Pedestrian road safety: Advances in measuring pedestrian risk exposure

Keywords: Pedestrian risk exposure, Pedestrian volume, Video analytics, Pedestrian-vehicle crashes, Crash modelling

Extended Abstract

Despite the well-known benefits of walking for health and transport aspects (i.e., active lifestyle, low carbon emission, etc.), pedestrians are the most vulnerable among all road users. Focusing on the conflicts between pedestrians and vehicles, several scholars have argued that pedestrian risk exposure (PRE) needs to be comprehensively estimated in quantitative modelling (Sze et al., 2019).

Several pedestrian risk exposure measures have been proposed for pedestrian safety analysis, such as population and walking trip characteristics (e.g., frequency, distance and duration) (Lam et al., 2014; Yao et al., 2015). However, pedestrian volume data have seldomly been collected officially on a large scale and in a time-variant manner. To this end, this study attempts to leverage the dashcam videos from buses and computer vision techniques to estimate hourly pedestrian volume in a time-variant manner. The pedestrian volume (PV_BDV) is then incorporated in crash prediction models as a PRE measure. Two other PRE measures, namely number of walking trips extracted from Travel Characteristics Survey (WT_TCS) and the number of pixels identified as pedestrians from Google Street View images (P_GSV) are also estimated.

Negative binomial (NB) regression and geographically weighted random forest (GWRF) regression are used to model pedestrian-vehicle crash frequency at the road segment level. The main objective is to test the robustness and compare the explanatory power of the proposed PRE measures. Moreover, this study discusses the relationships between various predictors and the target variable (i.e., pedestrian-vehicle crash frequency). Following the aforementioned research objectives, Figure 1 summarizes the main bodies of this study.

This case study in Hong Kong uses multiple data sources including pedestrian-vehicle crashes from 2016 to 2021 that are extracted from the police-reported Traffic Road Accident Database System (TRADS), bus dashcam videos and bus GPS data, TomTom road network data, and other open source spatial datasets.

Three PRE measures are first generated. To estimate ***pedestrian volume from bus dashcam videos (PV_BDV)***, a method combining Fast R-CNN and Deepsort algorithm is applied. Fast R-CNN (Girshick, 2015) detects objects from video frames, and Deepsort functions as an object tracker. Pedestrian counts from video frames are aggregated to GPS points according to timestamps. Then, GPS points are spatially matched to road segments with the help of a map-matching algorithm. Based on the number of pedestrians observed, hourly pedestrian volume can be estimate. ***Number of walking trips estimated from TCS 2011 (WT_TCS)*** and ***number of pixels identified as pedestrians from GSV images (P_GSV)*** are two other PRE measures. WT_TCS is estimated based on trip origins and destinations recorded in TCS 2011 using network analysis in ArcGIS software. P_GSV is generated using based on image segmentation method.

Table 1 shows part of observations from the regression analysis based on negative binomial (NB) model. The model including PV_BDV as PRE measure has the best model fit, indicating pedestrian volume estimated from videos can better explain the variance of

pedestrian-vehicle crash frequency. Moreover, the z-value of PV_BDV is the highest in the model. Other factors that are significantly associated to crash frequency include traffic speed, number of cautionary crossings, number of food related facilities, and vehicular traffic volume. There are 12 other factors that significantly affect the number of pedestrian-vehicle crashes but are not included in Table 1.

In terms of scientific contributions, this study elaborates the feasibility of measuring pedestrian risk exposure from bus dashcam videos with high efficiency and low cost. It can be applied in other geographical contexts and for long-term pedestrian risk exposure monitoring. Furthermore, the results shed light on the spatial-temporal variations of pedestrian activities and pedestrian-vehicle crashes.

References

Girshick, R. (2015). Fast R-CNN. *Proceedings of the IEEE International Conference on Computer Vision*, 1440–1448.

Lam, W. W. Y., Yao, S., & Loo, B. P. Y. (2014). Pedestrian exposure measures: A time-space framework. *Travel Behaviour and Society*, *1*(1), 22–30. https://doi.org/10.1016/j.tbs.2013.10.004

Sze, N. N., Su, J., & Bai, L. (2019). Exposure to pedestrian crash based on household survey data: Effect of trip purpose. *Accident Analysis & Prevention*, *128*, 17–24. https://doi.org/10.1016/j.aap.2019.03.017

Yao, S., Loo, B. P. Y., & Lam, W. W. Y. (2015). Measures of activity-based pedestrian exposure to the risk of vehicle-pedestrian collisions: Space-time path vs. Potential path tree methods. *Accident Analysis & Prevention*, *75*, 320–332. https://doi.org/10.1016/j.aap.2014.12.005

Table 1. Results of NB models incorporating different PRE measures (only shows top 5 high-impact factors)

Variables	Model 1 (PV_BDV)		Model 2 (WT_TCS)		Model 3 (P_GSV)	
	Coef. (S.E.)	z.	Coef. (S.E.)	z.	Coef. (S.E.)	z.
PRE measure	0.00*** (0.00)	7.65	0.00** (0.00)	2.80	0.00 (0.00)	0.86
Traffic speed	-0.03*** (0.00)	-7.58	-0.03*** (0.00)	-8.31	-0.03*** (0.00)	-8.27
Number of cautionary crossings	0.20*** (0.03)	6.26	0.21*** (0.03)	6.42	0.20*** (0.03)	6.16
Number of food related facilities	0.02*** (0.00)	5.05	0.02*** (0.00)	5.63	0.02*** (0.00)	5.60
Traffic volume	-0.00*** (0.00)	-4.47	-0.00*** (0.00)	-4.60	-0.00*** (0.00)	-4.59
N	24918		24918		24918	
AIC	8131.2		8173.7		8179.2	
BIC	8415.5		8458.0		8463.5	

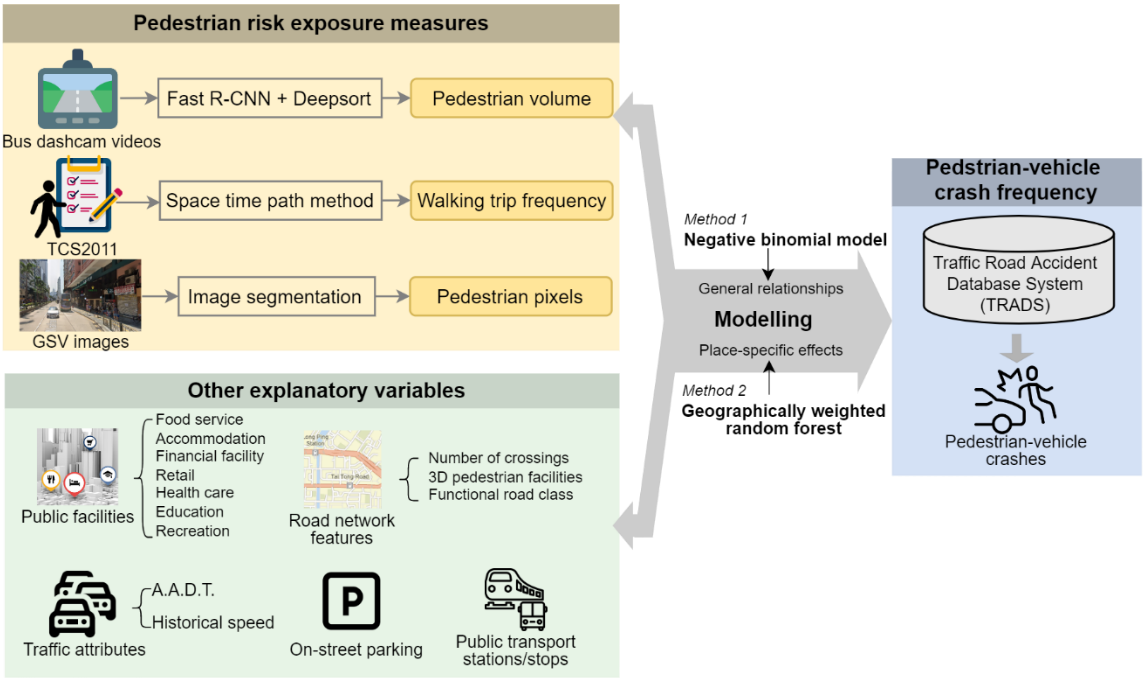

Figure 1. Research framework