CETRA2024 15-17 May 2024, Cavtat, Croatia
8th International Conferenceon Road and Rail Infrastructure

PREDICTING PEDESTRIAN CROSSING SPEED AT UNSIGNALIZED INTERSECTIONS WITH XGBOOST

Pelin Onelcin, Yalcin Alver
Ege University, Civil Engineering Department, Türkiye

Abstract

Understanding pedestrian behavior at unsignalized intersections becomes crucial for improving traffic safety. In this study, the crossing speeds of pedestrians at four unsignalized intersections in Izmir, Türkiye were investigated. Data were collected on weekdays for one hour each during peak and off-peak hours using unmanned aerial vehicles and cameras. The study incorporated pedestrian crossings executed on the designated crosswalks, as well as those occurring within 30 meters from the crosswalk. Two of the intersections did not have crosswalks, thus pedestrian crossings within the initial 40 meters from the roadway were considered. This study employed the eXtreme Gradient Boosting (XGBoost) machine learning algorithm to predict pedestrian crossing speed based on a comprehensive set of factors. Hyperparameter tuning was performed to optimize the model's performance. Various features, including the existence of a crosswalk, pedestrian age (young, adult, elderly), gender, group crossings, and load-carrying situations were considered in the model. The results of the study provide valuable insights for traffic management. This research contributes to the ongoing efforts to enhance pedestrian safety.

Keywords: pedestrian crossing speed, unsignalized intersections, XGBoost

1 Introduction

Exploring crossing speed is crucial as it is significantly different from walking speeds [1]. Montufar et al. [2] compared the normal walking speed and crossing speed of pedestrians at signalized intersections and found that the normal walking speed of pedestrians at signalized intersections is less than the crossing speed, and younger pedestrians walk faster than older pedestrians, regardless of the season and gender.

Signalized crosswalks prioritize pedestrian safety by consistently giving them the right of way. In contrast, unsignalised crosswalks increase pedestrian risk and require faster crossing speeds. Goh et al. [3] verified this assertion by observing 1579 pedestrian crossings at both signalized and un-signalized crosswalks in Malaysia. Pedestrians at non-signalized crosswalks had significantly faster crossing speed than at signalized crosswalks. The Bivariate analysis results showed that crosswalk type, age, and gender significantly contributed to pedestrian crossing speed.

Govinda et al. [4] investigated the pedestrian crossing behavior at both uncontrolled intersections and mid-blocks. ANOVA test was used to find out the parameters that affect the crossing speed. A notable difference in crossing speeds was observed between intersections and mid-block locations, with the average pedestrian crossing speed at mid-block (1.26 m/s) exceeding that at intersections (1.13 m/s). Another finding was that male and middle-aged pedestrians had higher crossing speeds than females and young/old pedestrians.

Aghabayk et al. [5] studied 552 pedestrian crossings at signalized and unsignalized crosswalks. The researchers examined the influence of gender, age group, group-crossing, technological devices, and carrying items on pedestrians crossing behaviors. They found that male and spatially violating pedestrians cross faster and older and group-crossing pedestrians move slower. In the study of [6] pedestrian crossing speeds were analyzed by age group at un-signalized mid-block crosswalks. Young adults walked faster than middle-aged and older pedestrians. Alver et al. [7] observed 498 pedestrian crossings in Izmir, Türkiye at mid-block crossings. They indicated that pedestrians who crossed individually walked faster than pedestrians who crossed in a group. This study investigates the critical aspect of understanding pedestrian behavior at unsignalized intersections, focusing on pedestrian crossing speeds in Izmir, Türkiye. The investigation covers four intersections, utilizing unmanned aerial vehicles and cameras during both peak and off-peak hours on weekdays. The study encompasses designated crosswalks and areas within 30 meters from the crosswalk, with additional consideration for intersections lacking crosswalks within the initial 40 meters from the roadway. Employing the eXtreme Gradient Boosting (XGBoost) machine learning algorithm, the research predicts pedestrian crossing speed based on an extensive set of factors, including crosswalk existence, pedestrian age, gender, group crossings, and load-carrying situations. The study's importance lies in addressing the challenges of unsignalised intersections, where pedestrian risk increases, and faster crossing speeds are required. By identifying factors that influence pedestrian behavior, this research will assist in the development of targeted strategies to mitigate safety concerns at unsignalised intersections and contribute to the broader goal of improving overall road safety.

2 Study area and data collection

The study aimed to comprehensively investigate pedestrian behavior at unsignalized intersections, considering different pedestrian crossing scenarios. To achieve this, the study focused on two main types of locations: intersections with designated crosswalks, and intersections lacking crosswalks. At intersections with designated crosswalks, pedestrian crossings were observed on the crosswalk itself, as well as within of 30 meters from the crosswalk. This approach allowed for the examination of pedestrian behavior not only at the designated crossing points but also in the vicinity of these points, where pedestrians may choose to cross for reasons such as proximity to their destination. To capture this aspect of pedestrian behavior, pedestrian crossings within 30 meters from designated crosswalks were meticulously considered, as well as intersections lacking crosswalks within the initial 40 meters from the roadway.
Data were collected on weekdays for one hour each during peak and off-peak hours using unmanned aerial vehicles and cameras. Two cameras were strategically mounted on tripods at each side of the crosswalk/road, providing comprehensive coverage of pedestrian movements from multiple angles. Additionally, a third camera was positioned approximately in the middle of the observed section, situated between 15 - 20 meters from the roadway. With this configuration, high-quality footage of pedestrian activities was captured, ensuring sufficient visual data for accurate analysis. The use of a multi-camera setup facilitated the determination of various attributes such as gender and approximate age.

2.1 Study area

In Izmir, intersections with high pedestrian-vehicle activity were identified, and a study was conducted at four unsignalized intersections. Two of these intersections (Hipodrom and Sirinyer) are located in the Buca district, while the other two (Stadyum and Ozkanlar) are in the Bornova district. The photographs of the intersections at which the field study was carried out are shown in Fig. 1 below. Table 1 shows the geometric characteristics of the intersections.

Figure 1 The study areas a) Hipodrom b) Sirinyer c) Stadyum d) Ozkanlar

Table 1 Geometric characteristics of the intersections

Intersection	Existence of crosswalk	Crosswalk length/Road width [m]	Crosswalk width [m]
Hipodrom	No	8.75	-
Sirinyer	No	6.95	-
Stadyum	Yes	11.10	4.40
Ozkanlar	Yes	6.10	4.10

3 Method

Pedestrian data were extracted manually from the video recordings. Spatial reference points were strategically positioned at one-meter intervals within the field, subsequently serving as a foundation for the creation of a 1 x 1 m grid using the AutoCAD program. The overlay of this grid onto the video footage facilitated precision in measurements. Fig. 2 illustrates an example of an applied grid system.

For each observed intersection, attributes of pedestrians, including gender, approximate age, load-carrying status, and individual or group crossings, were meticulously recorded in an Excel spreadsheet. Pedestrians crossing with objects such as backpacks, shopping bags, suitcases, and strollers are classified as pedestrians carrying items. Age was grouped into three categories with the best estimate from the recordings. The pedestrian crossing speed at each observation site was determined by dividing the distance traveled by each pedestrian by their crossing time.

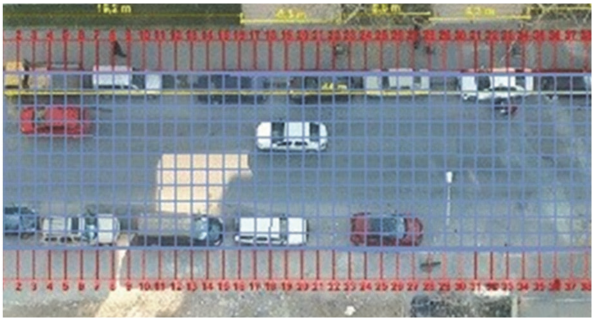

Figure 2 The grid system applied in the video footage

XGBoost, which is short for eXtreme Gradient Boosting, package was employed as the machine learning algorithm for predictive modeling in this study. XGBoost is a powerful and widely used ensemble learning technique that combines the strengths of gradient boosting and decision tree-based models [8]. It is particularly suitable for regression and classification tasks, making it well-suited for the objective of predicting pedestrian crossing speeds. Parameters that were used in the analysis were: learning rate (learning_rate) = 0.01, number of weak learners (n_estimators) = 500, maximum depth (max_depth) = 4. learning_rate determines how fast the model converges, n_estimator is the number of gradient-boosted classification trees in the XGBoost model, max_depth is the maximum depth of one classification tree in the XGBoost model.

4 Findings

The pedestrian crossing speeds were determined for all intersections. The average and the 15^{th} percentile pedestrian walking speed was found to be 1.34 m/s and 1.02 m/s, respectively. The 15th percentile of pedestrian speed can be used as design speed, by assuming that 85% of pedestrians walk faster than this speed.
Across all demographic and situational categories, a total of 2,312 pedestrians were recorded. This comprehensive sample size enables a robust analysis of pedestrian behaviors and characteristics across different variables such as gender, group size, age, and whether they are carrying items or not. The average walking speed, standard deviation, and 15th percentile speed values of pedestrians are shown in Table 2 based on the examined characteristics.
Females account for 46.32% of the sample, with an average speed of 1.29 m/s, while males make up 53.68% with an average speed of 1.38 m/s. The majority of pedestrians (66.87%) are individuals, with an average speed of 1.37 m/s, while the remaining 33.13% are part of groups of 2 or more, with a slightly lower average speed of 1.28 m/s. The data reveals that the adult group comprises the largest portion, representing 71.28% of the total sample. Their average walking speed is 1.37 m/s, with a standard deviation of 0.33 m/s, and a 15th percentile speed of 1.04 m/s. The results suggest that younger pedestrians walk faster than adults and the elderly. Those who carry items while crossing represent 42.26% of the sample. Despite carrying additional loads, their average walking speed remains the same as those without items, at 1.34 m/s. The standard deviation and 15th percentile speed for this group are also similar, indicating that carrying items does not significantly affect their walking speed. This suggests that, on average, pedestrians maintain a consistent walking speed regardless of whether they are carrying items or not. However, individual variations may exist within each group.

Table 2 Geometric characteristics of the intersections

		Count	**Percentage** [%]	**Average speed** [m/s]	**Std. deviation**	**15th percentile speed** [m/s]
Gender	Female	1071	46.32	1.29	0.32	1.00
	Male	1241	53.68	1.38	0.35	1.03
Group size	Individual	1546	66.87	1.37	0.34	1.02
	Group of 2 and more	766	33.13	1.28	0.32	0.99
Age	Young	292	12.63	1.39	0.34	1.07
	Adult	1648	71.28	1.37	0.33	1.04
	Elderly	372	16.09	1.17	0.31	0.87
Items-carrying	Without items	1335	57.74	1.34	0.34	1.02
	With items	977	42.26	1.34	0.33	1.02
Total		2312	100.00	1.34	0.34	1.02

Apart from the above results in intersections with crosswalks, the average pedestrian crossing speed was found to be 1.38 m/s, whereas in intersections without crosswalks, it was found to be 1.27 m/s.
In this study, the authors utilized an XGBRegressor, a variant of the popular XGBoost algorithm specifically tailored for regression tasks, to analyze the dataset. XGBRegressor is well-suited for predicting continuous numerical outcomes, making it an ideal choice for modeling pedestrian walking speeds. By employing XGBRegressor, the researchers aimed to accurately predict pedestrian walking speeds based on various demographic and situational factors.
Fig. 3 shows the feature importances. Higher feature importance values indicate that a feature has a greater impact on the model's predictions, while lower values suggest less influence. Interpreting feature importance values can provide insights into which factors are most influential in determining pedestrian crossing speeds.

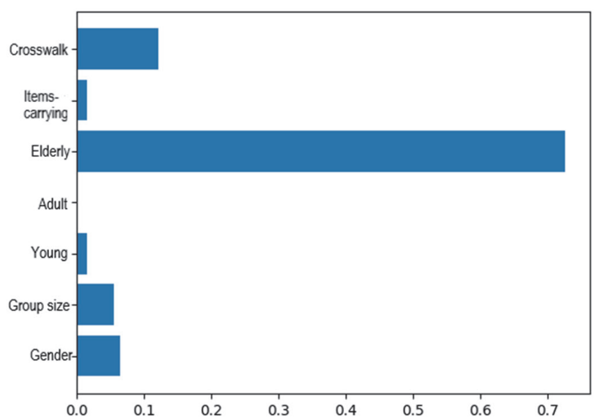

Figure 3 Feature importances

With an importance value of approximately 0.0645, gender demonstrates a modest influence on predicting pedestrian crossing speeds within the model. Group size, represented by an importance value of around 0.0559, also shows a moderate influence on pedestrian crossing speeds. This indicates that the number of individuals walking together can affect their overall speed to a certain degree. With an importance value of approximately 0.7272, age (elderly pedestrians) emerges as one of the most influential factors in predicting pedestrian crossing speeds within the model. This indicates that this particular age group has a significant impact on determining crossing speeds compared to other variables considered. Whether pedestrians are carrying items or not, represented by an importance value of approximately 0.0149, appears to have a minimal impact on their crossing speeds compared to other factors. The presence or absence of a crosswalk, with an importance value of around 0.1219, emerges as a significant predictor of pedestrian walking speeds. This suggests that designated crossing areas play a notable role in influencing how pedestrians adjust their speeds. Among the features considered, variables such as the presence of crosswalks and demographic factors like age demonstrate significant influence. Gender and group size also contribute to predicting walking speeds, albeit to a lesser extent.

5 Conclusions

Several critical aspects of pedestrian behavior at unsignalized intersections were illuminated by the findings of this study. In intersections with crosswalks, the average pedestrian crossing speed was found to be 1.38 m/s, whereas in intersections without crosswalks, it was found to be 1.27 m/s. This situation may arise from pedestrians feeling less secure, leading to faster crossings in intersections without crosswalks. The XGBoost analysis results have also emphasized the importance of the presence of crosswalks.

It was observed that age, particularly among elderly pedestrians, was identified as one of the most influential factors in predicting crossing speeds, thus highlighting the importance of age-related factors in traffic safety measures. While gender and group size were found to contribute to predicting walking speeds, their impact was relatively moderate compared to age and the presence of crosswalks. Interestingly, whether pedestrians were carrying items or not was observed to have minimal influence on their crossing speeds.

Through the identification of key factors influencing pedestrian crossing speeds, the development of targeted interventions to mitigate safety concerns and improve road safety for pedestrians in urban environments is supported by this study. At intersections where the elderly population is high crosswalks with features such as adequate width and clear markings can be designed. With technological enhancement, the use of smart crosswalks can serve as a solution to safety concerns.

Acknowledgement

The financial support of TUBITAK is greatly acknowledged. The authors also sincerely thank the support of Ömercan Ekinci, Handenur Keser, and Merve Azmaner during the fieldwork and data extraction.

References

[1] Chandra, S., Bharti, A.K.: Speed distribution curves for pedestrians during walking and crossing, Procedia - social and behavioral sciences, 104 (2013), pp. 660-667, DOI: 10.1016/j.sbspro.2013.11.160

[2] Montufar, J., Arango, J., Porter, M., Nakagawa, S.: Pedestrians' normal walking speed and speed when crossing a street, Transportation research record, 2002 (2007) 1, pp. 90-97, DOI: 10.3141/2002-12

[3] Goh, B.H., Subramaniam, K., Wai, Y.T., Mohamed, A.A.: Pedestrian crossing speed: The case of Malaysia, International journal for traffic and transport engineering, 2 (2012), pp. 323-332, DOI: 10.7708/ijtte.2012.2(4).03

[4] Govinda, L., Abhigna, D., Nair, P.M., Shankar, K.R.: Comparative Study of Pedestrian Crossing Behaviour at Uncontrolled Intersection and Midblock Locations, Transportation Research Procedia, 48 (2020), pp. 698-706

[5] Aghabayk, K., Esmailpour, J., Jafari, A., Shiwakoti, N.: Observational-Based Study to Explore Pedestrian Crossing Behaviors at Signalized and Unsignalized Crosswalks, Accident Analysis and Prevention, 151 (2021), 105990

[6] Forde, A., Daniel, J.: Pedestrian walking speed at un-signalized midblock crosswalk and its impact on urban street segment performance, Journal of traffic and transportation engineering, 8 (2021), pp. 57-69, DOI: 10.1016/J.JTTE.2019.03.007

[7] Alver, Y., Onelcin, P., Cicekli, A., Abdel-Aty, M.: Evaluation of Pedestrian Critical Gap and Crossing Speed at Midblock Crossing Using Image Processing, Accident Analysis and Prevention, 156, pp. 106127, 2021.

[8] Chen, T., Guestrin, C.: XGBoost: A Scalable Tree Boosting System, 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA, 13 - 17 August 2016.