Computers in Industry 167 (2025) 104259

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.sciencedirect.com/journal/computers-in-industry

Evaluating unsignalized crosswalk safety in the age of autonomous vehicles ☆

Andrea Avignone [a],*, Marco Bassani [b], Beatrice Borgogno [b], Brunella Caroleo [c], Silvia Chiusano [a], Federico Princiotto [c]

[a] Department of Control and Computer Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24, Torino, 10129, Italy
[b] Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24, Torino, 10129, Italy
[c] Fondazione Links, Via Pier Carlo Boggio, 61, Torino, 10138, Italy

A R T I C L E I N F O

Keywords:
Conflict measures
Autonomous vehicle
Road safety
Pedestrian–vehicle interaction

A B S T R A C T

As autonomous vehicles are poised to enter public roadways, a major concern is their interaction with pedestrians. It requires attention and ability for pedestrians to interact correctly and for autonomous vehicles to detect pedestrians hence avoiding collisions. We propose a complete pipeline to collect, process and elaborate video data to quantitatively assess the possible occurrence of conflicts. It integrates computer vision techniques and a conflict detection system to evaluate these interactions by rigorously implementing the theoretical formulation of two primary metrics: Time-to-Collision (TTC) for the pre-event phase and Post Encroachment Time (PET) for the post-event phase. This study is conducted in a real-world setting with mixed traffic conditions to analyse the differences in pedestrian interactions with both human-operated and autonomous vehicles during daytime. The computation of conflict measures allowed us to identify possible conflicts and assess the safety at an unsignalized crossing, in which pedestrians are exposed to more risky conflicts. The results obtained show a higher incidence of more severe conflicts for interactions between pedestrians and human-operated vehicles, which highlights the caution taken in programming the autonomous vehicle.

1. Introduction

The introduction of Autonomous Vehicles (AV) on public roads represents a paradigm shift in transportation technology, with potential implications for both operational efficiency and safety (Bagloee et al., 2016). Experiences and testing with AVs, especially autonomous shuttles for public transportation services, have been largely confined to controlled environments, such as closed areas or dedicated lanes (Chaalal et al., 2023; Razmi Rad et al., 2020). However, the real challenge lies in seamlessly integrating AVs into mixed traffic conditions, with the coexistence of Human-operated Vehicles (HVs) (Parks-Young and Sharon, 2022; Barthauer and Friedrich, 2019; Wu et al., 2021) and Vulnerable Road Users (VRU) (Gasper et al., 2018), such as pedestrians and cyclists (Jafarya et al., 2018). In our study, we distinguish between Human-operated Vehicles (HVs), encompassing all vehicles driven by humans regardless of power-train type, and Autonomous Vehicles (AV).

As municipalities begin the deployment of AVs for collective transport (Antonialli, 2019), a crucial component of the investigation is their interaction with pedestrians (Zhu et al., 2022). The most common case in which pedestrians interact with HV and AV is at pedestrian crossings, and of these, the most challenging interactions occur at unsignalized crosswalks where, for various reasons, a driver may not give pedestrian priority.

To investigate how risky interactions between road users can be, Conflict Measures (CM) have been proposed (Tarko, 2019a; Zheng et al., 2021). They are used to quantify the severity of the interaction between two road users distinguishing regular interactions (undisturbed passages) from conflicts of different magnitude (i.e., slight, serious), with the most severe conflict being a collision.

For addressing these concerns and contributing to the ongoing research on AV safety, in this study we investigated the interactions between pedestrians and autonomous shuttles on an urban road section with dense mixed traffic in an Italian city. This setting provides an opportunity to compare the dynamics of AV-pedestrian interactions with those involving HVs, offering valuable insights into the differences in safety and behaviour.

We propose a suitable integrated processing pipeline to automatically perform conflict analysis starting from videos recorded by on-street cameras. It integrates all the required steps, including (1) video

☆ This work was partially supported by the funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875530 (SHOW); and partially supported by the SmartData@PoliTO centre on Big Data and Data Science.

* Corresponding author.
E-mail address: andrea.avignone@polito.it (A. Avignone).

https://doi.org/10.1016/j.compind.2025.104259
Received 4 October 2024; Received in revised form 9 January 2025; Accepted 20 January 2025
Available online 3 February 2025

elaboration for *road user detection and tracking* to compute kinematic information; (2) *data analysis* to model the vehicle–pedestrian interaction and *detect potential conflicts*; and (3) *conflict measures computation* to evaluate the severity of the detected conflict. Here, we considered the reference conflict measures Time-To-Collision (TTC) and Post Encroachment Time (PET) as the benchmark indicators for the pre-event phase and post-event phase, respectively (Hydén, 1996). Their statistical distributions were used to compare the interaction of AV and HV with the pedestrian. In addition, we quantified and evaluated the pedestrian decision-making process by analysing crossing uncertainty and how the speed profile is influenced by the type of vehicle. This included assessing the total stop time during crossing, and examining variations in pedestrian behaviour when interacting with AVs compared to HVs.

Our contributions bring a methodology to move from the real-world scenario to a more consistent representation with the theoretical formulation of the conflict measures. The proposed conflict detection relies on a vehicle box-based representation for a 2D bird's eye view. By taking into consideration the spatial occupation of the vehicles, we gain a more accurate conflict measures evaluation. We consider the different vehicle sizes by automatically mapping the type of the considered vehicle with a set of reference spatial measures collected in a specifically defined catalogue. The data analysis block of the pipeline for conflict measures computation was developed in Python (it will be made available upon request).

We validated our pipeline by conducting *on field observations* at an unsignalized pedestrian crosswalk on different days and under various traffic conditions, comparing the conflict measures values of HV-pedestrian interactions (baseline) with those observed for AV-pedestrian interactions.

The video data used in this study was collected during an in-field experiment that took place in an Italian city within a funded research project to enhance sustainable urban mobility. The experiments involved an on-demand transport service using two autonomous shuttles (SAE level 4), operating in mixed traffic. Videos were collected during the technical validation and verification activities, when passengers were not allowed to board the shuttles.

This study provided valuable insights about the AV-pedestrian interactions and the integration of the autonomous vehicles in a mixed traffic condition with different types of road users. The observations showed that the autonomous shuttle has in general safer interactions with the pedestrians if compared to conventional vehicles. These results may support proper shuttle parameters configuration, enhancing the urban area integration.

The rest of this paper is structured as follows: Section 2 recalls the main related works, whereas Section 3 presents the adopted methodology for conflict measures computation. The proposed pipeline is explained in Section 4. The obtained results and the discussions are reported in Section 5 and Section 6 respectively.

2. Related works

Interactions between road users occur when they share areas where their trajectories intersect and evasive manoeuvres are necessary to avoid collisions, by altering speed and trajectory. Even if measures based on crash data are still the most prevalent, considering non-crash events and using conflict measures provides useful insights about the safety condition (Laureshyn et al., 2017).

Conflict Measures (CM) have been used to assess the impacts on road safety (Tarko, 2019a; Zheng et al., 2021). They are evaluated through field observations, as well as traffic and driving simulations (Tarko, 2012). In field observations, conflicts between road users are recorded on video, and image analysis techniques are then employed to derive their temporal and spatial trajectories (Kastrinaki et al., 2003). In Battiato et al. (2018), a framework was developed using a cascade of classifiers to evaluate road safety through traffic conflict analysis with computer vision techniques. Their approach, however, focuses on pre-event situations using on-board images under general traffic conditions. Our work extends this by specifically examining pedestrian crosswalks, where we evaluate both Time-to-Collision (TTC) and Post-Encroachment Time (PET). Additionally, we provide a comparative analysis between HVs and AVs, offering a more nuanced understanding of traffic conflicts in these critical environments.

Early literature includes studies about the interaction of AV and HV with pedestrians through the use of *micro-simulation techniques* (Papadoulis et al., 2019). These studies suffer from significant shortcomings due to the limited interpretative capacity of behavioural models, their inability to account for chains of events leading to risky traffic conflicts, and the impossibility of simulating the action of AVs, which operate based on sensors conditioned by several environmental factors (e.g., weather conditions, infrastructure characteristics, traffic conditions).

At present, only a few studies have reported information on interactions between pedestrians and AV in real environment settings (Kalantari et al., 2023). This is mainly due to the high costs associated with real–world observation and the legal authorizations required to allow AVs on roads open to traffic. Therefore, many studies report observations of AV interacting with Vulnerable Road Users (VRU) in simulated environments (Kalantari et al., 2023; Parkin et al., 2023).

The recent evolution of technologies in video recording equipment and automated video analysis supported the study and understanding of safety-related events (Yang et al., 2021). In one of the first field observations with autonomous shuttle operation (Madigan et al., 2019), interactions with VRU were studied in relatively high density traffic conditions, deriving qualitative outcomes across all possible interaction scenarios for these public transportation units. The behaviour of VRU was detected and quantified through video analysis, enabling the identification of several risk conditions (severe conflicts). In shared areas, pedestrians exhibited significant discomfort when sharing spaces with vehicles that do not communicate intentions at conflict points.

In Merat et al. (2018), qualitative measures from questionnaires were used to understand the perceived safety of cyclists and pedestrians during demonstrations of autonomous shuttles in shared areas of three European cities. Pedestrians expressed the need to be informed about the actions and manoeuvres of AVs.

In Beauchamp et al. (2022), the authors collected quantitative measures in analysing the AV interactions with road users in a lower traffic density environment compared to previous studies (Merat et al., 2018; Madigan et al., 2019). Their observations indicated that AV behaved more safely than HV, driving slowly and carefully while correctly giving right of way to pedestrians at crosswalks. The results also showed that CM were highly site–dependent, which currently complicates the interpretation of the factors governing the interactions between VRU and AVs.

Existing research has explored pedestrian uncertainty in various contexts. Pedestrian behaviour while crossing is influenced by various parameters, including the presence of conflicting vehicles, pedestrian speeds, pedestrian speed–flow–density relationships, pedestrian compliance with traffic signals, and pedestrian gap acceptance when crossing the road (Ishaque and Noland, 2008; Prédhumeau et al., 2023; Guan et al., 2024). Furthermore, pedestrians are more likely to cross in front of an autonomous vehicle (AV) when they could predict the AV's behaviour (Löcken et al., 2019).

Overall, the scientific literature reveals several knowledge gaps regarding how VRUs and AVs interact with each other. Therefore, our work proposes a complete pipeline to process video footage and rigorously implement the theoretical formulations of both TTC and PET. Our study is based on a real-world environment considering mixed traffic conditions to understand the differences in pedestrian interactions with both HVs and AVs. The computation of conflict measures allowed us to identify potential conflicts and assess the safety at an unsignalized crossing, in which pedestrians are exposed to higher–risk conflicts (Noh et al., 2021).

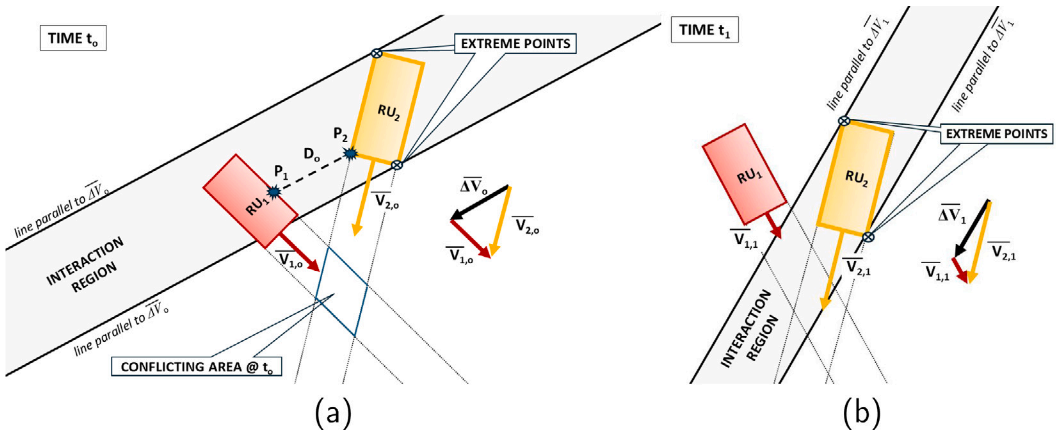


Fig. 1. Computation scheme for TTC: (a) reports two RUs in a collision course; (b) reports two RUs not in a collision course.

3. Adopted methodology for conflict measures computation

Interactions between road users are divided into *undisturbed passages*, when no changes in direction and/or speed are required due to sufficient distance between users in the conflicting area, and *conflicts*, when at least one evasive manoeuvre is required to avoid a collision. Among conflicts, those considered to be "minor" occur when a slight evasive manoeuvre (a small reduction in speed and/or a slight change of trajectory) can nullify the risk of a collision. Conversely, in "serious" conflicts, a collision will occur if no strong evasive action is taken by at least one conflicting user (Hydén, 1987).

The two most common measures of conflict before and after an event are Time-To-Collision (TTC) and Post Encroachment Time (PET). Starting from their original definition, this study implements a rigorous implementation to fully characterize the spatial occupation of interacting road users for pre-event and post-event conflict evaluation.

Time-To-Collision (TTC) is the time that separates two road users (e.g., two vehicles or a vehicle and a pedestrian) from a collision in the pre-event phase if the collision course and speed difference are maintained (Hydén, 1996); the collision can be avoided if at least one road user makes an evasive manoeuvre. Considering two general Road Users RU_1 at speed V_1 and RU_2 at speed V_2, they are on a collision course if there is at least a straight parallel to the vector of speed difference $\Delta V = V_2 - V_1$ that from the target user RU_2 intersects RU_1. In this case, RU_1 and RU_2 follow trajectories that generate a *conflicting area*, representing the predicted location where the incident is expected to occur if neither road user takes an evasive manoeuvre. Since speed and position change over time, the *Instantaneous Time-To-Collision (ITTC)* is calculated by evaluating the TTC at each instant t.

Starting from the most widely accepted definition of (I)TTC, in this study we considered a rigorous mathematical implementation that considers, frame by frame, (i) the *spatial dimension and orientation* of the interacting road users, and (ii) the *interaction region* that indicates if the two users are on a collision course or not.

Each road user RU is described in a two-dimensional space by the following elements:

1. *2D box-based* representation simulating a bird's-eye view of the road user;
2. *RU reference point*, given by the (X,Y) coordinates of the centre of front head;
3. two *RU extreme points*, given by the highest and lowest y-coordinates values of the box.

We denote as *interaction region* the area describing the impact of the spatial occupation of the target user on the interaction. This region is used to determine whether the two road users are on a collision course, and consequently, whether it is possible to calculate the ITTC.

Table 1
Summary of the conflict range for each CM.

Type of interaction	Pre-event	Post-event
No conflict	$ITTC_{min} \geq 3$ s	$PET > 3$ s
Slight conflict	1.5 s $\leq ITTC_{min} < 3$ s	$PET \leq 3$ s
Serious conflict	$ITTC_{min} < 1.5$ s	

Fig. 1 illustrates the general case of two road users RU_1 and RU_2 with their box-based representation: when considering RU_2 as target user, the *interaction region* is the region enclosed between two lines parallel to ΔV, each passing through the extreme points of RU_2. The presence or absence of RU_1 within the interaction region determines whether RU_1 and RU_2 are on a collision course.

In Fig. 1(a), at time t_o, RU_1 and RU_2 are on a collision course since RU_1 falls within the interaction region generated from RU_2. Without any evasive manoeuvres, a conflict is expected to occur in the conflicting area after a time period given by Eq. (1):

$$ITTC_o = \frac{D_o}{\|\Delta V_o\|} \tag{1}$$

where D_o is the current distance between users, defined as the shortest segment parallel to ΔV (now ΔV_o) and connecting the *potential collision points* P_1 and P_2. We denoted this time instant as *Interaction Time* (IT).

Fig. 1(b), at time t_1, RU_1 has reduced the speed with respect to t_0 ($V_{1,1} < V_{1,0}$), thus producing a reorientation and a change in the module of ΔV (now ΔV_1). The newly interaction region differs and no longer includes RU_1. In this case, there is not a collision course, ITTC $\rightarrow \infty$, and as a result, no actual conflicting area exists. We denoted this time instant as *no-Interaction Time* (no-IT).

The minimum ITTC value ($ITTC_{min}$) achieved is representative of the severity of the whole conflict. The lower the $\mathrm{ITTC_{min}}$, the higher the conflict severity (Tarko, 2019b; Arun et al., 2021; Ezzati Amini et al., 2022). Based on the literature review (Sacchi et al., 2013; Sacchi and Sayed, 2016; El-Basyouny and Sayed, 2013), here we assume that $ITTC_{min} \geq 3$ s indicates undisturbed passage (i.e., *no conflict*), $1.5 \leq ITTC_{min} < 3$ s indicates *slight conflicts*, whereas $ITTC_{min} < 1.5$ s a *serious conflict* (see Table 1). As noted in the literature, establishing reliable thresholds remains a challenge due to variability across studies and the lack of a universally accepted method to distinguish conflicts of varying severity. Different studies often use different thresholds for similar interactions, such as pedestrian–vehicle conflicts. In this study, we followed established conventions and considered the $ITTC_{min}$ threshold of 1.5 s as an indicator of serious conflicts. For mild conflicts, we used the range of 1.5 to 3 s (Arun et al., 2021; Gettman et al., 2008).

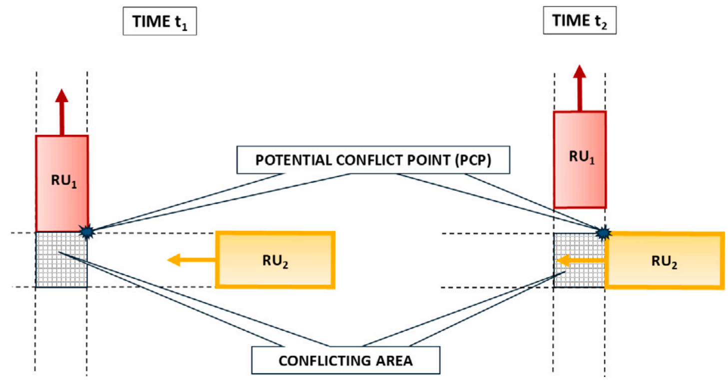


Fig. 2. Calculation scheme for PET.

Post Encroachment Time (PET) measures the temporal distance between two road users passing in a conflicting area when one of them has already left the area (i.e., the post-event phase). The proposed implementation uses the same 2D box-based representation for ITTC estimation. Here, the *conflicting area* was identified by the region enclosed by lines drawn parallel to the trajectories of the road users, with their width defining its boundaries.

According to Fig. 2, in the conflicting area it is necessary to determine the *Potential Conflict Point* (PCP). PET is given by Eq. (2):

$$PET = t_2 - t_1 \tag{2}$$

where t_2 is the time when the second user RU_2 encounter the PCP, and t_1 is the time when the first user RU_1 leaves the PCP (Laureshyn et al., 2010). According to Lord (1996), a $PET > 3\,\mathrm{s}$ indicates undisturbed passage (i.e., *no conflict*), whereas $PET \leq 3\,\mathrm{s}$ a *conflict* (see Table 1).

4. Proposed pipeline

The pipeline proposed in this study for detecting potential conflicts from a collection of videos consists of several stages, including *data acquisition*, video elaboration for *road user detection and tracking*, data analysis for *conflict detection*, data analysis for *pedestrian decision-making* process, and the final *outcome*. First, the acquired videos undergo specific processing to correct distortions and apply an appropriate object detection and tracking algorithm. Once processed, the videos are analysed to estimate the trajectories and speeds of the involved objects, such as vehicles and pedestrians. Finally, potential conflicts are identified, and the associated conflict measures are calculated. Fig. 3 shows the proposed pipeline with its main blocks described in detail in the following subsections. It is based on models and formulas defined in Section 3.

4.1. Shooting area selection and data acquisition

The unsignalized crosswalk shown in Fig. 4 was selected as a reference scenario for traffic scene recording. The elected area, close to an hospital's area, was characterized by mixed traffic, thus including different types of conventional vehicles like cars, buses and vans. While we could not identify the specific powertrain or driving mode (manual or assisted) for every vehicle analysed, we compare the autonomous shuttle directly with all others. This is justified by the absence of vehicles with an SAE automation level above 2 in the study area, as confirmed through site observations and the traffic context.

Scenes were recorded by alternating two low-cost action cams (Garmin VIRB™, 1080p HD, 30 frame/s). They were placed on the top of a carbon fibres telescopic pole at 10.80 m to the ground. The cameras were placed outside the roadway in a position that was difficult for drivers to detect.

4.2. Video elaboration for road user detection and tracking

This section describes the process of automatically detecting and tracking road users. To ensure accurate analysis of pedestrian–vehicle interactions, the collected video footage was first corrected to eliminate distortion errors, and the kinetic data were denoised to enhance precision.

Removing camera distortion error. Initially, distortion caused by the camera's wide-angle lens was corrected to prevent errors in spatial measurements. This distortion affects the measurement of the spatial variables used to analyse the interaction between pedestrians and vehicles, such as their position, relative distance and speed. This error was corrected using the MATLAB Camera Calibrator App, which requires us to load the camera parameters and its distortion matrix. As an example, in Fig. 4 we report an initial raw video frame with distortion error (Fig. 4(a)), along with the corresponding final frame after distortion correction (Fig. 4(b)).

Pedestrian and vehicle detection. The detection of both pedestrians and vehicles in the videos was carried out using YOLOv5 (Redmon et al., 2016) with 80 objects included in COCO classes (Lin et al., 2015). However, since the autonomous shuttle is not part of these classes, we manually labelled the videos containing interactions with the shuttle. Each detected pedestrian and vehicle is assigned the same identifier (ID) across all frames. The tracking of the detected objects was performed using StrongSORT (Du et al., 2023). Only the video sections involving an interaction between pedestrian and AV/HV were considered for further analysis.

Tracking pedestrians and vehicles. For measuring the spatial variables describing the interaction between pedestrians and vehicles, we migrated the frame representation from a pixel-based system to a georeferenced X-Y system. As shown in Fig. 4(c), the X-axis was drawn parallel to the road travelled by vehicles taking as a reference the road centre-line, while the Y-axis was drawn parallel to the pedestrian crossings. The axis intersection, i.e. the origin point (0,0), was placed in the bottom left corner of the zebra crossing.

During the on field inspection, the (X,Y) coordinates for a set of reference points were collected. Having these reference points, their corresponding values in pixels, and the intrinsic matrix of the camera, we were able to map the position of any pixel in the video to the corresponding position in the real-world (X-Y) reference system using the PnP (Perspective-n-Point) OpenCV tool (Bradski, 2000).

To measure the position in X,Y coordinates of pedestrians and vehicles at each time instant of the video, we considered the following representative points related to the YOLO bounding boxes: the centre point between the feet for the pedestrian, whereas the centre of the frontal number plate for the vehicles. The position was tracked every 34 ms. Fig. 5 shows an example of the result obtained using the adopted detection and tracking approach.

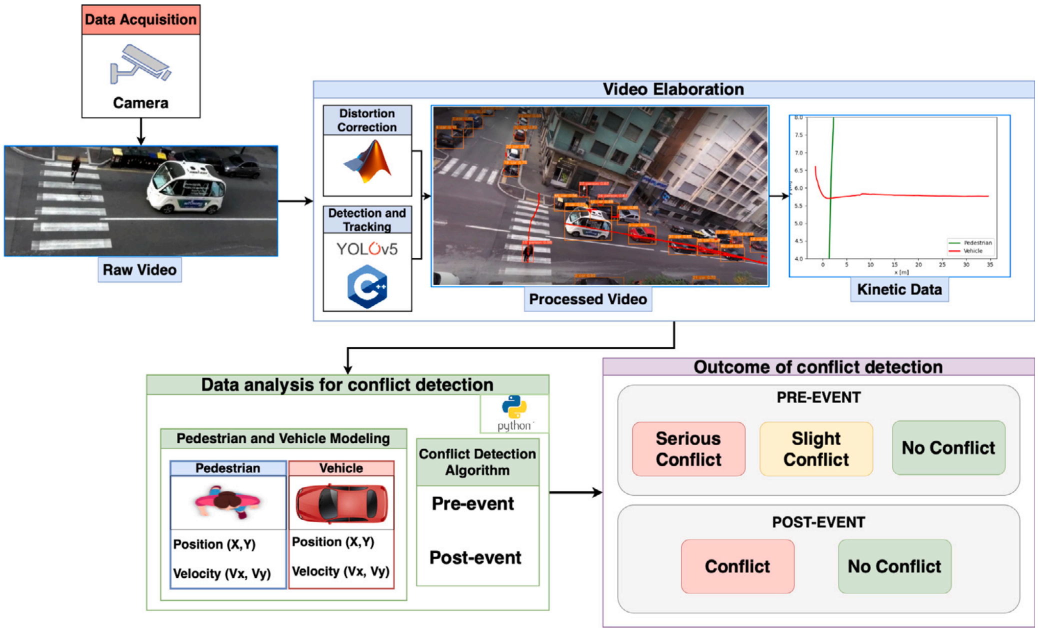


Fig. 3. Overview of the proposed pipeline, including blocks from video data acquisition to conflict measures computation.

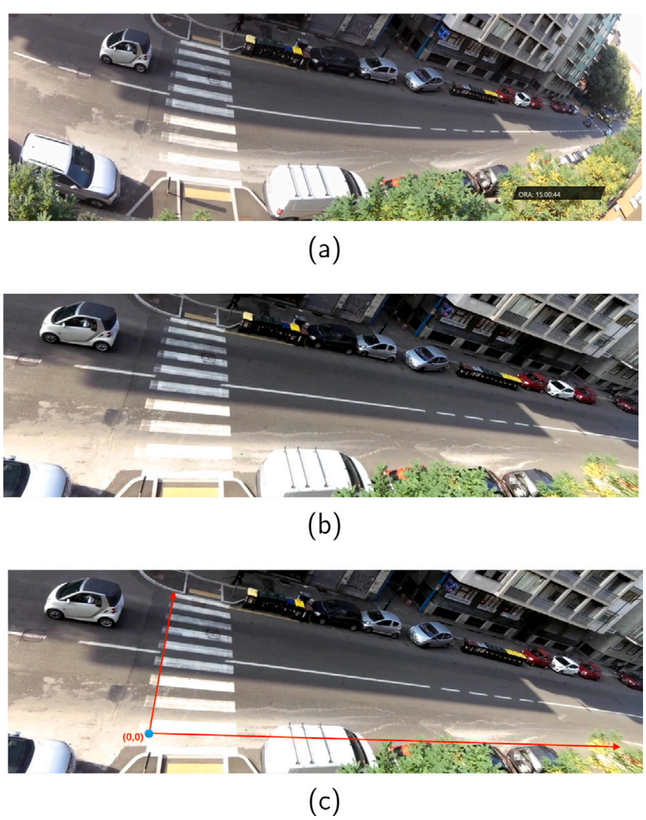


Fig. 4. Comparison between (a) the original video frame; (b) the distorsion-free video frame; (c) the distorsion-free video frame with the introduced geo-reference reference system.

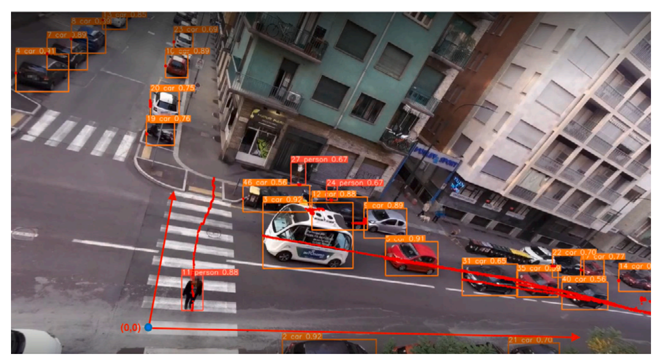


Fig. 5. A pedestrian is crossing in front of the autonomous shuttle, with automatic tracking and trajectory detection.

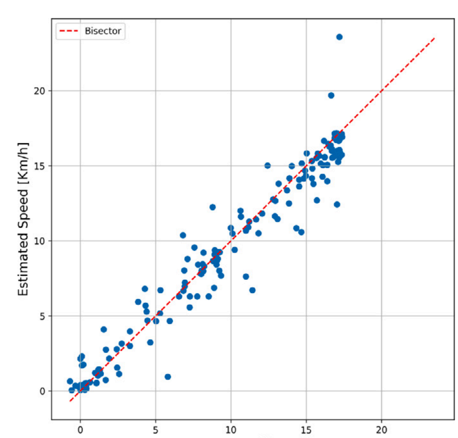


Fig. 6. Comparison between AV true speed and estimated speed with SMA method.

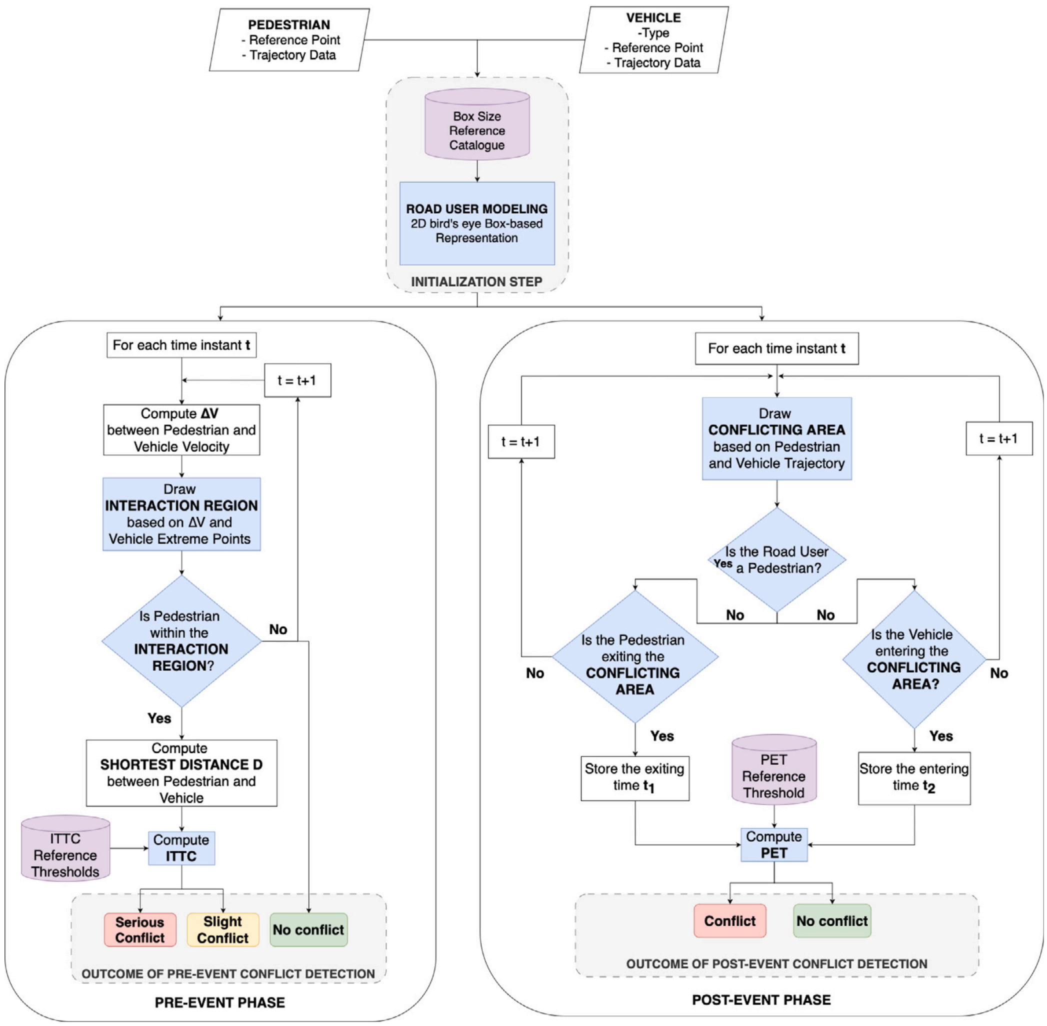


Fig. 7. Schema of the conflict detection algorithm for data analysis.

Trajectory smoothing and denoising. The resulting kinetic data was denoised to enhance the accuracy of spatial and motion-related variables. The (X,Y) coordinates of both pedestrians and vehicles extracted from the video were affected by noise due to camera oscillation and the discretization of the image into uniform colour pixels.

To overcome this problem, the Simple Moving Average (SMA) was employed. It returns the unweighted mean of the k data selected, according to Eq. (3):

$$SMA_k = \frac{1}{k} \sum_{i=n-k+1}^{n} p_i \tag{3}$$

where p_i is the i_{th} data point (i.e., the position), n the total number of data points, and k the window width. The higher the k, the smoother the curve; however, increasing k implies a decrease in the accuracy. After a sensitivity analysis with different k values, a moving window of 30 frames, i.e., equivalent to $1s$, was adopted. The window was centred and moved frame-by-frame.

We conducted a validation analysis using a sample of on-board sensor data from the AV close to the unsignalized crossing. Specifically, we compared the true speed recorded by the shuttle's on-board sensors with the estimated speed obtained from our methodology. The results of this comparison are presented in Fig. 6, with the scatter plot of true speed versus estimated speed showing a strong alignment along the bisector line. The computed metrics further support the validity of our approach: $R^2 = 0.95$, Mean Absolute Error (MAE) = 0.92 Km/h, and Root Mean Square Error (RMSE) = 1.4 Km/h. These results demonstrate a high level of agreement between the true and estimated speeds, indicating that the smoothing process effectively captures the underlying speed profile while mitigating the influence of noise.

4.3. Data analysis for conflict detection

We considered one-to-one interactions between two road users, specifically a pedestrian and a vehicle. Based on the methodology in Section 3, in a first initialization step the road users are modelled using a 2D box-based representation. Then, for each time frame, the potential conflict between pedestrian and vehicle is detected for both the pre-event and the post-event phases. Fig. 7 illustrates the overall structure of the proposed algorithm, with its main blocks described below; they were developed in Python.

4.3.1. Initialization step: vehicle and pedestrian modelling

Once obtained the coordinates and the kinematic information (trajectory and speed) at each timestamp, we filtered the portion of data in which the pedestrian and the AV/HV were actually interacting.

We then moved from the original camera image (Fig. 5) to the proper 2D perspective of the scene as formulated in Section 3 (Fig. 8). In fact, the camera images and standard bounding boxes obtained with the object detection system are not adequate to simulate a 2D bird's-eye view consistent with the theoretical formulation considered. For the vehicle, the corresponding 2D-box is placed starting from the vehicle representative point (i.e., centre of the frontal plate). To adapt the box

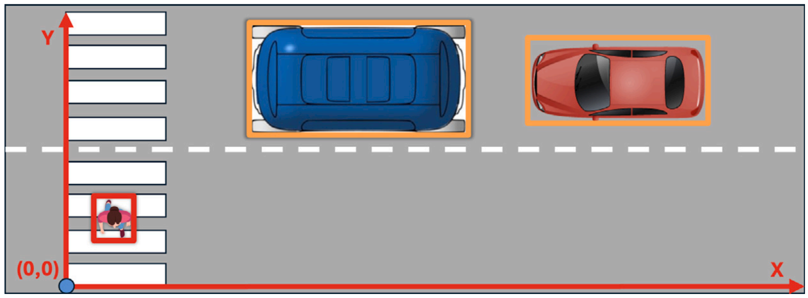


Fig. 8. Crosswalk representation using a 2D bird's eye view.

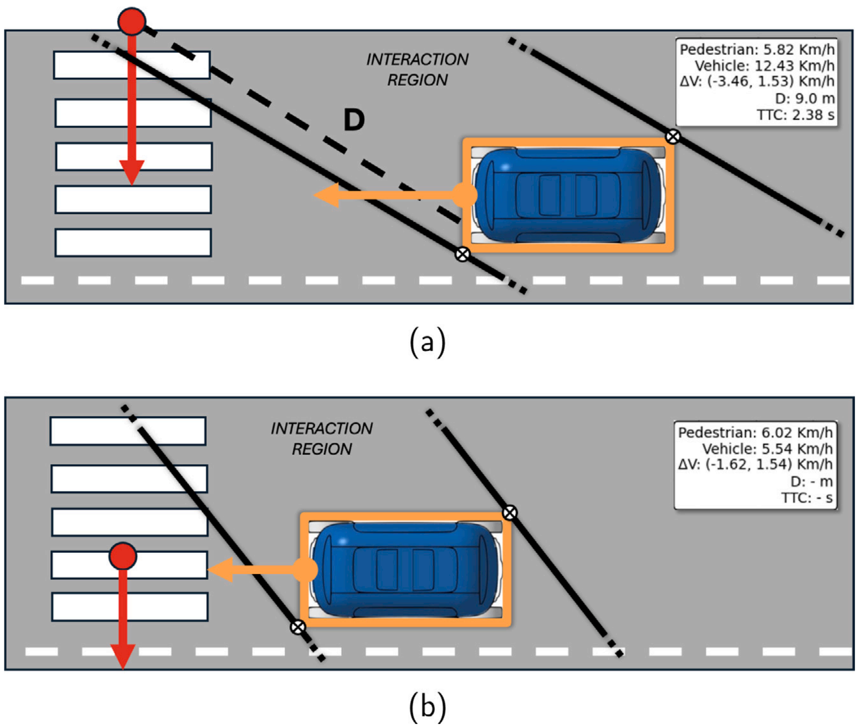


Fig. 9. Pre-event conflict detection. Representation of pedestrian–vehicle interaction using a 2D bird's-eye view for ITTC calculation, according to velocity and position of vehicle and pedestrian: (a) sample of time frame with actual conflict (b) sample of time frame with no conflict detected.

size to the vehicle type, we defined a *reference catalogue* of common standard dimensions for different vehicle types (Table 2) and a procedure to automatically map each vehicle type with the corresponding box in the catalogue. Due to the camera's positioning, which does not provide a fully top-down view of the scene, accurately estimating vehicle dimensions is challenging. Geometric distortions introduced by the perspective view can lead to inaccuracies, particularly in estimating the width of the vehicle. To address this issue and reduce measurement noise, we opted to use standardized dimensions for each vehicle type. This approach ensures consistency in the analysis, eliminates the variability from dynamic estimations, and provide a reliable baseline for our conflict-based analysis.

Given the relative proportions with the vehicles, the pedestrian was instead safely approximated to a point in the 2D bird's-eye view, corresponding to the bottom centre point of the reference bounding box.

4.3.2. Pre-event conflict detection

A potential conflict between a pedestrian and a vehicle is identified by determining whether the pedestrian falls within the interaction region defined by the vehicle's path. Once this condition is met, the ITTC is computed according to Eq. (1).

Table 2
Reference catalogue adopted for building the vehicle 2D box-based representation.

Box size	Vehicle type			
	Human-operated (HV)			Autonomous (AV)
	Car	Van	Bus	Shuttle
Length [m]	4.50	5.40	12.20	4.75
Height [m]	2.00	2.10	2.55	2.11

Fig. 9 shows the graphical representation of a 2D scene where a vehicle is interacting with a pedestrian in the pre-event conflict phase. Starting from the 2D vehicle box-based representation, we define the interaction region by two straight lines drawn at each time instant from the extreme points of the vehicle, thus the bottom-left point and top-right point. The two lines are parallel to the speed difference vector between the pedestrian and the vehicle (ΔV) and tangent to the vehicle contour. If a conflict is detected, we compute the minimum distance D between the two road users along ΔV, and then ITTC value according to Eq. (1).

Fig. 9(a) shows a possible conflict condition since the pedestrian falls within the lines generated by the vehicle contours. On the contrary, in Fig. 9(b) the road users are not in a potential collision course

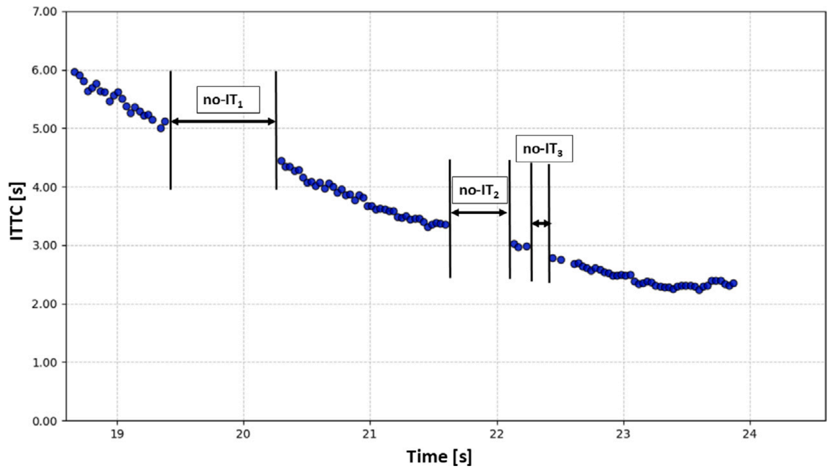


Fig. 10. Example of ITTC curve with three gaps due to no-Interaction Times.

since the pedestrian is outside the interaction region defined by the two lines. This is due to the change in the relative velocities, which affects the slope of ΔV. This procedure provides a convenient mechanism to investigate the evolution of the whole interaction depending on the changes in speed and direction, thus identifying the most dangerous instants.

Based on the considerations of Section 3, the rigorous evaluation of the TTC measure instant-by-instant can lead to gaps in the ITTC curve, as shown in Fig. 10 with three main no-ITs highlighted. These gaps are due to evasive actions related to the actors' decision-making process such as pedestrian hesitations or evasive manoeuvres of the vehicle (changes in direction or speed). Given a threshold of interest Thr_{gap} to filter the portion of interaction in which the actors are closer, we denote *Sum of no-Interaction Times* (SUM-no-IT) with Eq. (4):

$$\text{SUM-no-IT} = \sum_{i} \left(\text{no-IT}_i\right), \quad \text{when ITTC} < \text{Thr}_{\text{gap}} \tag{4}$$

In our study, we set $Thr_{gap} = 7$ s to filter the ITTC events of interest, including the first *No Conflict* phase.

4.3.3. Post-event conflict detection

The Post-Encroachment Time (PET) computation process is illustrated in Fig. 11(a) and Fig. 11(b) where we can observe a post-event scenario involving a pedestrian and a vehicle. In this case, we use the 2D box-based representation for the vehicle to define the conflicting area. It is delimited by two lines parallel to the vehicle box. Since the pedestrian is approximated to a point, the conflicting area in this scenario is actually a line. First we identify the time instant t_1 when the pedestrian exits the conflict area (Fig. 11(a)). Then, we determine the time instant t_2 when the vehicle enters the same area (Fig. 11(b)). Using these time instants, t_1 and t_2, we calculate the PET value based on Eq. (2).

4.4. Data analysis for pedestrian decision-making

The absence of interaction times between pedestrians and vehicles often indicates an evasive manoeuvrer by one or both parties. This study focuses on the pedestrian's behaviour to examine differences in decision-making and interactions with HVs and AVs.

In our work, we detect moments when pedestrians either slow down or halt their crossing movement (Jay et al., 2020; Onelcin and Alver, 2017). By analysing pedestrian speed profiles, we identify two distinct conditions:

1. **Stop Event**: This occurs when the pedestrian's speed falls below a defined threshold Thr_{stop}, indicating that the pedestrian is essentially stationary.
2. **Long Stop Event**: This is a prolonged stop, defined as a stop event lasting longer than Thr_{long_stop}.

The *Stop Event* reflects a moment of pedestrian's hesitation or uncertainty about the decision to cross. It may occur due to perceived risks, such as the speed or proximity of an approaching vehicle, or uncertainty about the vehicle's intentions. In contrast, the *Long Stop Event* represents a more substantial interruption to the crossing movement, suggesting a deliberate decision to delay or abandon crossing.

For this study, based on the literature (e.g., Noland (2008)) and an exploratory analysis of the speed profiles, we define $Thr_{stop} = 0.3$ m/s (1.08 km/h) to identify near-zero motion in walking profiles. This value accounts for small fluctuations in the estimated speed due to noise or minor body adjustments while effectively capturing moments when pedestrians come to a practical stop. However, it is an experimentally determined parameter and can be adjusted based on specific cases or contextual considerations. It is chosen because it reflects a pedestrian speed significantly lower than the average crossing speed of approximately 1.3 m/s (Ishaque and Noland, 2008), effectively indicating a stationary state. Similarly, we set $Thr_{long\_stop} = 1$ s, considering the average crossing time for pedestrians typically ranges from 6–8 s. These thresholds enable us to systematically identify and analyse instances of hesitation and prolonged stops in pedestrian behaviour.

Based on the previous characterized events, we define the following measures for evaluating each crossing: (i) *Total Stop Time* during crossing, representing the sum of interruptions occurred during the pedestrian crossing and (ii) *Number of Long Stops*.

4.5. Outcome of conflict detection

Based on the detected interaction patterns and the computed ITTC and PET values, the proposed methodology returns the following possible outcomes: (i) pre-event conflict, (ii) post-event conflict, (iii) both pre-event and post-event conflict, (iv) no conflict. In case of pre-event conflict, also the magnitude is notified, i.e., if a slight or serious conflict.

By way of example only, two outputs are reported in Fig. 12 for outcome types. A pre-event conflict was detected when the autonomous shuttle interacted with a pedestrian. Since the corresponding $ITTC_{min} = 2.15$ s falls in the range between 1.5 s and 3 s, the conflict was classified as slight according to the reference thresholds in Table 1. A post-event conflict event was instead notified between a pedestrian and a bus. Being the PET value lower then 3s (i.e., $PET = 2.35$ s), based on thresholds in Table 1, the pipeline pointed out the occurrence of the conflict.

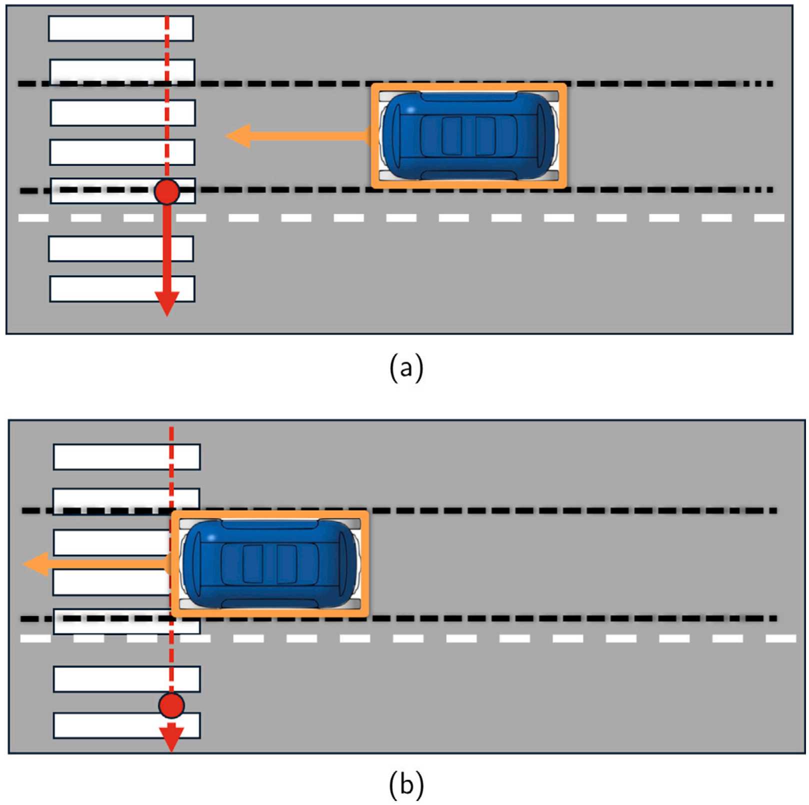


Fig. 11. Post-event conflict detection. Representation of pedestrian–vehicle interaction using a bird's-eye view for PET calculation: (a) post-event conflict at t_1 (pedestrian leaves the conflict area); (b) post-event conflict at t_2 (vehicle enters the conflict area).

Pre-event Conflict	
Pedestrian ID:	8
Vehicle ID:	30 (shuttle)
Initial time interaction:	24.446 s
End time interaction:	29.784 s
ITTC_min:	Slight Conflict detected (2.150 s)
PET:	No conflict detected (6.936 s)
PET time:	$t_1 = 29.444$ s, $t_2 = 36.380$ s

(a)

Post-event Conflict	
Pedestrian ID:	10
Vehicle ID:	35 (bus)
Initial time interaction:	1.904 s
End time interaction:	4.114 s
ITTC_min:	No conflict detected
PET:	Conflict detected (2.346 s)
PET time:	$t_1 = 3.808$ s, $t_2 = 6.154$ s

(b)

Fig. 12. Example of obtained report for (a) detection of a pre-event conflict and (b) detection of a post-event conflict.

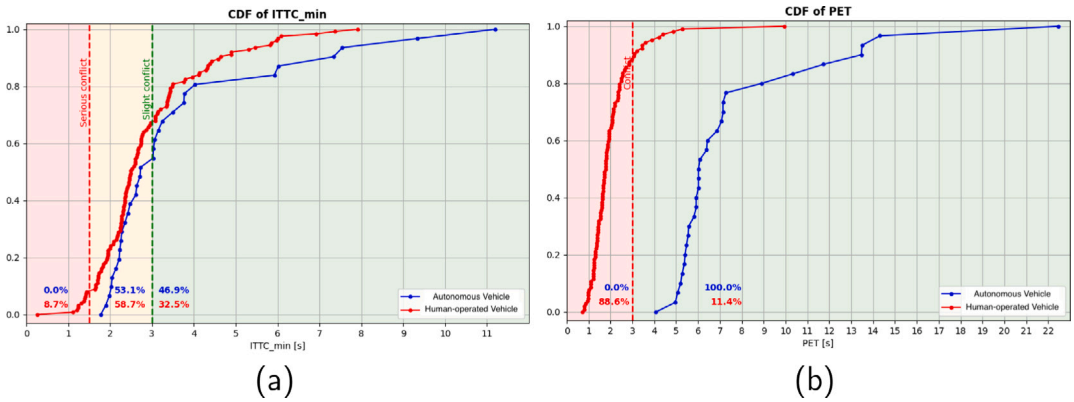


Fig. 13. Cumulative Density Function (CDF) comparison between Human-operated and Autonomous Vehicles (HV and AV) of experimental results distributions: (a) ITTC_{min}; (b) PET. The conflict thresholds are represented with the percentage of interactions falling into each single region for both AV and HV.

5. Experimental results

In this section, we present the results of the conflict analysis obtained from the collected videos which were processed using the proposed pipeline. The following interactions were considered:

1. conflicts between a pedestrian and a Human-operated Vehicle (HV);
2. conflicts between a pedestrian and an Autonomous Vehicle (AV).

Types of interactions and the corresponding potential conflicts were analysed based on ITTC and PET measures. To further inspect differences in AV/HV and pedestrian interactions in terms of statistical distributions, the proposed pipeline also supports the evaluation of both the ITTC and PET distributions using the Kolmogorov–Smirnov (K–S) statistical test (Massey, 1951).

5.1. Experimental setting for video collection

The non-signalized crosswalk considered in this study was part of the route covered by the autonomous shuttles under analysis. It is located in an Italian city, close to an hospital's area. The selected area featured mixed traffic, including various types of conventional HVs like cars, buses and vans. The road is a urban two-lane collector road, however the camera setup was optimized to capture a clear and detailed view of the pedestrian crossing and the traffic in one direction only. This choice was made to ensure the quality and consistency of the data collected, as the geometry of the road and the placement of the camera limited the visibility of traffic in the opposite direction. As a result, we focused exclusively on the lane used by the AVs.

Videos were recorded over 8 days across 3 different weeks, 3 h per day since 2:30 p.m. to 5:30 p.m., resulting in a total of 24 h of records. Pedestrians were aware of the possible presence of the autonomous shuttle as they have been informed through advertising campaigns around the neighbourhood. Blurring technologies were applied to anonymize human faces and car licence plates before processing the videos. The original video files were deleted from the servers. We analysed a total of 168 pedestrian–vehicle interactions, 33 of which involved the autonomous shuttle, while the remaining 135 involved human-operated vehicles.

5.2. Detecting pre-event and post-event conflicts

Table 3 reports statistical features for the ITTC and PET conflict measures. For each measure we show the mean value (M), standard deviation (*SD*), and minimum and maximum value (*MIN* and *MAX*). The conflict/no conflict condition and the severity degree of a possible conflict was assessed based on the measure values and the reference thresholds in Table 1.

Table 3
Summary of descriptive statistics of computed conflict measures.

CM	TYPE	M	SD	MIN	MAX
ITTC_{min} [s]	HV	2.885	1.306	0.331	7.952
	AV	3.660	2.278	1.819	11.204
PET [s]	HV	1.996	0.942	0.646	5.440
	AV	7.650	3.802	4.182	22.304

Pre-event conflicts. The experimental results show that AV-pedestrian interactions mostly fell within the *no conflict* range since the ITTC_{min} mean value is above the 3 s threshold (ITTC_{min} with $M = 3.66$ s, $SD = 2.278$ s). Most important, the ITTC_{min} minimum value for the AV case is still above the *serious conflict* threshold (i.e., $ITTC_{min} = 1.5$ s), indicating the absence of events with severe conflict.

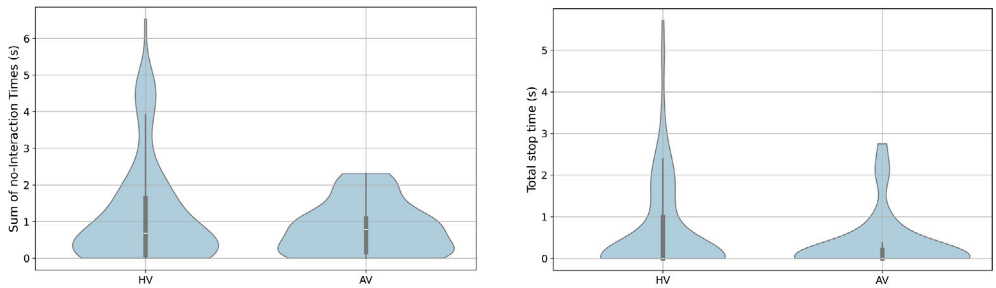
Conversely, HV-pedestrian interactions are on average considered *actual conflicts* since the ITTC_{min} mean value is below the 3 s conflict threshold ($M = 2.885$ s, $SD = 1.306$ s). The HV most hazardous event is fair below the dangerous threshold and in the order of a few tenths of a second (*MIN* $= 0.331$ s), representing a significantly unsafe situation.

It is worth noting that the Italian Highway Code[1] establishes that pedestrians have priority at zebra crossings. In all observed cases, AVs consistently yielded to pedestrians, resulting in 100% yielding rate and ensuring their right of way. In contrast, in 20% of the total records involving HVs, drivers failed to yield to pedestrians, highlighting a discrepancy in how the two types of vehicles behave.

Post-event conflicts. AV-pedestrian interactions are characterized by significantly higher mean PET values, always above the safe threshold. The behaviour of the AV proved to be extremely cautious, with a very slow pick-up compared to human-operated vehicles ($M = 7.65$ s, $SD = 3.802$ s). The event with a particularly high PET value (*MAX* $= 22.304$ s) is very far from the HV counterpart. The large difference between the AV minimum and maximum values also explains the rationale behind the higher standard deviation values.

Instead, HV-pedestrian interactions fall into the potentially dangerous post-events interactions, with PET mean value under the 3 s conflict threshold ($M = 1.996$ s, $SD = 0.942$ s). Specifically, the minimum PET value is significantly low, thus indicating a serious conflict in the post-event phase.

[1] https://www.aci.it/i-servizi/normative/codice-della-strada.html in Italian.



(a) Sum of no-Interaction Times (b) Total Stop Time

Fig. 14. Violin plot describing the interaction between pedestrian and HV/AV: comparison of Sum of no-Interaction Times and pedestrian stop behaviour.

5.3. Conflict measures distribution analysis

Pre- and post-events characterized by low ITTC_{min} and PET values are crucial to detect the most dangerous conflicts. In the following, we analyse the cumulative density functions (CDF) of both ITTC_{min} and PET using the K–S test to provide a complete comparison of pedestrian interactions with autonomous and human-operated vehicles.

ITTC_{min} Distribution Analysis. A first inspection using the K–S test reported that the two distributions are similar ($D = 0.168$, $p = .415$, thus the null hypothesis cannot be rejected). The lognormal distribution proved to be the best one to model data for both HV-pedestrian interactions ($D = 0.110$, $p = .087$) and AV-pedestrian interactions ($D = 0.076$, $p = .985$).

A further graphical exploration, shown in Fig. 13(a), highlighted that: (i) the curve referring to the AV exhibited higher ITTC_{min} values compared to HV; (ii) the two distributions significantly differed in their tails. Since lower ITTC values indicate dangerous conflicts, the lower tail of the ITTC_{min} distributions is particularly relevant due to the high-risk interactions it represents. The left tail for HV proved the presence of cases in which the interaction was considered a serious conflict, i.e., below the threshold of $1.5s$ (see Table 1). No similar events were observed for the AV.

PET Distribution Analysis. Similar to pre-event analysis, we examined the PET cumulative density functions (CDF) for HVs and AVs. In this case, the K–S test confirmed that the two distributions were statistically different ($D = 0.943$, $p < .001$). In fact, the difference between the two curves is evident in Fig. 13(b). The PET values for AVs were significantly higher than those for HVs, indicating that the AV-pedestrian interactions were the safest for the post-event phase. However, this also suggests that AVs generally took a remarkable amount of time to restart after a pedestrian had crossed.

5.4. Pedestrian decision-making behaviour

TTC inherently accounts for the relative speeds of both the pedestrian and the vehicle. Therefore, the observed gap time between interactions is a combined result of these dynamics.

Fig. 14(a) presents the comparison of the *Sum of no-Interaction Times* between pedestrian-HV and pedestrian-AV. Notably, the pedestrian-HV interaction exhibits longer average values ($M = 1.55$ s, $SD = 1.48$ s) than pedestrian-AV ($M = 0.94$ s, $SD = 0.75$ s), coupled with higher variability and maximum values. The maximum *Sum of no-Interaction Times* are particularly noteworthy as they may indicate possible evasive manoeuvrers ($MAX = 6.53$ s for HV and $MAX = 2.31$ s for AV). These discontinuities indicate the need for further analysis to evaluate how pedestrians respond to different types of vehicles.

To delve deeper into pedestrian decision-making, we focus on the pedestrian's perspective to evaluate how crossing behaviour adapts to the type of approaching vehicle. We analyse stop patterns and speed profiles during interactions according to Section 4.4. Stops, particularly at the beginning of crossings, serve as indicators of hesitation and decision-making uncertainty.

Fig. 14(b) compares the distribution of total stop time for pedestrian-HV and pedestrian-AV interactions. The violin plots are consistent with the *Sum of no-Interaction Times* distribution of Fig. 14(a) and reveal distinct trends in pedestrian behaviour. Pedestrians interacting with AVs tend to exhibit shorter and more consistent stop durations. The interquartile range (IQR) for AVs is narrower ($IQR_{AV} = 0.204$ s), while interactions with HVs show a wider spread in stop times, with a broader IQR ($IQR_{HV} = 0.99$s) and several outliers up to 5.7 s. This indicates greater variability in pedestrian behaviour when interacting with HVs, possibly due to the more unpredictable and inconsistent nature of human-driven vehicle movements. The presence of extreme outliers in the HV group further suggests that, in some cases, pedestrians may experience significant hesitation or uncertainty when deciding whether to cross in front of HVs. The stops duration tend to be more limited in the case of AVs, which aligns with the idea that pedestrians feel more confident and assured due to the predictability of AV behaviour.

Further analysis reveals that the percentage of maximum stop durations exceeding 1 s is lower for interactions with AV compared to HV. Specifically, only 12.12% of interactions with AVs presented at least a *Long Stop*, while this percentage rises to 23.53% for HVs. We also observed that, in the case of HVs, there were three instances where the $Number\ of\ Long\ Stops = 2$ within the same pedestrian–vehicle interaction.

These results demonstrate that AVs promote quicker and more predictable crossing decisions, as reflected by the shorter total stop times and limited maximum stop durations. In contrast, HVs are associated with longer and more variable stop durations, suggesting that pedestrians exhibit greater uncertainty and hesitation when crossing in front of HVs.

Exploratory example. Fig. 15 provides an illustrative example of pedestrian decision-making during a single interaction with an HV. The ITTC evolution (Fig. 15(a)) and the corresponding speed profile (Fig. 15(b)) highlight two distinct long stops at the start of the crossing. These long stops coincide with the deliberation phase, during which the pedestrian is assessing whether it is safe to cross. This hesitation contributes to prolonged no-Interaction Times, as seen in the ITTC curve. The pedestrian slows down and even comes to a complete stop, which is a key indicator of the decision-making process. These prolonged pauses may be a natural response to the uncertainty regarding the vehicle's intentions.

The observed behaviour reinforces the earlier findings that pedestrian hesitation is concentrated at the beginning of the crossing, where the pedestrian is making a critical decision about safety. These long stops represent a period of evaluation and waiting for an opportunity to cross. It is worth noting that these decisions are not only influenced by the pedestrian's own perception but also by the behaviour of the vehicle, particularly in unregulated crossings where the pedestrian must assess whether the vehicle will yield or stop.

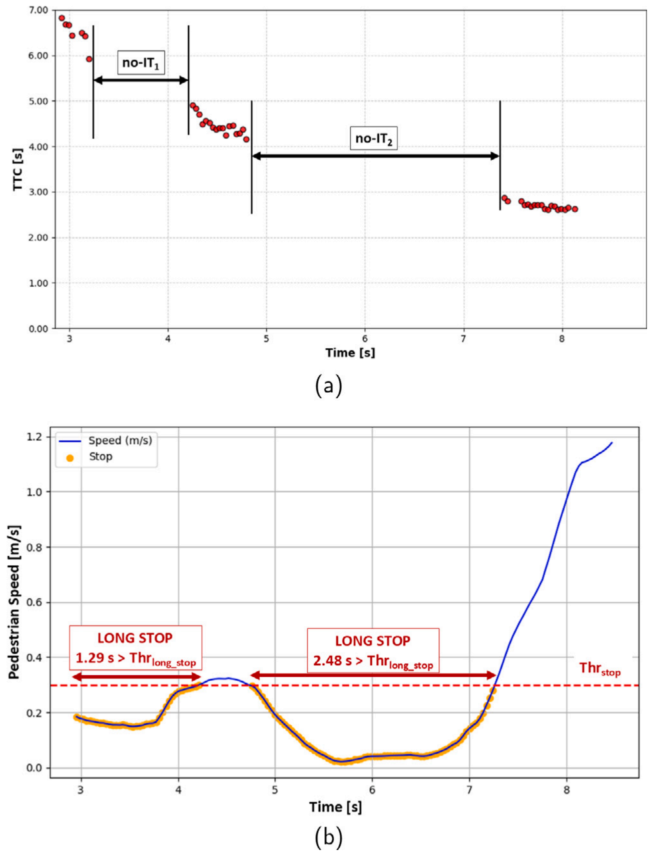


Fig. 15. Example of ITTC and speed profile for pedestrian decision-making during an interaction with an HV.

6. Discussion and conclusion

The introduction of autonomous vehicles on roads must be carefully evaluated to avoid uncomfortable and unsafe situations. With this motivation, the present study set out to assess the safety of pedestrian–vehicle interaction with both autonomous and human-operated vehicles in mixed traffic conditions during daytime. In our study, we observed a relevant number of interactions between pedestrians and vehicles in a real-world scenario. It is worth noting that our observations are among those few already carried out with real data collected in urban context with autonomous shuttles operating in roads actually open to traffic.

Proposed pipeline and real data analysis. We propose an automated pipeline to integrate all the necessary steps to process the recorded videos. Using the pipeline, we obtain the spatial–temporal trajectories and derive the common CMs ($ITTC_{min}$, PET) to evaluate the severity of the interactions between HV and AV with pedestrians. While our computation of conflict measures (CMs) relies on a 2D box-based representation derived from real-world video, this approach introduces assumptions about vehicle dimensions and positions to ensure consistency with the theoretical formulation presented in Section 3. Alternative methods, such as using dynamic measurements could be explored to further validate and enhance the robustness of the CM computation without introducing additional noise. The data analysis block of the pipeline for conflict measures computation was developed in Python.

Interaction AV/HV-pedestrian. Referring to our results on $ITTC_{min}$ in the pre-event phase, i.e., when vehicles and pedestrians have not yet occupied the conflict area, the AV-pedestrian interactions did not indicate critical situations. Instead, serious conflict events occured only in the HV case, highlighting more risky interactions.

In the post-event phase, we found conflicts only in HV-pedestrian interactions. In contrast, PET values recorded for AV-pedestrian interactions were considerably higher, indicating much safer situations compared to HVs. The AV under study showed a longer waiting time before restarting after the pedestrian had crossed.

Based on our observations, we can state that, under mixed traffic conditions, both in the pre-event and post-event phases, the current AVs operating way guarantees significantly lower risk of collision than HVs. It is also worth noting that interactions with AVs did not present any critical situations throughout the entire test period in mixed traffic: there were neither collisions nor situations considered unsafe.

Our results on Sum of no-Interaction Times and Total Stop Time (Section 5.4) highlighted the dependency of pedestrian decision-making on the type of interacting vehicle. Pedestrian interactions with HVs were characterized by longer stops and extended no-Interaction Times, reflecting a greater level of hesitation and uncertainty. This cautious behaviour is likely a response to the unpredictable nature of HVs, which demand more deliberation from pedestrians to determine the safety of crossing.

In contrast, pedestrian interactions with AVs exhibited shorter and more consistent stops, indicating increased confidence in crossing decisions. The predictable and non-aggressive behaviour of AVs appears to facilitate smoother pedestrian crossings, reducing hesitation and fostering a sense of safety.

These findings underscore the critical role of vehicle behaviour in shaping pedestrian decision-making processes. Specifically, the results emphasize the importance of designing AV systems that communicate their intentions clearly and predictably to pedestrians. Ensuring such transparency and maintaining non-aggressiveness in AV behaviour are essential for improving pedestrian safety and enhancing the overall efficacy of AV-human interactions.

Integration of AV in mixed traffic. In general, the autonomous vehicle proved to be safer during one-to-one interactions with the pedestrian. This is promising for the full integration of AVs in mixed traffic within urban areas. However, the impact of AVs on the other road actors, such as the human-operated vehicles, should also be considered to completely understand the effects of the coexistence of all types of vehicles. In our observations, the AV always stopped to yield priority to pedestrians, while 20% of HVs continued without giving pedestrians their right of way. Considering the post-event phase in our study, the AV waited for the pedestrian to have entirely completed the crossing, and during this time, no other pedestrians occupied the same crosswalk. On the contrary, HVs tended to depart immediately after the pedestrian had left the conflicting area, reflecting the behavior of HV drivers. This suggests that the autonomous vehicle may be considered overly cautious, which could impact the coexistence of different vehicle types by slowing down traffic.

AV settings improvement. The above assessments suggest that the autonomous shuttle configuration settings should be tuned to improve the urban area integration without affecting the safety of interactions. The differences on the conflict measurements we investigated are mainly due to the low operating speed and low accelerating/decelerating rates that were established by the vehicle provider during commissioning activities. This tuning operation might consider the general response of HVs depending on the geographic location and the specific area of the city.

It should be noted that this is the first autonomous driving experiment on public roads under mixed traffic in the city under study. Therefore, a cautious approach was adopted, assuming very low risk. It is precisely through such analyses and ongoing collection of data on vehicle-pedestrian interactions that valuable insights and recommendations can be derived to improve the behaviour of autonomous vehicles.

Impact on micro-simulations assessment. Due to the complexity of working with real-world data, traffic analyses are often conducted in a controlled environment with micro-simulations. For this reason, the assumptions made in these studies must closely reflect real-world conditions. The proposed pipeline enables the extraction of data and its statistical distributions to have a more accurate modelling of autonomous vehicle behaviour and their interactions with pedestrians in traffic micro-simulations. By computing the distributions of parameters for micro-simulation setting and validating them by vehicle type, particularly focusing on speed profiles, this approach contributes to a deeper comprehension of road users' patterns. Therefore, the results here presented can serve as a reference for traffic micro-simulation studies where outcomes need to be validated.

Final remarks and implications. Safer interactions are the lever of success expected by promoters and AV companies. In order for these vehicles to be used in public transport services in the near future, it is necessary to demonstrate their increased safety in the different environmental conditions in which they will operate. While this study confirms that AVs tend to exhibit safer behaviour than HVs in unsignalized crossing scenarios, it is important to consider the potential trade-off between safety and efficiency. AVs, due to their cautious driving behaviour, may experience delays in areas with high pedestrian activity, such as commercial zones. Investigating this trade-off in detail would require additional data collection and analysis across a wider range of scenarios, which represents an important avenue for future research.

Our analyses demonstrate that the pipeline we developed can provide valuable insights for (i) transport operators seeking to validate the feasibility of integrating AV in the urban context; and (ii) AV providers to define the most suitable configuration settings for their vehicles.

This study focused on evaluating safety of an unsignalized crosswalk considering a single lane direction. Future studies could extend this approach to other designs, such as those with varying geometric features (e.g., number of lanes, median refuges) or different warning systems (e.g., Rectangular Rapid-Flashing Beacons, RRFB, or Pedestrian Hybrid Beacons, PHB,). Other possibilities also include exploiting the flexibility of the pipeline to evaluate a broader range of interaction scenarios involving vulnerable road users and environmental conditions (e.g., nighttime crossings). These extensions would enhance the generalizability and applicability of the methodology to a wider variety of real-world contexts.

The future mobility is indeed moving towards the usage of more eco-friendly transportation systems, like bikes and scooters. However, their dynamic behaviour within traffic differs from pedestrians, implying the risk of more hazardous interactions with vehicles. The exploration and safety assessment of autonomous shuttles in interacting with these road users could be an effective indication about the importance of new smart mobility. Another promising future opportunity lies in adopting a multi-modal approach for the detailed analysis of one-on-one interactions between vehicles and other road users. This could involve supplementing video data with additional input from on-board sensors, such as LiDAR and telemetry systems. By integrating these data sources, a more comprehensive and accurate understanding of interactions could be achieved.

CRediT authorship contribution statement

Andrea Avignone: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Data curation. **Marco Bassani:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Beatrice Borgogno:** Writing – original draft, Software, Investigation, Data curation, Conceptualization. **Brunella Caroleo:** Resources, Project administration, Investigation, Formal analysis, Conceptualization. **Silvia Chiusano:** Writing – review & editing, Validation, Supervision, Methodology, Formal analysis. **Federico Princiotto:** Software, Methodology, Data curation.

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.

Acknowledgements

The authors declare that this revised version of the manuscript is the authors' original work and has not been published nor it has been submitted simultaneously elsewhere. Furthermore, all authors have checked the manuscript and have agreed to the submission.

Data availability

The data that has been used is confidential.

References

Antonialli, F., 2019. International benchmark on experimentations with Autonomous Shuttles for Collective Transport. In: 27th International Colloquium of Gerpisa. Paris, France, URL https://centralesupelec.hal.science/hal-02489797.

Arun, A., Haque, M.M., Bhaskar, A., Washington, S., Sayed, T., 2021. A systematic mapping review of surrogate safety assessment using traffic conflict techniques. Accid. Anal. Prev. 153, 106016. http://dx.doi.org/10.1016/j.aap.2021.106016, URL https://www.sciencedirect.com/science/article/pii/S0001457521000476.

Bagloee, S.A., Tavana, M., Asadi, M., Oliver, T., 2016. Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. J. Mod. Transp. 24, 284–303.

Barthauer, M., Friedrich, B., 2019. Presorting and presignaling: A new intersection operation mode for autonomous and human-operated vehicles. Transp. Res. Procedia 37, 179–186. http://dx.doi.org/10.1016/j.trpro.2018.12.181, URL https://www.sciencedirect.com/science/article/pii/S2352146518305969, 21st EURO Working Group on Transportation Meeting, EWGT 2018, 17th – 19th September 2018, Braunschweig, Germany.

Battiato, S., Farinella, G.M., Gallo, G., Giudice, O., 2018. On-board monitoring system for road traffic safety analysis. Comput. Ind. 98, 208–217. http://dx.doi.org/10.1016/j.compind.2018.02.014, URL https://www.sciencedirect.com/science/article/pii/S0166361517305353.

Beauchamp, E., Saunier, N., Cloutier, M.-S., 2022. Study of automated shuttle interactions in city traffic using surrogate measures of safety. Elsevier, Transp. Res. Part C 135, http://dx.doi.org/10.1016/j.trc.2021.103465.

Bradski, G., 2000. The OpenCV Library. Dr. Dobb's J. Softw. Tools.

Chaalal, E., Guerlain, C., Pardo, E., Faye, S., 2023. Integrating connected and automated shuttles with other mobility systems: Challenges and future directions. IEEE Access 11, 83081–83106. http://dx.doi.org/10.1109/ACCESS.2023.3294110.

Du, Y., Zhao, Z., Song, Y., Zhao, Y., Su, F., Gong, T., Meng, H., 2023. StrongSORT: Make deepsort great again. arXiv:2202.13514.

El-Basyouny, K., Sayed, T., 2013. Safety performance functions using traffic conflicts. Saf. Sci. 51 (1), 160–164. http://dx.doi.org/10.1016/j.ssci.2012.04.015, URL https://www.sciencedirect.com/science/article/pii/S0925753512001671.

Ezzati Amini, R., Yang, K., Antoniou, C., 2022. Development of a conflict risk evaluation model to assess pedestrian safety in interaction with vehicles. Accid. Anal. Prev. 175, 106773. http://dx.doi.org/10.1016/j.aap.2022.106773, URL https://www.sciencedirect.com/science/article/pii/S0001457522002081.

Gasper, R., Beutelschieß, S., Krumnow, M., Simon, L., Baksa, Z., Schwarzer, J., 2018. Simulation of autonomous RoboShuttles in shared space. EPiC Ser. Eng. 2, 183–193.

Gettman, D., Pu, L., Sayed, T., Shelby, S.G., Energy, S., et al., 2008. Surrogate Safety Assessment Model and Validation. Tech. Rep., Turner-Fairbank Highway Research Center.

Guan, F., Whitley, T., Xu, H., Wang, Z., Chen, Z., Hui, T., Tian, Y., 2024. Vehicle-pedestrian interaction analysis for evaluating pedestrian crossing safety at uncontrolled crosswalks a geospatial approach using multimodal all-traffic trajectories. J. Saf. Res. 91, 326–341. http://dx.doi.org/10.1016/j.jsr.2024.09.005, URL https://www.sciencedirect.com/science/article/pii/S0022437524001245.

Hydén, C., 1987. The development of a method for traffic safety evaluation: The Swedish Traffic Conflicts Technique. Lund Institute of Technology, Department of Traffic Planning and Engineering.

Hydén, C., 1996. Traffic conflicts technique: state-of-the-art. In: Traffic Safety Work with Video-processing, Green Series, vol. 43, University Kaiserlauten Transportation Department.

Ishaque, M.M., Noland, R.B., 2008. Behavioural issues in pedestrian speed choice and street crossing behaviour: A review. Transp. Rev. 28 (1), 61–85. http://dx.doi.org/10.1080/01441640701365239.

Jafarya, B., Rabieib, E., Diaconeasab, M.A., Masoomib, H., Fiondellaa, L., Moslehb, A., 2018. A Survey on Autonomous Vehicles Interactions with Human and other Vehicles. Probabilistic Saf. Assess. Manag. PSAM URL http://www.iapsam.org/psam14/proceedings/paper/paper_431_1.pdf.

Jay, M., Régnier, A., Dasnon, A., Brunet, K., Pelé, M., 2020. The light is red: Uncertainty behaviours displayed by pedestrians during illegal road crossing. Accid. Anal. Prev. 135, 105369. http://dx.doi.org/10.1016/j.aap.2019.105369, URL https://www.sciencedirect.com/science/article/pii/S0001457519303082.

Kalantari, A.H., Yang, Y., Garcia de Pedro, J., Lee, Y.M., Horrobin, A., Solernou, A., Holmes, C., Merat, N., Markkula, G., 2023. Who goes first? A distributed simulator study of vehicle–pedestrian interaction. Accid. Anal. Prev. 186, 107050. http://dx.doi.org/10.1016/j.aap.2023.107050, URL https://www.sciencedirect.com/science/article/pii/S0001457523000970.

Kastrinaki, V., Zervakis, M., Kalaitzakis, K., 2003. A survey of video processing techniques for traffic applications. Image Vis. Comput. 21 (4), 359–381. http://dx.doi.org/10.1016/S0262-8856(03)00004-0, URL https://www.sciencedirect.com/science/article/pii/S0262885603000040.

Laureshyn, A., De Ceunynck, T., Karlsson, C., Svensson, Å., Daniels, S., 2017. In search of the severity dimension of traffic events: Extended delta-V as a traffic conflict indicator. Accid. Anal. Prev. 98, 46–56. http://dx.doi.org/10.1016/j.aap.2016.09.026, URL https://www.sciencedirect.com/science/article/pii/S0001457516303566.

Laureshyn, A., Svensson, Å., Hydén, C., 2010. Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation. Elsevier, Accid. Anal. Prev. 42, 1637–1646.

Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C.L., Dollár, P., 2015. Microsoft COCO: Common objects in context. arXiv:1405.0312.

Löcken, A., Wintersberger, P., Frison, A.K., Riener, A., 2019. Investigating User Requirements for Communication Between Automated Vehicles and Vulnerable Road Users. http://dx.doi.org/10.1109/IVS.2019.8814027.

Lord, D., 1996. Analysis of pedestrian conflicts with left-turning traffic. Transp. Res. Rec. 1538 (1), 61–67. http://dx.doi.org/10.1177/0361198196153800108.

Madigan, R., Nordhoff, S., Fox, C., Amini, R.E., Louw, T., Wilbrink, M., Merat, N., 2019. Understanding interaction between Automated Road Transport System and other road users: A video analysis. Elsevier, Transp. Res. Part F 66, 196–213. http://dx.doi.org/10.1016/j.trf.2019.09.006.

Massey, Jr., F.J., 1951. The Kolmogorov-Smirnov test for goodness of fit. J. Amer. Statist. Assoc. 46 (253), 68–78. http://dx.doi.org/10.1080/01621459.1951.10500769, arXiv:https://www.tandfonline.com/doi/pdf/10.1080/01621459.1951.10500769, URL https://www.tandfonline.com/doi/abs/10.1080/01621459.1951.10500769.

Merat, N., Louw, T., Madigan, R., Wilbrink, M., Schieben, A., 2018. What externally presented information do VRUs require when interacting with fully Automated Road Transport Systems in shared space? Elsevier, Accid. Anal. Prev. 118, 244–252. http://dx.doi.org/10.1016/j.aap.2018.03.018.

Noh, B., Ka, D., Lee, D., Yeo, H., 2021. Analysis of vehicle–pedestrian interactive behaviors near unsignalized crosswalk. Transp. Res. Rec. 2675 (8), 494–505. http://dx.doi.org/10.1177/0361198121999066.

Noland, R., 2008. Behavioural issues in pedestrian speed choice and street crossing behaviour: A review. Transp. Rev. 28, 61. http://dx.doi.org/10.1080/01441640701365239.

Onelcin, P., Alver, Y., 2017. The crossing speed and safety margin of pedestrians at signalized intersections. Transp. Res. Procedia 22, 3–12. http://dx.doi.org/10.1016/j.trpro.2017.03.002, URL https://www.sciencedirect.com/science/article/pii/S2352146517301369, ‘‘19th EURO Working Group on Transportation Meeting, EWGT2016, 5-7 September 2016, Istanbul, Turkey’’.

Papadoulis, A., Quddus, M., Imprialou, M., 2019. Evaluating the safety impact of connected and autonomous vehicles on motorways. Accid. Anal. Prev. 124, 12–22. http://dx.doi.org/10.1016/j.aap.2018.12.019, URL https://www.sciencedirect.com/science/article/pii/S0001457518306018.

Parkin, J., Crawford, F., Flower, J., Alford, C., Morgan, P., Parkhurst, G., 2023. Cyclist and pedestrian trust in automated vehicles: An on-road and simulator trial. Int. J. Sustain. Transp. 17 (7), 762–774. http://dx.doi.org/10.1080/15568318.2022.2093147.

Parks-Young, A., Sharon, G., 2022. Intersection management protocol for mixed autonomous and human-operated vehicles. IEEE Trans. Intell. Transp. Syst. 23 (10), 18315–18325. http://dx.doi.org/10.1109/TITS.2022.3169658.

Prédhumeau, M., Spalanzani, A., Dugdale, J., 2023. Pedestrian behavior in shared spaces with autonomous vehicles: An integrated framework and review. IEEE Trans. Intell. Veh. 8 (1), 438–457. http://dx.doi.org/10.1109/TIV.2021.3116436.

Razmi Rad, S., Farah, H., Taale, H., van Arem, B., Hoogendoorn, S.P., 2020. Design and operation of dedicated lanes for connected and automated vehicles on motorways: A conceptual framework and research agenda. Transp. Res. C 117, 102664. http://dx.doi.org/10.1016/j.trc.2020.102664, URL https://www.sciencedirect.com/science/article/pii/S0968090X20305799.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: Unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition. CVPR, pp. 779–788. http://dx.doi.org/10.1109/CVPR.2016.91.

Sacchi, E., Sayed, T., 2016. Conflict-based safety performance functions for predicting traffic collisions by type. Transp. Res. Rec. 2583 (1), 50–55. http://dx.doi.org/10.3141/2583-07.

Sacchi, E., Sayed, T., deLeur, P., 2013. A comparison of collision-based and conflict-based safety evaluations: The case of right-turn smart channels. Accid. Anal. Prev. 59, 260–266. http://dx.doi.org/10.1016/j.aap.2013.06.002, URL https://www.sciencedirect.com/science/article/pii/S0001457513002303.

Tarko, A.P., 2012. Use of crash surrogates and exceedance statistics to estimate road safety. Accid. Anal. Prev. 45, 230–240. http://dx.doi.org/10.1016/j.aap.2011.07.008, URL https://www.sciencedirect.com/science/article/pii/S000145751100193X.

Tarko, A., 2019a. Measuring Road Safety with Surrogate Events. Elsevier.

Tarko, A., 2019b. Measuring Road Safety with Surrogate Events. Elsevier.

Wu, M., Jiang, H., Tan, C.-A., 2021. Automated parking space allocation during transition with both human-operated and autonomous vehicles. Appl. Sci. 11 (2), http://dx.doi.org/10.3390/app11020855, URL https://www.mdpi.com/2076-3417/11/2/855.

Yang, D., Ozbay, K., Xie, K., Yang, H., Zuo, F., Sha, D., 2021. Proactive safety monitoring: A functional approach to detect safety-related anomalies using unmanned aerial vehicle video data. Transp. Res. C 127, 103130. http://dx.doi.org/10.1016/j.trc.2021.103130, URL https://www.sciencedirect.com/science/article/pii/S0968090X21001492.

Zheng, L., Sayed, T., Mannering, F., 2021. Modeling traffic conflicts for use in road safety analysis: A review of analytic methods and future directions. Anal. Methods Accid. Res. 29, 100142. http://dx.doi.org/10.1016/j.amar.2020.100142, URL https://www.sciencedirect.com/science/article/pii/S2213665720300324.

Zhu, Z., Hu, Z., Dai, W., Chen, H., Lv, Z., 2022. Deep learning for autonomous vehicle and pedestrian interaction safety. Elsevier, Saf. Sci. 145, http://dx.doi.org/10.1016/j.ssci.2021.105479.