kalman gain has a relatively low value. This process is repeated with the *a posteriori* estimates to predict the new a priori estimates (Hamuda et al., 2018; Pathan et al., 2009; Bishop and Welch, 2001).

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \tag{8}$$

$$\widehat{x}_k = \widehat{x}_k^- + K_k(z_k - H\widehat{x}_k^-) \tag{9}$$

$$P_k = (I - K_k H)P_k^- \tag{10}$$

In SORT algorithm, the state of each target (**x**) is a seven-dimensional state space. In Eq. 11, u and v are the bounding box center position, s is the scale, r is the aspect ratio of the target's bounding box. The other three elements are the respective velocities in image coordinates. The aspect ratio is assumed to be constant. The target state is updated using the detected box (Bewley et al., 2016).

$$\mathbf{x} = [u, v, s, r, u', v', s']^T \tag{11}$$

The intersection-over-union distance between the detection and predicted bounding boxes constitutes the assignment cost matrix. Hungarian algorithm is used to solve the assignment problem. Hence, the predicted Kalman states and the measurements are associated using the Hungarian Algorithm (Bewley et al., 2016).

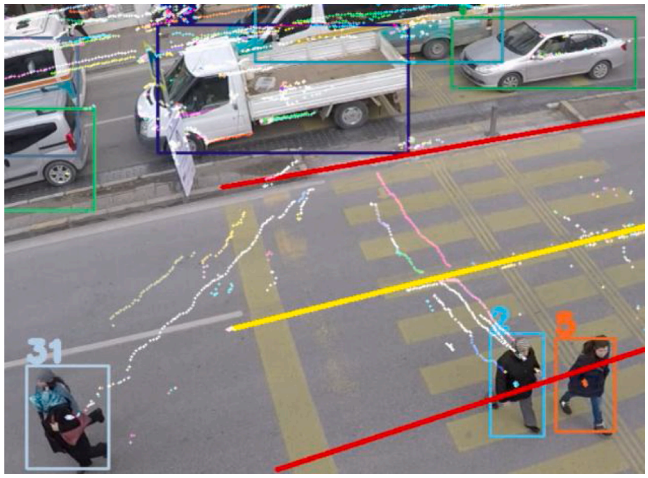


Fig. 5. Pedestrian and vehicle trajectories.

Trajectories are the points belonging to the center positions of the moving pedestrian and vehicle. Trajectories were extracted for each frame. Fig. 5 shows an example of the printed trajectories.

The accepted and rejected time gaps were extracted for pedestrians, excluding the following ones:

- Pedestrians who crossed between stopped vehicles.
- Pedestrians who crossed when no vehicles were approaching within 100 m from the midblock crossing.

Pedestrian and vehicle trajectories were recorded for each frame. The rejected gaps were the gaps collected until the pedestrian started to move. As the crossing began, the trajectories started to change for each frame. The trajectories of the vehicle approaching the pedestrian were recorded, as well. The frame difference between the pedestrians' arrival to the middle of the lane and the vehicle reaching the same point was calculated. Then using the fps, the frame difference was converted to time.

In Sirinyer, there were two lanes that the pedestrian could be in an interaction with a vehicle. Thus, to observe any difference related to gap acceptance in lanes, gap analyses were made on a lane-based.

4.5. Multiple linear regression (MLR) model

Multiple linear regression is a type of multivariable analysis. A multivariable analysis is defined as "a statistical tool for determining the relative contributions of different causes to a single event or outcome" (Katz, 2011). The type of the outcome determines which type of multivariable analysis will be used. The linear regression model is used with continuous outcomes. The MLR model equation is given below:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_n X_n \tag{12}$$

In Eq. 12, Y is the outcome (dependent variable), $X_{1\text{-}n}$ represents the explanatory (independent variables), β_0 is the y-intercept of the regression line, and $\beta_{1\text{-}n}$ are the estimated coefficients.

5. Results

During the video recordings, 644 pedestrians were counted. However, pedestrian speeds and trajectories were extracted, excluding the pedestrians who crossed when the vehicle was not moving and crossed when there was no vehicle approaching the pedestrian. In the end, 498 pedestrians were evaluated for this study.

In Bornova, 167 out of 169 pedestrians and in Sirinyer, 315 out of

Table 2
Comparison of pedestrian counts.

		Pedestrian counts			
		From median to curbside	From curbside to median	Total counts	% difference
Sirinyer	**Image processing counts**	201	233	434	**4.83**
	Manual counts	212	244	456	
Bornova	**Image processing counts**	78	105	183	**2.66**
	Manual counts	80	108	188	

Table 3
Comparison of detection results.

Factor	Performance (%)	
	YOLACT	YOLOv3
Items detection	84.21	66.67

Table 4
Pedestrian descriptives.

Factor		Count	
		Sirinyer	Bornova
Gender	Female	175	102
	Male	154	67
Group size	Individual	88	79
	2+	241	90
Carrying items	Yes	169	56
	No	160	113
Total		329	169

321 pedestrians' accepted gap values could be extracted. When a person is detected and starts to be tracked, the center coordinates are recorded for each frame. However, in some cases, due to the losses in detection or tracking, coordinates could not be recorded, and thus the gap values could not be extracted. In Sirinyer, there were two lanes that pedestrians encountered vehicular traffic. Pedestrians crossing the first lane (either from the curbside or the median) finding an available gap could cross the second lane between the stopped vehicles. There were also cases in the second lane; there was no approaching vehicle to the pedestrian. Thus, the number of pedestrians was reduced to 221.

5.1. Detection results

Pedestrians were detected using YOLOv3, and items were detected using YOLACT. In order to verify the detection results, the number of pedestrians crossing the midblock was manually counted, and meanwhile, the gender of the pedestrians was recorded in an Excel file. Pedestrian counts are compared in Table 2 for both locations and both directions (from the median to curbside, and from curbside to median). YOLOv3 could detect more than 95 % of the pedestrians that crossed. In total, 644 pedestrians crossed during the observation period.

In items detection, YOLACT performed better. The results are given in Table 3. YOLACT could detect approximately 85 % of the items, while YOLOv3 could only detect 67 %.

Table 6
Comparison of pedestrian' critical gaps.

Midblock	Critical Gap (s)		
	Image Processing Results	Manual Results	% difference
Bornova	4.1	3.9	4.88
Sirinyer first lane	6.2	5.9	4.84
Sirinyer second lane	5.7	5.5	3.51

5.2. Pedestrian crossing speed

At both midblock crossings, a total of 498 pedestrians have been tracked. The descriptive statistics are presented in Table 4.

Pedestrian speeds extracted from videos using YOLOv3-SORT algorithms and the speeds computed manually are compared in Table 5. Extracting pedestrian speed manually from the videos is time-consuming and requires too much effort. Thus, automating this process is valuable. As can be seen in Table 6 image processing algorithm resulted in similar crossing speed values extracted manually. The average crossing speed resulted in a 4% difference at most.

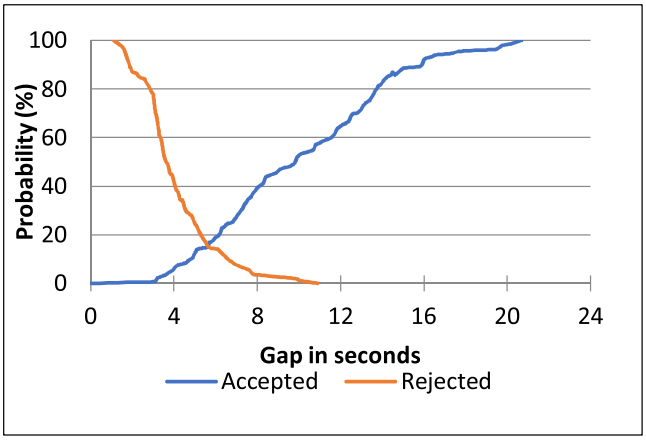
The average crossing speed was 1.00 m/s in Bornova, where pedestrians crossed only one lane due to the parking vehicles. In Sirinyer, the average crossing speed increased in the second lane, for which the pedestrians used rolling gaps. At both midblock crossings, the 15th percentile crossing speeds (0.78 m/s for Bornova, 0.79 m/s for Sirinyer first lane, and 0.80 m/s for Sirinyer second lane) were found to be similar. Pedestrians who crossed individually walked faster than the pedestrians who crossed in a group.

5.3. Pedestrian critical gap

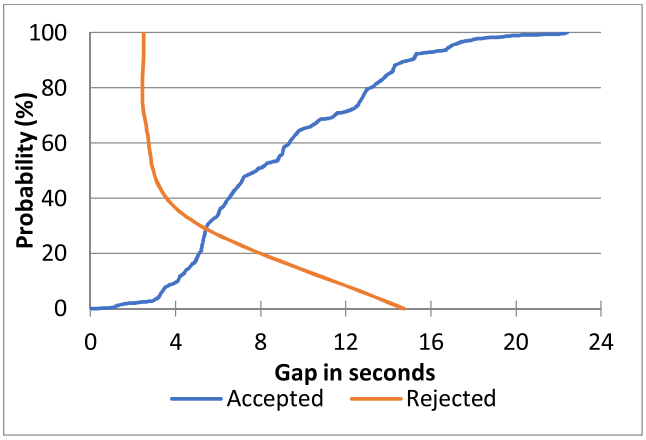
The critical gap was estimated using accepted/rejected gaps by Raff's method (Raff and Hart, 1950). Fig. 6 shows the results of the method. The critical gap is 6.2 s for the first lane and 5.7 s for the second lane in Sirinyer. In Bornova, the critical gap is found to be 4.1 s. In Bornova, the observed traffic volume is higher. The minibusses exhibit a stop-and-go movement because they are allowed to pick up passengers anywhere along the road. For this reason, the vehicles move at low speeds. Hence, pedestrians may perceive that the drivers could brake and stop easily even if they do not yield the way to the pedestrians. It is also expected that pedestrians become impatient and choose smaller gaps to cross

Table 5
Comparison of pedestrian crossing speeds.

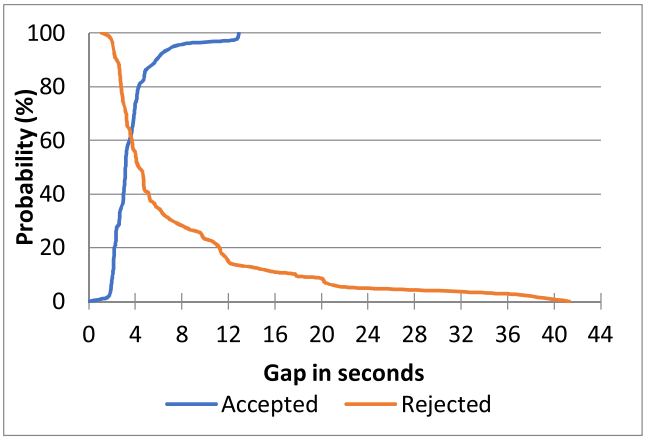
Factor		Avg. Crossing Speed (m/s) (Image Processing Result) Bornova	Avg. Crossing Speed (m/s) (Manual Result)	% Difference	Avg. Crossing Speed (m/s) (Image Processing Result) Sirinyer Lane 1	Avg. Crossing Speed (m/s) (Manual Result)	% Difference	Avg. Crossing Speed (m/s) (Image Processing Result) Sirinyer Lane 2	Avg. Crossing Speed (m/s) (Manual Result)	% Difference
Gender	Female	1.00	1.04	4.00	1.09	1.13	3.67	1.11	1.14	2.70
	Male	1.00	1.04	6.00	1.16	1.21	4.31	1.19	1.22	2.52
Group size	Individual	1.03	1.07	3.88	1.22	1.28	4.92	1.31	1.33	1.53
	2+	0.98	1.03	5.10	1.14	1.18	3.51	1.17	1.20	2.56
Carrying items	Yes	1.00	1.04	4.00	1.04	1.08	3.84	1.01	1.03	1.98
	No	0.99	1.03	4.04	1.12	1.15	2.68	1.12	1.15	2.68
Average		1.00	1.04	4.00	1.13	1.17	3.54	1.15	1.18	2.61



(a) Sirinyer first lane



(b) Sirinyer second lane



(c) Bornova

Fig. 6. Critical gap at midblock crossings.

when the area is congested.

In Fig. 6, the blue curve represents the accepted gap probability, and the red curve represents the rejected gap probability. The critical gap is where the two curves intersect.

The critical gap values obtained from image processing algorithms and manually are compared in Table 6. The difference between the results has less than 5% difference.

Table 7
List of independent variables.

Variable	Definition (unit)
Gender	Male: 0; Female: 1
Group size	Individual: 0; Two or more:1
Carrying items	Without items:0; With items:1
Vehicle type (Car; Minibus; Heavy vehicle; Motorcycle)	Car:1; otherwise:0 Minibus:1; otherwise:0 Heavy vehicle:1; otherwise:0
Existence of physical barrier	No: 0; Yes:1

Table 8
MLR results of gap size.

Variable	β	Std. error	t-Value	p-Value
Constant	8.125	0.337	24.142	0.000
Physical barrier	−4.934	0.629	−7.847	0.000
Vehicle Type (Car)	1.768	0.302	5.858	0.000

5.4. Multiple linear regression (MLR) model results

SPSS v25 was used to model the relation between the gap and the explanatory variables. Pearson's correlation was used to check the collinearity between the variables. The confidence level was chosen as 95 %.

The list of the independent variables used in the model is given in Table 7.

The results of the MLR test are summarized in Table 8. The existence of a physical barrier and vehicle type are found to be significant on the accepted gap.

The final MLR equation is as follows:

$$Gap = 8.125 - 4.934 PhysicalBarrier + 1.768 VehicleType \quad (13)$$

Physical barriers are built to avoid pedestrian crossings except for designated facilities. In Bornova, the physical barrier forced pedestrians to cross only at the midblock section. The model shows that pedestrians accepted smaller gaps where a physical barrier was built and is reflected in the critical gap (4.1 s) found by Raff's method. On the other hand, in Sirinyer, pedestrians could cross at midblocks and anywhere along the road. Here, pedestrians accepted larger gaps compared to Bornova.

Vehicle size is another parameter affecting gap acceptance. Pedestrians accepted smaller gaps when the approaching vehicle was a car, which may be related to the perception that a collision with a bigger sized vehicle might cause more severe consequences.

6. Discussion

In this study, pedestrians' crossing speed, factors affecting the gap acceptance, and critical gap were investigated using tracking by detection method for two mid-block crossings in Izmir, Turkey using image processing. Midblock crossings, in particular when they are located on a high-speed road, are risky for pedestrians if the drivers do not yield to them. Unfortunately, pedestrians face hard times at midblock crossings in developing countries due to drivers' unyielding behavior.

Image processing is an ever-evolving field with numerous applications in different branches. In the last decade, deep learning models offered efficient methods in the field of object detection. The detection results are encouraging in the vehicle and pedestrian detection. In this study, pedestrian and item detection was performed using YOLOv3 and YOLACT. Both models belong to the YOLO family. It is highly possible to obtain accurate results using image processing, and YOLOv3 resulted in a highly accurate detection rate (more than 95 %) in pedestrian detection. Computing pedestrian speed and gap acceptance require a lot of time and effort; thus, image processing is a useful tool to minimize

human effort.

Pedestrian crossing speeds and gaps were extracted using SORT, which is the tracking by detection method. Once the pedestrians and vehicles are detected, they are tracked in every frame, and the coordinates of their centers are recorded, which are used in further analysis. The critical gap was found using Raff's method. Pedestrian safety and behaviors are studied at different locations, such as signalized intersections, midblock crossings, and overpass locations. However, a limited number of studies used image processing in pedestrian detection, speed, and gap estimation at midblock crossings.

The results of the study showed that the average crossing speed is 1.00 m/s in Bornova, where the pedestrians encounter vehicle flow in the lane close to the median. In Sirinyer, the average crossing speed was comparatively higher, which was found to be 1.13 m/s for the first lane and slightly higher for the second lane. Pedestrians crossing in groups had lower crossing speed. This may be related to the fact that they feel safer when they move in groups and do not increase their speed because the oncoming vehicle will recognize them and brake. A significant difference between pedestrians carrying items and crossing without any items has been observed in Sirinyer, where a supermarket was located just behind the midblock. Most pedestrians (67 %) crossed with bags in their hands, leading to a deceleration in their crossing speeds. The average speed obtained for the observed midblocks is relatively lower compared with the previous findings. Bennett et al. (2001) estimated the average walking speed of 1.24 m/s in Melbourne. Demiroz et al. (2015) found a crossing speed of 1.21 m/s in Turkey.

Critical gaps were estimated for each lane with Raff's critical gap method. In Bornova, the critical gap was found to be 4.1 s. The critical gap was 6.2 s for the first lane and 5.7 s for the second lane in Sirinyer. Koh and Wong (2014) found 6.3 s for near-end crossing and 5.2 s for far end crossing, similar to the observed results in this study. Pedestrians arriving at the curbside start to look for available gaps. Zhang et al. (2018) indicated that the rolling gap is the predominant crossing strategy among pedestrians. The rolling gap is observed when a pedestrian crosses a street discontinuously and adjusts his/her speed to avoid conflict with the vehicle. The results showed that pedestrians accepted a lower gap in the second lane.

The posted speed limit at the study sites was 50 km/h. However, the maximum observed vehicle speed was 25 km/h (std.dev. is 4.30 km/h). In the literature, higher crossing speeds were observed, however considering the vehicle speeds in the study sites; it is assumed that pedestrians perceived it safe to cross, assessing a low risk for crash involvement. At both midblock crossings, the 15th percentile crossing speeds were similar, ranging between 0.78 m/s and 0.80 m/s.

Factors affecting gap acceptance were found by the MLR model. The results showed that the approaching vehicle's size and the existence of a physical barrier affected gap acceptance.

7. Conclusions

Vulnerable road users' safety is a primary concern for transport planners and policymakers. Drivers who do not yield the right of way to pedestrians at midblock crossings have to pay an extremely high administrative fine compared to other fines. Increasing fines without efficient control will not make a noticeable difference in the yielding behavior of drivers. Thus, control strategies should be implemented to make drivers obey the yielding the right of way rule.

The results of this study present pedestrian speed and the critical gap at midblock crossings using computer-vision techniques. Pedestrian counts, crossing speed, and gaps were extracted using image processing and compared with the real-world data. The used algorithms have given promising results. In this study, gender was recorded manually. Further improvements will be made to detect pedestrians' gender automatically.

Driver warning systems can be used to prevent any pedestrian-vehicle conflict where pedestrians accept small gaps. Cameras could be implemented to pedestrian crossings, and pedestrians' crossing intention can be estimated by image processing methods (Zhang et al., 2020). The information related to the pedestrian movement can be transferred to drivers via their cellphones. In order to make the best use of advanced driver assistance systems, pedestrian behaviors should be well studied and identified.

Author contributions

Conception and design of the Study: Yalcin Alver, Pelin Onelcin

Acqusition of Data: Yalcin Alver, Pelin Onelcin, Ahmet Cicekli

Sofware: Yalcin Alver, Pelin Önelcin, Ahmet Cicekli

Analysis and/or interpretation of data: Yalcin Alver, Pelin Onelcin, Mohamed Abdel-Aty.

Drafting the manuscript: Yalcin Alver, Pelin Onelcin, Mohamed Abdel-Aty.

Writing reviewing and editing: Yalcin Alver, Pelin Önelcin, Mohamed Abdel-Aty.

Declaration of Competing Interest

We have no conflicts of interest to disclose.

Acknowledgements

The first author would like to kindly acknowledge the support of The Scientific and Technological Research Council of Turkey (TUBITAK) through grant number 2219.

References

Alhajyaseen, W., Iryo-Asano, M., 2017. Studying critical pedestrian behavioral changes for the safety assessment at signalized crosswalks. Saf. Sci. 91, 351–360.
Alkahtani, K.F., Abdel-Aty, M., Lee, J., 2019. A zonal level safety investigation of pedestrian crashes in Riyadh, Saudi Arabia. Int. J. Sustain. Transp. 13 (4), 255–267.
Alver, Y., Onelcin, P., 2018. Gap acceptance of pedestrians at overpass locations. Transp. Res. Part F Traffic Psychol. Behav. 56, 436–443.
Atev, S., Arumugam, H., Masoud, O., Janardan, R., Papanikolopoulos, N.P., 2005. A vision-based approach to collision prediction at traffic intersections. Ieee Trans. Intell. Transp. Syst. 6 (4), 416–423.
Avinash, C., Jiten, S., Arkatkar, S., Gaurang, J., Manoranjan, P., 2019. Evaluation of pedestrian safety margin at mid-block crosswalks in India. Saf. Sci. 119, 188–198.
Bennett, S., Felton, A., Akçelik, R., 2001. Pedestrian movement characteristics at signalised intersections. 23rd Conference of Australian Institutes of Transport Research (CAITR 2001).
Bewley, A., Ge, Z., Ott, L., Ramos, F., Upcroft, B., 2016. Simple online and real-time tracking. In: Proceedings of the International Conference on Image Processing. ICIP, IEEE Computer Society, pp. 3464–3468.
Bishop, G., Welch, G., 2001. An introduction to the kalman filter. Proc of SIGGRAPH. Course 8 (27599-23175), 41.
Bolya, D., Zhou, C., Xiao, F., Lee, Y.J., 2019. Yolact: Real-time instance segmentation. Proceedings of the IEEE International Conference On Computer Vision 9157–9166.
Brunetti, A., Buongiorno, D., Trotta, G.F., Bevilacqua, V., 2018. Computer vision and deep learning techniques for pedestrian detection and tracking: a survey. Neurocomputing 300, 17–33.
Cai, Q., Abdel-Aty, M., Sun, Y., Lee, J., Yuan, J., 2019. Applying a deep learning approach for transportation safety planning by using high-resolution transportation and land use data. Transp. Res. Part A Policy Pract. 127, 71–85.
Demiroz, Y.I., Onelcin, P., Alver, Y., 2015. Illegal road crossing behavior of pedestrians at overpass locations: factors affecting gap acceptance, crossing times and overpass use. Accid. Anal. Prev. 80, 220–228.
Essa, M., Sayed, T., 2018. Traffic conflict models to evaluate the safety of signalized intersections at the cycle level. Transp. Res. Part C Emerg. Technol. 89, 289–302.
Farahi, F., Yazdi, H.S., 2020. Probabilistic Kalman filter for moving object tracking. Signal Process. Image Commun. 82, 115751.
Fu, T., Miranda-Moreno, L., Saunier, N., 2016. Pedestrian crosswalk safety at nonsignalized crossings during nighttime: use of thermal video data and surrogate safety measures. Transp. Res. Rec. 2586 (1), 90–99.
Fu, T., Hu, W., Miranda-Moreno, L., Saunier, N., 2019. Investigating secondary pedestrian-vehicle interactions at non-signalized intersections using vision-based trajectory data. Transp. Res. Part C Emerg. Technol. 105, 222–240.
Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Garcia-Rodriguez, J., 2017. A review on deep learning techniques applied to semantic segmentation. arXiv preprint arXiv 1704, 06857.
General Directorate of Highways, 2018. Summary of Traffic Accidents. Traffic Safety Division. http://www.kgm.gov.tr/SiteCollectionDocuments/KGMdocuments/Istatistikler/TrafikveUlasimBilgileri/17TrafikUlasimBilgileri.pdf.

Guo, H., Gao, Z., Yang, X., Jiang, X., 2011. Modeling pedestrian violation behavior at signalized crosswalks in China: a hazard-based duration approach. J. Traffic Injury Prevention 80, 220–228.
Guo, Y., Essa, M., Sayed, T., Haque, M.M., Washington, S., 2019. A comparison between simulated and field-measured conflicts for safety assessment of signalized intersections in Australia. Transp. Res. Part C Emerg. Technol. 101, 96–110.
Hamuda, E., Mc Ginley, B., Glavin, M., Jones, E., 2018. Improved image processing-based crop detection using Kalman filtering and the Hungarian algorithm. Comput. Electron. Agric. 148, 37–44.
Ismail, K., Sayed, T., Saunier, N., 2010. Automated analysis of pedestrian–vehicle conflicts: context for before-and-after studies. Transp. Res. Rec. 2198 (1), 52–64.
Kadali, B.R., Vedagiri, P., Rathi, N., 2015. Models for pedestrian gap acceptance behaviour analysis at unprotected mid-block crosswalks under mixed traffic conditions. Transp. Res. Part F Traffic Psychol. Behav. 32, 114–126.
Katz, M.H., 2011. Multivariable Analysis: a Practical Guide for Clinicians and Public Health Researchers. Cambridge university press.
Koh, P., Wong, Y., 2014. Gap acceptance of violators at signalized pedestrian crossings. Accid. Anal. Prev. 62, 178–185.
Noh, B., No, W., Lee, J., Lee, D., 2020. Vision-based potential pedestrian risk analysis on unsignalized crosswalk using data mining techniques. Appl. Sci. 10 (3), 1057.
Pathan, S.S., Al-Hamadi, A., Michaelis, B., 2009. Intelligent feature-guided multi-object tracking using kalman filter. 2nd International Conference on Computer, Control and Communication 1–6.
Pawar, D.S., Patil, G.R., 2016. Critical gap estimation for pedestrians at uncontrolled mid-block crossings on high-speed arterials. Saf. Sci. 86, 295–303.
Raff, M.S., Hart, J.W., 1950. A Volume Warrant for Urban Stop Signs. Eno Foundation for Highway Traffic Control., Connecticut.
Rahman, M.S., Abdel-Aty, M., Hasan, S., Cai, Q., 2019. Applying machine learning approaches to analyze the vulnerable road-users' crashes at statewide traffic analysis zones. J. Safety Res. 70, 275–288.
Ravishankar, K.V.R., Nair, P.M., 2018. Pedestrian risk analysis at uncontrolled midblock and unsignalised intersections. J. Traffic Transp. Eng. 5 (2), 137–147.
Redmon, J., Farhadi, A., 2017. YOLO9000: better, faster, stronger. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 7263–7271.
Redmon, J., Farhadi, A., 2018. Yolov3: an incremental improvement. arXiv preprint arXiv 1804, 02767.
Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 779–788.
Rosenbloom, T., 2009. Crossing at a red light: behaviour of individuals and groups. Transp. Res. Part F Traffic Psychol. Behav. 12 (5), 389–394.
Saunier, N., Sayed, T., 2007. Automated analysis of road safety with video data. Transp. Res. Rec. 2019 (1), 57–64.
Shaaban, K., Abdel-Warith, K., 2017. Agent-based modeling of pedestrian behavior at an unmarked midblock crossing. Procedia Comput. Sci. 109, 26–33.
Siddiqui, N.A., Chu, X., Guttenplan, M., 2006. Crossing locations, light conditions, and pedestrian injury severity. Transp. Res. Rec. 1982 (1), 141–149.
St-Aubin, P., Saunier, N., Miranda-Moreno, L., 2015. Large-scale automated proactive road safety analysis using video data. Transp. Res. Part C Emerg. Technol. 58, 363–379.
TSI, 2019. Turkish Statistical Institute. https://biruni.tuik.gov.tr/medas/?kn=89&locale=tr.
Wang, W., Guo, H., Gao, Z., Bubb, H., 2011. Individual differences of pedestrian behaviour in midblock crosswalk and intersection. Int. J. Crashworthiness 16 (1), 1–9.
WHO, 2013. Supporting a Decade of Action. World Bank, and World Health Organisation, pp. 1–318.
Wu, X., Sahoo, D., Hoi, S.C., 2020. Recent advances in deep learning for object detection. Neurocomputing 396, 39–64.
Yannis, G., Papadimitriou, E., Theofilatos, A., 2013. Pedestrian gap acceptance for mid-block street crossing. Transp. Plan. Technol. 36 (5), 450–462.
Zaki, M.H., Sayed, T., Tageldin, A., Hussein, M., 2013. Application of computer vision to diagnosis of pedestrian safety issues. Transp. Res. Rec. 2393 (1), 75–84.
Zhang, C., Zhou, B., Qiu, T.Z., Liu, S., 2018. Pedestrian crossing behaviors at uncontrolled multi-lane mid-block crosswalks in developing world. J. Safety Res. 64, 145–154.
Zhang, S., Abdel-Aty, M., Yuan, J., Li, P., 2020. Prediction of pedestrian crossing intentions at intersections based on long short-term memory recurrent neural network. Transp. Res. Rec., 0361198120912422
Zhuang, X., Wu, C., 2012. The safety margin and perceived safety of pedestrians at unmarked roadway. Transp. Res. Part F Traffic Psychol. Behav. 15 (2), 119–131.