Pedestrian Planet: What YouTube Driving from 233 Countries and Territories Teaches Us About the World

Md Shadab Alam*
m.s.alam@tue.nl
Eindhoven University of Technology
Eindhoven, The Netherlands

Marieke H. Martens
m.h.martens@tue.nl
Eindhoven University of Technology
Eindhoven, The Netherlands

Pavlo Bazilinskyy
p.bazilinskyy@tue.nl
Eindhoven University of Technology
Eindhoven, The Netherlands

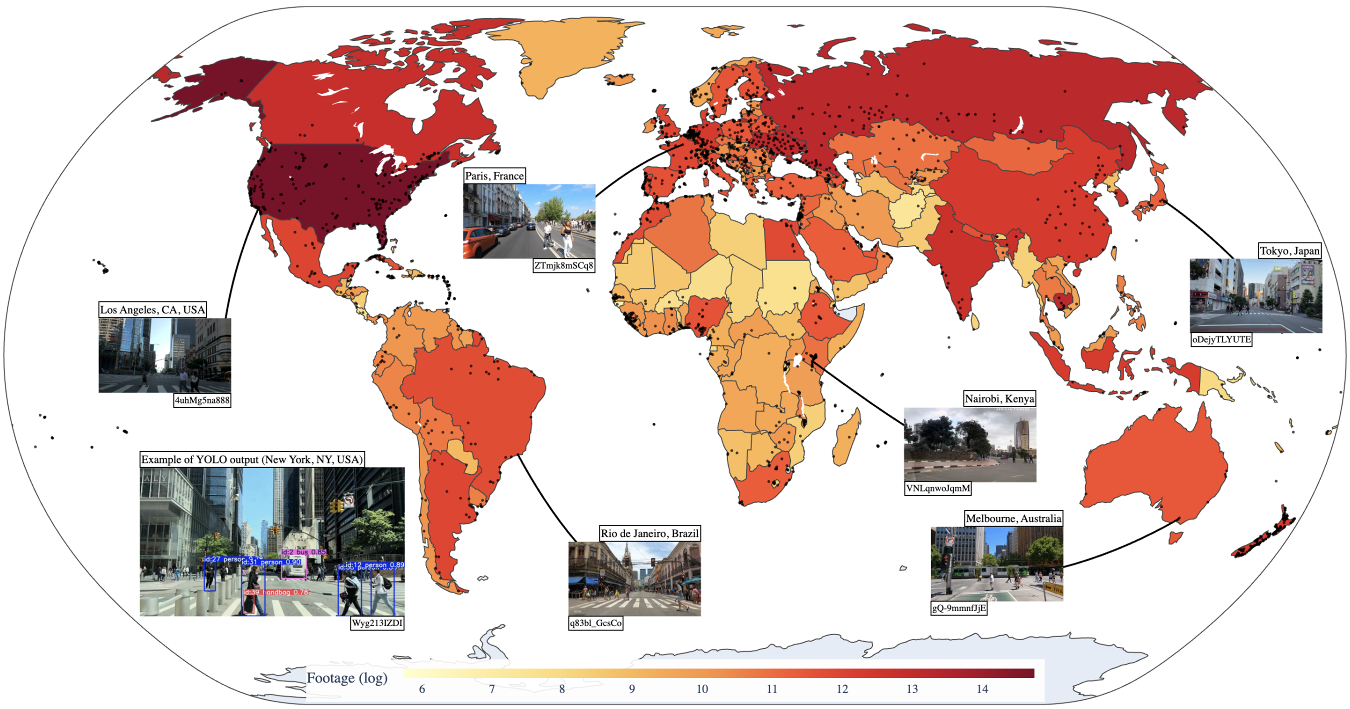

Figure 1: The 233 countries and territories with dashcam footage in CROWD dataset [2]. The colouration is based on the logarithm of the total recorded time per country or territory, calculated as $\log_e(1 + \text{time in seconds})$, to reduce the skew from outliers such as the United States (with 707.76 hours available). The black dots show the 2,495 cities in the dataset. The labels under images show the corresponding YouTube video ID. The frame on the bottom left shows an example of object detection using YOLOv11x with identified objects such as pedestrians, vehicles, and traffic signs; in this image, the labels 'id' refer to the unique ID of the detected object with the type mentioned later and end with the confidence of detection of the object.

Abstract

Pedestrian crossing behaviour varies globally. This study analyses dashcam footage from the CROWD dataset, covering 233 countries and territories, to examine crossing initiation time, crossing speed, and contextual variables, including detected vehicles, traffic mortality, GDP, and Gini coefficient. Qatar had the longest mean crossing initiation time (6.44 s), while China exhibited the fastest crossing speed (1.69 m/s). On average, worldwide, pedestrians exhibited a crossing initiation time of 3.18 s and crossing speed 1.20 m/s. Crossing speed and crossing initiation time are negatively correlated ($r = -0.18$), indicating slower crossings after longer hesitation. Crossing speed is negatively correlated with Gini coefficient ($r = -0.19$) and positively correlated with traffic mortality ($r = 0.18$). Similar crossing times in countries with different infrastructures, such as Bangladesh (3.42 s) and the Netherlands (3.40 s), underscore the complex interaction between infrastructure and behavioural adaptation. These findings emphasise the importance of culturally aware road design and the development of adaptive interfaces for vehicles.

*Corresponding Author

AutomotiveUI '25, Brisbane, QLD, Australia

ACM ISBN 979-8-4007-2013-0/2025/09
https://doi.org/10.1145/3744333.3747827

CCS Concepts

• **Computing methodologies** → **Cross-validation**; **Object detection**; **Tracking**; • **Mathematics of computing** → *Probability*

and statistics; • **Human-centered computing** → Human computer interaction (HCI).

Keywords

Dashcam Videos, Cross-country Analysis, Pedestrian Behaviour, Crossing Speed, Crossing Initiation Time

ACM Reference Format:

Md Shadab Alam, Marieke H. Martens, and Pavlo Bazilinskyy. 2025. Pedestrian Planet: What YouTube Driving from 233 Countries and Territories Teaches Us About the World. In *17th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '25), September 21–25, 2025, Brisbane, QLD, Australia.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3744333.3747827

1 Introduction

Every year, road traffic accidents claim more than 1.19 million lives around the world. In particular, more than half of global road traffic fatalities occur among pedestrians, cyclists, and motorcyclists, especially in low- and middle-income countries [42]. This highlights that the burden of road traffic fatalities is not evenly distributed throughout the world. Significant differences in traffic safety exist between countries, influenced by factors such as infrastructure quality, enforcement of regulations, and cultural norms [3, 40]. To illustrate an example, one may compare two countries: India and The Netherlands. The Netherlands, characterised by advanced traffic management systems, comprehensive pedestrian infrastructure, and strict road safety regulations, exhibits substantially lower traffic mortality rates compared to India. India's rapid urbanisation combined with less consistent enforcement of traffic regulations contributes to significantly higher mortality rates, with reported figures of 4.0 deaths per 100,000 inhabitants in the Netherlands compared to 15.6 in India, respectively [43]. India alone accounts for approximately 10% of fatalities in road crashes worldwide. In this context, analysing why things go wrong on the road in various parts of the world is crucial to improving global traffic safety.

Research has shown that a significant proportion of road accidents occur at pedestrian crossings, often due to complex interactions between drivers and pedestrians. For example, studies have found that pedestrian errors—such as unpredictable crossing decisions, misjudging vehicle speed, or not using designated crossings—are key contributors to accidents [31, 33]. Similarly, driver failure to yield, poor visibility, and inadequate infrastructure can further increase the risk of collisions [9, 22]. Understanding these dynamics highlights that pedestrian crossing behaviour is an important factor influencing road safety in diverse urban settings. Therefore, a comprehensive investigation of pedestrian behaviour, particularly in crossing situations, is essential for developing effective interventions and policies that reduce accidents and save lives. To mitigate these risks, recent developments in automated vehicles (AVs) and advanced driver assistance systems (ADAS) leverage smart sensors and active safety features to better detect and respond to VRUs. By compensating for potential human errors, these technologies can help prevent accidents; however, this requires a deep understanding of pedestrian behaviour under a wide range of environmental and cultural conditions [24, 30, 33, 48]. Therefore, analysing how pedestrians behave in diverse real-world contexts is essential for developing robust, context-aware AV and ADAS systems.

Although research has been done on pedestrian crossing behaviour [16, 37], much of this knowledge is limited to controlled environments, single-city studies, or limited cultural contexts [31]. Shi et al. studied pedestrian behaviour, such as walking speed, waiting delay, and clustering, at a single crossing in Beijing, China, through which it is obtained by counting and measuring with a video camera [35]. Similarly, Deb et al. conducted a survey in the US with 50 survey items that allowed respondents to rate the frequency with which they engage in different types of road use behaviour as pedestrians [9]. The validation study was conducted on 425 participants (228 males and 197 females) with the age between 18 and 71. Rasouli et al. gathered data from 5 different countries, collecting a total of 240 hours of footage, and identified 2,400 pedestrians to understand how pedestrians communicate intent (especially before crossing) [30]. The narrow geographic coverage, small sample sizes, and use of artificial or highly controlled settings in these studies restrict the ability to develop truly generalisable algorithms for pedestrian detection, tracking, and prediction.

The researchers have previously studied specific variables associated with pedestrian behaviour, such as speed of crossing. For example, Goh et al. conducted a study in Kuala Lumpur, Malaysia, with 1,579 participants in 4 different locations [15]. They concluded that children pedestrians (<20 years) are the fastest group, and elderly pedestrians (>55 years) are the slowest group in terms of pedestrian crossing speed with a population average speed of 1.31 m/s at signalised crossing and 1.39 m/s at unsignalised crossing. At the same time, Duim et al. conducted a study with 1,911 participants with an average age of 70.1 years in São Paulo, Brazil [10]. They found that their average walking speed was 0.75 m/s (95%CI 0.73; 0.84). However, while valuable in understanding pedestrian behaviour in specific locations and age groups, these studies are limited to particular cities or populations. As a result, they do not enable a comprehensive comparison of pedestrian behaviour across different countries, cultures, and infrastructure contexts, highlighting the need for broader, cross-national analyses such as the one presented in this paper.

Another variable of interest is the crossing initiation time, which in this study is defined as the time interval between when a pedestrian is identified as intending to cross the road and when they actually begin to step onto the roadway. This metric captures the delay between the formation of the intention to cross and the initiation of the crossing action. Previous studies have typically anchored the crossing initiation time to external events, such as the onset of a traffic signal change [39] or the appearance of a safe gap in traffic [38]. Wickramasinghe et al. studied pedestrians who crossed the road in three different pedestrian crossings, namely, signal configurations: (1) traditional red and green phase, (2) countdown with time to start crossing phase, and (3) countdown without time to start crossing phase [41]. The authors then provided a multiregression model to estimate the time to start crossing the road. Lobjois & Viola examined the effects of age, vehicle speed, and time constraints on the selection of gaps in crossing decisions, finding that older pedestrians (aged 60 to 80 years) tend to choose larger gaps to compensate for longer crossing times, yet may experience reduced safety margins at higher car speed, for example, unsafe decision

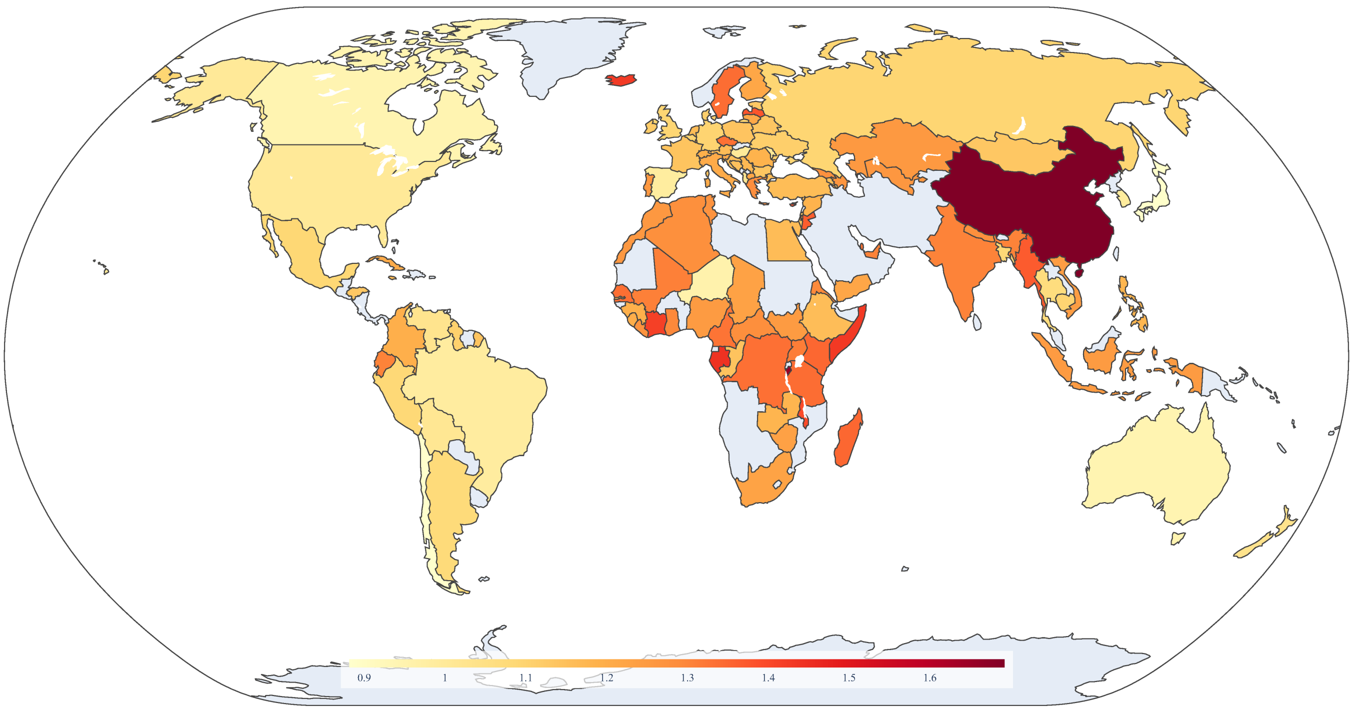

Figure 2: Mean pedestrian crossing speed (in m/s). For each country or territory, values represent the average of daytime and nighttime speeds (if both are present); if only daytime or nighttime data are available, the respective value is shown.

rates (i.e., accepting a gap shorter than their own crossing time) increased from 3.1% at 40 km/h to 8.6% at 60 km/h [22]. Furthermore, Simeunović et al. found that the presence of a countdown timer at pedestrian crossings, which tells people exactly how many seconds remain before vehicles get a green light, actually encourages more pedestrians to start crossing after the "walk" signal has ended (during what is called the "clearance phase"). Specifically, 65.7% of pedestrians began crossing during this period when a timer was present, compared to only 22% at crossings without a timer. This behaviour also resulted in pedestrians spending 1.3 times longer on the road during the red signal, increasing their exposure to traffic [36]. Their analysis further suggests that, although showing pedestrians exactly how much time remains to cross (for example, with a countdown timer) can help people cross the street with less waiting, it can also unintentionally encourage riskier behaviour, such as starting to cross when there is not actually enough time left, if pedestrians overestimate how quickly they can cross before the light changes.

These approaches, although methodologically sound for their specific contexts, do not capture the rich diversity of global urban environments, thereby hampering the development of truly generalisable models for pedestrian detection, tracking, and prediction. The increasing ubiquity of video hosting platforms, notably YouTube (https://www.youtube.com), has fundamentally transformed the means by which researchers can acquire real-world data. These platforms offer a large repository of publicly available content that serves a multitude of analytical purposes. Recent studies have shown that large-scale dashcam video data, as employed by Franchi et al. [13], can effectively capture the diverse characteristics of urban traffic in different cities, providing compelling motivation for our approach using footage obtained from YouTube. Similarly, video-based research such as that conducted by Rao et al. [29] has successfully delineated the subtleties of pedestrian crossing decisions, thus underscoring the value of dashcam footage in analysing cross-cultural behaviours. These platforms afford a unique opportunity to record and scrutinise authentic urban scenes ranging from pedestrian dynamics to traffic behaviour across varied geographical and cultural contexts. In contrast to controlled field studies, freely available videos on platforms such as YouTube encapsulate spontaneous and natural interactions, rendering them invaluable for the development of scalable and cost-effective methodologies to study complex urban phenomena. In this context, Alam et al. [2] have compiled a dataset, entitled "City Road Observations With Dashcams (CROWD)", applying computer vision techniques to dashcam videos, allowing a comparative analysis of pedestrian behaviour in different cities around the world.

1.1 Aim of Study

The purpose of the study is to provide a comprehensive global analysis of pedestrian behaviour using the **City Road Observations With Dashcams (CROWD)** dashcam footage from 233 countries and territories. Using its YOLO-based object tracking, it quantifies

Hyperparameter	Value	Description
track high thresh	0.25	Threshold for the first association during tracking.
track low thresh	0.1	Threshold for the second association during tracking.
new track thresh	0.25	Threshold to initialise a new track if the detection does not match any existing tracks.
track buffer	30	Number of frames lost tracks should be kept alive before getting removed.
match thresh	0.8	Threshold for matching tracks.
fuse score	True	Determines whether to fuse confidence scores with IoU distances before matching.

Table 1: Tracking hyperparameters and their descriptions.

Parameter	Value	Description
min shared frames	5	Minimum number of frames where both person and vehicle are present for comparison.
dist thresh	80	Maximum distance (in pixels) between person and vehicle centers to be considered "moving together".
similarity thresh	0.8	Minimum cosine similarity threshold for movement direction to be considered "similar" (range: -1 to 1).
overlap ratio	0.7	Fraction of overlapping frames where proximity and movement similarity must be satisfied.

Table 2: Parameters for trajectory-based filters. The first three parameters correspond to the filter that determines whether a detected person is likely a rider (bicycle or motorcycle) and should be excluded from pedestrian analysis. `overlap ratio` is used by the filter that removes pedestrians whose trajectories may mimic a crossing event due to camera movement.

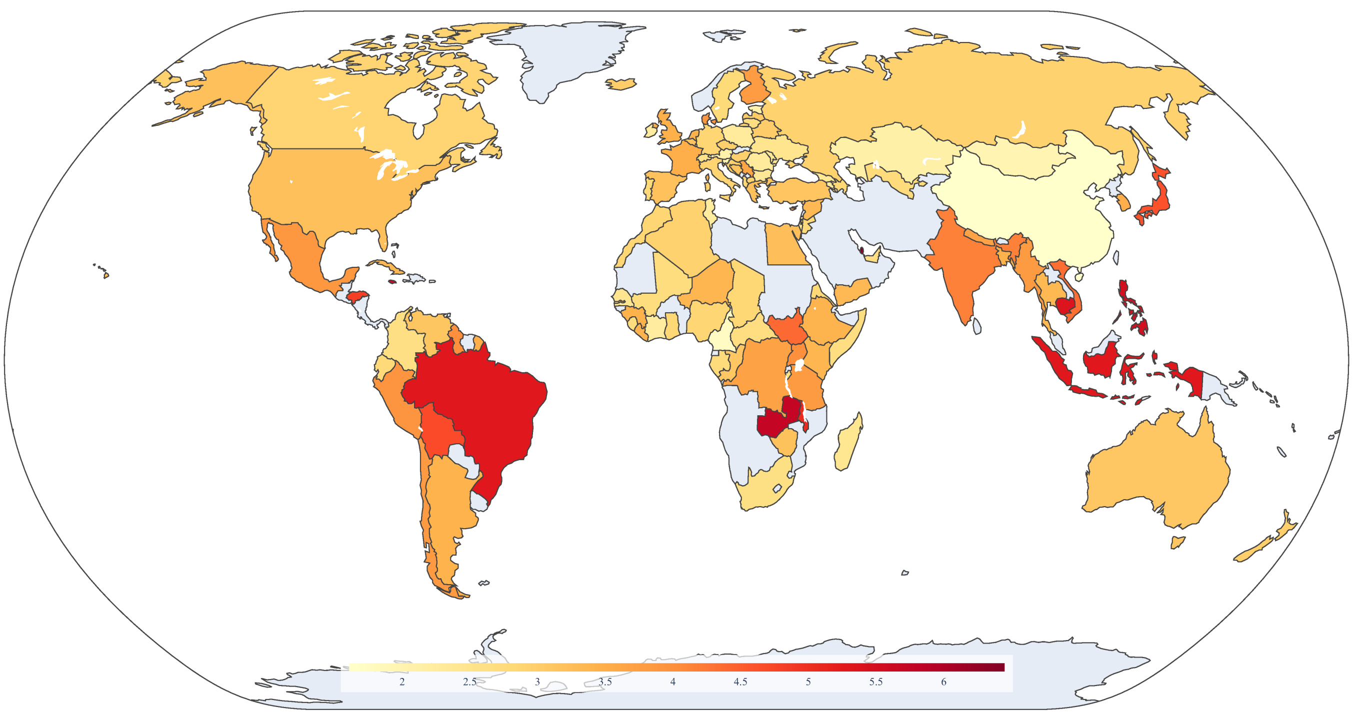

Figure 3: Mean pedestrian crossing initiation time (in s). For each country or territory, values represent the average of daytime and nighttime times (if both are present); if only daytime or nighttime data are available, the respective value is shown.

key metrics, such as the crossing initiation time and the speed of crossing, to capture diverse behaviour across varied cultural, socioeconomic, and infrastructural environments. By investigating the relationships and distributions of these metrics worldwide, this study demonstrates how the CROWD dataset can be used as a powerful tool for cross-cultural pedestrian research. Our work serves as a use case that shows how such large-scale real-world video data can generate valuable insights for urban planning and for the future integration and safety of AVs in diverse global contexts.

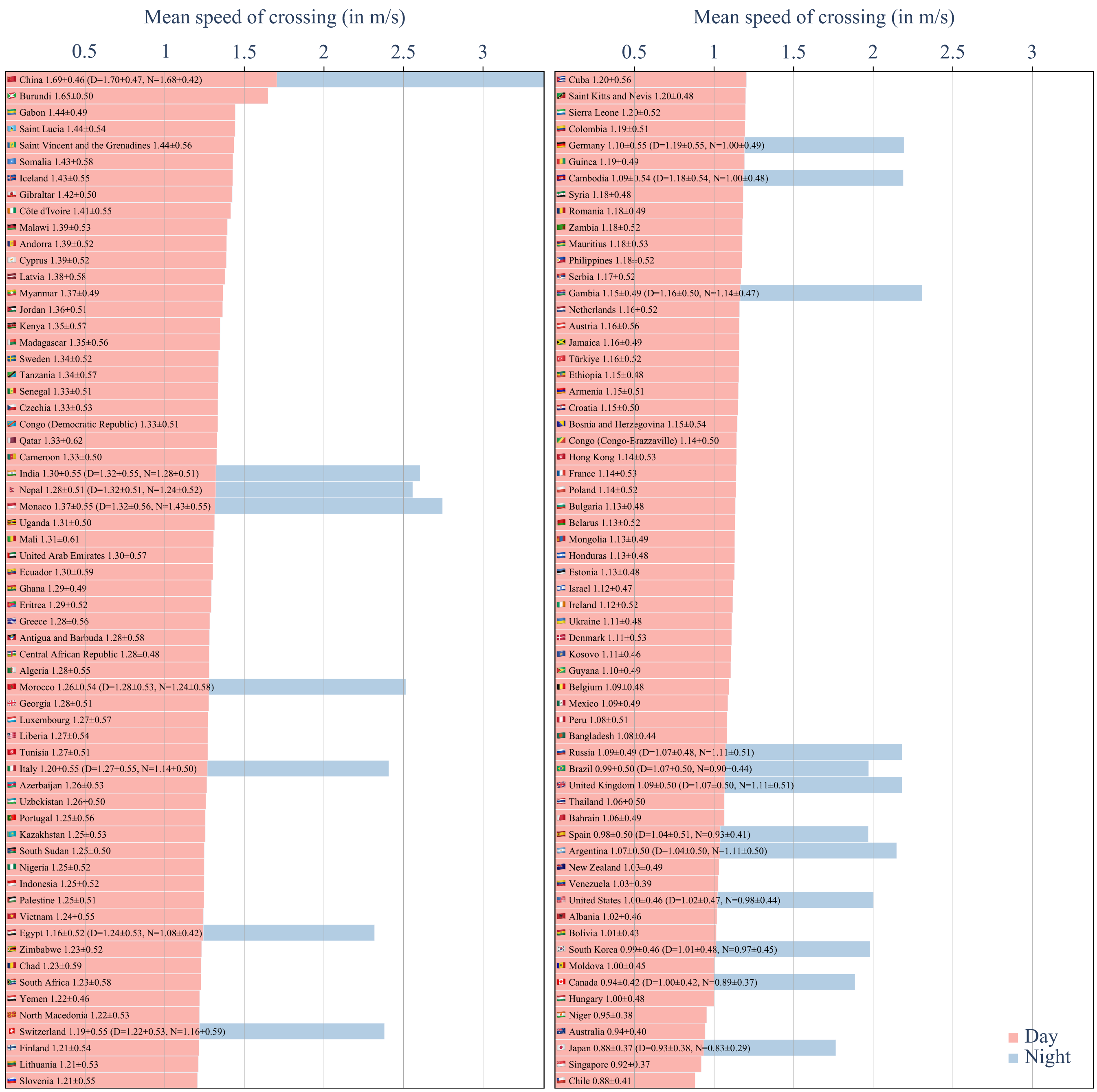

Figure 4: Mean pedestrian crossing speed by country or territory, shown separately for daytime (red) and nighttime (blue) observations. The mean and standard deviation (SD) of the overall crossing speed are displayed next to the country name, followed by the daytime (D) and nighttime (N) values (mean ± SD) in parentheses. Bars are sorted in ascending order based on the daytime mean speed for each country; if daytime data are unavailable, nighttime values are used to order.

2 Method

The study uses the videos referenced in the CROWD dataset [2] (version as of 24 July 2025), which includes dashcam footage sourced from publicly accessible YouTube content, to examine pedestrian behaviour in urban driving environments. The dataset provides information for each video or video segment, whether it was captured during the day or at night, allowing for analyses that directly compare pedestrian behaviour under different lighting conditions.

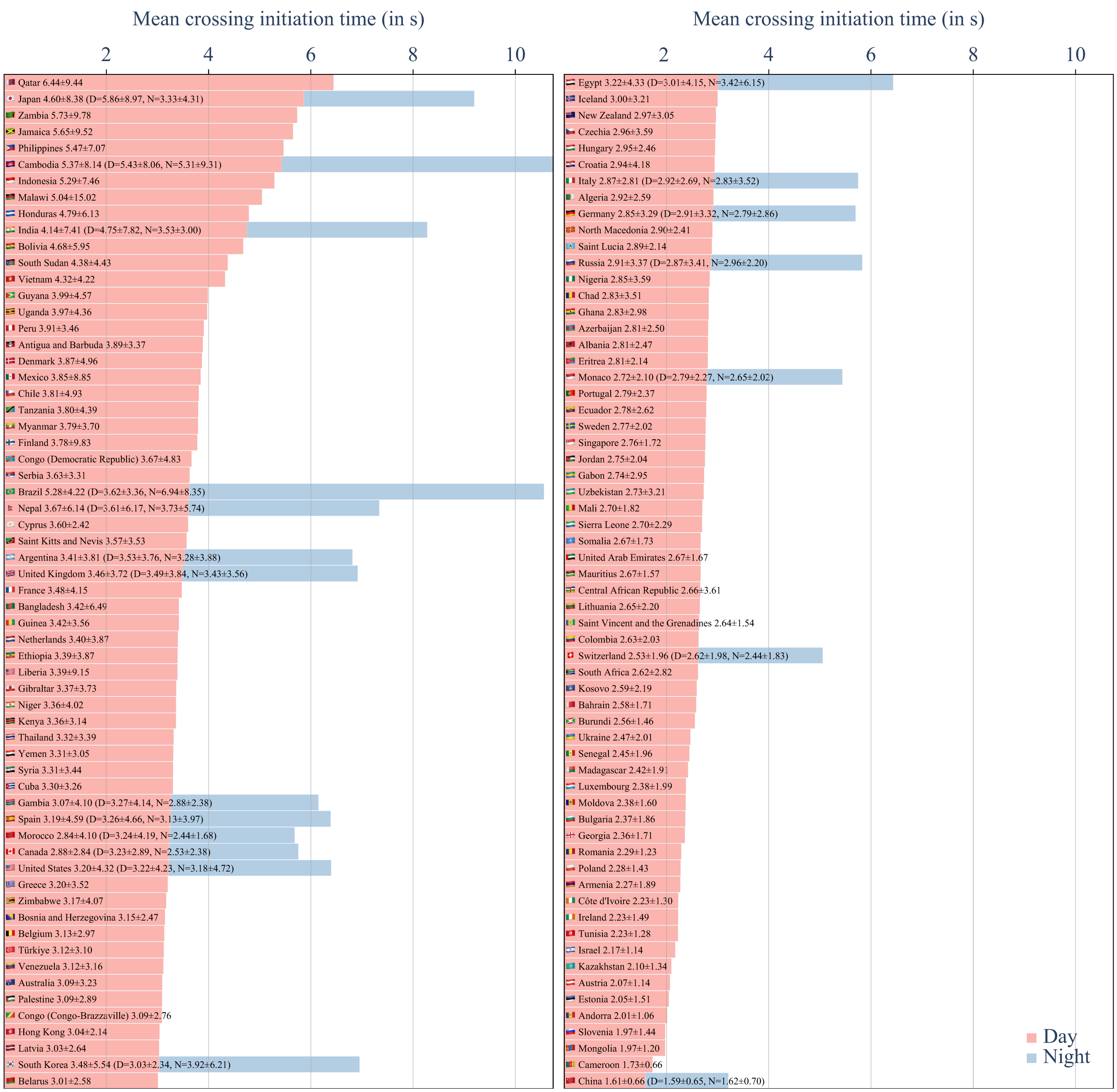

Figure 5: Mean pedestrian crossing initiation time by country or territory, shown separately for daytime (red) and nighttime (blue) observations. The mean and standard deviation (SD) of the overall crossing initiation time are displayed next to the country name, followed by the daytime (D) and nighttime (N) values (mean ± SD) in parentheses. Bars are sorted in ascending order based on the daytime mean speed for each country; if daytime data are unavailable, nighttime values are used to order.

In total, 8,494 videos were extracted with a total of 4,122.28 hours of footage that span 233 countries and territories (see Figure 1).

To analyse pedestrian behaviour, we first applied the You Only Look Once (YOLOv11x) object detection algorithm [17, 32] with the ByteTrack tracker [49] (see Table 1), using a confidence score threshold of 0.7. This enabled robust detection and tracking of various objects, including pedestrians, in video frames. Only video clips captured in cars were retained for analysis, as pedestrian reactions can vary by vehicle type [11].

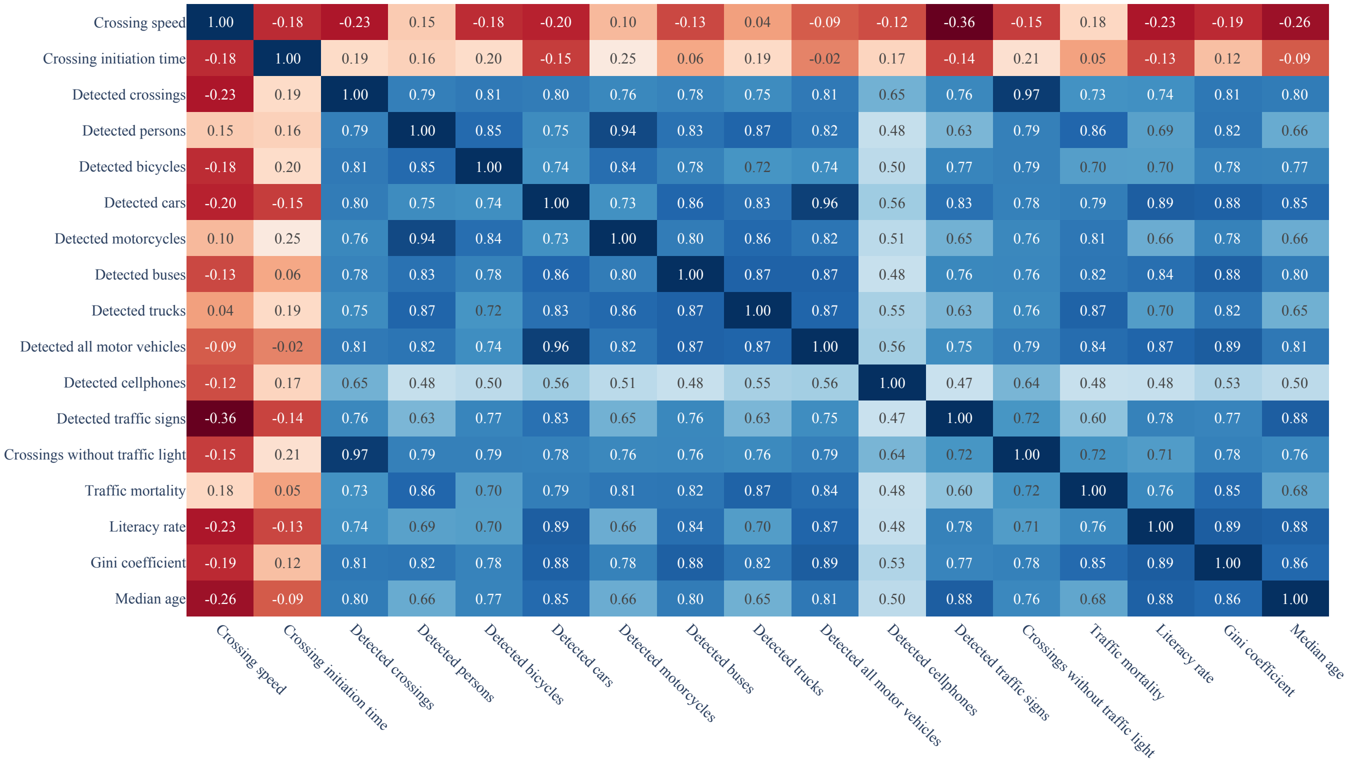

Figure 6: Spearman correlation matrix of pedestrian behaviour, traffic characterises, and socio-economic factors for 124 countries and territories included in analysis.

Using the generated object tracks, we identified pedestrian crossing events. A crossing was operationally defined as an event in which a pedestrian's horizontal trajectory moved from below 45% to above 55% of the frame width, or vice versa, indicating a lateral traversal across the road. To reduce the inclusion of irrelevant movements, filters were applied to the detected crossings. Events were excluded if the crossing was made by a rider (motorbike or bicycle), or if the apparent crossing was caused by camera movement—such as when a car turns and the resulting pixel movement of nearby individuals mimics a crossing event—as these can result in false positives. It should be noted that the YOLOv11x model detects "person", "bicycle", and "motorcycle" as separate object classes, rather than explicitly identifying riders. Therefore, additional trajectory-based filtering was necessary to distinguish between actual pedestrians and individuals riding bicycles or motorcycles. The specific logic and implementation details of these filters are further described by Alam et al. [2], with the corresponding parameter settings summarised in Table 2.

For all events passing these filters, the distances based on pixels were translated into real-world units using the national average human height as a reference scale in each frame [51]. This conversion allowed for a more meaningful estimation of parameters such as the crossing distance and speed, although individual height variability and camera perspective can still introduce some uncertainty. Next, thresholds were applied to exclude implausible measurements: crossing speed values lower than 0.5 m/s and greater than 2.5 m/s were excluded, effectively removing outliers such as skateboarders or cyclists, as recommended in previous work [7, 20, 26]. Extremely slow crossings could reflect rare cases, such as individuals with severe mobility impairments [4], people carrying heavy loads [7], or distracted pedestrians, but are likely to result from tracking errors in the context of the dataset. For the analysis of the crossing initiation time, only values between 1 s and 150 s (following [12]) were retained. To verify movement at the beginning of the crossing, we checked for movement at three evenly spaced intervals within the first second (e.g., at the 10th, 20th, and 30th frame for a video with 30 frames per second). A margin of ±0.1 times the mean height was allowed when assessing movement at these intervals to compensate for potential camera motion.

Only after all relevant crossing events and their metrics were determined did we apply additional filters based on duration of footage and vehicle type. Dashcam footage from cities from the country or territory was included if there was at least 30 minutes of available driving footage to adequately represent a range of urban scenarios. Using the CROWD dataset metadata, the analyses were stratified by time of day (day versus night). Inclusion for each period was determined independently: a country or territory was included in the daytime or nighttime analysis if at least 100 valid pedestrian crossings were detected during that respective period. For example, if a location met the threshold for daytime crossings but not for nighttime, only its daytime data were included. This approach guarantees that each subset of the analysis—daytime

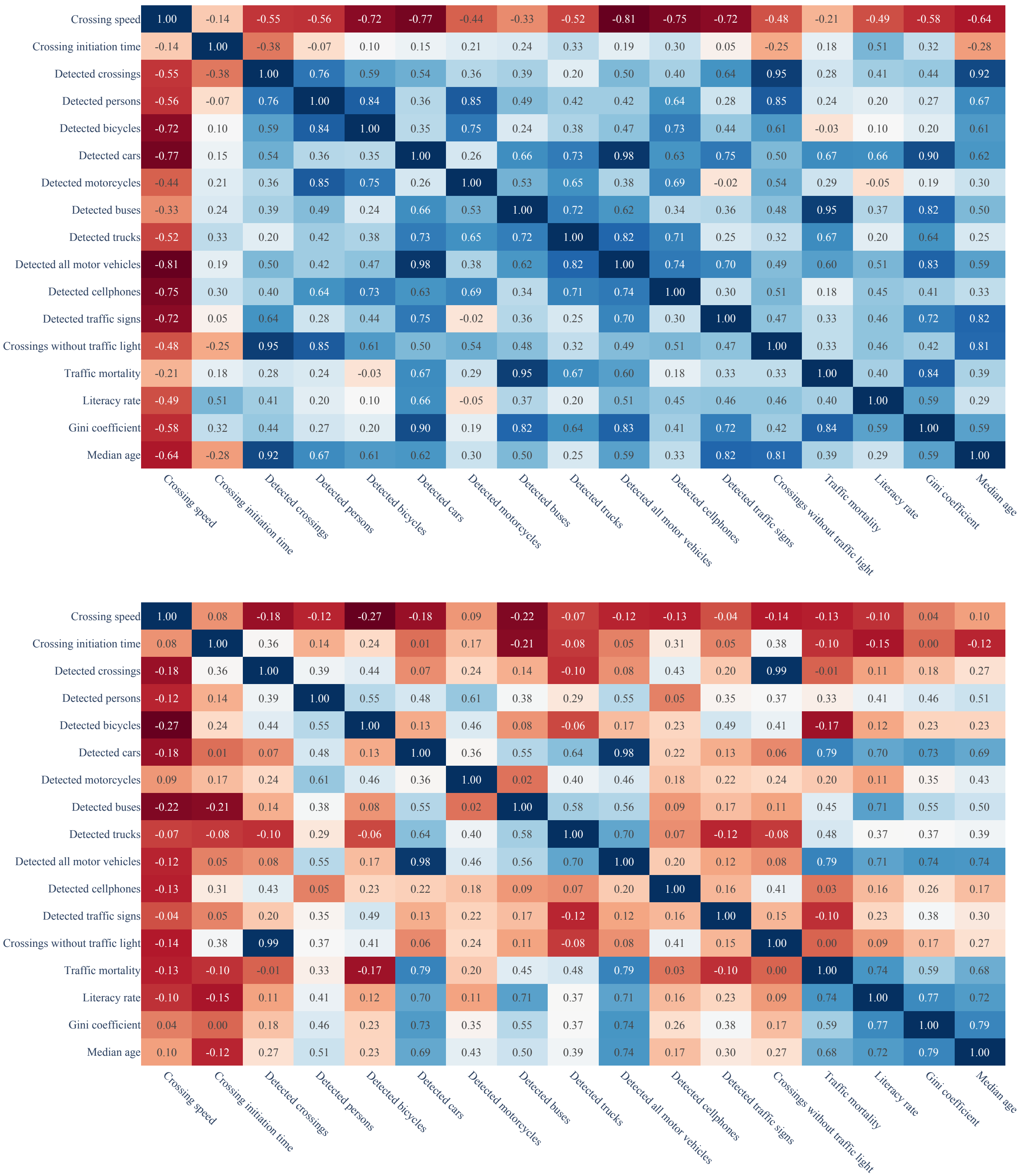

Figure 7: Spearman correlation matrix of pedestrian behaviour, traffic characterises, and socio-economic factors in North America (10 countries and territories, top) and Europe (40 countries and territories, bottom).

and nighttime—is based on robust and sufficiently sized samples, supporting meaningful cross-period comparisons.

The administrative and dependent territories were treated as separate analysis units whenever they possessed unique ISO 3166-1 alpha-3 codes (i.e., "ISO-3", https://www.iso.org/iso-3166-country-codes.html). For the purposes of continent-based analysis, each territory was grouped according to its geographic location, not its sovereign affiliation (for example, French Guiana was considered part of South America rather than Europe, despite being an overseas department of France).

To enhance our analysis, we incorporated several key statistical indicators for each country or territory. These include population size (sourced from https://restcountries.com), road traffic mortality rates measured as the number of traffic-related deaths per 100,000 inhabitants (https://data.worldbank.org/indicator/SH.STA.TRAF.P5), and income inequality as reflected by the Gini coefficient (https://restcountries.com). We also included regional literacy rates (https://data.worldbank.org/indicator/SE.ADT.LITR.ZS) and the median age of the population (https://simplemaps.com/data/countries). Together, these variables provide socioeconomic context for interpreting patterns in pedestrian behaviour across different regions.

3 Results

To supplement the results below, the distributions of the crossing speed and the crossing initiation time in the unfiltered dataset are presented in Figure A1 (Appendix). The results for the crossing speed and the crossing initiation time with stricter filter values (at least 10 hours of footage and at least 500 valid pedestrian crossing detections, either daytime or nighttime, per country or territory) are shown in Figure A2 and Figure A3, respectively.

Following the filtering process detailed in section 2, the resulting data based on CROWD comprise 3,388.88 hours of dashcam video from 124 countries and territories. The distributions of the crossing speed and the crossing initiation time for the filtered dataset are shown in Figure A4.The country with the maximum duration of the video is the United States, with 707.76 hours of footage, while the country with the minimum duration is Niger, with 0.65 hours. The average video duration per country is 27.32 hours (SD = 67.59).

The analysis of the filtered subset yielded several notable results. The mean pedestrian crossing speed among the countries and territories included is 1.20 m/s (SD = 0.14), where each country or territory contributes a single mean value, regardless of the number of pedestrian crossings detected within it. In our analysis, China exhibited the highest pedestrian crossing speed observed (1.69 m/s), while Chile showed the lowest (0.88 m/s). Figure 2 illustrates the average speed of pedestrian crossings in the countries and territories analysed. Figure 4 shows the distribution of the crossing speed during the day and night in these locations. The largest absolute differences between daytime and nighttime crossing speeds were observed in Germany (0.18 m/s, *daytime faster*), Cambodia (0.18 m/s, *daytime faster*), Egypt (0.16 m/s, *daytime faster*), Brazil (0.16 m/s, *daytime faster*), and Italy (0.12 m/s, *daytime faster*). These values represent the absolute differences between the mean speeds recorded during the day and at night, with the direction of the difference indicated for each country. In contrast, countries or territories with minimal differences between daytime and nighttime crossing speeds include the United States (0.04 m/s, *nighttime faster*), India (0.03 m/s, *daytime faster*), Russia (0.03 m/s, *nighttime faster*), China (0.02 m/s, *daytime faster*) and Gambia (0.02 m/s, *daytime faster*).

Similarly, in our analysis, the average pedestrian crossing initiation time is 3.18 s (SD = 0.87). Among the countries and territories included, pedestrians in Qatar have the longest observed crossing initiation time (6.44 s), while those in China have the shortest (1.61 s). Figure 3 illustrates the average crossing initiation time for pedestrians before starting the crossing in the locations analysed, and Figure 5 shows the distribution of the initiation time during the day and night in these countries and territories. The locations with the greatest absolute differences between daytime and nighttime initiation times include Brazil (3.31 s, *nightime longer*), Japan (2.53 s, *daytime longer*), India (1.22 s, *daytime longer*), South Korea (0.89 s, *nighttime longer*) and Morocco (0.80 s, *daytime longer*). In contrast, the smallest differences were found in Italy (0.09 s, *daytime longer*), Russia (0.09 s, *nighttime longer*), the United Kingdom (0.06 s, *daytime longer*), the United States (0.04 s, *daytime longer*) and China (0.03 s, *nighttime longer*).

Figure 6 presents the correlation matrix based on the Spearman rank correlation coefficient among pedestrian characteristics, detected objects, and socioeconomic factors. The heatmap shows that the detected crossings are highly positively correlated with the detected persons ($r = 0.79$), bicycles ($r = 0.81$), cars ($r = 0.80$), motorcycles ($r = 0.76$), buses ($r = 0.78$), trucks ($r = 0.75$) and traffic signs ($r = 0.76$). The detected cars are strongly correlated with the total number of motor vehicles ($r = 0.96$) and with the traffic signs ($r = 0.83$). Traffic mortality exhibits high positive correlations with detected persons ($r = 0.86$), cars ($r = 0.79$), total number of motor vehicles ($r = 0.84$), and the Gini coefficient ($r = 0.85$). The Gini coefficient is also positively correlated with detected persons ($r = 0.82$) and cars ($r = 0.88$). The median age shows a positive correlation with the detected persons ($r = 0.68$) and the cars ($r = 0.87$). Crossing speed is positively correlated with detected persons ($r = 0.15$), but negatively correlated with cellphones ($r = -0.12$), traffic signs ($r = -0.36$), the Gini coefficient ($r = -0.19$), median age ($r = -0.31$), and crossing initiation time ($r = -0.18$). Crossing initiation time is negatively correlated with the detected cars ($r = -0.15$) and the median age ($r = -0.09$), but positively correlated with detected motorcycles ($r = 0.25$).

Furthermore, Figure 7 presents the correlation matrices for the same behavioural and contextual attributes in North America and Europe, respectively. In North America, the correlation between median age and detected crossing is ($r = 0.92$), and with detected persons ($r = 0.67$); In Europe, these values are ($r = 0.24$) and ($r = 0.51$), respectively. The correlation between crossing speed and detected motorcycles is ($r = -0.44$) in North America and ($r = 0.09$) in Europe; with detected bicycles, ($r = -0.72$) in North America and ($r = -0.27$) in Europe; and with detected cars, ($r = -0.77$) in North America and ($r = -0.18$) in Europe. For crossing initiation time, its correlation with the detected bicycles is ($r = 0.10$) in North America and ($r = 0.24$) in Europe; with detected motorcycles, ($r = 0.21$) and ($r = 0.17$), respectively; and with the median age, ($r = -0.28$) in North America and ($r = -0.12$) in Europe. Furthermore, the correlation between crossing speed and crossing initiation time is ($r = -0.14$) in North America and ($r = 0.08$) in Europe, and the correlation between traffic mortality and detected cars is ($r = 0.67$) in North America and ($r = 0.79$) in Europe.

In addition, the scatter plots presented in Figure 8 and Figure 9 reveal significant variations between countries and territories with respect to socio-economic and safety indicators. Figure 8 shows a negative relationship ($r = -0.19$) between the Gini coefficient and the crossing speed. Furthermore, Figure 9 illustrates the variability in crossing speed relative to the mortality rates of national traffic with a positive correlation ($r = 0.18$).

4 Discussion

The findings presented in this study highlight significant global variations in pedestrian crossing behaviour and underline the influence of regional and country-specific socioeconomic and infrastructure factors. The mean crossing speed (1.20 m/s) aligns closely with previous studies of pedestrian behaviour [4, 27, 44], but the substantial deviations observed in countries and territories such as China (highest speed, 1.67 m/s) and Chile (lowest speed, 0.88 m/s) reflect considerable cultural and infrastructure differences.

A global analysis of pedestrian crossing speed reveals that neighbouring countries and territories often cluster around a similar average walking speed, a pattern strongly influenced by comparable economic status and median age profiles (see Figure 4). For example, in West and Central Africa, nations such as Nigeria (1.25 m/s), Cameroon (1.33 m/s), and the Central African Republic (1.28 m/s) show brisk walking speed values, consistent with their young median age (19.2, 18.8, and 20.2 years, respectively) and developing economies. Western European neighbours like Germany (1.10 m/s), France (1.14 m/s), the Netherlands (1.16 m/s), and Belgium (1.09 m/s) fall into a moderate-speed group, reflecting both higher-income economies and older median age (46.7, 42.4, 42.2, and 41.9 years, respectively), which are associated with more cautious pedestrian movement. A similar pattern is observed in Asia, where Japan shows one of the lowest speed values worldwide at 0.88 m/s, closely matched by other developed and ageing societies such as South Korea (0.99 m/s) and Singapore (0.92 m/s), all of which have a high median age (49.5 for Japan, 43.7 for South Korea, and 38.9 for Singapore). In contrast, emerging Asian economies with younger populations—such as India (1.30 m/s), Nepal (1.28 m/s), and Vietnam (1.24 m/s)—cluster at higher crossing speed values, reflecting their demographic profiles and different traffic conditions (median age: 29.5 for India, 27.1 for Nepal, and 32.5 for Vietnam).

Pedestrian crossing initiation times reveal strong regional and economic similarities among neighbouring countries (see Figure 5). For example, Austria (2.07 s), Andorra (2.01 s), and Estonia (2.05 s) show remarkably close values, which can be linked to their advanced economies, well-planned pedestrian infrastructure, and similar median age (44.8, 48.1 and 44.7). A contrasting pattern emerges in Southeast Asia, where Indonesia (5.29 s), Philippines (5.47 s), and Cambodia (5.37 s) display some of the slowest pedestrian initiation speed values, likely due to dense traffic environments and limited infrastructure. Likewise, South Korea (3.48 s) and Japan (4.60 s) show moderate initiation times, which may be related to their structured traffic systems and ageing populations that tend to exercise greater caution (49.5 for Japan, 43.7 for South Korea). Qatar has the longest recorded initiation time at 6.44 s, a result that may be influenced by extremely car-centric urban design, limited pedestrian infrastructure, high vehicle speed, and socioeconomic factors such as low pedestrian activity due to a predominantly high-income, vehicle-owning population.

The negative correlation between crossing speed and crossing initiation time ($r = -0.18$) suggests that pedestrians who hesitate longer before crossing tend to move more slowly, potentially reflecting greater caution or uncertainty in more challenging or ambiguous traffic conditions. In addition, the presence of traffic signs shows a negative association with crossing speed ($r = -0.36$), which could suggest that pedestrians are more cautious and reduce speed in regulated or more complex environments. Higher income inequality, measured by the Gini coefficient ($r = -0.19$) and median age ($r = -0.31$) are also associated with slower crossing speed values, pointing to broader demographic and social influences, such as increased vulnerability, lower mobility, or differences in pedestrian infrastructure. In contrast, the positive correlation with the number of detected persons ($r = 0.15$) indicates that pedestrians are likely to cross faster when more people are present, perhaps due to a sense of collective safety or social facilitation.

Similarly, the crossing initiation time is negatively correlated with the number of detected cars ($r = -0.15$) and the median age ($r = -0.09$), suggesting that pedestrians are likely to make quicker crossing decisions in environments with more vehicular traffic or where populations are younger, possibly due to increased urgency or greater mobility. In contrast, a positive correlation with the presence of motorcycles ($r = 0.25$) indicates that pedestrians tend to wait longer before crossing when motorcycles are present, which may reflect the perceived unpredictability or risk associated with these vehicles.

Interestingly, our study reveals that certain countries and territories exhibit remarkably similar pedestrian crossing times despite differing infrastructure contexts. For example, Bangladesh and the Netherlands have average crossing times of approximately 3.42 and 3.40 s, respectively. In the Netherlands, well-developed road infrastructure and clear pedestrian priority at crossings contribute to reduced crossing initiation time [18]. In contrast, in Bangladesh, the lack of formal pedestrian infrastructure often requires an opportunistic approach to crossing the road [34]. Pedestrians in Bangladesh typically assess traffic conditions looking left and right before deciding to cross when they perceive a sufficient gap [46]. This behaviour includes strategies such as accelerating your pace or stopping mid-crossing to allow vehicles to pass, as noted in studies on pedestrian behaviour in developing countries and territories [47]. Such adaptive behaviour can introduce variability and noise into the measurements of crossing initiation time, highlighting the complex interaction between the quality of the infrastructure and the decision-making processes of pedestrians.

The correlation matrices for North America and Europe reveal clear regional differences in how contextual and behavioural attributes interact. In North America, the median age is strongly positively correlated with both detected crossings ($r = 0.92$) and detected persons ($r = 0.67$), indicating that areas with older populations generally experience higher levels of pedestrian activity. In contrast, these relationships are much less pronounced in Europe ($r = 0.24$ and $r = 0.51$, suggesting that age exerts a weaker influence on pedestrian patterns there. Crossing speed in North America also

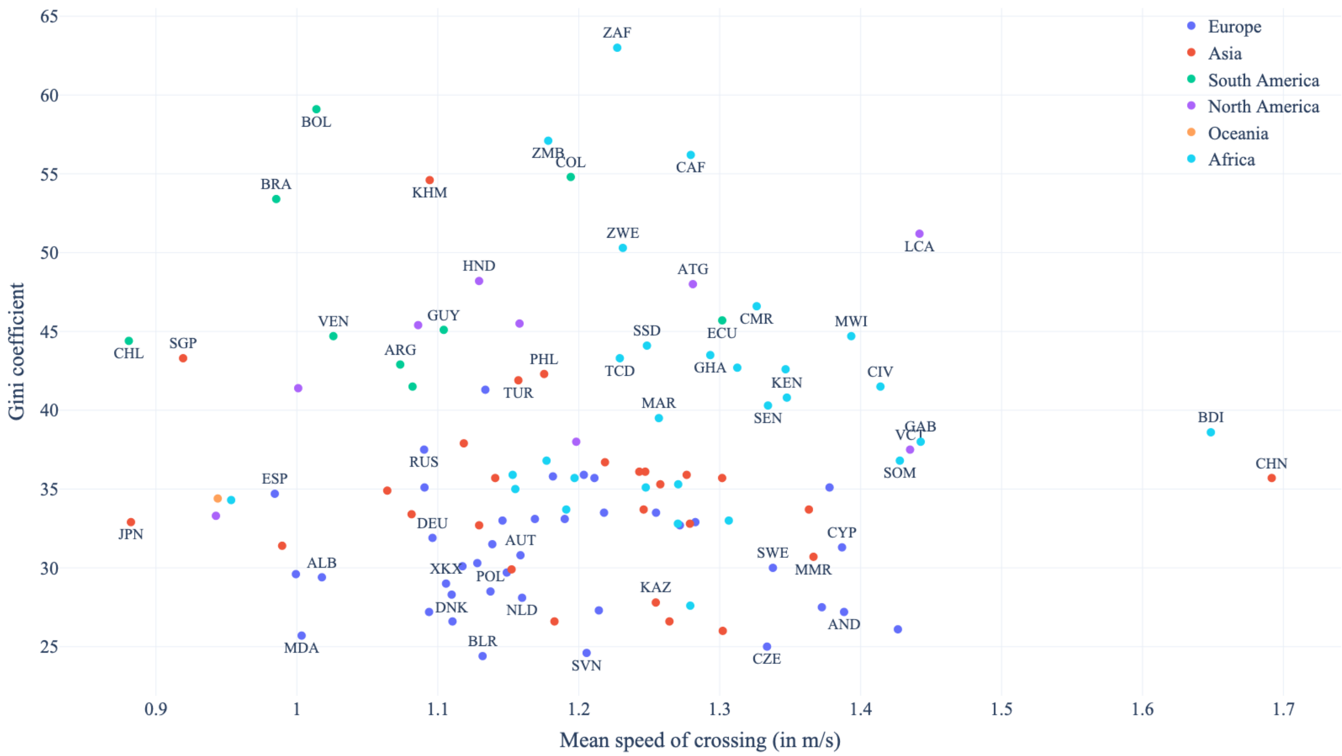

Figure 8: Relation between Gini coefficient and speed of crossing. Labels show the ISO-3 codes of countries and territories.

shows strong negative correlations with the detected motorcycles ($r = -0.44$), bicycles ($r = -0.72$), and cars ($r = -0.77$), meaning that an increased presence of vehicles is associated with slower pedestrian movement. In Europe, these associations are weaker or almost negligible, especially for motorcycles ($r = 0.09$), highlighting a decreased impact of vehicular context on pedestrian speed. Furthermore, the link between crossing speed and crossing initiation time is only slightly negative in North America ($r = -0.14$) and almost absent in Europe ($r = 0.08$), indicating a minimal association between how quickly pedestrians begin to cross and the speed at which they cross.

Substantial correlations were found between traffic mortality and detected cars in both North America ($r = 0.67$) and Europe ($r = 0.79$), underscoring the connection between higher vehicle volumes and traffic deaths. Together, these results illustrate that demographic and traffic-related factors are more closely interrelated in North America than in Europe. This may reflect less pedestrian-orientated infrastructure or a higher perception of risk in North America, making behavioural attributes more sensitive to contextual variables. In contrast, the weaker correlations observed in Europe could be attributed to a more uniform pedestrian infrastructure, a stronger safety culture, or other social factors that buffer the influence of context on behaviour.

The scatter plots in Figure 8 and Figure 9 reveal a notable variation between countries in pedestrian behaviour in relation to socioeconomic and safety indicators. The weak negative correlation ($r = -0.19$) between the Gini coefficient and the speed of crossing suggests that in countries with higher income inequality, pedestrians may cross more slowly, possibly due to poorer infrastructure, less investment in pedestrian facilities, or a greater sense of vulnerability among disadvantaged groups. In contrast, the weak positive correlation ($r = 0.18$) between national traffic mortality rates and crossing speed indicates that in countries with higher rates of traffic deaths, pedestrians may cross the roads more quickly, potentially due to a greater perception of danger or urgency to avoid traffic risks. Although these correlations are modest, they point to underlying contextual factors, such as economic disparity and road safety, that can subtly influence how pedestrians navigate urban environments in different national settings.

4.1 Limitations and Future Work

This study has several key limitations that must be considered when interpreting the results. The primary limitation arises from the inherent bias and uneven distribution of available dashcam footage across different countries and territories, with some countries or territories being heavily represented while others have much less data, potentially skewing global analyses of pedestrian behaviour.

Another limitation is the focus exclusively on urban environments, as the analysis is restricted to footage from cities, thereby excluding suburban and rural settings, where pedestrian behaviour can differ significantly due to variation in infrastructure and traffic patterns. Additionally, the conversion of pixel-based measurements into real-world distances is based on national average human

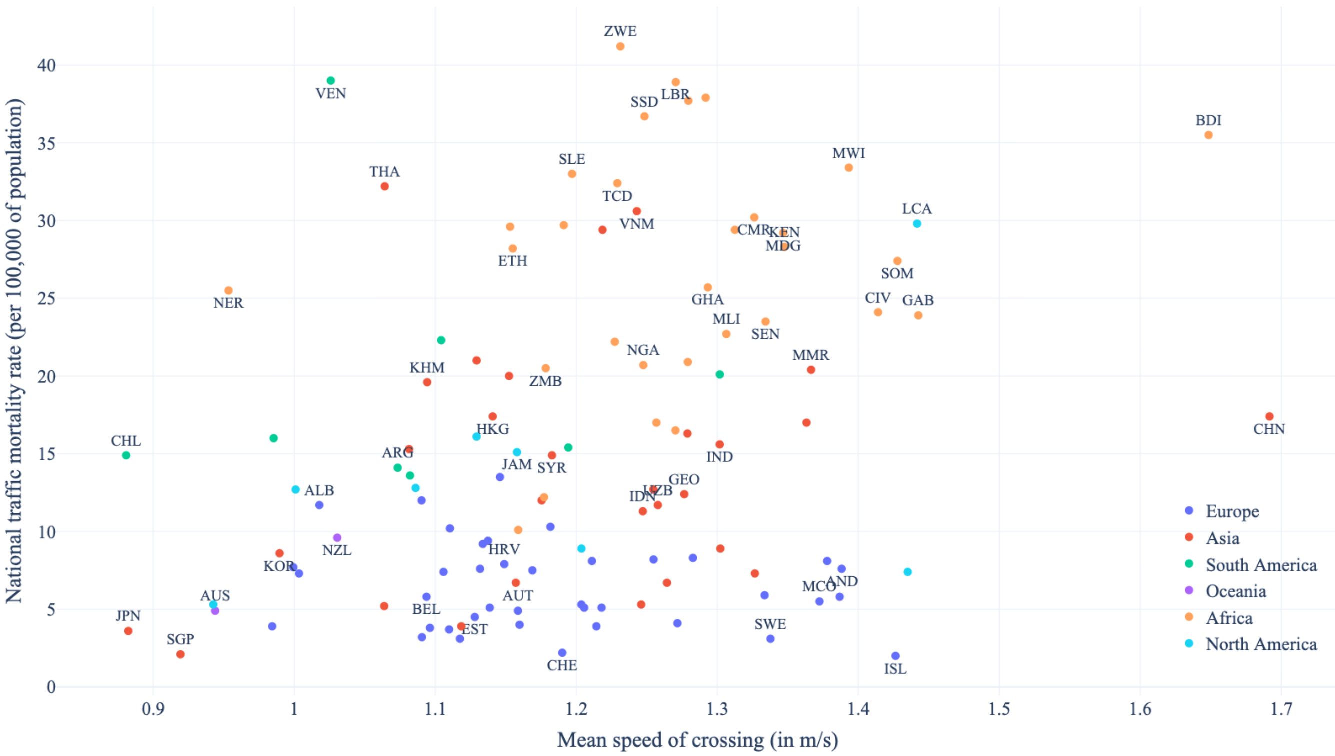

Figure 9: Relation between traffic mortality rate and speed of crossing. Labels show the ISO-3 codes of countries and territories.

heights, which introduces uncertainty due to demographic variation and camera perspective.

From a technical perspective, our reliance on the YOLOv11x object detection model, which is trained on the COCO dataset [21], restricts the analysis to basic object classes and relatively simple pedestrian behaviours. Many context-specific objects or subtle cues, such as road boundaries, lane markings, gestures, or posture changes, are not detected by YOLOv11x due to limitations in the COCO label set. As illustrated in Figure 10, sample frames of the slowest and fastest detected crossings highlight several challenges of the current approach. In the slowest crossing (top left), the reduced speed of a pedestrian is due to pushing a cart; in the second slowest (top right), the person approaches the ego vehicle before starting to cross, resulting in a longer crossing time. For the fastest crossings (bottom row), camera rotation or vehicle turning causes pedestrians to be detected as crossing in front of the vehicle, even if they are merely walking nearby. Since YOLOv11x cannot explicitly recognise road boundaries, such scenarios can introduce additional errors.

Furthermore, object identifiers are reset between daytime and nighttime segments, disrupting the continuity of pedestrian tracking and preventing longitudinal behavioural analysis. The exclusive use of fixed parameter values, such as the detection confidence threshold, may also exclude relevant cases or allow false positives. Future research should systematically investigate the effects of these parameters and consider advanced techniques - including pose estimation methods such as OpenPifPaf [19] and improved multi-object trackers like Bot-SORT [1]—to enable richer behavioural annotation and improve reliability. Incorporating lane detection and additional contextual understanding may also reduce false positives and enable a more accurate interpretation of pedestrian environments. Finally, expanding the dataset to include suburban and rural footage, as well as balancing representation across countries and territories, will be crucial to improving the robustness and generalisability of future analyses.

The current approach also faces challenges in accurately distinguishing between pedestrians and other road users. YOLOv11x sometimes mislabels cyclists and motorcyclists as pedestrians, leading to false positives that can skew crossing speed data toward higher values. In some cases, pedestrian IDs are reassigned incorrectly, for example, when a pedestrian is temporarily obscured by a static object. Although the applied filters eliminate most false positives, the parameter choices do not remove all of them; additional filtering techniques are needed to better ensure that true crossing events are retained without inadvertently rejecting valid cases.

Another limitation is that, for certain countries and territories, the estimated crossing speed and the crossing initiation time do not align with previously published results. This suggests that current detection and tracking algorithms may require further tuning and validation. Moreover, it is important to recognise that the application of a single set of parameters across all countries and territories may not yield optimal results, given the substantial diversity in environmental conditions, pedestrian behaviour, and video characteristics worldwide. In future work, we plan to address this

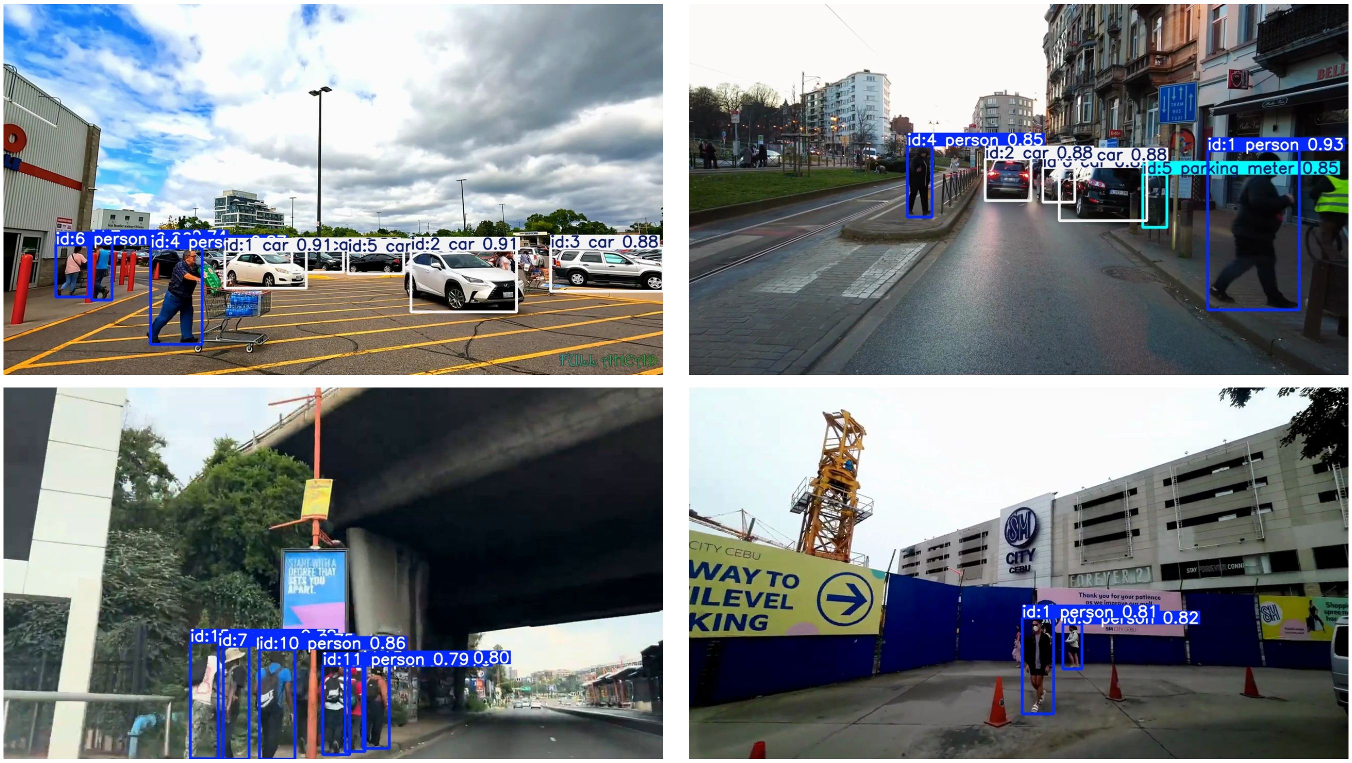

Figure 10: Sample frames illustrating detected pedestrian crossings with YOLOv11x-assigned IDs and confidence values from four selected videos. The top row shows the two slowest detected crossings. Left: the slowest crossing, with the person marked as ID 4 moving at 0.5 m/s(YouTube ID: `STbVEZMJdC0`); right: the second slowest crossing, with the person marked as ID 4 at 0.5 m/s in Brussels, Belgium (YouTube ID: `P00zkZ3_yYY`). The bottom row shows the two fastest detected crossings. Left: the fastest crossing, with the person marked as ID 1 at 2.5 m/s in Johannesburg, South Africa (YouTube ID: `rnTPlWqduEg`); right: the second fastest crossing, with the person marked as ID 1 at 2.5 m/s in Cebu City, Philippines (YouTube ID: `Wc6_DYNNzqQ`).

limitation by employing advanced deep learning models for lane detection (such as SCNN [28]) and incorporating speed approximation techniques. These approaches can improve the contextual understanding of pedestrian environments and provide more accurate speed estimates by leveraging detected lane boundaries and perspective signals. Future work should therefore explore adaptive or region-specific parameter settings and algorithmic adjustments to more accurately capture local variations and ensure more reliable cross-national comparisons.

Broader research directions should include expanding the dataset for better balance between countries and territories and incorporating data from suburban and rural environments. Integrating additional contextual factors, such as weather, time of year, intersection design, and local traffic policies, could provide a more robust and comprehensive model of pedestrian behaviour. Advanced computer vision methods that enable detailed recognition of pedestrian attributes (e.g., intention, gait, and accessories) would further enrich the analysis.

Future work should incorporate an analysis of non-standard behaviour such as jaywalking [45, 50], cyclist dynamics [8, 14], and interactions between other vulnerable road users [23]. Investigating how these behaviours vary with cultural norms, traffic regulations, and infrastructure quality across different regions could provide valuable information for urban planning and public safety initiatives [25]. Then, future studies must address the critical need to develop effective external human-machine interfaces (eHMIs) for AVs and interactive interfaces for modern cars [5, 6]. Such interfaces should consider global variations in pedestrian behaviour to effectively communicate the intentions of an AV to pedestrians. Given that pedestrian responses and expectations can vary greatly in cultural contexts, designing universally comprehensible or culturally adaptive user interfaces could significantly improve pedestrian safety and the acceptance of AV technology worldwide.

5 Dataset and Code Availability

A maintained version of the code and additional figures of analysis are available at https://github.com/bazilinskyy/youtube-national. The dataset is available at https://doi.org/10.4121/fe366b3a-5053-4b90-9f78-cc6d3056aaa2.

References

[1] Nir Aharon, Roy Orfaig, and Ben-Zion Bobrovsky. 2022. BoT-SORT: Robust Associations Multi-Pedestrian Tracking. doi:10.48550/arXiv.2206.14651

[2] Md Shadab Alam, Marieke H. Martens, Olena Bazilinska, and Pavlo Bazilinskyy. 2025. Understanding global pedestrian behaviour in 565 cities with dashcam

videos on YouTube. (2025). Under review.

[3] Qamar Ali, Muhammad Rizwan Yaseen, and Muhammad Tariq Iqbal Khan. 2019. Road traffic fatalities and its determinants in high-income countries: A continent-wise comparison. *Environmental Science and Pollution Research* 26, 19 (2019), 19915–19929. doi:10.1007/s11356-019-05410-9

[4] Jorge Arango and Jeannette Montufar. 2008. Walking speed of older pedestrians who use canes or walkers for mobility. *Transportation Research Record* 2073, 1 (2008), 79–85. doi:10.3141/2073-0

[5] Pavlo Bazilinskyy, Dimitra Dodou, and J. C. F. De Winter. 2022. Crowdsourced assessment of 227 text-based eHMIs for a crossing scenario. In *Proceedings of International Conference on Applied Human Factors and Ergonomics (AHFE)*. New York, USA. doi:10.54941/ahfe1002444

[6] Pavlo Bazilinskyy, Lars Kooijman, K. P. T. Mallant, V. E. R. Roosens, M. D. L. M. Middelweerd, Dodou Dimitra Overbeek, L. D., and J. C. F. De Winter. 2022. Get out of the way! Examining eHMIs in critical driver-pedestrian encounters in a coupled simulator. In *Proceedings of International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutoUI)*. Seoul, South Korea. doi:10.1145/3543174.3546849

[7] Raymond C. Browning, Emily A. Baker, Jessica A. Herron, and Rodger Kram. 2006. Effects of obesity and sex on the energetic cost and preferred speed of walking. *Journal of Applied Physiology* 100, 2 (2006), 390–398. doi:10.1152/japplphysiol.00767.2005

[8] Tim De Ceunynck, G. J. Wijlhuizen, Aslak Fyhri, Regine Gerike, Dagmar Köhler, Alice Ciccone, Atze Dijkstra, Emmanuelle Dupont, and Mario Cools. 2023. Behavioural profiling of cycling and walking in nine European cities. *Sustainability* 15, 3 (2023), 1996. doi:10.3390/su15031996

[9] Shuchisnigdha Deb, Lesley Strawderman, Janice DuBien, Brian Smith, Daniel W. Carruth, and Teena M. Garrison. 2017. Evaluating pedestrian behavior at crosswalks: Validation of a pedestrian behavior questionnaire for the US population. *Accident Analysis & Prevention* 106 (2017), 191–201. doi:10.1016/j.aap.2017.05.020

[10] Etienne Duim, Maria L. Lebrão, and José L. F. Antunes. 2017. Walking speed of older people and pedestrian crossing time. *Journal of Transport & Health* 5 (2017), 70–76. doi:10.1016/j.jth.2017.02.001

[11] Victor Fabricius, Azra Habibovic, Daban Rizgary, Jonas Andersson, and Pontus Wärnestål. 2022. Interactions between heavy trucks and vulnerable road users—A systematic review to inform the interactive capabilities of highly automated trucks. *Frontiers in Robotics and AI* 9 (2022), 818019. doi:10.3389/frobt.2022.818019

[12] Nicholas N. Ferenchak. 2016. Pedestrian age and gender in relation to crossing behavior at midblock crossings in India. *Journal of Traffic and Transportation Engineering (English Edition)* 3, 4 (2016), 345–351. doi:10.1016/j.jtte.2015.12.001

[13] Matt Franchi, Debargha Dey, and Wendy Ju. 2024. Towards instrumented fingerprinting of urban traffic: A Novel methodology using distributed mobile point-of-view cameras. In *Proceedings of the 16th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutoUI)*. 53–62. doi:10.1145/3640792.3675740

[14] Rahul Goel, Anna Goodman, Rachel Aldred, Ryota Nakamura, Lambed Tatah, Leandro Martin Totaro Garcia, Belen Zapata-Diomedi, Thiago H. De Sa, Geetam Tiwari, Audrey De Nazelle, Marko Tainio, Ralph Buehler, Thomas Götschi, and James Woodcock. 2022. Cycling behaviour in 17 countries across 6 continents: levels of cycling, who cycles, for what purpose, and how far? *Transport Reviews* 42, 1 (2022), 58–81. doi:10.1080/01441647.2021.1915898

[15] Hoe B. Goh, Kulanthayan Subramaniam, Tuck Y. Wai, and Ali A. Mohamed. 2012. Pedestrian crossing speed: The case of Malaysia. *International Journal for Traffic and Transport Engineering* 2, 4 (2012), 323–332. doi:10.7708/ijtte.2012.2(4).03

[16] Akash Jain, Ankit Gupta, and Rajat Rastogi. 2014. Pedestrian crossing behaviour analysis at intersections. *International Journal for Traffic and Transport Engineering* 4, 1 (2014), 103–116. doi:10.7708/ijtte.2014.4(1).08

[17] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. 2023. *YOLO by Ultralytics*. https://github.com/ultralytics/ultralytics

[18] P. Knight and K. Millard. 2013. *Literature review looking at Dutch style signalised junctions*. Project Report PPR716. Transport Research Laboratory. https://www.trl.co.uk/Uploads/TRL/Documents/TfL%20Cycling%20trials/PPR716%20-%20Literature%20review%20looking%20at%20Dutch%20style%20signalised%20junctions.pdf

[19] Sven Kreiss, Lorenzo Bertoni, and Alexandre Alahi. 2019. Pifpaf: Composite fields for human pose estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 11977–11986. doi:10.48550/arXiv.1903.06593

[20] Robert V. Levine and Ara Norenzayan. 1999. The pace of life in 31 countries. *Journal of Cross-Cultural Psychology* 30, 2 (1999), 178–205. doi:10.1177/0022022199030002003

[21] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. 2015. Microsoft COCO: Common Objects in Context. doi:10.48550/arXiv.1405.0312 arXiv:1405.0312 [cs.CV]

[22] Régis Lobjois and Viola Cavallo. 2007. Age-related differences in street-crossing decisions: The effects of vehicle speed and time constraints on gap selection in an estimation task. *Accident Analysis & Prevention* 39, 5 (2007), 934–943. doi:10.1016/j.aap.2006.12.013

[23] Wei Lv, Zhiming Fang, Xiaoge Wei, Weiguo Song, and Xuan Liu. 2013. Experiment and modelling for pedestrian following behavior using velocity-headway relation. *Procedia Engineering* 62 (2013), 525–531. doi:10.1016/j.proeng.2013.08.096

[24] Yi Ma, Eric Waiming Lee, Zuoan Hu, Meng Shi, and Richard Kowkkit Yuen. 2019. An intelligence-based approach for prediction of microscopic pedestrian walking behavior. *IEEE Transactions on Intelligent Transportation Systems* 20, 10 (2019), 3964–3980. doi:10.1109/TITS.2019.2931892

[25] Michael Mögele and Henrike Rau. 2020. Cultivating the "car state": A culturally sensitive analysis of car-centric discourses and mobility cultures in Southern Germany. *Sustainability: Science, Practice and Policy* 16, 1 (2020), 15–28. doi:10.1080/15487733.2020.1756188

[26] Betty J. Mohler, William B. Thompson, Sarah H. Creem-Regehr, Herbert L. Pick, and William H. Warren. 2007. Visual flow influences gait transition speed and preferred walking speed. *Experimental Brain Research* 181 (2007), 221–228. doi:10.1007/s00221-007-0917-0

[27] Jeannette Montufar, Jorge Arango, Michelle Porter, and Satoru Nakagawa. 2007. Pedestrians' normal walking speed and speed when crossing a street. *Transportation Research Record* 2002, 1 (2007), 90–97. doi:10.3141/2002-12

[28] Xingang Pan, Jianping Shi, Ping Luo, Xiaogang Wang, and Xiaoou Tang. 2018. Spatial as deep: Spatial CNN for traffic scene understanding. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence*. Article 891, 8 pages. doi:10.48550/arXiv.1712.06080

[29] Shruti Rao, Surjya Ghosh, Gerard Pons Rodriguez, Thomas Röggla, Abdallah El Ali, and Pablo Cesar. 2022. Investigating affective responses toward in-video pedestrian crossing actions using camera and physiological sensors. In *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutoUI)*. 226–235. doi:10.1145/3543174.3546842

[30] Amir Rasouli, Iuliia Kotseruba, and John K. Tsotsos. 2017. Understanding pedestrian behavior in complex traffic scenes. *IEEE Transactions on Intelligent Vehicles* 3, 1 (2017), 61–70. doi:10.1109/TIV.2017.2788193

[31] Amir Rasouli and John K. Tsotsos. 2020. Autonomous Vehicles That Interact With Pedestrians: A Survey of Theory and Practice. *IEEE Transactions on Intelligent Transportation Systems* 21, 3 (2020), 900–918. doi:10.1109/TITS.2019.2901817

[32] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 779–788. doi:10.48550/arXiv.1506.02640

[33] Daniela Ridel, Eike Rehder, Martin Lauer, Christoph Stiller, and Denis Wolf. 2018. A literature review on the prediction of pedestrian behavior in urban scenarios. In *21st International Conference on Intelligent Transportation Systems (ITSC)*. 3105–3112. doi:10.1109/ITSC.2018.8569415

[34] Mohammad Shaheen Sarker, Oliver Carsten, Yue Huang, and Foroogh Hajiseyedjavadi. 2023. Motivations of pedestrians for safe use of highway crossing: an application of the behaviour change model COM-B in Bangladesh. *Traffic Safety Research* 4 (2023). doi:10.55329/apul4688

[35] Jiangang Shi, Yanyan Chen, Futian Ren, and Jian Rong. 2007. Research on pedestrian behavior and traffic characteristics at unsignalized midblock crosswalk: Case study in Beijing. *Transportation Research Record* 2038, 1 (2007), 23–33. doi:10.3141/2038-04

[36] Milan Simeunović, Andrijana Jović, Pavle Pitka, and Mladen Dobrić. 2021. The impact of uncertainty on pedestrians' decision to start roadway crossing during the clearance phase. *Complexity* 2021, 1 (2021), 6769688. doi:10.1155/2021/6769688

[37] Hüseyin O. Tezcan, Mahmoud Elmorssy, and Göker Aksoy. 2019. Pedestrian crossing behavior at midblock crosswalks. *Journal of Safety Research* 71 (2019), 49–57. doi:10.1016/j.jsr.2019.09.014

[38] Kai Tian, Gustav Markkula, Chongfeng Wei, Yee M. Lee, Ruth Madigan, Toshiya Hirose, Natasha Merat, and Richard Romano. 2024. Deconstructing pedestrian crossing decisions in interactions with continuous traffic: An anthropomorphic model. *IEEE Transactions on Intelligent Transportation Systems* 25, 3 (2024), 2466–2478. doi:10.1109/TITS.2023.3323010

[39] Thi M. T. Tram, Callum Parker, Xinyan Yu, Debargha Dey, Marieke H. Martens, Pavlo Bazilinskyy, and Martin Tomitsch. 2024. Evaluating autonomous vehicle external communication using a multi-pedestrian VR simulator. *Proceedings of the ACM on Interactive Mobile Wearable and Ubiquitous Technologies* 8 (2024), 2. doi:10.1145/3678506

[40] Chinebuli Uzondu, Samantha Jamson, and Daryl Hibberd. 2020. Can infrastructure improvements mitigate unsafe traffic safety culture: a driving simulator study exploring cross cultural differences. *Transportation Research Part F: Traffic Psychology and Behaviour* 73 (2020), 205–221. doi:10.1016/j.trf.2020.06.022

[41] Vasantha Wickramasinghe, P. A. Wijesinghe, and K. G. M. Lakmali. 2021. Start-up time of crossing pedestrians with different signal configurations. *Engineer: Journal of the Institution of Engineers, Sri Lanka* 54, 4 (2021). doi:10.4038/engineer.v54i4.7475

[42] World Health Organization. 2023. Global status report on road safety 2023. https://www.who.int/publications/i/item/9789240086517 Accessed: 10 July 2025.

[43] World Health Organization. 2025. Road traffic mortality. https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/road-traffic-mortality Accessed: 10 July 2025.

[44] Hang Wu, Jiajie Tan, and S-H Gary Chan. 2021. Pedometer-free geomagnetic fingerprinting with casual walking speed. *ACM Transactions on Sensor Networks (TOSN)* 18, 1 (2021), 1–21. doi:10.1145/3470850

[45] S. Q. Xie, S. C. Wong, Tsz M. Ng, and William H. K. Lam. 2017. Pedestrian crossing behavior at signalized crosswalks. *Journal of Transportation Engineering, Part A: Systems* 143, 8 (2017), 04017036. doi:10.1061/JTEPBS.0000055

[46] Niaz Mahmud Zafri, Rashada Sultana, Md Rakibul Hasan Himal, and Tanzila Tabassum. 2020. Factors influencing pedestrians' decision to cross the road by risky rolling gap crossing strategy at intersections in Dhaka, Bangladesh. *Accident Analysis & Prevention* 142 (2020), 105564. doi:10.1016/j.aap.2020.105564

[47] Jakub Zębala, Piotr Ciępka, and Adam Reza. 2012. Pedestrian acceleration and speeds. *Problems of Forensic Sciences* 91 (2012), 227–234. https://arch.ies.gov.pl/images/PDF/2012/vol_91/91_Zebala.pdf

[48] Chi Zhang and Christian Berger. 2023. Pedestrian behavior prediction using deep learning methods for urban scenarios: A review. *IEEE Transactions on Intelligent Transportation Systems* 24, 10 (2023), 10279–10301. doi:10.1109/TITS.2023.3281393

[49] Yifu Zhang, Peize Sun, Yi Jiang, Dongdong Yu, Fucheng Weng, Zehuan Yuan, Ping Luo, Wenyu Liu, and Xinggang Wang. 2022. Bytetrack: Multi-object tracking by associating every detection box. In *Proceedings of European Conference on Computer Vision (ECCV)*. 1–21. doi:10.48550/arXiv.2110.06864

[50] Yinan Zheng, Thomas Chase, Lily Elefteriadou, Bastian Schroeder, and Virginia P. Sisiopiku. 2015. Modeling vehicle–pedestrian interactions outside of crosswalks. *Simulation Modelling Practice and Theory* 59 (2015), 89–101. doi:10.1016/j.simpat.2015.08.005

[51] Dilara Özcerit. 2022. Average height of men and women by country. Kaggle dataset. https://www.kaggle.com/datasets/dilaraozcerit/averageheightdata Accessed: 10 July 2025.

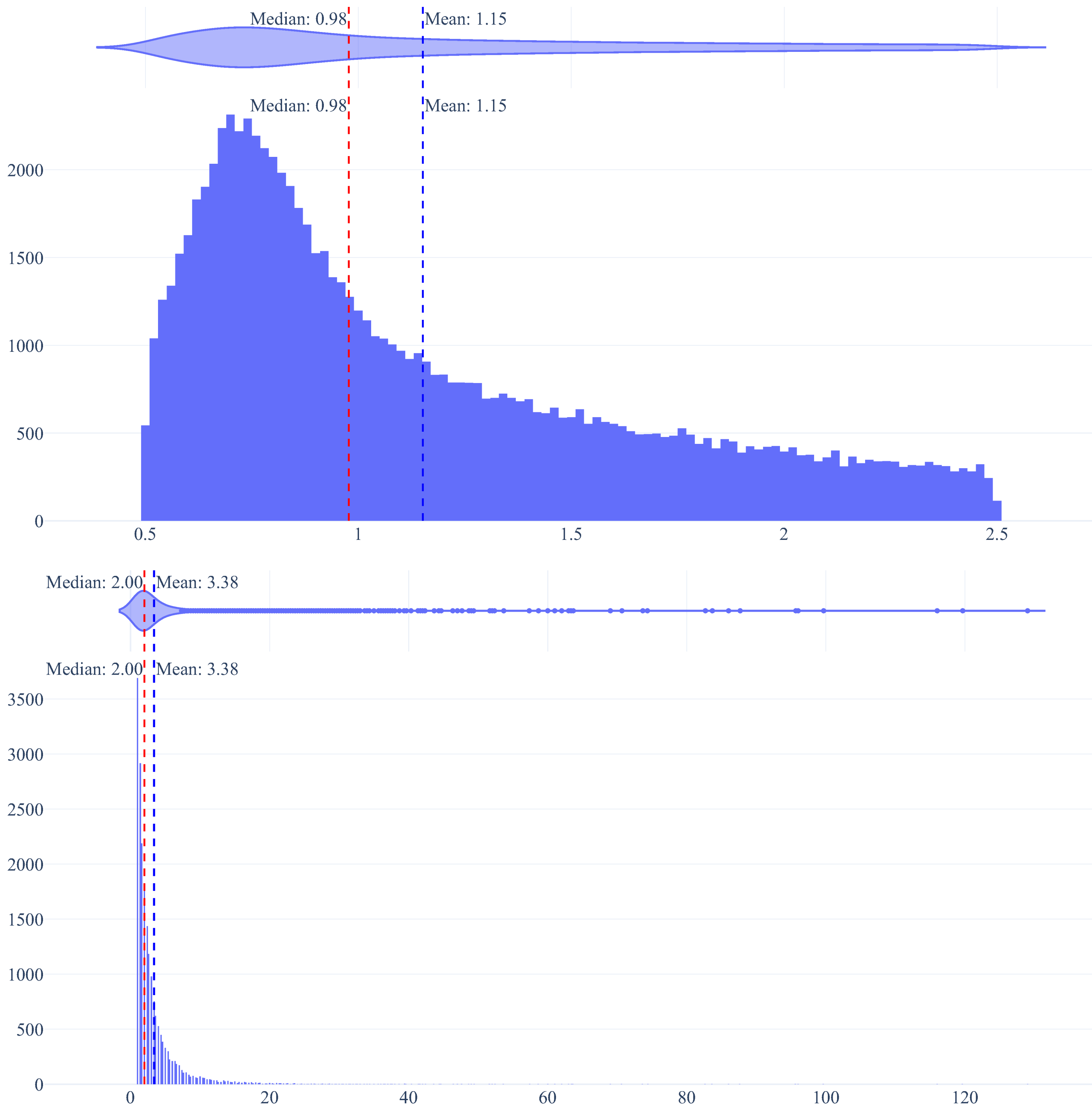

Figure A1: Histograms of pedestrian crossing speed (top) and crossing initiation time (bottom). Data are shown prior to population and video duration thresholds, but after filtering out irrelevant movements, riders (bicycle/motorcycle), and applying minimum and maximum limits for speed (0.5–2.5 m/s) and crossing initiation time (1–150 s). 233 countries are considered with 112,299 crossings in total. Each value in the histograms represents an individual pedestrian crossing detected in the dataset, regardless of which country or territory it occurred in. The displayed mean and median are calculated across all individual crossings, not as averages per country.

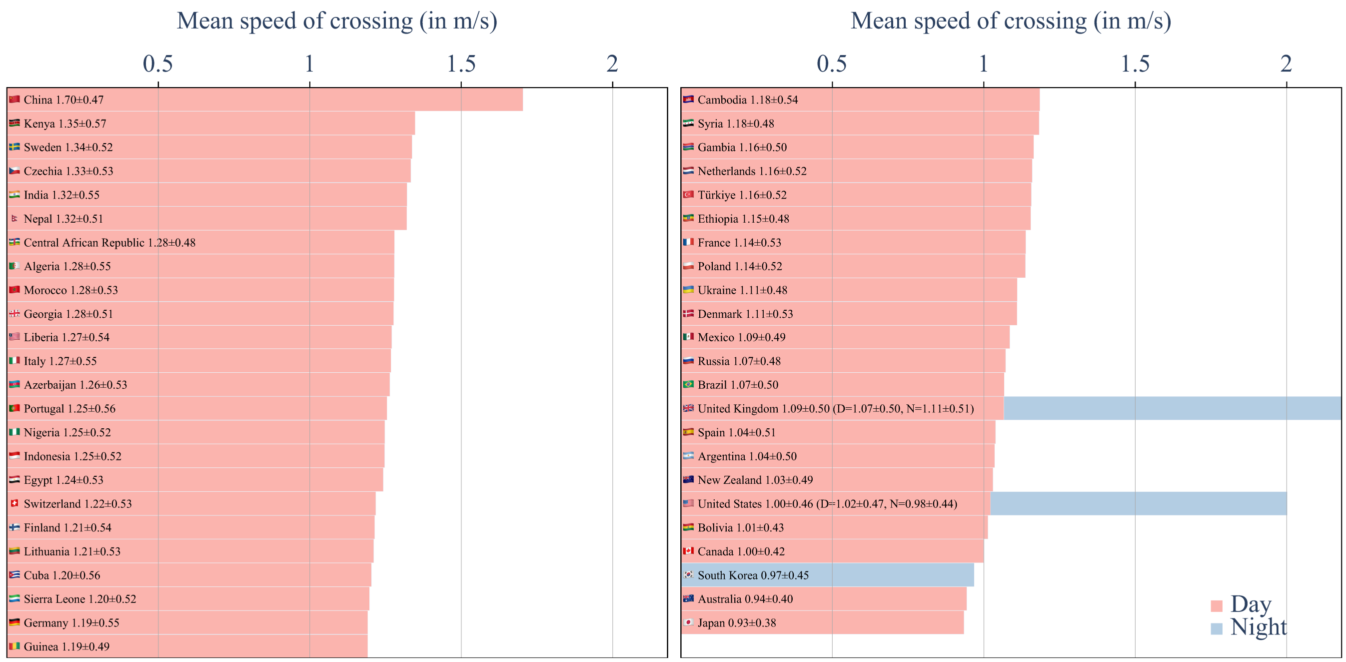

Figure A2: Mean pedestrian crossing speed by country or territory, shown separately for daytime (red) and nighttime (blue) observations. 47 countries and territories with at least 10 hours of footage and at least 500 valid pedestrian crossing detections (either daytime or nighttime) are considered. The mean and standard deviation (SD) of the overall crossing speed are displayed next to the country name, followed by the daytime (D) and nighttime (N) values (mean ± SD) in parentheses. Bars are sorted in ascending order based on the daytime mean speed for each country; if daytime data are unavailable, nighttime values are used to order.

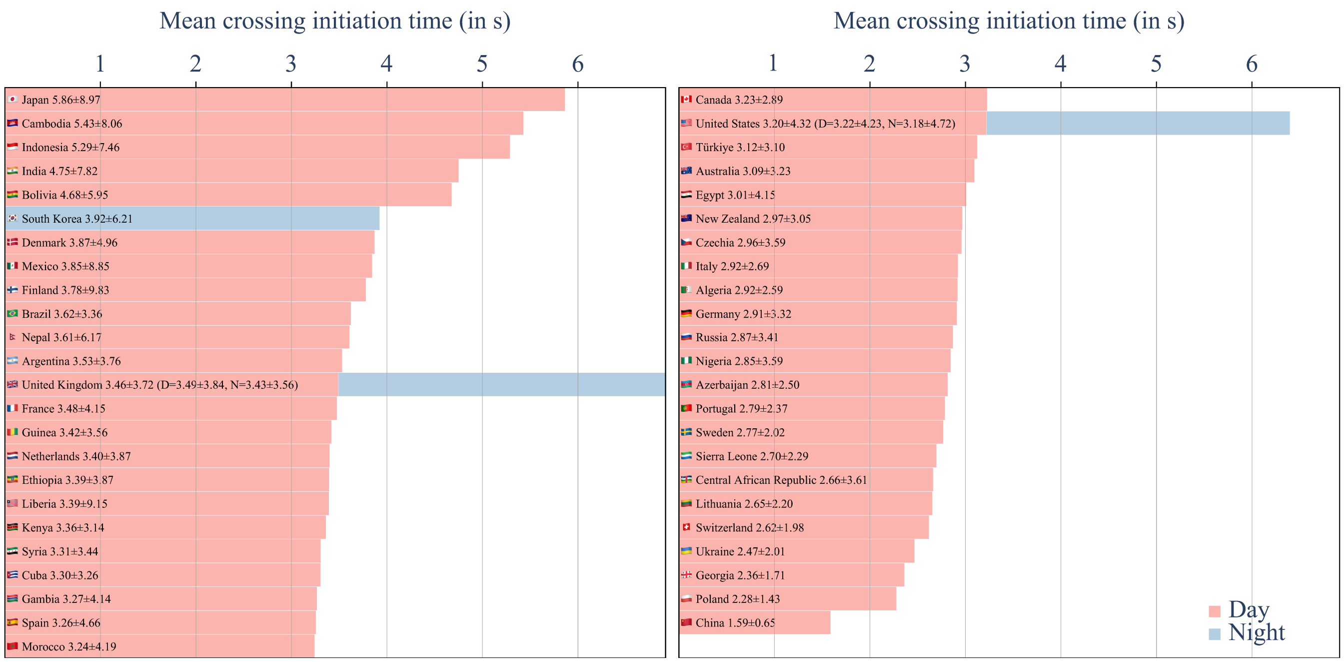

Figure A3: Mean pedestrian crossing initiation time, presented by country or territory, with daytime (red) and nighttime (blue) observations. 47 countries and territories with at least 10 hours of footage and at least 500 valid pedestrian crossing detections (either daytime or nighttime) are considered. The mean and standard deviation (SD) of the overall crossing initiation time are displayed next to the country name, followed by the daytime (D) and nighttime (N) values (mean ± SD) in parentheses. Bars are sorted in ascending order based on the daytime mean speed for each country; if daytime data are unavailable, nighttime values are used to order.

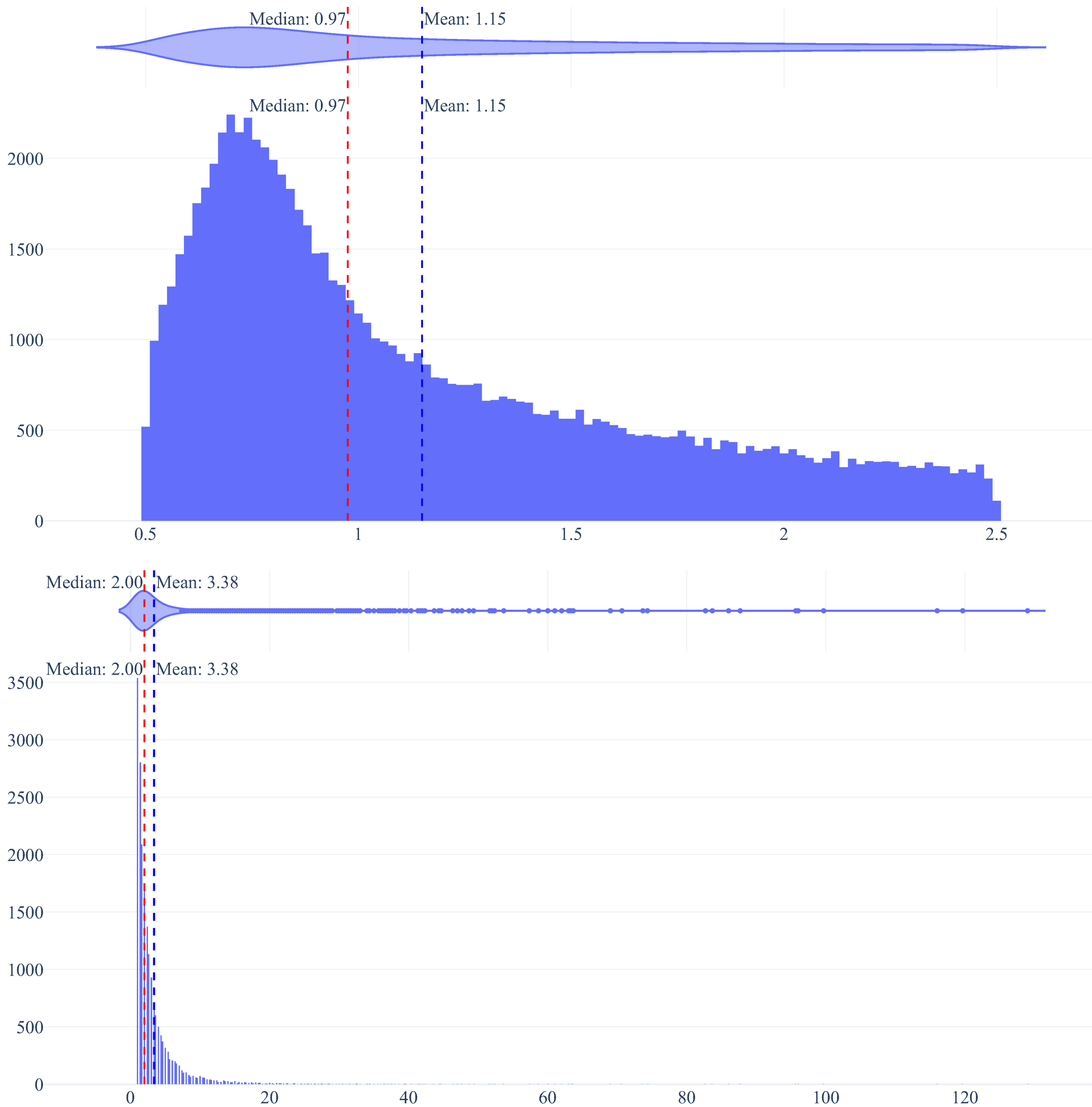

Figure A4: Histograms of pedestrian crossing speed (top) and crossing initiation time (bottom). Data are shown after applying population and video duration thresholds, filtering out irrelevant movements, riders (bicycle/motorcycle), and applying minimum and maximum limits for speed (0.5–2.5 m/s) and crossing initiation time (1–150 s). 124 countries are considered with 109,245 crossings in total. Each value in the histograms represents an individual pedestrian crossing detected in the dataset, regardless of which country or territory it occurred in. The displayed mean and median are calculated across all individual crossings, not as averages per country.