Pedestrian Planet: What 1,609 Hours of YouTube Driving from 133 Countries Teaches Us About the World

MD SHADAB ALAM*, Eindhoven University of Technology, The Netherlands

MARIEKE H. MARTENS, Eindhoven University of Technology, The Netherlands and TNO, The Netherlands

PAVLO BAZILINSKYY, Eindhoven University of Technology, The Netherlands

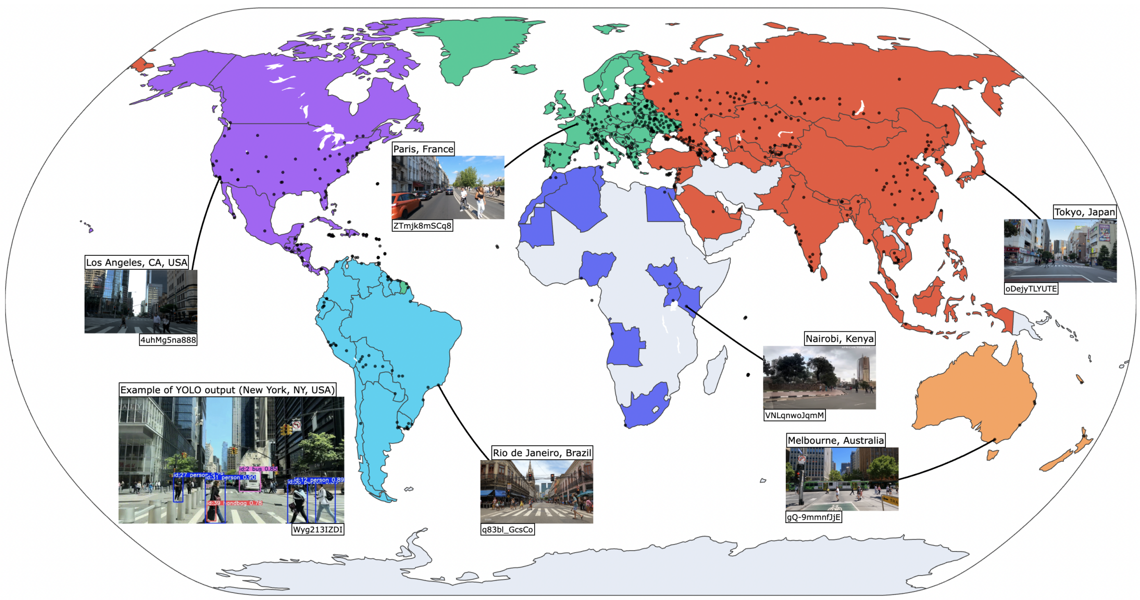

Fig. 1. The 133 countries with dashcam footage included in analysis on the political map (coloured by continent). The labels under images show the corresponding YouTube video ID. The frame on the bottom left shows an example of object detection using YOLOv11 with identified objects such as pedestrians, vehicles, and traffic signs. The labels 'id' refer to the unique ID of the detected object with the type mentioned later. The labels end with the confidence of detection of the object.

Pedestrian crossing behaviour varies globally. This study analyses dashcam footage from the PYT dataset, covering 133 countries, to examine decision time to cross, crossing speed, and contextual variables, including detected vehicles, traffic mortality, GDP, and Gini. Bulgaria had the longest decision time (10.50 s), while San Marino exhibited the fastest crossing speed (1.14 m/s). A global negative correlation between speed and decision time (r = -0.54) suggests that more cautious or uncertain pedestrians cross more

*Corresponding Author

Authors' Contact Information: 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 and TNO, Helmond, The Netherlands; Pavlo Bazilinskyy, p.bazilinskyy@tue.nl, Eindhoven University of Technology, Eindhoven, The Netherlands.

Manuscript submitted to ACM

slowly. Regional differences reveal stronger inverse correlations in Europe and North America, likely due to varying infrastructure, regulation, and cultures. Pedestrian decision time is positively correlated with the presence of other road users, especially bicycles (r = 0.35). Similar crossing times in countries with different infrastructures, such as Belgium and India, 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.

CCS Concepts: • **Computing methodologies** → **Cross-validation**; **Object detection**; **Tracking**; • **Mathematics of computing** → *Probability and statistics*; • **Human-centered computing** → Human computer interaction (HCI).

Additional Key Words and Phrases: Dashcam videos, Pedestrian behaviour, Cross country, Dataset analysis

ACM Reference Format:
Md Shadab Alam, Marieke H. Martens, and Pavlo Bazilinskyy. 2025. Pedestrian Planet: What 1,609 Hours of YouTube Driving from 133 Countries Teaches Us About the World. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX).* ACM, New York, NY, USA, 15 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Every year, road traffic accidents claim more than 1.19 million lives, making road injuries the leading killer of children and young people aged 5 to 29 years [37]. More than half of fatalities occur among pedestrians, cyclists, and motorcyclists, particularly those living in low- and middle-income countries. 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. In contrast, India's rapid urbanisation combined with less consistent enforcement of traffic regulations contributes to significantly higher mortality rates, with reported figures of 4.0 versus 15.6, respectively (source: https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/road-traffic-mortality). In this context, a comprehensive understanding of pedestrian behaviour is essential in diverse urban settings. Automated vehicles (AV) and advanced driver assistance systems (ADAS) have evolved rapidly in the past decade, driven by breakthroughs in computer vision, machine learning, and sensor technologies [17]. It has enabled complex scene understanding capabilities for AVs, including pedestrian recognition, intention estimation, and trajectory prediction [7, 31]. Despite these advances, a persistent challenge is modelling pedestrian behaviour under diverse environmental and cultural conditions. Accurate prediction of how pedestrians will act in different traffic contexts is crucial to ensure safe human-vehicle interactions and efficient transportation systems around the world [23, 30, 33, 41].

However, much of the existing research on pedestrian behaviour is limited to controlled environments, single-city studies, or limited cultural contexts [31]. Shi et al. (2007) 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 [34]. Similarly, Deb et al. (2017) conducted a survey in the US with 50 survey items that allow respondents to rate the frequency with which they engage in different types of road use behaviour as pedestrians [10]. The validation study was conducted on 425 participants (228 males and 197 females) between the ages of 18 and 71. Rasouli et al. (2017) collected data from 5 different countries, collecting a total of 240 hours of footage, and identified 2400 pedestrians to understand how pedestrians communicate intent (especially before crossing) [30]. These restrictions hamper the development of truly generalisable algorithms for pedestrian detection, tracking, and prediction.

Similarly, the variables associated with pedestrians, such as crossing speed or hesitation, have previously been studied by researchers. For example, Goh et al. (2021) conducted a study in Kuala Lumpur, Malaysia, with 1,579 participants in 4 different locations [15]. They concluded that 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 average speed of population as 1.31 m/s at signalised crossing and 1.39 m/s at unsignalised crossing. At the same time, Duim et al. [12] conducted a study with 1,911 participants with an average age of 70.1 years in São Paulo, Brazil. They found that their average walking speed was 0.75 m/s (95%CI 0.73; 0.84) (median = 0.72 m/s). These findings highlight the need to consider demographic characteristics in the modelling of pedestrian crossing times [7].

Another variable is the time to start crossing the road. Wickramasinghe et al. (2021) 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 and gave a multiregression model to estimate the start-up time of the pedestrian crossing [36]. Lobjois et al. (2007) examined the effects of age, vehicle speed and time constraints on gap selection in crossing decisions, finding that older pedestrians tend to choose larger gaps to compensate for longer crossing times, yet they may experience reduced safety margins at higher speeds - for example, unsafe decision rates increased from 3.1% at 40 km/h to 8.6% at 60 km/h [21]. Furthermore, Simeunovi et al. (2021) investigated pedestrian decision making during the clearance phase and found that the presence of a countdown timer, by reducing the uncertainty about the available crossing time, significantly increases the probability of initiating a crossing, with 65.7% of pedestrians starting their crossing when a timer is present compared to only 22% without it, resulting in a delay rate of 1.3 times during the red signal [35]. Their analysis further suggests that while clearer temporal information can improve crossing efficiency, it can inadvertently promote riskier behaviour if pedestrians overestimate their ability to complete the crossing in time.

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-sharing platforms, notably `YouTube` (https://www.youtube.com/), has fundamentally transformed the means by which researchers 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 diverse urban traffic characteristics in different cities, providing compelling motivation for our approach using YouTube-sourced footage. 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.* [1] have compiled a dataset, entitled "Pedestrian in YouTube (PYT)", by applying computer vision techniques to dashcam videos, thus enabling 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 "Pedestrian in YouTube (PYT)" dashcam dataset from 133 countries. Using its YOLO-based object tracking, it quantifies key metrics, such as crossing decision times, and speed of crossing, to capture diverse behaviours across varied cultural, socioeconomic, and infrastructural environments. The research investigates why pedestrians behave differently worldwide and how these variations impact traffic safety today, while exploring the relationship between factors such as decision time and speed with traffic density and infrastructure quality. In contrast to traditional lab or single-city studies, the PYT dataset

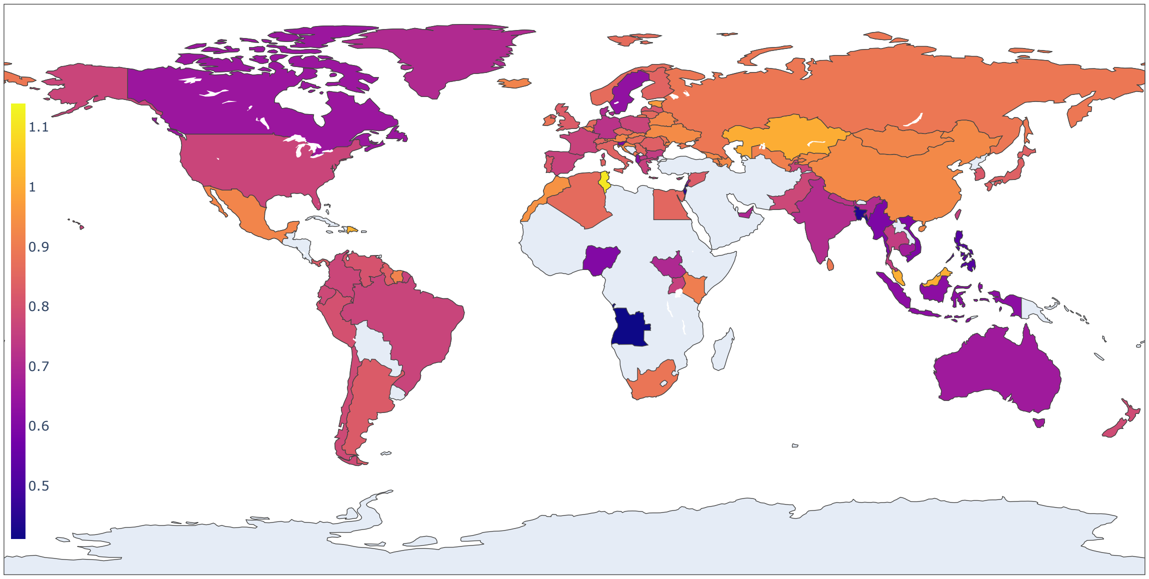

Fig. 2. Mean pedestrian crossing speed (in m/s) by country.

enables a broad cross-cultural evaluation that offers valuable insights for urban planning and future integration and safety of automated vehicles (AVs).

2 Method

This study used a dataset of dashcam videos named `Pedestrian in Youtube (PYT)` sourced from publicly accessible platforms, specifically YouTube, to investigate pedestrian behaviour under normal urban driving conditions. To ensure the representativeness and relevance of the collected footage, Alam et al. [1] established rigorous inclusion criteria. First, selected videos consistently featured urban environments, explicitly excluding highways, rural roads, and parking lots, capturing pedestrian interactions typical of everyday urban settings. Second, each video was required to have a minimum duration of ten minutes to sufficiently cover various urban scenarios. Third, footage depicting atypical events such as accidents, emergency responses, or special events, ensuring that the data represented routine pedestrian and traffic dynamics. Lastly, videos were targeted from cities with populations exceeding 20,000 inhabitants; however, an exception was made to include capital cities or major urban centres in countries where populations were below this threshold, thereby ensuring comprehensive geographic representation, particularly for smaller countries (e.g. Aruba, San Marino).

The application of these criteria resulted in the inclusion of 2,949 videos from 650 cities in 133 countries. The collected footage represented varied urban environments on multiple continents, totalling approximately 1,609 hours of driving recordings, segmented into daytime, nighttime, and combined footage scenarios. All videos were downloaded in a standardised resolution of 1280 x 720 pixels using pytubefix (https://pytubefix.readthedocs.io) and yt-dlp (https://github.com/yt-dlp/yt-dlp) libraries, facilitating consistent video quality and analytical robustness.

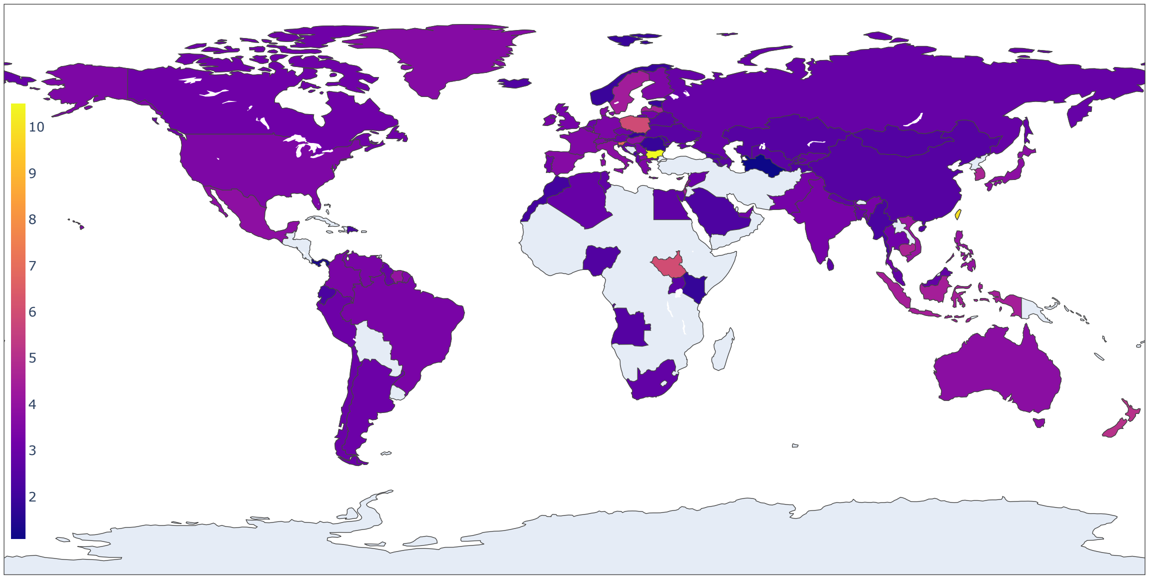

Fig. 3. Mean pedestrian time to start crossing (in s) before crossing.

To annotate the footage, the You Only Look Once (YOLOv11x) [18, 32] object detection algorithm was used. YOLOv11 provides simultaneous predictions of bounding boxes and class probabilities for objects, enabling efficient processing of large datasets. Each frame was analysed to detect and classify pedestrians, bicycles, motorcycles, cars, buses, trucks, traffic lights, and stop signs with a confidence threshold of at least 70% [8]. YOLOv11 generated unique tracking identifiers for each object, allowing longitudinal tracking across sequential frames. However, unique identifiers were reset in transitions between daytime and nighttime segments within the same video to avoid misidentification during drastically changing lighting conditions.

Pedestrian crossing events were identified using an algorithm derived from Alam et al. [1]. 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. The algorithm filtered out irrelevant movements, ensuring that only genuine crossing events were recorded and analysed.

To accurately calculate pedestrian crossing speed, pixel-based displacement data extracted from YOLO annotations were converted to real-world measurements. The conversion used national average human heights as reference lengths, enabling the determination of pixel-to-metre scaling ratios specific to each geographic region. Pedestrian speeds were calculated by dividing the displacement in the real world by the time taken to cross, measured by frame timestamps. Speeds greater than 1.42 m/s were excluded from the analyses, effectively removing outliers such as individuals who move on skateboards or bicycles, which could bias the modelling of pedestrian behaviour [6, 20, 27].

Furthermore, pedestrian crossing decision time was meticulously measured, defined as the duration required by pedestrians to initiate crossing movement after reaching the road's edge. This measurement relied on detecting periods of positional stability within a defined margin of 10% of the pedestrian's height, which lasted at least one second before initiating the crossing. Such detailed quantification allowed for the exploration of the behavioural patterns associated with decision making in various urban contexts.

3 Results

The analysis of the data yielded several notable results. The mean pedestrian crossing speed is 0.81 m/s (SD = 0.13). San Marino exhibited the highest pedestrian crossing speed (1.14 m/s), while Angola is (0.41 m/s). Figure 2 illustrates the average speed of pedestrian crossing in various countries around the world. Figure 4 shows the distribution of speed during the day and night in different countries. The countries demonstrating the largest discrepancies between daytime and nighttime crossing speeds are Albania (0.688 m/s), Uzbekistan (0.451 m/s), Guyana (0.417 m/s), Germany (0.227 m/s), and Venezuela (0.202 m/s). In contrast, countries with minimal speed differences between daytime and nighttime include Cambodia (0.020 m/s), Egypt (0.016 m/s), India (0.015 m/s), Brazil (0.011 m/s), and Japan (0.003 m/s).

Similarly, the average time to start crossing before pedestrians cross is 3.31 s (SD = 1.39). Pedestrians in Bulgaria have the longest waiting period before crossing (10.50 s), while Turkmenistan has the lowest time to start crossing across the world (1.09 s). Figure 3 illustrates average time to start crossing for pedestrians before initiating the crossing of the road. Figure 5 shows the distribution of speed during the day and night in different countries. The countries with the greatest variations in time to start crossing between daytime and nighttime include Mexico (13.920 s), Taiwan (10.598 s), Latvia (4.741 s), Singapore (2.746 s), and Indonesia (2.263 s). The countries displaying the smallest differences include Argentina (0.074 s), Spain (0.065 s), Azerbaijan (0.038 s), Brazil (0.024 s), and the United States (0.022 s). Figure 5 illustrates average time to start crossing for pedestrians before initiating the crossing of the road.

Figure 6 presents the correlation matrix based on Spearman rank correlation coefficient among pedestrian characteristics, other detected objects, and socioeconomic factors such as the literacy rate, the traffic mortality rate and the Gini coefficient. The analysis reveals a negative correlation between the speed of the crossing and the decision time of the crossing ($r = -0.55$). Furthermore, the speed of crossing is negatively correlated with the presence of persons ($r = -0.52$), bicycles ($r = -0.41$), trucks ($r = -0.31$) and cars ($r = -0.43$). The matrix also provides information on traffic patterns, indicating a positive relationship between the frequency of pedestrian crossings without traffic lights and the number of trucks ($r = 0.41$).

Furthermore, Figure 7 and Figure 8 present the correlation matrices for the same attributes in North America and Europe, respectively. In both regions, the crossing speed and the crossing decision time exhibit an inverse relationship ($r = -0.10$ for North America and $r = -0.74$ for Europe). However, the relationship between crossing speed and literacy rate differs markedly between the two continents; in Europe, there is virtually no correlation ($r = -0.09$), while in North America the correlation is strongly negative ($r = -0.70$). A similar disparity is observed in other comparisons: the crossing speed versus the mortality rate shows a correlation of $r = 0.90$ for North America compared to $r = -0.14$ for Europe; crossing speed versus the number of cars detected yields $r = 0.60$ for North America and $r = -0.35$ for Europe; and the crossing decision time versus the total number of motor vehicles detected presents $r = 0.80$ for North America and $r = 0.06$ for Europe. Furthermore, the relationship between crossing without a traffic light and decision time is strongly positive in North America, in contrast to Europe ($r = 0.80$ for North America and $r = 0.15$ for Europe). Taken together, these pronounced differences underscore why focussing on Europe versus North America yields a particularly insightful framework for understanding regional variations in pedestrian behaviour, as their distinct urban infrastructures, cultural norms, and policy environments reveal nuances that comparisons with other continents might obscure.

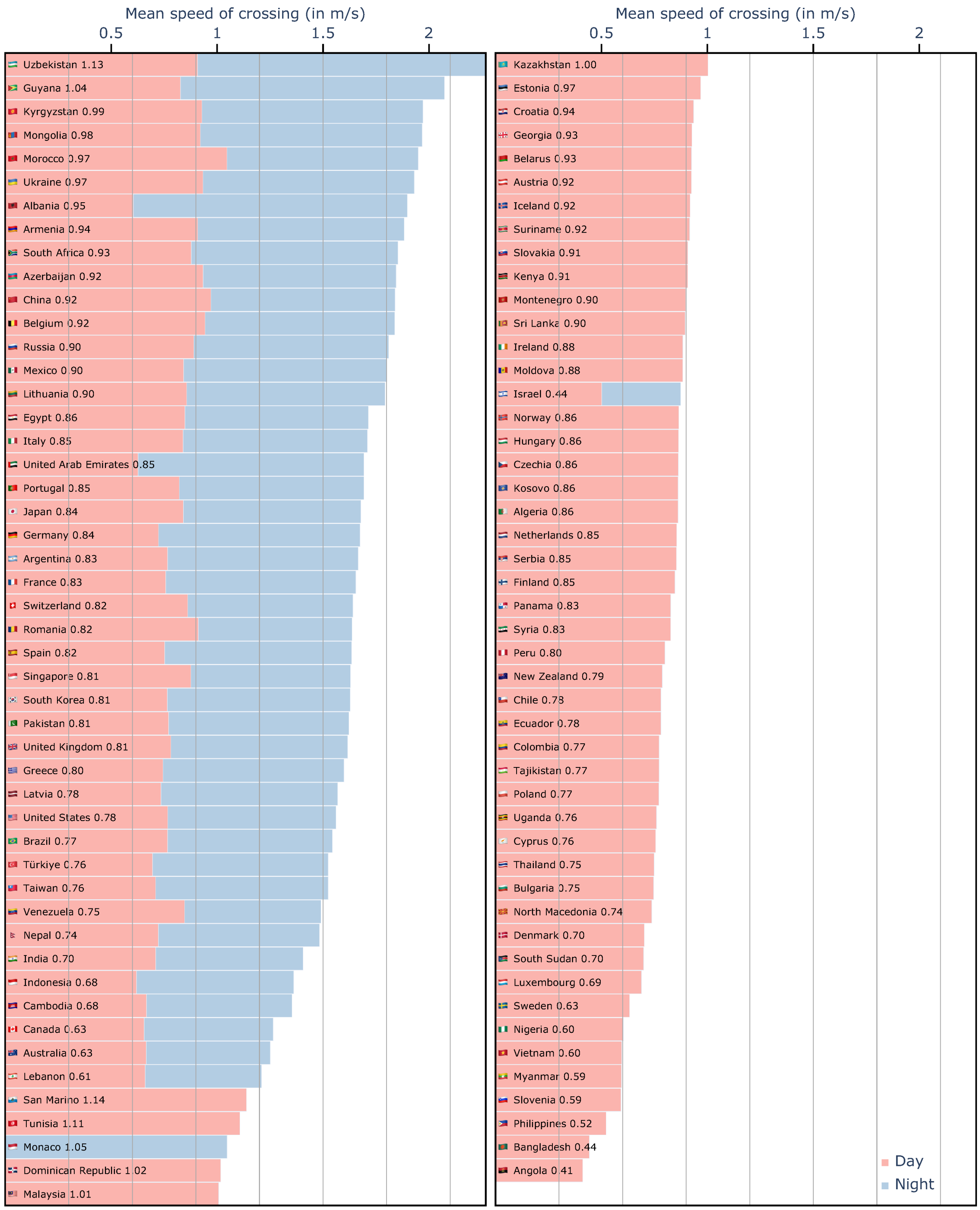

Fig. 4. Mean pedestrian crossing speeds by country are shown for daytime (red) and nighttime (blue) periods. Next to each country's name, the value represents the average speed calculated from both periods. The bars are arranged in ascending order based on the total (daytime plus nighttime) speed.

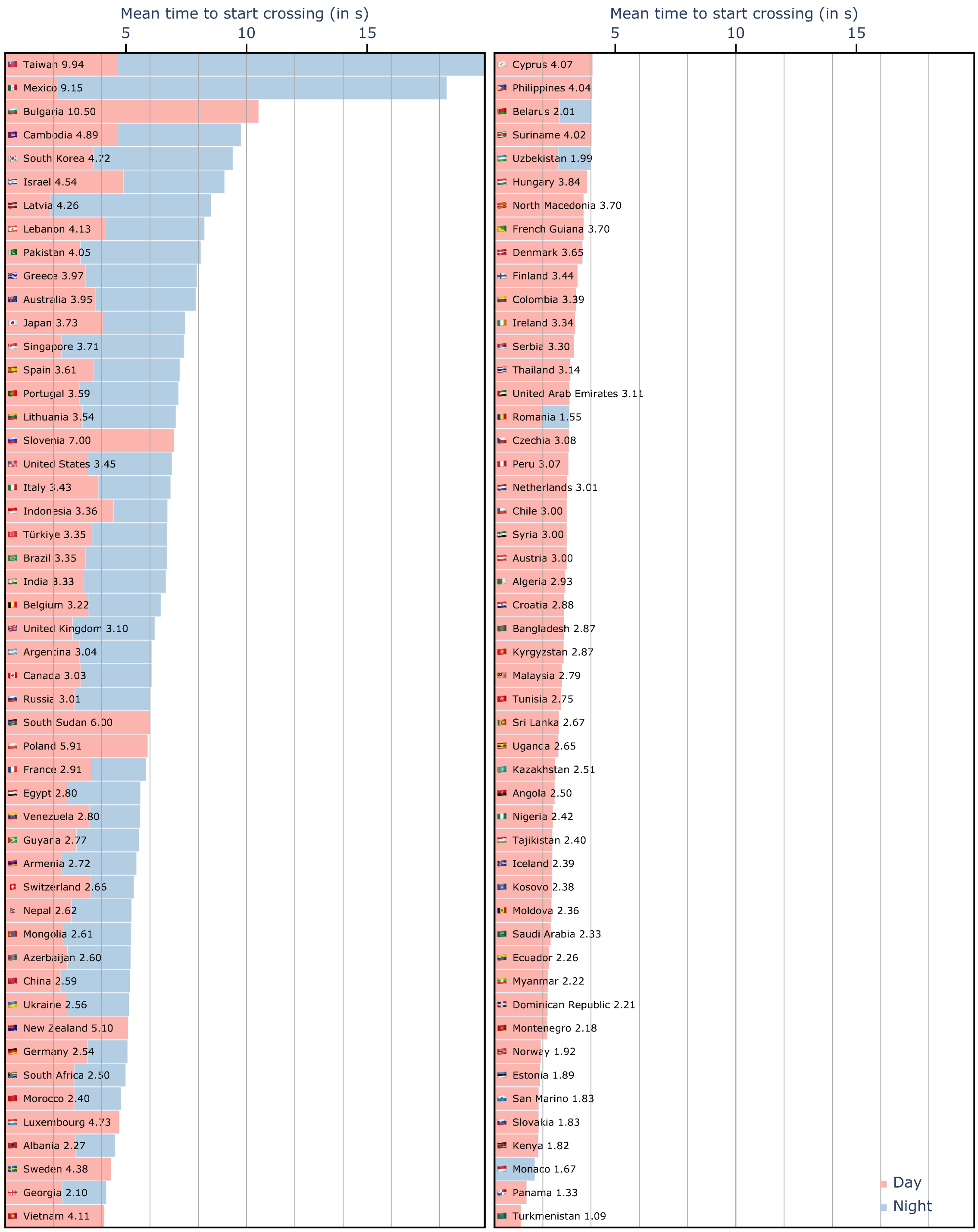

Fig. 5. Pedestrian time to start crossing before road crossing are presented by country, with daytime (red) and nighttime (blue) observations. Values displayed next to each country represent the average of the daytime and nighttime measurements, and the bars are arranged in ascending order based on their combined time to start crossing.

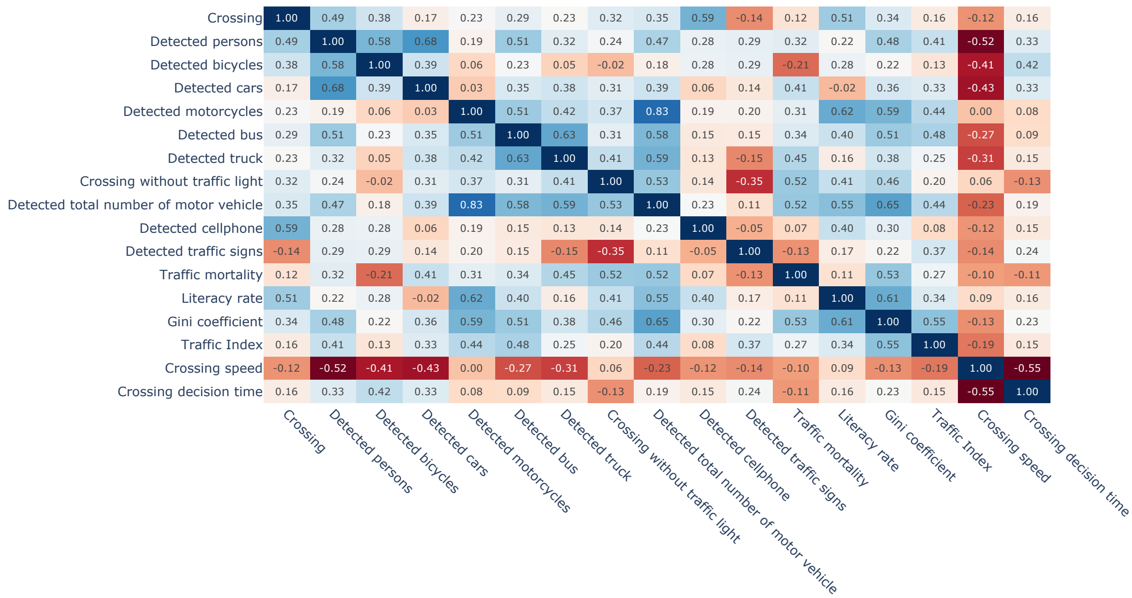

Fig. 6. Spearman correlation matrix of pedestrian behaviour, traffic characterises, and socio-economic factors for all countries.

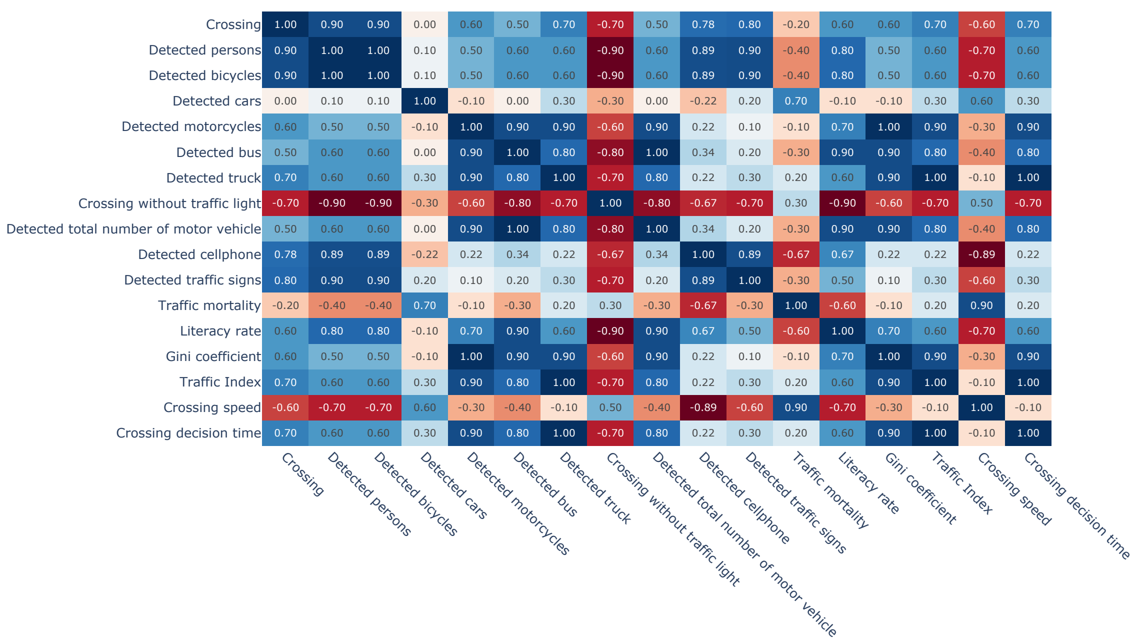

Fig. 7. Spearman correlation matrix of pedestrian behaviour, traffic characterises, and socio-economic factors in North America.

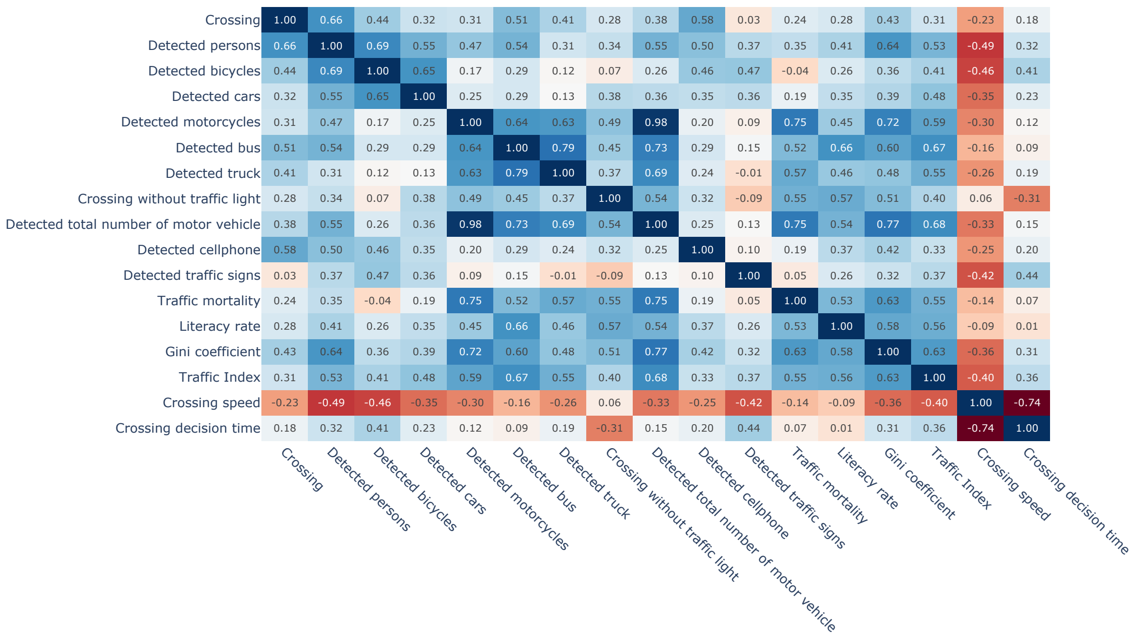

Fig. 8. Spearman correlation matrix of pedestrian behaviour, traffic characterises, and socio-economic factors in Europe.

In addition, the scatter plots presented in Figure 9 and Figure 10 reveal significant variations between countries with respect to socioeconomic and safety indicators. Figure 9 shows a positive relationship between the Gini coefficient and pedestrian decision time, suggesting longer decision times in countries with higher income inequality. In contrast, Figure 10 illustrates the variability in crossing decision times relative to national traffic mortality rates, highlighting that countries with higher mortality rates tend to experience longer pedestrian decision times before crossing.

4 Discussion

The findings presented in this study highlight significant global variations in pedestrian crossing behaviours and underline the influence of regional and country-specific socioeconomic and infrastructure factors. The mean crossing speed (0.81 m/s) aligns closely with previous studies of pedestrian behaviour [2, 28, 38], but the substantial deviations observed in countries such as San Marino (highest speed, 1.14 m/s) and Angola (lowest speed, 0.41 m/s) reflect considerable cultural and infrastructure differences.

The pronounced negative correlation between the speed of the crossing and the decision time to cross (r = -0.55 globally) indicates that pedestrians who take longer to begin crossing tend to move more slowly across the street. This relationship suggests increased caution or uncertainty among these pedestrians, possibly influenced by heavier traffic conditions, inadequate infrastructure, or lack of effective pedestrian safety regulations. Furthermore, negative correlations between pedestrian speed and the presence of other road users (people, bicycles, cars, and trucks) underscore how the complexity of traffic potentially affects pedestrian confidence and behaviour.

Our analysis also reveals that geographically close countries exhibit similar pedestrian behaviours, as illustrated in Figure 4. For example, neighbouring countries such as Germany (0.72 m/s) and France (0.76 m/s) demonstrate

Fig. 9. Relation between Gini coefficient and crossing decision time. Labels show the ISO-3 codes of countries.

Fig. 10. Relation between traffic mortality rate and crossing decision time. Labels show the ISO-3 codes of countries.

comparable crossing speeds (5.26% difference). Similarly, South American countries like Brazil (0.76 m/s), Chile (0.78 m/s), and Colombia (0.77 m/s) exhibit closely aligned crossing speeds. Comparable patterns are also evident in pedestrian time to start crossing; India (3.26 s) and Pakistan (3.23 s) display almost identical average time to start crossing (0.92% difference), as do Armenia (2.37 s) and Georgia (2.35 s) with a difference of only 0.84% (see Figure 3). However, some exceptions such as Bulgaria's notably high time to start crossing (10.5 s) could be influenced by a smaller sample of footage available from this country.

Interestingly, our study reveals that certain countries exhibit remarkably similar pedestrian crossing times despite differing infrastructure contexts. For example, Belgium and India have average crossing times of approximately 3.27 and 3.26 seconds, respectively. In Belgium, well-developed road infrastructure and clear pedestrian priority at crossings contribute to reduced time to start crossing [25]. In contrast, in India, the lack of formal pedestrian infrastructure often requires an opportunistic approach to crossing the road [3, 16]. Pedestrians typically assess traffic conditions looking left and right before deciding to cross when they perceive a sufficient gap [24]. 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 [40]. These adaptive behaviours can introduce variability and noise into crossing time measurements, highlighting the complex interaction between infrastructure quality and pedestrian decision-making processes.

The regional comparison between Europe and North America is particularly illuminating because both continents include numerous developed countries, providing a meaningful framework for analysing pedestrian behaviours in economically similar but culturally distinct contexts. In Europe, the strong inverse correlation between crossing speed and decision time ($r = -0.74$) reflects careful and deliberate pedestrian behaviour, which could be due to dense urban infrastructures and comprehensive pedestrian safety measures. In contrast, the relatively weaker relationship in North America ($r = -0.10$) may suggest different behavioural dynamics influenced by vehicle-dominated infrastructures and different pedestrian education or awareness programmes.

In addition, stark contrasts in the correlations between speed and literacy rates, mortality rates, and vehicle numbers between these two regions suggest complex interplays of educational, cultural, infrastructural, and regulatory factors. The strong negative correlation between literacy rate and speed of crossing in North America ($r = -0.70$) could indicate that areas with lower educational attainment could face challenges in awareness of pedestrian safety, which influence pedestrian decision-making processes. In contrast, Europe's negligible correlation ($r = -0.09$) may reflect more homogeneous educational outreach and standardised pedestrian safety measures [11].

The significant positive correlation between pedestrian crossing without traffic lights and the presence of trucks worldwide ($r = 0.41$) raises questions about pedestrian risk taking behaviours in commercial traffic environments. This finding could reflect inadequate pedestrian infrastructure or a cultural norm of assertive crossing behaviours in regions with substantial freight traffic.

The scatter plot analyses further reinforce the link between socioeconomic conditions, such as income inequality (Gini coefficient), and pedestrian behaviour. Countries with higher economic disparities showed longer pedestrian decision times, probably due to limited access to safe crossing infrastructure, poorer pedestrian education, or increased caution due to increased traffic risks. Similarly, the observed variability in crossing decision times relative to traffic mortality rates emphasises the need for targeted pedestrian safety interventions in countries experiencing high traffic-related deaths.

5 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 the available dashcam footage in different countries. Countries such as Denmark have significantly more footage compared to similarly sized nations like the Netherlands, potentially skewing pedestrian behaviour analyses globally.

Another limitation is the focus exclusively on urban environments, specifically cities with populations over 20,000, with exceptions for capital cities. This excludes potentially valuable information from suburban and rural settings, where pedestrian behaviour can differ significantly due to varying infrastructure and traffic dynamics. Furthermore, using national average human heights to convert pixel measurements into real-world distances introduces uncertainty, as individual heights vary substantially across different demographics, and camera angles might distort these measurements.

Additionally, the exclusive use of the YOLOv11 object detection model limits the analysis to basic pedestrian behaviours, potentially missing subtle but critical pedestrian signals such as gestures or posture changes. One such tool to study gestures and postures of pedestrians can be OpenPifPaf [19]. Resetting object identifiers between daytime and nighttime segments also disrupts the continuity of pedestrian tracking data, limiting longitudinal behavioural analyses.

Crucially, this study highlighted instances where significantly different contexts, such as Belgium and India, yielded almost identical average crossing times. This finding suggests the need for further investigation into metrics that more effectively capture nuanced behavioural differences and decision-making processes in diverse infrastructural and cultural contexts. Future studies should develop new context-sensitive metrics that can clearly distinguish these subtleties in behaviour.

In addition, the present analysis overlooks potential confounders such as weather conditions, seasonal variations, intersection design, and local traffic policies. Incorporating these variables into future studies could improve the robustness and applicability of pedestrian behaviour models.

Future research should also expand the dataset to ensure balanced representation in countries and integrate diverse urban, suburban, and rural environments. Furthermore, adopting advanced computer vision techniques capable of detailed recognition of pedestrian attributes such as intention, gait, and accessory analysis could significantly enrich our understanding of pedestrian behaviour. Future work should incorporate an analysis of nonstandard behaviours such as jaywalking [39, 42], cyclist dynamics [9, 14], and interactions between other vulnerable road users [22]. 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 [26]. Then, future studies must address the critical need to develop effective external human-machine interfaces (eHMIs) for (AVs) and interactive interfaces for modern cars [4, 5]. Such interfaces should consider global variations in pedestrian behaviour to effectively communicate the intentions of 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.

6 Dataset and Code Availability

The code is available at https://github.com/bazilinskyy/youtube-national. The PYT dataset is available at https://www.kaggle.com/datasets/anonymousauthor123/pedestrian-in-youtubepyt.

Acknowledgments

References

[1] 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.

[2] 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.

[3] Chaudhari Avinash, Shah Jiten, Shriniwas Arkatkar, Joshi Gaurang, and Parida Manoranjan. 2019. Evaluation of pedestrian safety margin at mid-block crosswalks in India. *Safety Science* 119 (2019), 188–198.

[4] 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

[5] 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

[6] 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

[7] Francisco Camara, Nicolai Navarro-Guerrero, Felipe Silva, and Norbert Krüger. 2020. Pedestrian models for autonomous driving part I: A study of pedestrian behavior in urban scenarios. In *2020 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 1361–1368.

[8] Zayed Uddin Chowdhury, Atiqur Rahman Chowdhury, Ahmed Al Jawad, Reaz Murshed, Aminur Rashid, Mohammed Mynuddin, Reza Ahmari, and Ahmad Mohammadi. 2024. Performance comparison of YOLO models for safety helmet detection: Insights from YOLOv5 to YOLOv10 with transfer learning. *Authorea Preprints* (2024). doi:10.36227/techrxiv.172789022.28202678/v1

[9] Tim De Ceunynck, Gert Jan 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

[10] 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

[11] Cong Ding, Xiaolei Ma, Yinhai Wang, and Yunpeng Wang. 2019. Urban traffic congestion estimation and prediction based on floating car trajectory data. *Future Generation Computer Systems* 96 (2019), 596–608.

[12] Etienne Duim, Maria Lucia Lebrão, and José Leopoldo Ferreira Antunes. 2017. Walking speed of older people and pedestrian crossing time. *Journal of Transport & Health* 5 (2017), 70–76.

[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*. 53–62. doi:10.1145/3640792.36757

[14] Rahul Goel, Anna Goodman, Rachel Aldred, Ryota Nakamura, Lambed Tatah, Leandro Martin Totaro Garcia, Belen Zapata-Diomedi, Thiago Herick de Sa, Geetam Tiwari, Audrey de Nazelle, et al. 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] B Hoe Goh, Kulanthayan Subramaniam, Y Tuck Wai, A Ali Mohamed, and A Ali. 2012. Pedestrian crossing speed: the case of Malaysia. *International Journal for traffic and transport engineering* 2, 4 (2012), 323–332.

[16] Atif Hussain and Sanjeev Sinha. 2021. Investigation of pedestrian crossing behavior at crossing facilities in India. *Civil Engineering and Architecture* 9, 5 (2021), 1442–1455.

[17] Rasheed Hussain and Sherali Zeadally. 2019. Autonomous Cars: Research Results, Issues, and Future Challenges. *IEEE Communications Surveys and Tutorials* 21, 2 (2019), 1275–1313. doi:10.1109/COMST.2018.2869360

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

[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*. 11977–11986.

[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] 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.

[22] 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

[23] 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

[24] Maharashtra Highway Traffic Police. 2025. Road Safety Guidelines. https://highwaypolice.maharashtra.gov.in/road-safety-guideline/ Accessed: 2025-04-10.

[25] Hilde Meersman and Marzieh Nazemzadeh. 2017. The contribution of transport infrastructure to economic activity: The case of Belgium. *Case studies on transport policy* 5, 2 (2017), 316–324.

[26] 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

[27] 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

[28] 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.

[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*. 226–235.

[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 *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 3105–3112. doi:10.1109/ITSC.2018.8569415

[34] 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

[35] 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.

[36] V Wickramasinghe, PA Wijesinghe, and KGM Lakmali. 2021. Start-Up Time of Crossing Pedestrians with Different Signal Configurations. *Engineer: Journal of the Institution of Engineers, Sri Lanka* 54, 4 (2021).

[37] World Health Organization. 2023. *Global Status Report on Road Safety 2023*. World Health Organization, Geneva. https://www.who.int/publications/i/item/9789240086517 Licence: CC BY-NC-SA 3.0 IGO.

[38] 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.

[39] SQ Xie, SC Wong, Tsz Man Ng, and William HK Lam. 2017. Pedestrian crossing behavior at signalized crosswalks. *Journal of transportation engineering, Part A: Systems* 143, 8 (2017), 04017036. doi:10.1061/JTEPBS.0000055

[40] Jakub Zębala, Piotr Ciępka, and Adam Reza. 2012. Pedestrian acceleration and speeds. *Problems of Forensic Sciences* 91 (2012), 227–234.

[41] 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

[42] 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

Received 10 April 2025; revised 30 April 2025; accepted 10 June 2025