TRANSPORTES
revistatransportes.org.br
anpet
associação nacional de pesquisa e ensino em transportes
anpet.org.br

Customization and application of an automated data collection tool of pedestrian crossing at signalized intersections

Customização e aplicação de ferramenta para coleta automatizada de dados de travessia de pedestres em interseções semaforizadas

Juliana de Abreu e Tréz[1], **Cornélio Albuquerque de Sousa**[1], **Alessandro Macêdo de Araújo**[1], **Manoel Mendonça de Castro Neto**[1]

[1]Universidade Federal do Ceará, Fortaleza, Ceará, Brasil

Contact: julianatrez@det.ufc.br, (JAT); cornelio.sousa@det.ufc.br, (CAS); alessandro.araujo@det.ufc.br, (AMA); manoel@det.ufc.br, (MMCN)

Submitted:
13 October, 2023
Revised:
18 January, 2024
Accepted for publication:
29 June, 2024
Published:
15 October, 2024
Associate Editor:
Sara Ferreira, Universidade do Porto, Portugal

Keywords:
YOLO.
Computer vision.
Signalized intersections.
Active transport.

Palavras-chave:
YOLO.
Visão computacional.
Interseções semaforizadas.
Transporte ativo.

DOI: 10.58922/transportes.v32i3.2961

OPEN ACCESS
CC BY

ABSTRACT

Pedestrian crossing during vehicular green time is a problem that still requires better understanding and investigation given the complexity of the variables involved and their interrelationships. Automated collection tools can be important for observing and for analyzing these variables and interrelationships. The main objective of this study is to customize and apply an automated tool to collect data of important variables in studies of pedestrian crossings at signalized intersections, such as vehicle headways, pedestrian delays, vehicle speeds, vehicle types, and crossing times, per lane. The tool, applied to a video of a signalized intersection in Fortaleza, consisted of the YOLOv7 and StrongSORT algorithms. The tool training mAP was approximately 90%. In total, 9427 vehicles and 723 pedestrians were tracked; the headways showed great amplitude, the average speed of vehicles was 28 km/h, and the average delay for pedestrians was 18 s. Validation with a collection tool (RUBA) showed that there were no significant differences in the two methods regarding the vehicle passage times and headways. For vehicle speeds, the differences were circa ± 6 km/h, and for the pedestrian variables, the mean of differences were up to 0.2 sec.

RESUMO

A travessia de pedestres durante o verde veicular é um problema que ainda necessita de maior compreensão e investigação, visto a complexidade das variáveis envolvidas e suas inter-relações. Ferramentas de coleta automatizada podem ser importantes aliadas na obtenção dessas variáveis e análise de suas inter-relações. O objetivo principal deste estudo é customizar e aplicar uma ferramenta automatizada para coletar variáveis importantes em estudos de travessias de pedestres em interseções semaforizadas, sendo estas os *headways* veiculares, os atrasos dos pedestres, as velocidades veiculares, os tipos de veículo e os instantes de travessia, por faixa. A ferramenta, aplicada em um vídeo de uma interseção semaforizada de Fortaleza, consistiu nas ferramentas YOLOv7 e StrongSORT. O mAP de treinamento da ferramenta foi de quase 90%. Ao todo, 9427 veículos e 723 pedestres foram rastreados; os *headways* mostraram grande amplitude, a velocidade média dos veículos foi de 28 km/h e o atraso médio dos pedestres foi de 18 seg. A validação com uma ferramenta de coleta (RUBA) apontou que não houve diferenças significativas nas coletas pelos dois métodos quanto aos instantes de passagem dos veículos e de seus *headways*; para as velocidades veiculares as diferenças foram entre ± 6 km/h, e para as variáveis dos pedestres, as médias das diferenças foram de até 0,2 seg.

1. INTRODUCTION

Pedestrian crossings on red, in other words, on green for vehicles, regardless of the existence of traffic lights for pedestrians, are a frequent and serious road safety problem (Onelcin and Alver, 2015). Studies have been conducted to understand the variables that affect pedestrians' acceptance of vehicular headways on red, including pedestrian delay, speed, and vehicle type (Zhu et al., 2021; Raoniar and Maurya, 2022; Nikolaou et al., 2023).

Collecting these variables manually can be very time-consuming and expensive. Therefore, the use of automated tools for data collection is relevant, as they offer faster and less laborious results (Noh et al., 2022). Thus, the main objective of this paper is to customize and to apply an automated tool to collect variables related to pedestrian crossings at intersections, such as pedestrian delay, vehicle speed, type, and headway per lane, from video images from monitoring cameras.

This automated tool can help planning the operation of signalized intersections in urban cities that have video surveillance cameras, as it allows, for example, the signal control to be optimized focused on the safety and quality of service of pedestrian movements, which is generally not given much priority (Marisamynathan and Vedagiri, 2019).

2. LITERATURE REVIEW

2.1. Acceptance of headways by pedestrians

In the pedestrian crossing literature, there are different concepts about vehicular headway. The most common is the time interval between the arrival of the pedestrian, i.e., the beginning of the wait, and the passage of the approaching vehicle (Chandra et al., 2014; Marisamynathan and Vedagiri, 2019). A second type considers the accepted headway as the time difference between the vehicle that approached the conflict area immediately before the pedestrian crossing and the vehicle that approached the conflict area immediately after (Koh and Wong, 2014). In this study, a third concept was adopted, known as *rolling-gap:* the accepted headway, which is obtained per lane, is the time interval between the moment the pedestrian reaches the center of the crossing lane and the passage of the approaching vehicle on that section.

2.2. Factors that influence pedestrian crossing on red

One of the key factors that can influence crossing on red is pedestrian delay, which is defined as the time difference between the moment when the pedestrian arrives at the intersection to cross it and when he/she starts crossing (Zhu et al., 2021). In studies that address this factor as relevant to red crossings (Dommes et al., 2015; Song et al., 2019, Afshari et al., 2021; Raoniar and Maurya, 2022), it was noted that the longer the pedestrians wait on red, the greater their tendency to look for vehicular headways. The work of Raoniar and Maurya (2022), in India, highlights that the probability of crossing on red for pedestrians who waited more than 40 s increased by 21.4% compared to those who waited until 20 s.

Another relevant factor found at crossings on red studies was the vehicular flow, which, the higher it is, the less likely pedestrians are to accept headways (Dommes et al., 2015; Ma et al., 2020). Ma et al. (2020), in a study at 3 signalized intersections in China, concluded that 40% of pedestrians who crossed on red did so when the volume of vehicles was moderate (between 250 and

550 pcu/h/lane), and when it was high (above 500 pcu/h/lane) the proportion of crossings on red light decreased to 30%.

The vehicle type can also affect pedestrians' crossing decisions, according to Zhu et al. (2021), who evaluated 6 signalized intersections in Hong Kong, and inferred, by a logistic regression model that, an increase of 1% in the volume of heavy vehicles reduced the probability of crossing on red became by 0.48%, as the chance of serious injuries is greater when there are conflicts involving heavy vehicles (buses and trucks).

Finally, vehicle speed was found significant in the studies by Onelcin and Alver (2017), Mukherjee and Mitra (2022), and Nikolaou et al. (2023). Mukherjee and Mitra (2022) examined 55 signalized intersections in India using historical data, applying questionnaires and observing crossings, and the results showed that a 10 km/h increase in vehicle speed reduced the chances of crossings on red by 10%.

2.3. Use of computer vision in studies of pedestrian crossings

Computer vision, if well used, can be very helpful in pedestrian crossing analysis, as it allows for quick and accurate data collection of the necessary variables, such as the headway accepted by the pedestrian, their walking speed and delay (Sayed et al., 2016; Castro Jr. et al., 2023), providing possibilities for a variety of studies on pedestrian behavior.

Such variables can be estimated from pedestrian and vehicle trajectories data. However, to collect this type of data, it is necessary to use auxiliary tools to automate computer routines, given the substantial number of observations to be made at each moment. One of the areas of application of computer vision is object tracking, which seeks automated methods for extracting the trajectories of objects moving through an environment. For the specific case of tracking objects in videos, the detection tracking approach is one of the most common (Luo et al., 2022), which employs object detection models and tracking algorithms together.

The detection models are trained to locate and classify objects in an environment. *You Only Look Once (YOLO)* (Redmon et al., 2016) is one of the object detection models most found in the literature, given its simplicity, efficiency, computational performance and easy customization. Over the years, several improvements and modifications have been implemented, with *YOLOv8* being one of the most recent versions (Ultralytics, 2023). Objects are commonly located using bounding boxes that frame the "body" of the object. For each bounding box found in an image, the model returns the coordinates and dimensions of the box, the class/type of the framed object and the degree of confidence that it really is an object.

Detection tracking algorithms, in turn, estimate entity trajectories based on the data provided by the detection models. An example of this type of algorithm is *Simple Online and Realtime Tracking (SORT)*, which stands out for its simplicity and computational efficiency (Bewley et al., 2016). Being an open-source and easily expandable algorithm, *SORT* has served as the basis for several modifications and improvements over time (Wojke et al., 2017; Du et al., 2023). According to Bewley et al. (2016), the *SORT* algorithm receives object detection data referring to the current frame and attempts to associate them with existing numerical identifiers (identities (ID), each corresponding to a trajectory). If a detection is not associated with an active ID, the algorithm creates a new ID and associates it with the detection (thus starting a new trajectory). If a trajectory is not associated with a detection for a specific period, it is deactivated, indicating that the object has left the scene.

To perform the association between detections and trajectories, the algorithm first estimates the future position of the tracked objects based on their locations in previous frames. The model then uses these estimates to calculate the proximity between detections and trajectories, optimizing the associations. *DeepSORT* (Wojke et al., 2017) is an extension of the *SORT* algorithm that seeks to improve object tracking after occlusion situations using metrics that quantify the appearance of objects. The *StrongSORT* algorithm (Du et al., 2023), in turn, implements a series of modifications to the *DeepSORT* algorithm, such as replacing the metric for calculating the distance between detections and trajectory and replacing the model that quantifies the appearance of objects. Finally, the result of these tracking algorithms is summarized in a tabular file, where each line corresponds to a bounding box that locates and classifies an object, a number that represents the ID of the located object, and the frame number associated with the detection. In this manner, a set of lines with the same ID number forms the trajectory of an entity.

The most common metrics for evaluating the localization and classification of objects in images are Intersection Over Union (IoU) , Precision, Recall, Confusion Matrix, Average Precision (AP) and Mean Average Precision (mAP). The IoU seeks to quantify the overlap between two regions (bounding boxes, in the case of object detection), and is calculated as the ratio between the intersection area and the union area of the regions involved (Equation 1). Precision and Recall are normally calculated by determining the true positives (TP), false positives (FP), and false negatives (FN) (Equations 2 and 3). However, the values of these three indicators depend on the choice of the correspondence threshold, which determines whether a predicted detection corresponds to a real object in the image. This threshold is usually based on the IoU metric, where, for example, only detections with an IoU above 50% in relation to a real object are considered true positives (or false positives, if the detector misses the class of the real object).

$$IoU = \frac{AI}{AU} \tag{1}$$

where: *AI:* area of intersection between two bounding boxes; and *AU:* area of union between two bounding boxes

$$Precision = \frac{VP}{VP + FP} \tag{2}$$

$$Recall = \frac{VP}{VP + FN} \tag{3}$$

where: *VP:* true positives; *FP:* false positives; and *FN:* false negatives

The Confusion Matrix, in turn, is a visual representation of the TP, FP and FN values, extremely useful for making considerations about the errors and successes of the detection system. Another threshold that influences the metrics of an object tracking system is the confidence threshold, which determines which predictions should be considered detections, based on the confidence level value. A common confidence threshold of 50% is used, where only detections with a confidence level above 50% are considered. Choosing a high confidence threshold implies a higher precision value but reduces the recall of the detection system. The AP metric seeks to quantify how much of the precision is influenced by this variation in the confidence threshold, being equivalent to the

area under the precision versus recall curve. The value of the AP metric ranges from 0 to 1, the higher the value the better (Padilla et al., 2020). AP values are commonly calculated for each object class and vary depending on the chosen IoU threshold. Finally, the mAP metric is the average of all the AP values for different IoU values and/or for different object classes and can be considered as the overall performance of the detector (Al-qaness et al., 2021).

Training the detection and tracking algorithms is important to achieve better results. It is necessary to separate the images into training (used for model learning), validation (for evaluation and improvement of parameters throughout training) and test groups (for final evaluation of model performance), in a stratified random manner, in a ratio that is typically 70:15:15 (Genc and Tunc, 2019). Another important aspect, highlighted by Al-qaness et al. (2021), is the retraining of models with more generic data (transfer learning), as it reduces training time and increases the accuracy of the final model.

3. METHOD

The method proposed in this work is divided into 5 steps: 1) sites selection; 2) customization of the automated tool; 3) tool application: collection of trajectories and variables of interest; 4) tool validation; and 5) exploratory data analysis of variables related to the headway acceptance.

3.1. Sites selection

The selection of the intersections was based on three characteristics that enabled the observation of variables relevant to the acceptance of headways on red: (i) The roads must have two or three lanes in one direction, due to their typical nature in the city of Fortaleza; (ii) enough vehicular flows, to enable the formation of acceptable headways by pedestrians; and (iii) presence of camera positioned in a way that allows observation of all vehicles passing over the stop bar, pedestrians waiting to cross on both sidewalks, and the crossings. The videos were provided by Fortaleza Advanced Traffic Control (CTAFOR).

3.2. Customization of the automated tool

The *YOLOv7* object detection model (Wang *et al.*, 2022) and the *StrongSORT* tracking algorithm (Du et al., 2023) were used as the basis for the trajectory extraction tool. Both were chosen due to their simplicity of use and customization, as well as their computational efficiency. A pre-trained *YOLOv7* model was retrained using a database of images extracted from videos of signalized intersections in Fortaleza-CE provided by CTAFOR. Objects referring to the classes' "bicycle", "bus", "car", "motorcycle", "pedestrian" and "truck" were manually labeled. The labeled images were separated into training, validation, and test groups. The mAP@0.5 metric (mAP metric using a fixed 50% matching threshold based on the IoU metric) was used to choose the best set of parameters during training from the validation group data. Post-training evaluation of the detection model consisted of analyzing the mAP@0.5 metric and the Confusion Matrix for the test group data. The tracking algorithm, in turn, was subjected to a trial-and-error calibration process, adjusting the parameters that control the model's operation. The final combination of parameters was chosen based on visual analysis of the trajectory results for the various experiments conducted on a sample video of the intersection chosen for analysis.

3.3. Application of the tool: collection of trajectories and variables of interest

After training and evaluation, the automated tool was applied to video, generating a file with the trajectories of all the objects tracked throughout each frame. This file was processed in computer codes to extract the volume of vehicles, by lane and type, the speed of the vehicles, the volume of pedestrians crossing, by area of origin, their delays, and the accepted headways. All these data were collected using trajectories in image coordinates based on the known field distance (4 m) between sections 1 and 2 of Figure 1.

The areas and sections of interest to pedestrians and vehicles for recording time instants to calculate the parameters were delimited as shown in Figure 1. In this way, the conflict area between pedestrians and vehicles (crosswalk), the reference for vehicle passage times (section 1), the references for the collection of vehicle speeds (section 1 and section 2), the numerical order considered for the vehicle lanes, the reference areas in each lane for the collection of pedestrian crossing times in the process of crossing (yellow rectangles), and the waiting areas for pedestrians on the sidewalks on both sides of the road (origins A and B) were determined.

With the trajectories of pedestrians and vehicles extracted, the variables of interest for the work were obtained: volumes, accepted headways, delays and speeds.

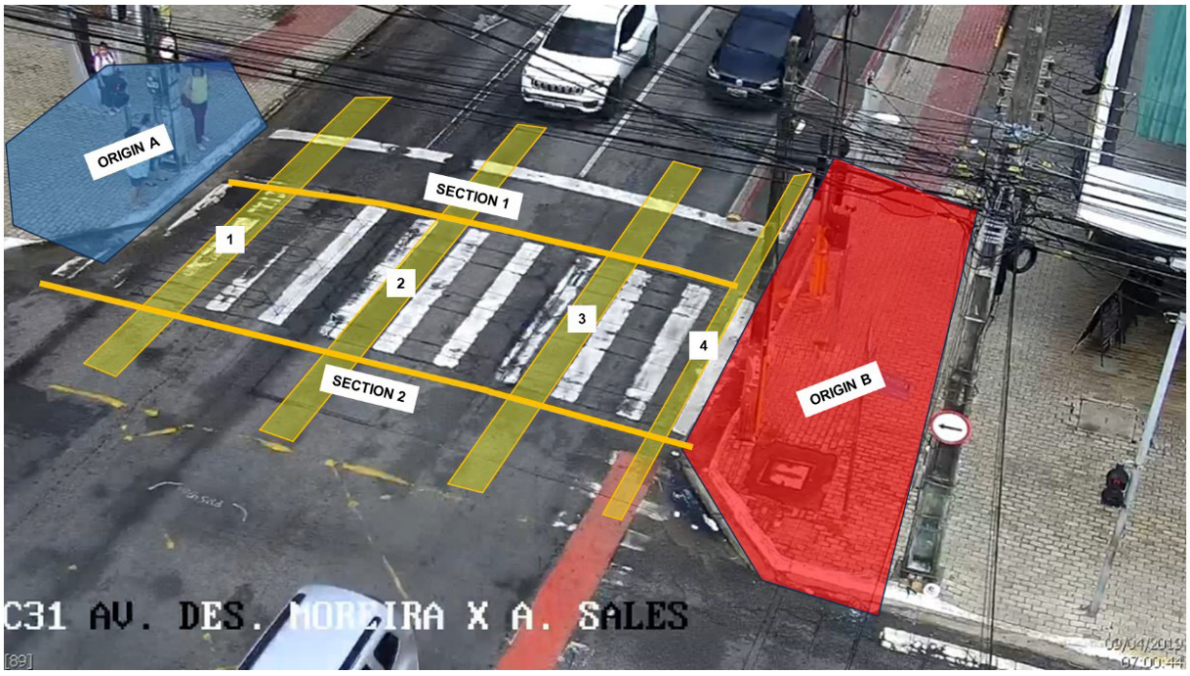


Figure 1. References adopted to record moments of pedestrians and vehicles.

3.4. Validation of the tool

To verify the detection quality of the computer vision tool in this study (hereinafter referred to as *YOLO*, for the sake of simplicity), a manual collection was conducted in the first 10 minutes of video to validate the vehicle and pedestrian volume results (by type and lane), vehicular headways, accepted headways, vehicular speeds, and pedestrian delays obtained by automated collection within the same time interval. The *Ground Truth Annotator* tool from *RUBA* (*Road User Behavior Analysis)* software was used (Tonning et al., 2017). As in *RUBA* the moments of passage are collected with the video reproduced in slow motion (10 frames per second), and each frame can be triggered

by click by the user, it can be assumed that the errors using this method are small compared to others (observation of videos in real time, traffic sensors, etc.). Therefore, we considered these data as reference (ground truth) to validate the tool developed in this paper.

The instants vehicles passed through the 2 sections were manually marked, to calculate the headways and vehicle speeds, as well as the instants at which pedestrians entered and left the origins, to calculate the delay. The comparison was made by creating scatter plots of the errors between the two methods (*YOLO – RUBA*) for the passage times through the sections, headways, speeds, and delays, and by analyzing the statistical confidence intervals of the error means of the start waiting times, end waiting times, and pedestrians' delay.

3.5. Exploratory analysis of variables related to the headway acceptance

With the results of applying the automated tool, it was possible to collect the headways accepted by pedestrians during red times. For each crossing, 4 accepted headways were obtained, one per lane, as illustrated in Figure 2, in which the pedestrian is represented by a red rectangle along the crossing, in each lane, until the opposite sidewalk. As shown in the figure, the accepted headways were computed as the time between the pedestrian passes the center of the traffic lane and the vehicle passes through the same point.

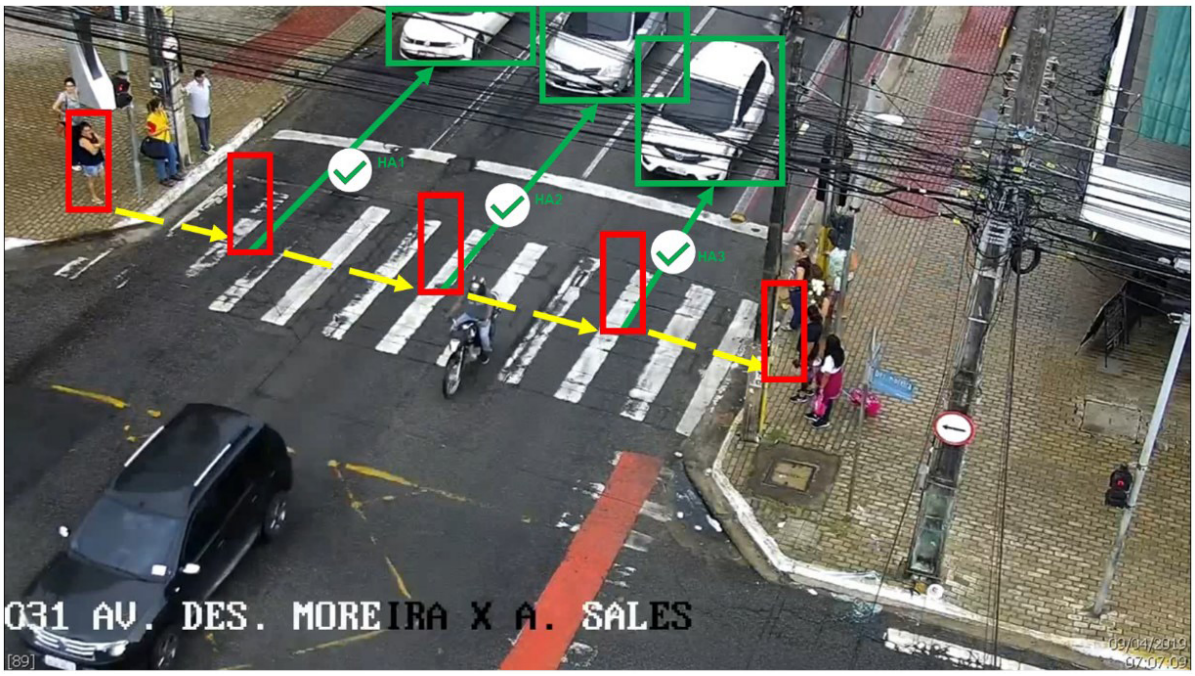


Figure 2. Illustrative example of headways acceptance in each lane.

In the analysis of crossing on red, only accepted headways lower than 10 sec were considered, because we assumed that only such headways cause assessment of headway acceptance by pedestrians. For the exploratory analysis of the variables, we used scatter plots that related the headways accepted by the pedestrian with their delays and the with speeds of the vehicles considering these accepted headways.

4. RESULTS

4.1. Selection of the intersections

After analyzing the videos made available by CTAFOR, we selected one for the application of the tool, which is the approach of Antônio Sales Avenue, at the intersection with Desembargador Moreira Avenue, in Fortaleza. This video was chosen because of its desirable characteristics.

4.2. Customization of the automated tool

The number of images labeled for training were 1,516, of which 1,140 (75.2%) were for training, 185 (12.2%) for validation and 191 (12.6%) for test, values close to the 70%/15%/ 15% suggested by Genc and Tunc (2019). These training images came from 6 signalized intersections in Fortaleza, including the intersection that is the subject of this study. Training was performed in 1,250 iterations, reaching a training/validation mAP of 94.4%, precision of 95.2%, and recall of 92.2%. The evaluation of the tool with test images resulted in a mAP test of 89.4%. However, there was a low AP for trucks (75.7%) due to the incidence of false positives, which suggests future studies to retrain the model with new images of this class to increase the tool's accuracy. According to the Confusion Matrix in Figure 3, which compares the success rates between the manual classification and the model's prediction, all classes had success rates greater than 85%, except for trucks, which reached 70%.

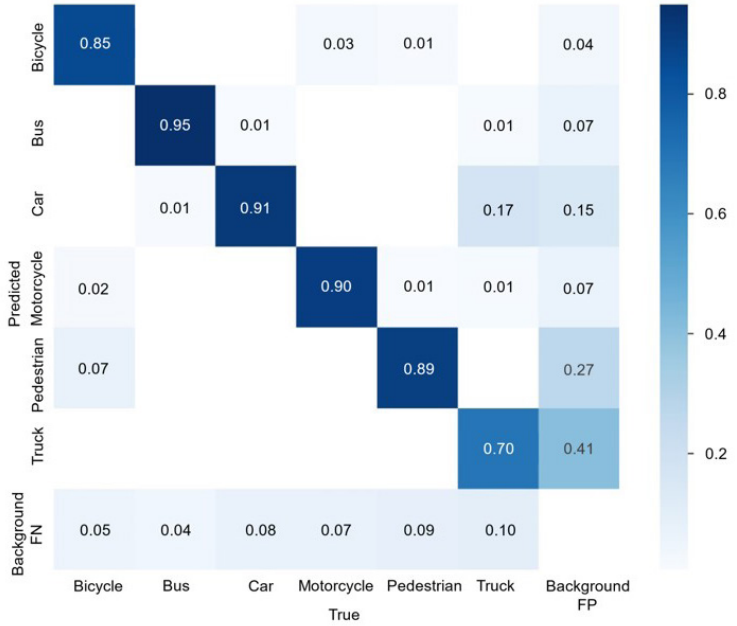


Figure 3. Confusion Matrix of objects in the test images.

4.3. Application of the tool: collection of trajectories and variables of interest

In total, 9,427 vehicles and 723 pedestrians were tracked. The vehicle composition included 366 bicycles, 125 buses, 7,044 cars, 1,767 motorcycles, and 125 trucks. The headways showed high values, reaching almost 600 s, due to the low flow of bicycles and buses in their respective exclusive lanes, in some periods of the filming, but practically all of them were under 100 s. The average speed of the vehicles at the intersection was 28 km/h, with bicycles showing the lowest average speed (13 km/h). The average pedestrian delay was 18 s.

4.4. Validation of the tool

4.4.1. Validation of object detection

Regarding the comparison between the tools (*YOLO* and *RUBA*), the manual counting of objects using *RUBA* in the initial 10 minutes of the video approaching Antônio Sales Avenue resulted in 343 vehicles and 37 pedestrians, while using *YOLO*, in the same period, 337 vehicles and 32 pedestrians were recorded, that is, 1.7% and 13.5% less than in *RUBA*. This difference between counts occurred because of detection problems in *YOLO* such as to noise in the images, occlusions due to poles, signs and wires, crowds of vehicles at closed traffic lights and pedestrians in crossing areas, etc.

4.4.2. Validation of vehicle crossing times and headways

Figure 4 shows the scatter plots of the differences (errors) between the passage instants in section 1 (Figure 4a) and between the instants in section 2 (Figure 4b) collected by the two methods (in s), considering their relationship with the speeds collected in *RUBA.* The results of the two graphs show that, for all classes, most of the errors are between -0.1 s and +0.1 s. Isolated points with discrepant values can also be noted in the two graphs, but these differences reach a maximum value of -0.5 s.

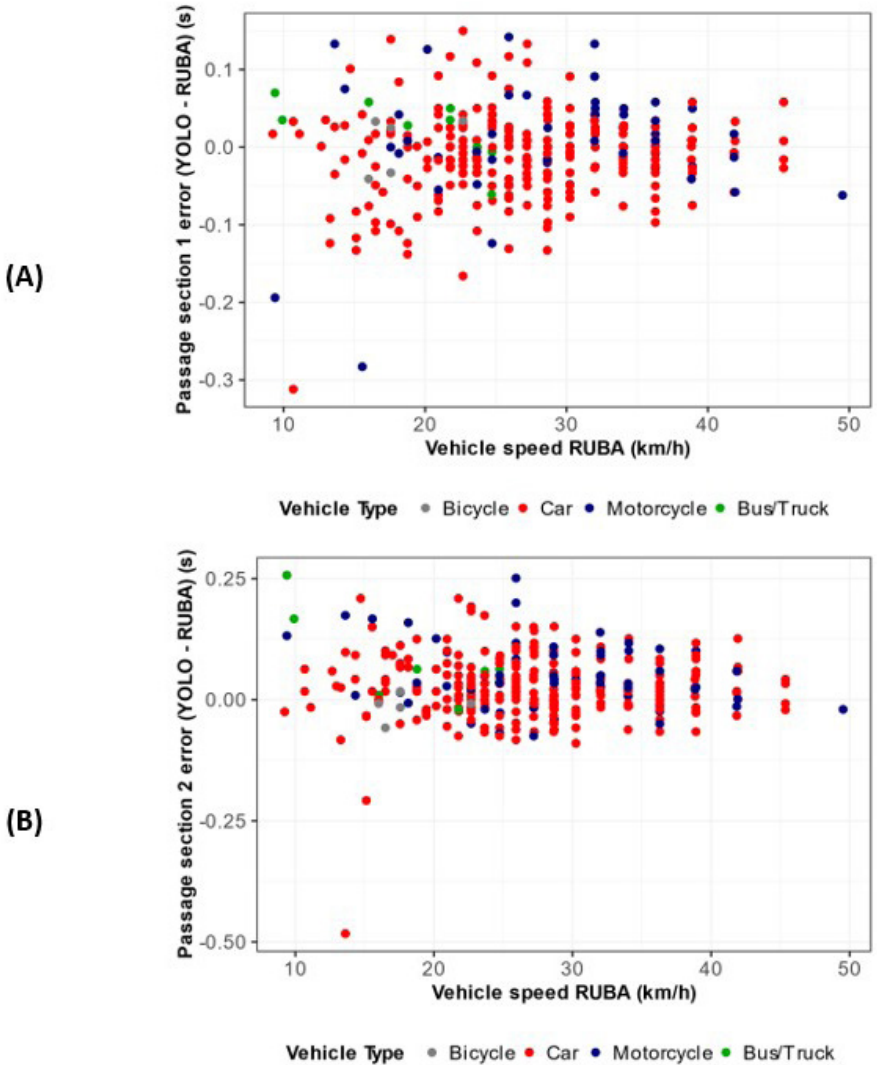


Figure 4. Errors between the instants of passage in section 1 (A) and between the instants of passage in section 2 (B) (in s) vs speeds collected in *RUBA* (in km/h).

The dispersion of errors between headways (in sec) considering their relationship with the speeds collected by *RUBA*, illustrated in Figure 5, shows that most values are concentrated between -0.1 and +0.1 s, with few observations reaching -0.2 s and +0.2 s. This is probably due to the differences in the collection of arrival times at the beginning of the stop bar, as the vehicle headway is calculated based on this parameter. An example is the car with a speed of 10 km/h and an error of -0.25 s, which appears with the largest error (-0.3 s) in the graph of arrival instant errors in section 1, in Figure 4a.

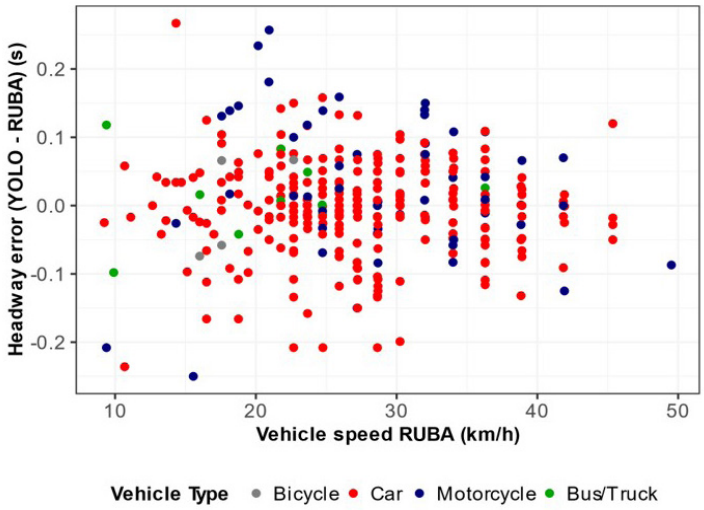


Figure 5. Headways errors (in s) vs speeds collected in *RUBA* (in km/h).

4.4.3. Validation of vehicle speeds

In the scatter plot of errors between speeds (in km/h) considering their relationship with the speeds collected by *RUBA,* in Figure 6, it is suggested that the dispersion of errors, which vary between ±6 km/h, increases with speed, which can be investigated in future work. The errors shown in the graph can be considered low because they correspond to approximately ±1.6 m/s difference.

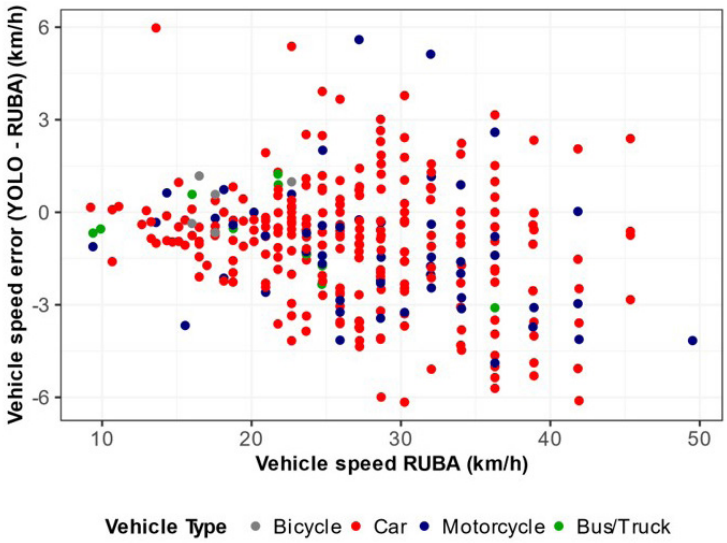


Figure 6. Speed errors vs speeds collected in *RUBA* (in km/h).

4.4.4. Validation of pedestrian delays

Figure 7 shows the pedestrian delay errors in ascending order. It is possible to observe that the lowest error was around -0.6 s and the highest was around +0.3 s.

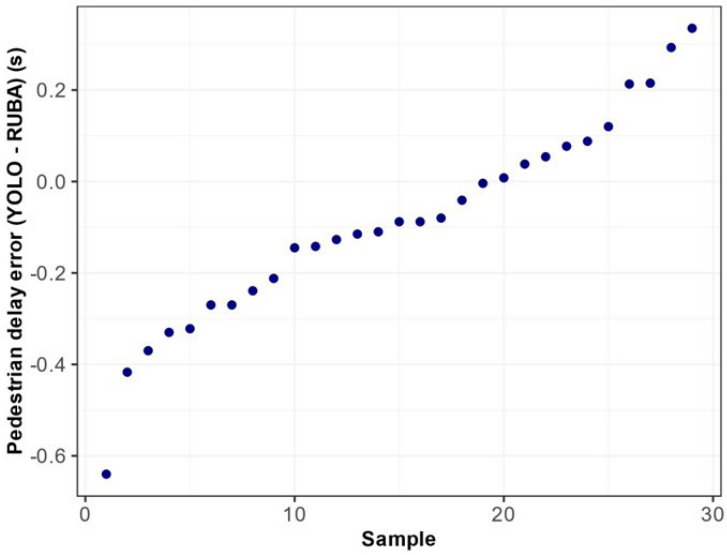


Figure 7. Pedestrian delay errors (in s).

The 95% t-Student confidence intervals for the mean errors are shown in Table 1, and we considered the average errors small. It is important to highlight that, although 37 pedestrians were identified by *RUBA* and 32 by *YOLO,* 29 pedestrians who appeared correctly in both collections were paired. It was not possible to pair the rest due to tracking problems caused by occlusions in the image and grouping of pedestrians in the waiting and crossing areas, for example, keeping the tool from obtaining waiting times and delays. Therefore, the average of the differences was calculated with the 29 pedestrians paired.

Table 1: CI (95%) of the mean differences for pedestrians (in s).

	N	Mean (s)	Standard Deviation (s)	CI (95%) (s)
Delay errors	29	–0.1	0.2	[–0.2; 0.0]
Start of wait errors	29	0.1	0.2	[0.0; 0.0]
End of wait errors	29	0.0	0.1	[–0.1; 0.0]

4.5. Exploratory analysis of variables related to the headway acceptance

From the video, two durations of the traffic signal cycles were identified, one in the first 3 h with a pedestrian red time of 70 s, and the other in the last 2 h with a pedestrian red time of 80 s, both with flashing red of 7 s. 45 observations of accepted headways lower than 10 s in red were collected, referring to 36 pedestrians: 9 in lane 1, 12 in lane 2, 7 in lane 3 and 17 in lane 4. As only accepted headways lower than 10 s were considered for the analysis, the amounts are different across lanes.

The relationships between the headways accepted by pedestrians in red and the speeds of vehicles corresponding to these headways by vehicle type, vehicle lane, and area of origin of the pedestrian are shown in Figure 8. The motorcycle with speed of 0 km/h and accepted headway of

5 s (Lane 1 Origin B) had its speed set as -1 due to tracking problems. The highest speeds of vehicles whose headways were accepted in red occurred in lanes 2 and 3, with emphasis on pedestrians from origin A who accepted headways of less than 5 s from motorcycles that were at speeds of 60 km/h, and there was also acceptance of *headways* of cars at speeds approaching 40 km/h.

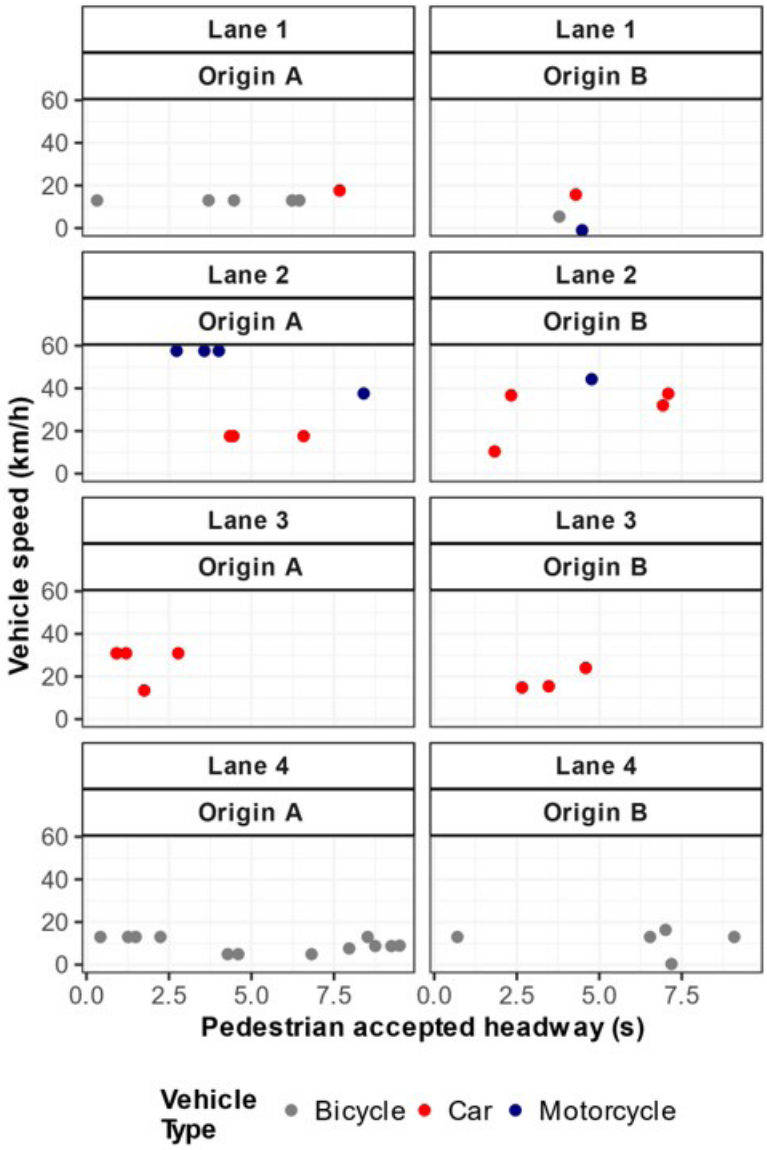


Figure 8. Relationship between accepted headways in red (in s) and vehicle speeds (in km/h), by vehicle type, vehicle lane and pedestrian origin zone.

When exploring the relationship between headways accepted on red by pedestrians and their delay, in Figure 9, there are more headways accepted by pedestrians from origin A in the 4 lanes, but we observed no pattern amongst them. In lane 1, for example, there are different waiting times, reaching 60 s, and there is a predominance of acceptance of headways by pedestrians from origin A when the vehicles were bicycles, while those from origin B show a balance in vehicle type. In lane 2, pedestrians from origin B accepted headways from cars and motorcycles after waiting for up to 20 s, while those from origin A waited longer, and, in lane 3, pedestrians from both origins waited for around 20 s, but

those from origin A accepted headways of less than 2.5 s, while those from origin B accepted headways between 2.5 s and 5 s. Finally, in lane 4, there was a divergence in the accepted headways of pedestrians from origin A, but those from origin B waited 20 s or less and accepted headways of 7.5 s or longer.

It should be noted that the focus of this work is on the data collection of the variables of interest, not on making inferences on the relationships among these variables. The collected data provided enough samples to illustrate the tool's potential. Larger samples are required to model and understand the interrelationships among the variables.

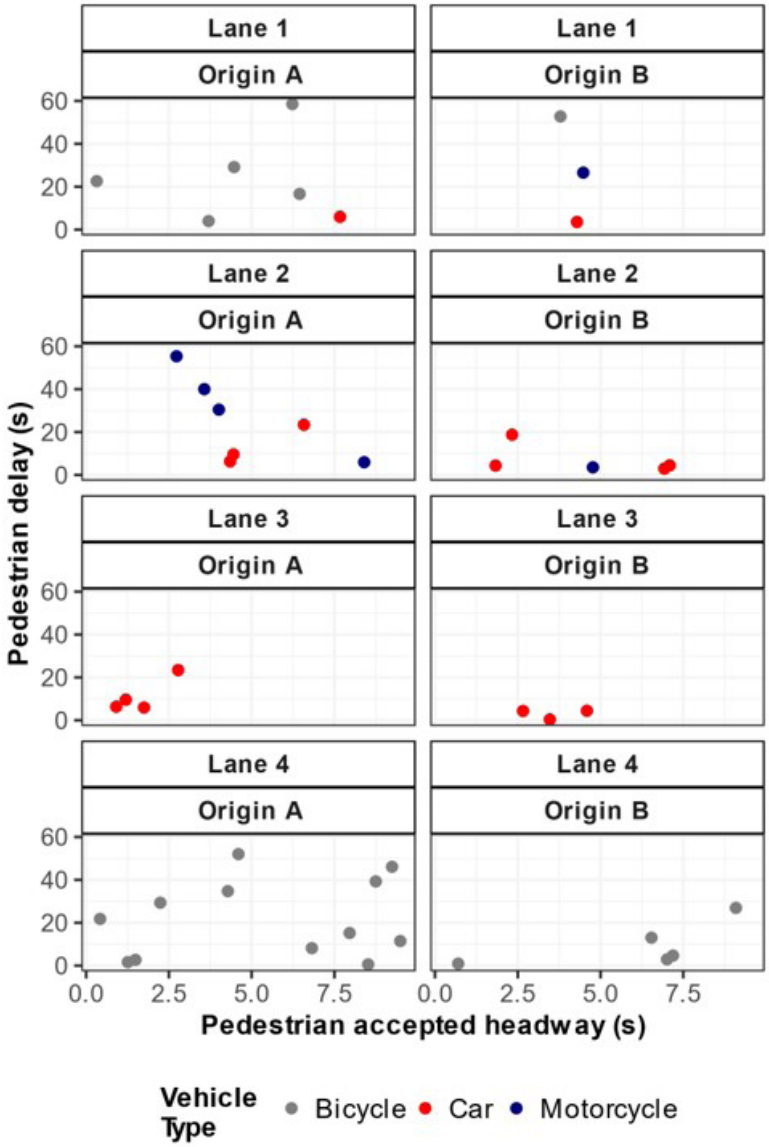


Figure 9. Relationship between accepted headways in red and pedestrian delays (in s), by vehicle type, vehicle lane and pedestrian origin zone.

5. CONCLUSIONS AND RECOMMENDATIONS

This work aimed to customize and apply an automated tool to collect important variables in studies of pedestrian crossings at signalized intersections. The selected variables were vehicle headways

per lane and type of vehicle, pedestrian delay, and vehicle speed. To collect these variables, we used the algorithms *YOLOv7* (Wang *et al.*, 2022) for detection and *StrongSORT* (Du et al., 2023) for tracking.

Training the tool using labeled images from other videos of signalized intersections in Fortaleza resulted in a validation mAP of 94.4% and a test mAP of almost 89.4%. However, there was an incidence of false positives for trucks, which had precision of 51% and AP of 76%. It is suggested to retrain the model with new images of this class, which would increase the tool's accuracy. Based on these results, we considered that the tool has been well trained and is, in general, calibrated for use.

Validation of the automated tool with *RUBA* (*Road User Behavior Analysis)* showed that errors in collecting passage times, headways and delays between the two methods were low. Speed errors seem to increase with increasing speed, which can be investigated in future work with statistical tests for homoscedasticity.

The last step stage was the exploratory analysis of the variables related to the acceptance of headways lower than 10 sec in red for pedestrians. The relationship among headways accepted on red, vehicle speeds and pedestrian delays shows that, in general, pedestrians from origin A waited longer to cross the road and accepted more headways of smaller sizes and of vehicles at higher speeds than those of origin B. Nevertheless, it is important to highlight the merely exploratory nature of the analyses, which means that larger samples need to be collected and more analyses conducted to reach robust conclusions.

An important limitation of this work was the sample sizes, making it impossible to statistically estimate the relationships among the variables explored here, such as the effect of pedestrian delay on the acceptance of headways, as only 36 pedestrians evaluated the headways and crossed in the red during the 5 hours of filming. It is therefore recommended to apply the tool to a new set of videos of the intersection approach examined in this study, to increase the sample size and enable more in-depth analysis. The tool could also be applied to videos of other signalized intersections with significant flow of pedestrians, to model their behavior in other situations.

Finally, we suggested the development, calibration and validation of statistical regression models, considering other variables, such as the arrival pattern of vehicles at the intersection, to predict the relationship between headways accepted on red by pedestrians and their delay at signalized intersections in Fortaleza.

ACKNOWLEDGEMENTS

This work was developed with the support of the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES). The authors are also grateful to Municipal Transit and Citizenship Authority (AMC) of Fortaleza for the images.

REFERENCES

Afshari, A.; E. Ayati and M. Barakchi (2021) Evaluating the effects of external factors on pedestrian violations at signalized intersections (a case study of Mashhad, Iran). *IATSS Research*, v. 45, n. 2, p. 234-240. DOI: 10.1016/j.iatssr.2020.10.004.

Al-qaness, M.A.A.; A.A. Abbasi; H. Fan et al. (2021) An improved YOLO-based road traffic monitoring system. *Computing*, v. 103, n. 2, p. 211-230. DOI: 10.1007/s00607-020-00869-8.

Bewley, A.; Z. Ge; L. Ott et al. (2016). Simple online and realtime tracking. In *2016 IEEE International Conference on Image Processing (ICIP)*. New York: IEEE, p. 3464-3468. DOI: 10.1109/ICIP.2016.7533003.

Bochkovskiy, A.; C.-Y. Wang and H.-Y.M. Liao (2020). YOLOv4: optimal speed and accuracy of object detection. *ArXiv*. DOI: 10.48550/arXiv.2004.10934.

Castro Jr., F.A.B.; M.M. Castro-Neto and F.J.C. Cunto (2023) Análise do atraso e da brecha aceita dos pedestres em travessias semaforizadas: um estudo na cidade de Fortaleza utilizando técnicas de visão computacional baseadas em deep learning. *Transportes*, v. 31, n. 2, e2845. DOI: 10.58922/transportes.v31i2.2845.

Chandra, S.; R. Rastogi and V.R. Das (2014) Descriptive and parametric analysis of pedestrian gap acceptance in mixed traffic conditions. *KSCE Journal of Civil Engineering*, v. 18, n. 1, p. 284-293. DOI: 10.1007/s12205-014-0363-z.

Dhoke, A.; A. Kumar and I. Ghosh (2021) Hazard-based duration approach to pedestrian crossing behavior at signalized intersections. *Transportation Research Record: Journal of the Transportation Research Board*, v. 2675, n. 9, p. 519. DOI: 10.1177/03611981211003102.

Dommes, A.; M.A. Granié; M.S. Cloutier et al. (2015) Red light violations by adult pedestrians and other safety-related behaviors at signalized crosswalks. *Accident; Analysis and Prevention*, v. 80, p. 67-75. DOI: 10.1016/j.aap.2015.04.002. PMid:25884542.

Du, Y.; Z. Zhao; Y. Song et al. (2023) StrongSORT: make DeepSORT great again. *IEEE Transactions on Multimedia*, v. 25, p. 8725-8737. DOI: 10.1109/TMM.2023.3240881.

Genç, B. and H. Tunc (2019) Optimal training and test sets design for machine learning. *Turkish Journal of Electrical Engineering and Computer Sciences*, v. 27, n. 2, p. 1534-1545. DOI: 10.3906/elk-1807-212.

Koh, P.P. and Y.D. Wong (2014) Gap acceptance of violators at signalized pedestrian crossings. *Accident; Analysis and Prevention*, v. 62, p. 178-185. DOI: 10.1016/j.aap.2013.09.020. PMid:24172084.

Leal-Taixé, L.; A. Milan; I. Reid et al. (2015). MOTChallenge 2015: towards a benchmark for multi-target tracking. *ArXiv*. DOI: 10.48550/arXiv.1504.01942.

Luo, W.; J. Xing; A. Milan et al. (2022). Multiple object tracking: a literature review. *ArXiv*. DOI: 10.48550/arXiv.1409.7618.

Ma, Y.; S. Lu and Y. Zhang (2020) Analysis on illegal crossing behavior of pedestrians at signalized intersections based on Bayesian network. *Journal of Advanced Transportation*, v. 2020, p. 1-14. DOI: 10.1155/2020/2675197.

Marisamynathan, S. and P. Vedagiri (2019) Pedestrian perception-based level-of-service model at signalized intersection crosswalks. *Journal of Modern Transportation*, v. 27, n. 4, p. 266-281. DOI: 10.1007/s40534-019-00196-5.

Mukherjee, D. and S. Mitra (2022) What affects pedestrian crossing difficulty at urban intersections in a developing country. *IATSS Research*, v. 46, n. 4, p. 586-601. DOI: 10.1016/j.iatssr.2022.10.002.

Nikolaou, D.; A. Ntontis; E. Michelaraki et al. (2023) Pedestrian safety attitudes and self-declared behavior in Greece. *IATSS Research*, v. 47, n. 1, p. 14-24. DOI: 10.1016/j.iatssr.2022.12.002.

Noh, B.; H. Park; S. Lee et al. (2022) Vision-based pedestrian's crossing risky behavior extraction and analysis for intelligent mobility safety system. *Sensors*, v. 22, n. 9, p. 3451. DOI: 10.3390/s22093451. PMid:35591139.

Onelcin, P. and Y. Alver (2015) Illegal crossing behavior of pedestrians at signalized intersections: factors affecting the gap acceptance. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 31, p. 124-132. DOI: 10.1016/j.trf.2015.04.007.

Onelcin, P. and Y. Alver (2017) The crossing speed and safety margin of pedestrians at signalized intersections. *Transportation Research Procedia*, v. 22, p. 3-12. DOI: 10.1016/j.trpro.2017.03.002.

Padilla, R.; S. Netto and E.A.B. Silva (2020). A survey on performance metrics for object-detection algorithms. In *Proceedings of the International Conference on Systems, Signals, and Image Processing (IWSSIP).* New York: IEEE. DOI: 10.1109/IWSSIP48289.2020.9145130.

Raoniar, R. and A.K. Maurya (2022) Pedestrian red-light violation at signalized intersection crosswalks: Influence of social and non-social factors. *Safety Science*, v. 147, p. 105583. DOI: 10.1016/j.ssci.2021.105583.

Redmon, J. and A. Farhadi (2016) YOLOv3: an incremental improvement. *ArXiv.* DOI: 10.48550/arXiv.1804.02767.

Redmon, J.; S. Divvala; R. Girshick et al. (2016) You only look once: unified, real-time object detection. *ArXiv.* DOI: 10.48550/arXiv.1506.02640.

Sayed, T.; M. Zaki and A. Tageldin (2016). Automated pedestrians data collection using computer vision. In: Leon-Garcia, A. (ed.) *Smart City 360°. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*. Cham: Springer International Publishing, v. 166, p. 31-43. DOI: 10.1007/978-3-319-33681-7_3.

Song, Y.; J. Wang and X. Long (2019) Analysis and modeling of pedestrian crossing behavior at intersections. In *19th International Conference of Transportation*. Reston: ASCE, p. 6083-6093. DOI: 10.1061/9780784482292.522.

Tonning, C.; T.K.O. Madsen; C.H. Bahnsen et al. (2017) Road user behavior analyzes based on video detections: Status and best practice examples from the RUBA software. In *Proceedings of the 24th ITS World Congress*. Dubai: ITS World, p. 1-10.

Ultralytics (2023) *YOLOv8 Documentation*. Available at: <https://docs.ultralytics.com/> (accessed 10/12/2023).

Wang, C.-Y.; A. Bochkovskiy and H.-Y.M. Liao (2022) YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. ArXiv. DOI: 10.48550/arXiv.2207.02696.

Wojke, N.; A. Bewley and D. Paulus (2017) Simple online and realtime tracking with a deep association metric. In *IEEE International Conference on Image Processing (ICIP)*. New York: IEEE, p. 3645-3649. DOI: 10.1109/ICIP.2017.8296962.

Zhu, D.; N.N. Sze and L. Bai (2021) Roles of personal and environmental factors in the red-light running propensity of pedestrians: case study at the urban crosswalks. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 76, p. 47-58. DOI: 10.1016/j.trf.2020.11.001.