Accident Analysis and Prevention 185 (2023) 107017

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Using computer vision and machine learning to identify bus safety risk factors

Becky P.Y. Loo [a,b], Zhuangyuan Fan [a,*], Ting Lian [a], Feiyang Zhang [a]

[a] Department of Geography, The University of Hong Kong, Hong Kong, China
[b] School of Geography and Environment, Jiangxi Normal University, Nanchang, China

ARTICLE INFO

Keywords:
Bus safety
Pedestrian behaviour
Video analytics
Crash modeling

ABSTRACT

In road safety research, bus crashes are particularly noteworthy because of the large number of bus passengers involved and the challenge that it puts to the road network (with the closure of multiple lanes or entire roads for hours) and the public health care system (with multiple injuries that need to be dispatched to public hospitals within a short time). The significance of improving bus safety is high in cities heavily relying on buses as a major means of public transport. The recent paradigm shifts of road design from primarily vehicle-oriented to people-oriented urge us to examine street and pedestrian behavioural factors more closely. Notably, the street environment is highly dynamic, corresponding to different times of the day. To fill this research gap, this study leverages a rich dataset - video data from bus dashcam footage - to identify some high-risk factors for estimating the frequency of bus crashes. This research applies deep learning models and computer vision techniques and constructs a series of behavioural and street factors: pedestrian exposure factors, pedestrian jaywalking, bus stop crowding, sidewalk railing, and sharp turning locations. Important risk factors are identified, and future planning interventions are suggested. In particular, road safety administrations need to devote more efforts to improve bus safety along streets with a high volume of pedestrians, recognise the importance of protection railing in protecting pedestrians during serious bus crashes, and take measures to ease bus stop crowding to prevent slight bus injuries.

1. Introduction

Road safety is a dire problem. Every year, about 1.3 million people died on the road; road traffic injuries are the eighth leading cause of death in the world in 2022 (WHO, 2022). Also, between 20 and 50 million more people suffered from non-fatal injuries, with many leading to disability (WHO, 2022). More worryingly, road traffic has been the leading cause of death among children (5–14 years old) and adolescents (15–29 years old) (WHO, 2022). The social costs of traffic fatalities and casualties are tremendous.

Hong Kong road safety records have improved over the last thirty years, with road deaths per 10,000 people dropping from 0.60 in 1983 to 0.14 in 2019 (Transport Department, 1983–2019). The corresponding figures per road km are 0.26 and 0.05, respectively (Transport Department, 1983–2019). Nonetheless, there have been major bus crashes in recent years (Road Safety Council, 2017–2021). In particular, crashes involving franchised buses need more focused research because of the large number of bus passengers involved and the challenge that it puts to the road network (with the closure of multiple lanes or entire roads for hours) and the public health care system (with multiple injuries that need to be dispatched to public hospitals within a short time). However, traditional methods of hazardous road locations focus primarily on road junctions; hence, the densest areas with dense road junctions have been identified as black spots (Transport Department, 2019). Moreover, research based on actual road crash frequency is biased towards urban areas with higher annual average daily traffic (AADT), when the exposure factor has not been considered (Kim et al., 2022; Yao et al., 2015).

A closer examination of fatal bus crashes in Hong Kong over the years shows that they generally did not happen at traffic black spots. Some notably serious bus crashes happened on highways and suburban roads in the last five years (Road Safety Council, 2017–2021). Moreover, studies of improving bus safety have focused mainly on the bus drivers, for example, whether they have sufficient rest, driving experience, driving attitudes, and even interactions with bus passengers (RSC, 2021). All these factors are indeed relevant. However, a whole series of factors related to bus safety is related to the dynamic road environment

* Corresponding author at: Room 1034, 10/F Jockey Club Tower, Pokfulam Road, Hong Kong Island, Hong Kong, China.
E-mail addresses: bpyloo@hku.hk (B.P.Y. Loo), yuanzf@connect.hku.hk (Z. Fan), tlian@connect.hku.hk (T. Lian).

https://doi.org/10.1016/j.aap.2023.107017
Received 29 October 2022; Received in revised form 11 February 2023; Accepted 27 February 2023
Available online 6 March 2023

and behaviour of other road users. These factors need to be considered specifically for bus safety, especially in Hong Kong, where buses are a major mode of transport; franchised buses alone carry around 4.4 million passengers per day (Transport Department, 2020).

Focusing on analysing bus crash frequency, this study proposes an approach to examine five risk factors that are rarely captured in previous studies: 1) pedestrian exposure; 2) pedestrian crowding at bus stops; 3) jaywalking; 4) railing protection; and 5) streets with sharp turns. These variables have long been considered important in dense urban areas while lacking empirical tests at high resolution and large geographical scale. To do so, this study leverages a unique data source, dashcam videos from buses in Hong Kong, to gather pedestrian behaviour data and railing along the bus road network. Then, three models are applied to estimate the relationships between the five risk factors and historical bus crash frequency by level of severity.

The remainder of this paper is structured as follows. The second section reviews the literature on bus crash studies. The third part explains detailed methods to derive risk factors and our model calibration process. Finally, the results are explained and discussed before a conclusion is made.

2. Literature review

This section reviews the related literature from three major aspects. The first subsection reviews studies on bus crash risk from the perspectives of crash severity, frequency, and exposure. Then, focusing on crash frequency analysis, three research gaps are identified. The second subsection summarises recent research progress on using computer vision to gather data for road safety analysis. The last part of the review summarises the statistical models and machine learning models frequently used to predict road crash frequency and to identify risk factors. The rationale of model selection in this study is also explained.

2.1. Bus crash risk factors

An event of road crash is generally viewed as a result of an interaction between the environment, vehicles, and drivers (Barabino et al., 2021; Goh et al., 2014). In the context of bus-related crashes, previous research has explored the effects of factors regarding the *frequency* and *severity* of road crashes and *risk exposure*. Regarding bus crash severity, drivers' profiles are most frequently studied. For instance, Kaplan and Prato (2012) found that bus crash severity increased with the presence of bus drivers under the age of 25 or older than 65. Other studies found that female, fatigued, and inexperienced bus drivers (Evans & Courtney, 1985; Huting et al., 2016; Samerei et al., 2021a) and those with a history of traffic violations are more likely to be involved in bus crash (Feng et al., 2016). Recent literature in bus crash severity prediction also included environmental variables in the models. These variables mostly focused on describing the road network, such as road intersection location (Samerei et al., 2021a), road types, speed limits, surface materials, etc. (Iranitalab & Khattak, 2017).

Another thread of studies focuses more on the analysis of bus crash frequency. Again, road environmental factors such as traffic volume, number and width of the lane(s), and type of road are widely recognised. Wider lanes and medians were found to reduce bus crashes, while more lanes with higher traffic volume increased bus-related crash occurrence (Chimba et al., 2010). High-speed zones are likely to be associated with higher bus crash rate (Samerei et al., 2021a). Curbside parking and loading are also positively correlated with more bus crashes (Barabino et al., 2021; Cheung et al., 2008).

The third thread of bus crash studies have put more attention to risk exposure, especially in dense urban areas. One of the biggest challenges this group of researchers faces is collecting pedestrian-related data. To overcome this challenge, some have relied on detailed bus crash records that directly describe pedestrian involvement in related bus crashes (Almasi et al., 2021; Samerei et al., 2021a; Samerei et al., 2021b). Others have tried to quantify pedestrian crash risk exposure by estimating pedestrian volume from household travel characteristics surveys (TCS), land uses, and points of interests (Almasi et al., 2021; Su & Sze, 2022; Yao et al., 2015).

There are three research gaps to be addressed in analysing bus crash frequency. First, a large body of studies has focused on driver-related risk factors. In contrast, pedestrian-related and environmental factors deserve closer examination with the availability of big data and advanced computing techniques. On the one hand, pedestrian activities recorded in police crash records only capture pedestrians involved in traffic crashes but not other pedestrians or the street scene in general. On the other hand, using TCS data to estimate pedestrian exposure is limited to the collection year. They are usually not up to date and have a limited sample size.

Second, in analysing bus crashes, bus drivers are generally familiar with the routes they typically drive along. Hence, they will get into a driving habit based on the usual road conditions. As a result, they will be more vulnerable to sudden events, such as jaywalking pedestrians. Such unexpected events are rarely captured in previous studies.

Lastly, environmental factors used in previous studies have rarely considered road safety-related installations. Features such as protection railings, safety islands, and extended sidewalks deserve further study to evaluate their effectiveness as buses are heavy vehicles, and people on unprotected walkways (including passengers waiting for buses at bus stops and pedestrians on the street) are particularly vulnerable to serious injury and fatalities.

Addressing these research gaps requires further data gathering and robust analysis. The following section first reviews the potential methods to gather data via video analysis in road safety studies.

2.2. Video analytics and its application in road safety analysis

One promising way of capturing street behaviours in cities is through video analytics. Researchers have used fixed cameras in cities to analyse pedestrian crossing behaviour (Avinash et al., 2019; Gitelman et al., 2019; Zhang et al., 2020), vehicle–pedestrian interaction (Beitel et al., 2018; Fu et al., 2019; Liang et al., 2021), and pedestrian gap acceptance (Gorrini et al., 2018; Sheykhfard & Haghighi, 2020). In parallel, recent developments in deep learning-based computer vision algorithms further enhance the capacity to leverage video data to capture pedestrian behaviour. Object detection models such as Fast R-CNN (Girshick, 2015) and Mask R-CNN (He et al., 2017), empowered with tracking algorithms such as Deep Sort (Wojke et al., 2017) can detect pedestrians, predict pedestrian behaviour, and identify pedestrian movement. However, these techniques have not yet been fully integrated into bus safety analysis.

2.3. Machine learning methods in road crash prediction

While a significant body of safety research has been dedicated to predicting road crash severity and frequency, here, the focus is on reviewing the methods used to conduct crash frequency prediction and risk factor analysis. A common approach in these studies is using a statistical modelling tool with crash frequency as the dependent variable and characteristics of the driver, roadway, time, etc., as independent variables. Developed from the regression approach, Negative Binomial (NB) (Hilbe, 2011) has been widely employed to overcome the problem of over-dispersion in the crash data (Abdulhafedh, 2016; Lord & Mannering, 2010; Mousavi et al., 2021). As such, NB models have a limitation when dealing with under-dispersion data (the mean of crash frequency is higher than the variance), leading to biased parameter estimates. Moreover, considering that crash events are rare, researchers also use zero-inflated negative binomial models and zero-inflated Poisson models to deal with this over-dispersion problem (Abdulhafedh, 2017). Still, with findings that the many underlying relationships between crash frequency and risk factors are non-linear, recent researchers have explored more non-linear approaches to predict the crash frequency. Machine learning models such as Artificial Neural Networks (ANN), Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) are applied as predictive tools in behavior-based risk analysis (Atumo et al., 2022; Chakraborty et al., 2019; Dong et al., 2015; Gu et al., 2023; Yang et al., 2022). Among these models, ANN and SVM, although strong in prediction, are less informative in explaining the importance of input variables. Neural Network models are generally still considered as a black box with feature transformation and dropout included in the process. Similarly, SVM conducts a kernel transformation of the original space depending on the kernel used. On the contrary, RF and XGBoost, although using different ensembled methods, are tree-based models that are easy to interpret. Recently RF models have been extended to discover causal effects (Wager & Athey, 2018). XGBoost has been used to examine the leading causes of traffic crashes (Yang et al., 2022).

3. Method

In view of the three research gaps, this study will focus on five bus crash risk factors that are rarely addressed in previous literature (Fig. 1). The following section will explain how these factors can be extracted via a unique dataset – bus dashcam video records. Then the importance of these factors is validated via three models selected per previous road crash frequency analysis literature.

3.1. Study area and data

The study area includes 244.36 km of street segments in Hong Kong, covered by 33 bus routes from morning to night. The road segment network was downloaded from the TomTom dataset. The dataset contains the road travel direction, name, and type (notably expressway, tunnel, etc.). Previous road crash studies have used both street intersections and street segments as the basic unit of analysis. Following previous work on pedestrian crash hot zones, this study converted the road network to Basic Spatial Unit (BSU) for detailed analysis (Yao & Loo, 2012). The length of BSUs included in this study ranges from 15 m to 100 m. Street segments that are expressways only (where pedestrians are not present) were removed from the sample collection.

The bus dashcam videos were collected from July 2021 to March 2022. All video data (Fig. 1 a) were first processed with an anonymiser algorithm[1] to blur individual faces for privacy protection. Ethical approval was obtained from the authors' university. All videos obtained are also accompanied by a GPS file. The GPS file contains the latitude, longitude, timestamp, bus ID, and route ID so that each video frame can be associated with a GPS point. Although GPS records give the estimated location of a bus, there are often missing data points or reported locations that deviate from the true location of the bus (Barabino et al., 2017; McLeod, 2007). To deal with this problem, all bus routes were manually matched to associated street segments via ArcGIS. Then applying a Map Matching algorithm (Lou et al., 2009), the GPS points were associated to those street segments of that particular bus route (known in advance). In addition, the video data are dropped for bus GPS records with known data gaps (missing location data identified by abnormal bus speed) to ensure accuracy and reliability. In this way, a video frame-GPS-street segment association is constructed for further analysis.

3.2. Risk factors

3.2.1. Pedestrian exposure factor

The pedestrian exposure factor is defined as the pedestrian volume at a BSU k through any given time period T (T is usually one hour) (Yao et al., 2015). Using video from the buses dashcam to obtain the P_{kT} involves two steps. The first step is to assign a unique ID for all pedestrians who appeared in the video during a bus trip. To do so, this study leverages a pedestrian tracking algorithm (Wang et al., 2020) that combines the Fast R-CNN (Girshick, 2015) and Deepsort (Wojke et al., 2017) to detect the unique number of visitors that appeared in video during any given time ΔT (Fig. 2a). The second step is to use the observed number of pedestrians P_{jk} to estimate the P_{kT}. This estimation adopts a commonly pedestrian average speed (disregarding direction) at $v_p = 1.203$ m/s (Chandra & Bharti, 2013). Then for any pedestrian, it takes $t = \frac{L_k}{v_p}$ to pass through a BSU with length L_k. For a bus j passing through a BSU k during time period Δt, its speed is captured as v_{bk}. Given that $\overline{v_p} \ll \overline{v_b}$, bus j could observe all pedestrians at BSU k during time T:

$$P_{jkT} = \frac{P_{jk} \times v_p}{L_k} \times T \times AF \tag{1}$$

$$AF = \begin{cases} 1, & if\ v_{bk} \geq \overline{v_b}/2 \\ \frac{v_{bk}}{\overline{v_b}}, & otherwise \end{cases} \tag{2}$$

where AF is an adjustment factor considering the variation of the bus speed through different videos. When the bus speed is much lower than the average speed, it will capture more pedestrians. P_{jkT} is adjusted downward by the ratio of v_{bk} and $\overline{v_b}$. The detailed process resembles a previous study (Lian et al, 2022). Since there are more than one videos taken by different buses for each BSU, the final pedestrian exposure P_{kT} is measured as the mean of the P_{jkT} for all n videos taken:

$$P_{kT} = \frac{1}{n}\sum_{j=1}^{n} P_{jkT} \tag{3}$$

3.2.2. Pedestrian jaywalking index

The process of generating the jaywalking index involves two steps. The first step trains a Mask R-CNN model using 200 manually labelled images and applies the model to all videos (He et al., 2017). These labelled samples were sufficient to train the model and achieved relatively good results compared to the original Mask R-CNN coco dataset's mask mAP. Details of the model can be found in Fan & Loo (2021). The manual labelling process identifies a pedestrian's location, shape, and if they are jaywalking or not. Then this model is applied to all bus videos, and the detection results are aggregated to road segments to derive the jaywalking index.

The pedestrian jaywalking index is defined as the likelihood of observing a jaywalking person at a road segment s through any given time period t. Given that a person could be jaywalking at a point in time but not during another time, this jaywalking measure avoids the absolute count of the number of people jaywalked and applied the following equation to calculate the jaywalking index for a BSU k.

$$J_k = \frac{\sum_{m=1}^{M_k} n_m}{M_k} \tag{4}$$

where the n_m is the total number of jaywalking people detected in any given video frame m, and M is the total number of video frames that have covered the BSU k in the study sample. Distribution of captured jaywalking is shown in Fig. 2b.

3.2.3. Pedestrian crowding at bus stops

Pedestrian crowding C_k measures the average pedestrian area occupancy (APAO) within a 10-meter buffer of a given bus stop (Fig. 2c). In order to calculate APAO, the sidewalk width information at the bus stop locations is needed to estimate the amount of space available for pedestrians and waiting bus passengers. In this study, the sidewalk width is estimated based on the surrounding land use of the road network, as the minimum width of the sidewalk has been specified for different types of land use in the *Hong Kong Planning Standards and Guidelines* (Planning Department, 2022). The spatial join function in ArcMap was used to find

[1] https://github.com/animikhaich/Real-Time-Face-Anonymizer.

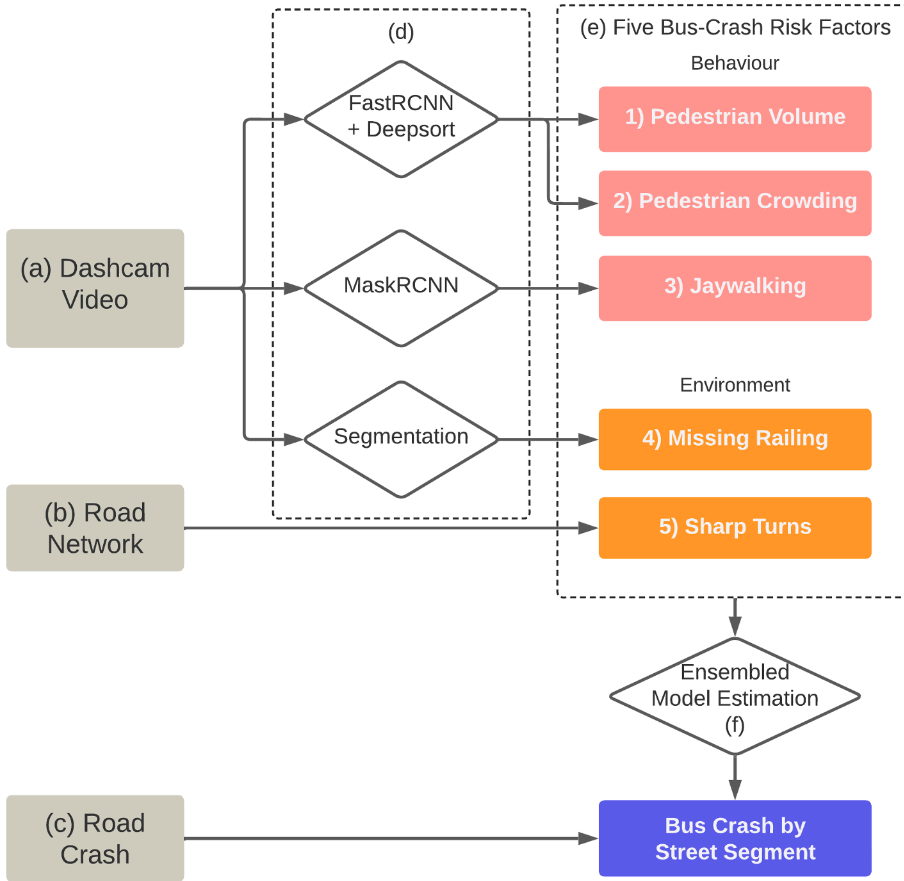

Fig. 1. The research structure. On the left, (a), (b), and (c) show the major data sources. Via methods in (d), four out of five risk factors are extracted. e) The five risk factors include 1) pedestrian volume, 2) pedestrian crowding, 3) jaywalking, 4) missing railing, 5) sharp turn. f) Final models are developed to estimate the bus crash frequency by each street segment.

the nearest land use information for each BSU. A distance threshold of 30 m is applied to ensure the accuracy of matching. This threshold was selected based on the maximum distance from any given street centerline to its adjacent defined land use block. Then, the corresponding sidewalk width w_k of BSU k is determined based on the identified nearest land use. For BSUs that have no land use value found, their sidewalk width is considered to be the same as the nearest street segments with the same street names. The number of pedestrians P_k near bus stops is derived by counting the number of pedestrians (with unique IDs) within 10 m buffers from each bus stop. Finally, C_k is calculated using equation (5):

$$C_k = \frac{P_k}{w_k \times 10} \tag{5}$$

where 10 m is the buffer length used in the measure. Fig. 3c illustrates the results captured in each video frame and the aggregated distribution of C_k through all observed BSUs. The distribution of pedestrian crowding at bus stops are visualized in Fig. 2c.

3.2.4. Sidewalk railings

The sidewalk railings are detected from the bus video frames using an image segmentation model, which was trained based on the ADE20K dataset (Zhou et al., 2017). Night videos (videos covering time after 18:00 pm), except for the case where alternative daytime videos cannot cover the same BSUs, are excluded from the analysis to ensure the accuracy of segmentation. Based on the image segmentation result, the percentage of railing pixels on each bus video frame can be calculated. Mathematically, the railing index R_k is defined as:

$$R_k = \begin{cases} \dfrac{P_{railing,k}}{bound}, & if\ P_{railing} \le bound; \\ 1 & otherwise \end{cases} \tag{6}$$

where $P_{Railing,k}$ is the average percentage of railing pixels on a BSU derived by calculating the mean value of the percentage of railing pixels of all bus video frames on the same BSU. Given other needs like pedestrian crossings and drop-off areas, a BSU is considered to be fully protected by railings if $P_{Railing}$>=bound (Fig. 3 right). Here the bound is a threshold to classify the BSU with railings. To identify the threshold, 100 images with railing detected were randomly selected. Then the images were labelled by three classes: 1) no railing; 2) some railing, but the BSU has opening; 3) continuous railing – the BSU has no opening between the sidewalk and the vehicle road. The bound is determined as 0.06, which is the median value of $P_{Railing}$ of all images with continuous railing detected.

3.2.5. Sharp-turn locations

Sharp turns can be a risk factor for buses, especially in Hong Kong, where most of them are double-deckers and over 12 m long. Based on the road network structure, a dummy variable S_k is used here to describe if a BSU k is associated with a sharp turn of fewer than 90 degrees or not, where 1 indicates that the street segment is associated with at least one sharp turn, whereas 0 indicates no sharp turn.

3.3. Road crash data

To systematically estimate these five risk factors, this study collected

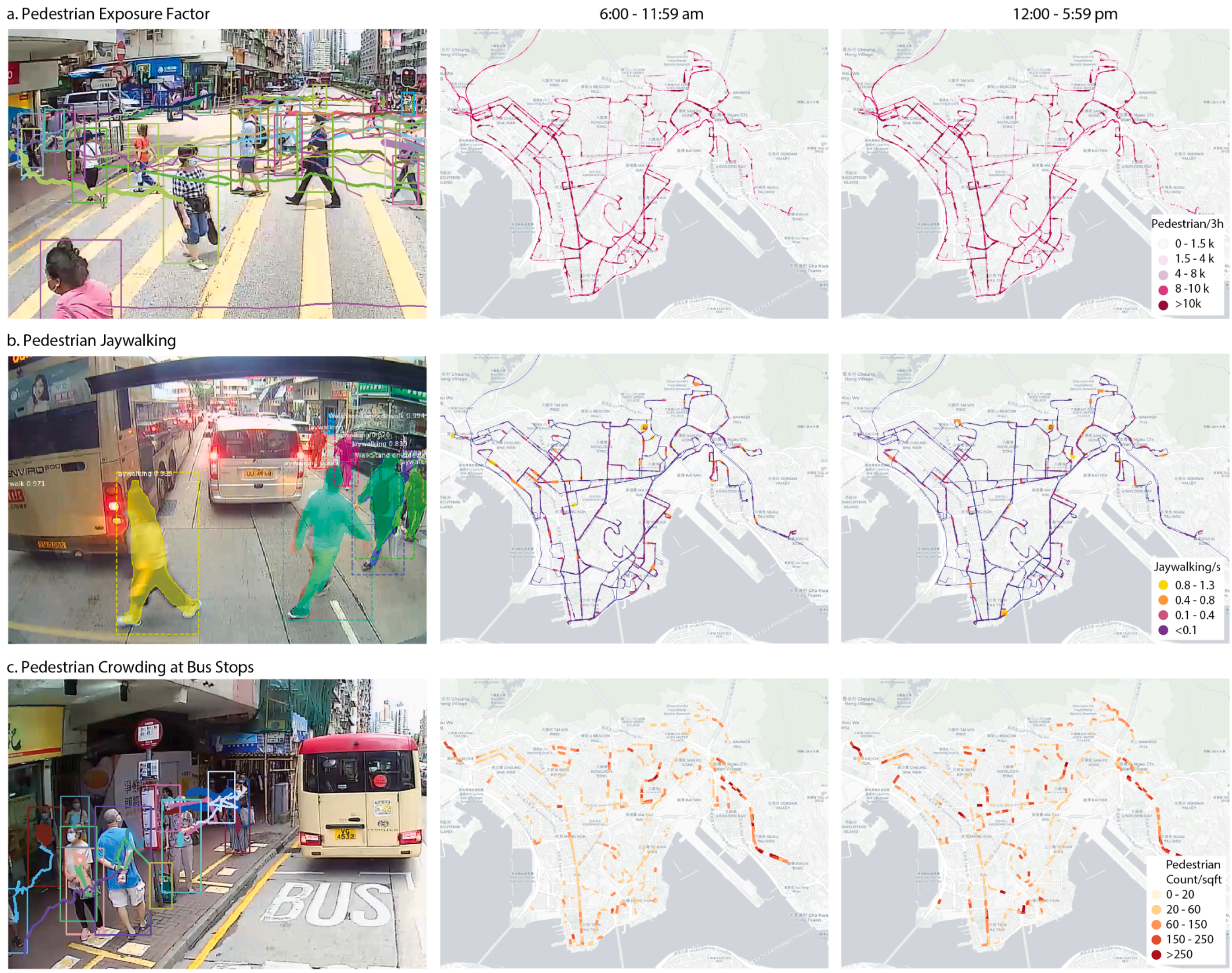

Fig. 2. Three behavioural risk factors constructed from the dashcam video. A) Pedestrian exposure. B) Pedestrian Jaywalking. C) Pedestrian crowding at bus stops. Maps of distribution are plotted accordingly at different times of a day.

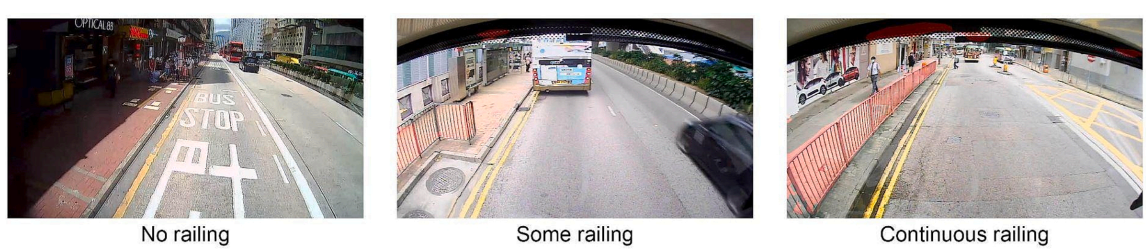

Fig. 3. Levels of railing presence at BSUs.

road crash data from 2015 to 2019 and aggregated them to the BSU level. The procedure of combining traffic crashes in five years is to recognise that bus crashes are rare events. The randomness of traffic crashes needs to be properly taken into account when identifying systematic bus crash risk factors. The original data was provided by the Transport Department of Hong Kong. The data contains the fields of location coordinates of crashes, the street address of crashes, the type of vehicle involved and the severity of crashes (slight, serious, and fatal), among others. With these parameters, the crash data were further aggregated into three groups: 1) bus-related serious crashes, which include both serious and fatal crashes; 2) bus-related slight crashes; and 3) non-bus-related crashes. From 2015 to 2019, a total of 77,060 crashes happened in Hong Kong. This study used 12,679 crashes that were bus-related. Among all bus crashes, 1,408 were fatal or severe, and 11,189 were slight crashes.

Table 1 shows the summary statistics of all continuous variables. It is observed that most of the features are not normally distributed. Hence, the transformation of the variables was conducted with least changes to the dataset to improve the model performance while maintaining the interpretability of the parameters. Fig. 4 plots the correlation among transformed risk factors and that there is no multi-collinearity among the features.

Table 1
Summary statistics of key parameters.

	count	mean	std	min	max
Log(Pedestrian Exposure)	3548	7.162	0.847	2.916	9.92
Log(Jaywalking)	3548	0.077	0.094	0.0000	0.964
Log(Pedestrian Crowding @ Bus Stop)	3548	0.645	1.645	0.0000	6.185
Railing Index	3548	0.212	0.2630	0.0000	1.000
Log(Segment Length)	3548	3.824	0.662	1.800	4.615

3.4. Risk feature importance

This section uses multiple models to evaluate the importance of each risk factor in predicting road crashes. As discussed in the literature review, NB is one of the widely used statistical models to analyse crash frequency. However, tree-based machine learning models are found with better model fit in comparison to statistical models and more explainable than other high-dimensional machine learning models such as ANN or SVM. Still, this study admits the fact that all these models could still have their own limitations. Therefore, it applies three different models and compares their results to verify the effects of the five selected risk factors. The three models are 1) NB, 2) RF (Breiman, 2001), and 3) XGBoost (Chen & Guestrin, 2016). Here, NB is used as a baseline to compare with the performance of the other two models. RF and XGBoost are both tree-based ensemble models. The former trains separate decision trees and output the mean of the prediction from all trees. The latter trains decision trees sequentially. Recent literature has demonstrated both models have out-of-sample prediction capability and can identify non-linear relationships between the features and target/dependent variables. Without previous knowledge of which one would outperform the other, the experiments were conducted with both models, and their performance is obtained. The following sub-section first explains the hyperparameter tuning process for the RF and XGB models. Then, the model performance and feature importance are compared.

3.4.1. Hyperparameter tuning for Random Forest regression

To select the appropriate hyperparameters for the RF model, the data were split into training and test datasets (70%: 30%). Then using a 5-fold cross validation (CV) from the training dataset, the following hyperparameters are selected:

1. n_estimator: number of trees in the forest
2. max_features: maximum number of features considered
3. max_depth: maximum number of levels in each decision tree
4. min_samples_split: minimum number of data points placed in a node before the node is split
5. min_sample_leaf: minimum number of data points allowed in a leaf node.

This experiment first used the randomised CV search method to narrow down the number of combinations of all parameters. Then the Grid Search method was used to evaluate all combinations defined. The evaluation process uses the Mean Absolute Error (MAE) given that many street segments do not have bus crashes recorded. Table 2 summarises the hyperparameters selected for the RF model.

3.4.2. Hyperparameter tuning for Extreme Gradient Boosting (XGBoost) regression

Similarly, the following experiment adopted a 5-fold cross-validation process with the training dataset and fine-tuned with the following parameters for the XGBoost model:

1. max_depth: maximum number of levels in each decision tree
2. colsample_bytree: the fraction of columns to be randomly sampled from each tree
3. subsample: the fraction of observations to be sampled for each tree.
4. learning rate: Step size shrinkage used in the update to prevent overfitting.
5. min_child_weight: Minimum sum of instance weight (hessian) needed in a child.

Table 3 summarizes the parameter selected for each objective.

3.4.3. Feature importance

Feature importance is one of the key criteria to explain machine learning models. There are many feature-importance calculation methods that capture different aspects of the model. Instead of the

Table 2
Hyperparameters for the RF model.

Parameters	Serious Bus Crashes	Slight Bus Crashes	Non-Bus Crashes
max_depth	4	7	6
max_features	sqrt	sqrt	sqrt
min_samples_leaf	5	5	3
min_samples_split	10	5	3
n_estimator	100	180	130

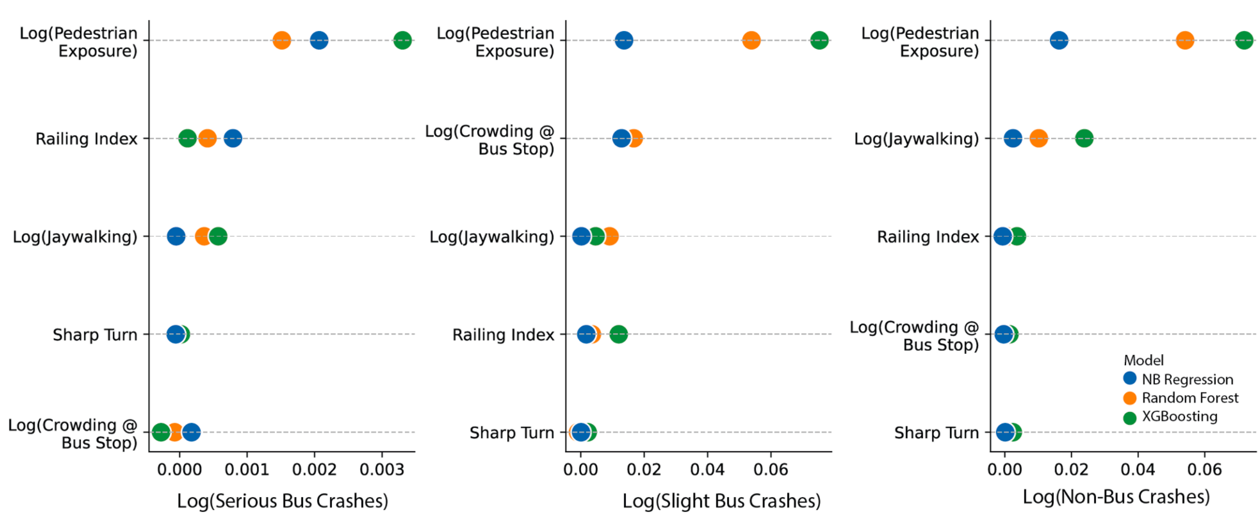

Fig. 4. Permutation Importance (based on negative MSE score predicting the log-transformed dependent variable) from three models predicting serious bus crashes, slight bus crashes, and non-bus crashes.

Table 3
Hyperparameter for the XGBoost model.

Parameters	Serious Bus Crashes	Slight Bus Crashes	Non-bus Crashes
max_depth	2	3	2
min_child_weight	4	7	7
Learning Rate	0.3	0.3	0.3
subsample	0.9	1	1
colsample_bytree	0.8	1	0.9

popular mean decrease in impurity (Gini importance) mechanism that is provided through the default scikit-learn function in python, this study adopted the permutation importance given that it is model agnostic and produces reliable results (Strobl et al., 2007). Impurity-based feature importance for trees is strongly biased, which can give high importance to features not predictive of unseen data when the model is overfitting. On the contrary, the permutation importance does not have such bias. In addition, the permutation importance was fit on the test dataset rather than the training dataset. In this way, one can further emphasise the feature importance in out-of-sample prediction. For each feature f_j, its permutation importance p_j is computed as:

$$p_j = s - \frac{1}{K}\sum_{k=}^{K} s_{k,j} \tag{7}$$

where s is the total score (negative MAE and negative MSE) of the models. The calculation was repeated for K = 5 times. Each time, column j of dataset D was shuffled to generate a corrupted version of the data $\widehat{D_{k,j}}$; then, the score $s_{k,j}$ of the model on the corrupted data $\widehat{D_{k,j}}$ is computed.

4. Results

4.1. Prediction results

Using the fine-tuned models, final prediction results as shown in Table 4. Both training results and testing results are listed in the table to show the models' out-of-sample prediction capability. One can observe that both RF and XGBoost outperform the NB model when estimating bus-related crashes, especially for the test dataset. Regarding the non-bus crash model, the RF model performs slightly better than the NB and XGBoost.

4.2. Feature importance

Fig. 4 plots the permutation importance based on MSE produced from NB, RF and XGBoost for all three target crash variables. It is observed that all three sets of models produce similar results in feature importance ranking. Moreover, the results produced with MAE is similar and shown in the appendix.

For all three models, the Pedestrian Exposure Factor is constantly the most important in estimating vehicle–pedestrian crash frequency. Nonetheless, Bus Stop Crowding ranks higher for slight bus crashes, while least important for serious bus crashes. Where bus stops are crowded, waiting bus pedestrians and passengers are more likely to overflow to roads and give rise to various slight injury bus crashes. The railing index is of much higher importance in predicting serious bus crashes. It suggests that the functions of railings in protecting pedestrians are especially important in case of serious and fatal bus crashes. Jaywalking, while often ignored in road safety analysis, is important in accounting for non-bus-related crashes. Last but not least, it is also observed that other than the Pedestrian Exposure Factor and Jaywalking, these other factors play minor roles in predicting non-bus crashes. The results highlight the value of differentiating crash analysis by transport modes.

Table 4
Model results on train and test datasets.

Y	Data Set	Negative Binomial (NB) (1) MAE	MSE	Random Forest (RF) (2) MAE	MSE	XGBoost (3) MAE	MSE
Log(Slight Bus Crashes)	train	0.565	0.441	0.465	0.320	0.389	0.238
	test	0.556	0.418	0.500	0.359	0.499	0.386
Log(Serious Bus Crashes)	train	0.133	0.055	0.130	0.051	0.103	0.031
	test	0.162	0.055	0.135	0.053	0.137	0.054
Log(Non-bus Crashes)	train	0.713	0.735	0.618	0.556	0.575	0.489
	test	0.686	0.699	0.641	0.596	0.676	0.659

4.3. Feature relationship

After evaluating the feature importance, a partial dependence plot from all three models visualises the specific relationship between bus crashes and the four risk factors found to be significant in the two bus crash models – Pedestrian Exposure, Jaywalking, Bus Stop Crowding and railing index (Fig. 5). Four main observations can be made.

First, it is evident that all crashes are positively correlated with Pedestrian Exposure Factors. However, It is realised that street segments with very few people have slightly higher chances of getting bus crashes (revealed by the RF and XGBoosting models). This result is not yet captured in previous studies on pedestrian exposure, given the lack of high-resolution data (Almasi et al., 2021; Lee & Abdel-Aty, 2005). It is likely that as the potential of overlook when buses travel through streets normally with fewer pedestrians.

Second, the increase in the railing index is correlated with a sharp drop in serious bus crashes. Some initiatives for increasing walkability have advocated for the removal of pedestrian railings. This finding suggests that the decision must be evaluated very carefully, especially if the relevant road segments are still having vehicular traffic with heavy vehicles such as buses. Other traffic calming measures, such as reducing traffic speed and bus route re-routing, should be implemented before the pedestrian railings are removed. Along corridors where buses traverse, these protection railings can save life.

Third, pedestrian crowding at bus stops shows a positive correlation with bus-related crashes, especially for slight bus crashes. This result resonates with previous literature on bus stop accident analysis (Truong & Somenahalli, 2011). In addition, detailed descriptions of bus crashes in the TRADS database also show that many slight accidents happening around bus stops were related to pedestrians falling on the ground as buses approached and/or - pulled off from bus stops.

Lastly, the pedestrian jaywalking indicator shows a similar non-linear relationship for all three sets of models. When jaywalking happens occasionally, it is likely to see an increase in bus crashes. In real life, drivers are more cautious on street segments where a lot of jaywalking happens. Similar to the permutation plot above, it is observed that the contribution of crowding at the bus stop and railing index is insignificant for non-bus crashes.

5. Discussion

In comparison to many global cities in the world, Hong Kong has many narrow and/or crowded pedestrian sidewalks, especially in downtown areas like Central and Wanchai. Although pedestrian railing hinders the vibrancy of some streets, results here suggest that it is still necessary to adopt railing as a protection against serious bus crashes. Also, results in this study demonstrate that pedestrian exposure is an important factor in predicting all types of crashes. The results are highly consistent with the existing literature and suggest that the pedestrian exposure factor needs to be included in scientific road safety models (Kim et al., 2022; Lam et al., 2014; Yao et al., 2015). As bus crashes in

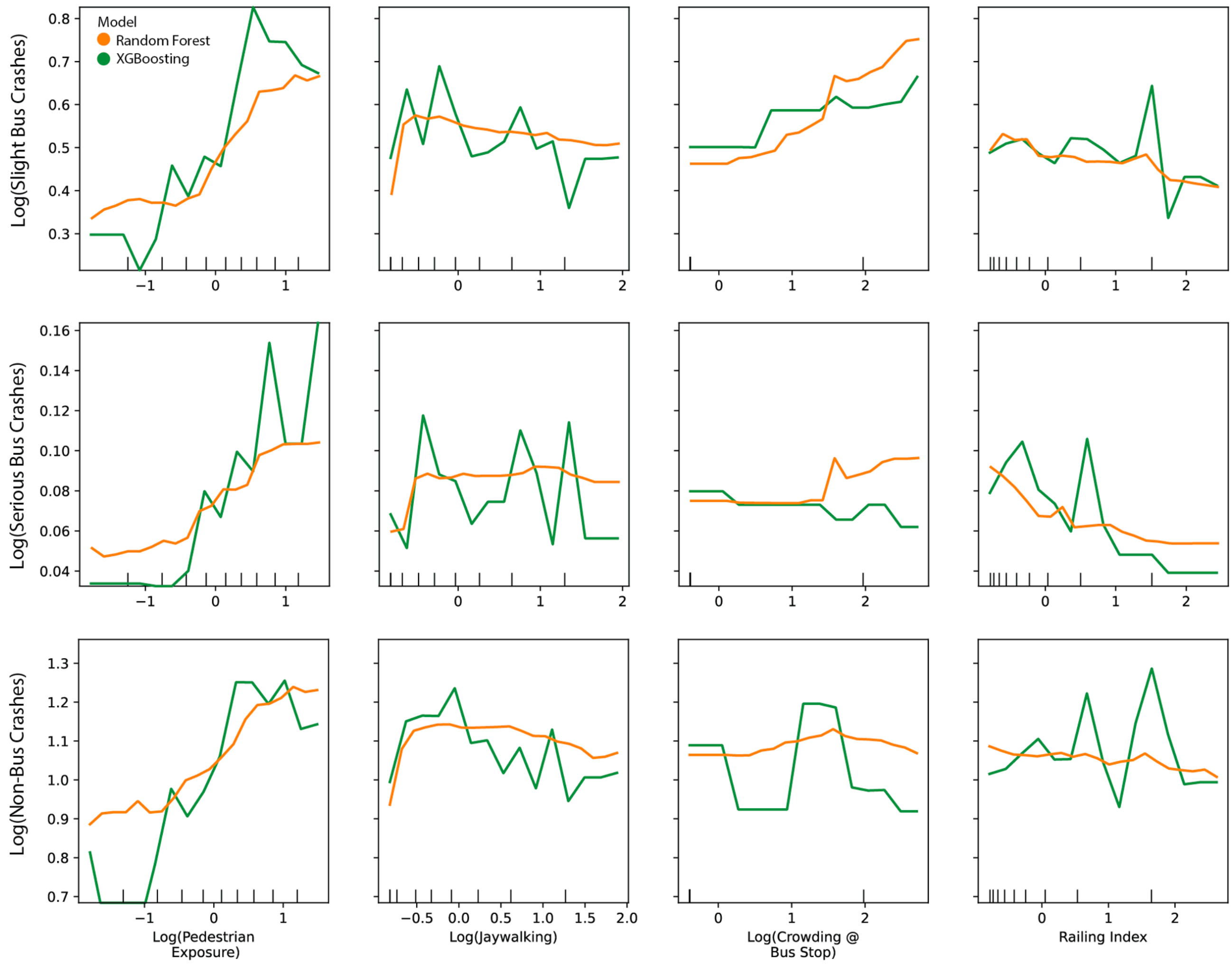

Fig. 5. Partial Dependence Plot for RF and XGBoost model results. NB model uses count as the dependent variable thus is presented separately in the ***S2. All features are standardized to improve model performance.

Hong Kong are already relatively rare events, it should be further targeted at reducing the number of severe crash events when individual mistakes, either caused by drivers or pedestrians, are largely unpredictable.

The results on five risk factors also help suggest several key points to be considered for future planning interventions. First, locations with pedestrian railing have seen fewer serious bus crashes. Although many walkability initiatives support removing pedestrian railings, this approach should be reconsidered in dense urban areas where other traffic calming interventions are not feasible to install. Second, bus stops tend to see more slight crashes. This result suggests better design considerations at bus stops are needed so that bus passengers can have a safer environment to wait for buses on the one hand, and to get on and off buses in an orderly manner on the other hand.

This study also shows the difference between bus crashes and other road crashes. Pedestrian crowding at bus stops and railing index is more critical in predicting bus crashes than other crashes. General crash models reveal underlying explanatory factors but are not enough to pinpoint safety measures for achieving road safety targets (Wong et al., 2006; Allsop et al., 2011). Particular groups of traffic crashes, such as those involving buses and heavy vehicles, need to be modelled separately to identify effective road safety measures. This study highlights the fact that transport planning can target different vehicle categories to better manage road safety. Multi-pronged action plans targeting specific crash types should be considered in a holistic road safety strategy (Loo et al., 2005; Petrov and Evtyukov, 2020).

6. Conclusion

This paper leverages a rich dataset - bus dashcam videos to capture specific risk factors for bus-pedestrian crashes: pedestrian exposure factor, pedestrian jaywalking, pedestrian crowding at the bus stops, and railings. Then, these factors are used to construct three sets of models estimating the 1) bus-related serious crashes, 2) bus-related slight crashes, and 3) non-bus related crashes. Both XGBoosting and RF models generated similar results on feature importance for three sets of models, implying the relative importance of these factors for bus-related crashes. The partial dependency plots show non-linear relationships of many risk factors with bus crashes. This research reveals that data collected by public transport companies primarily for operation reasons, such as bus dashcam videos, contains rich and valuable information for in-depth research and can help transport planning and urban management. Factors such as pedestrian exposure, pedestrian jaywalking, pedestrian crowding, and railings are operationalised factors in a dense urban environment that urban planners could manage through many pathways.

This study still has limitations. First of all, the videos were collected more recently than the crash data. The study was conducted with the assumption that the pedestrian use of streets and streets environment features are relatively stable. During the time when the videos were collected, Hong Kong was still under inbound travel restrictions. Therefore, there would not be as many tourists as before along the street sidewalk. Second, derived from the segmentation model, the railing index still hides information such as where the railing might need to be renovated or renewed. Lastly, as an index, the pedestrian jaywalking does not reveal the exact number of people jaywalked. These limitations mostly hinge upon the precision of the deep learning model, and tremendous improvements are foreseeable in the near future (Fan & Loo, 2021).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

This project is supported by the Research Grant Council General Research Fund (GRF) project on "Improving bus safety in Hong Kong: From advanced spatial analysis to artificial intelligence" (Project number: 17616221) and the "Using computer vision, machine learning, and artificial intelligence to improve bus safety in smart cities: A pilot study" funded by the Kowloon Motor Bus Company (1933) Limited (KMB), and the Research Matching Grant Scheme of the University Grant Council.

References

Abdulhafedh, A., 2016. Crash frequency analysis. J. Transp. Technol. 6 (04), 169.

Abdulhafedh, A., 2017. Road crash prediction models: Different statistical modeling approaches. J. Transp. Technol. 7 (02), 190.

Allsop, R.E., Sze, N.N., Wong, S.C., 2011. An update on the association between setting quantified road safety targets and road fatality reduction. Accident Anal. Prev. 43, 1279–1283.

Almasi, S.A., Behnood, H.R., Arvin, R., 2021. Pedestrian Crash Exposure Analysis Using Alternative Geographically Weighted Regression Models. J. Adv. Transp. 2021, 1–13. https://doi.org/10.1155/2021/6667688.

Atumo, E.A., Fang, T., Jiang, X., 2022. Spatial statistics and random forest approaches for traffic crash hot spot identification and prediction. Int. J. Inj. Contr. Saf. Promot. 29 (2), 207–216.

Avinash, C., Jiten, S., Arkatkar, S., Gaurang, J., Manoranjan, P., 2019. Evaluation of pedestrian safety margin at mid-block crosswalks in India. Saf. Sci. 119, 188–198. https://doi.org/10.1016/j.ssci.2018.12.009.

Barabino, B., Di Francesco, M., Mozzoni, S., 2017. Time reliability measures in bus transport services from the accurate use of automatic vehicle location raw data. Qual. Reliab. Eng. Int. 33 (5), 969–978.

Barabino, B., Bonera, M., Maternini, G., Olivo, A., Porcu, F., 2021. Bus crash risk evaluation: An adjusted framework and its application in a real network. Accid. Anal. Prev. 159, 106258 https://doi.org/10.1016/j.aap.2021.106258.

Beitel, D., Stipancic, J., Manaugh, K., Miranda-Moreno, L., 2018. Assessing safety of shared space using cyclist-pedestrian interactions and automated video conflict analysis. Transp. Res. Part D: Transp. Environ. 65, 710–724. https://doi.org/10.1016/j.trd.2018.10.001.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Chakraborty, A., Mukherjee, D., Mitra, S., 2019. Development of pedestrian crash prediction model for a developing country using artificial neural network. Int. J. Inj. Contr. Saf. Promot. 26 (3), 283–293.

Chandra, S., Bharti, A.K., 2013. Speed Distribution Curves for Pedestrians During Walking and Crossing. Procedia. Soc. Behav. Sci. 104, 660–667. https://doi.org/10.1016/j.sbspro.2013.11.160.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, pp. 785–794.

Cheung, C., Shalaby, A.S., Persaud, B.N., Hadayeghi, A., 2008. Models for safety analysis of road surface transit. Transp. Res. Rec. 2063 (1), 168–175.

Chimba, D., Sando, T., Kwigizile, V., 2010. Effect of bus size and operation to crash occurrences. Accid. Anal. Prev. 42 (6), 2063–2067. https://doi.org/10.1016/j.aap.2010.06.018.

Dong, N., Huang, H., Zheng, L., 2015. Support vector machine in crash prediction at the level of traffic analysis zones: Assessing the spatial proximity effects. Accid. Anal. Prev. 82, 192–198.

Evans, W.A., Courtney, A.J., 1985. An analysis of accident data for franchised public buses in Hong Kong. Accid. Anal. Prev. 17 (5), 355–366. https://doi.org/10.1016/0001-4575(85)90089-2.

Fan, Z., Loo, B.P., 2021. Street life and pedestrian activities in smart cities: Opportunities and challenges for computational urban science. Computational Urban Science 1, 1–17.

Feng, S., Li, Z., Ci, Y., Zhang, G., 2016. Risk factors affecting fatal bus accident severity: Their impact on different types of bus drivers. Accid. Anal. Prev. 86, 29–39. https://doi.org/10.1016/j.aap.2015.09.025.

Fu, T., Hu, W., Miranda-Moreno, L., Saunier, N., 2019. Investigating secondary pedestrian-vehicle interactions at non-signalized intersections using vision-based trajectory data. Transp. Res. Part C: Emerging Technol. 105, 222–240. https://doi.org/10.1016/j.trc.2019.06.001.

Girshick, R., 2015. Fast r-cnn. Proceedings of the IEEE International Conference on Computer Vision 1440–1448.

Gitelman, V., Levi, S., Carmel, R., Korchatov, A., Hakkert, S., 2019. Exploring patterns of child pedestrian behaviors at urban intersections. Accid. Anal. Prev. 122, 36–47. https://doi.org/10.1016/j.aap.2018.09.031.

Goh, K., Currie, G., Sarvi, M., Logan, D., 2014. Factors affecting the probability of bus drivers being at-fault in bus-involved accidents. Accid. Anal. Prev. 66, 20–26. https://doi.org/10.1016/j.aap.2013.12.022.

Gorrini, A., Crociani, L., Vizzari, G., Bandini, S., 2018. Observation results on pedestrian-vehicle interactions at non-signalized intersections towards simulation. Transport. Res. F: Traffic Psychol. Behav. 59, 269–285. https://doi.org/10.1016/j.trf.2018.09.016.

Gu, Y., Liu, D., Arvin, R., Khattak, A.J., Han, L.D., 2023. Predicting intersection crash frequency using connected vehicle data: A framework for geographical random forest. Accid. Anal. Prev. 179, 106880.

He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask r-cnn. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2961–2969.

Hilbe, J.M., 2011. Negative binomial regression. Cambridge University Press.

Huting, J., Reid, J., Nwoke, U., Bacarella, E., Ky, K.E., 2016. Identifying Factors That Increase Bus Accident Risk by Using Random Forests and Trip-Level Data. Transp. Res. Record: J. Transp. Res. Board 2539 (1), 149–158. https://doi.org/10.3141/2539-17.

Iranitalab, A., Khattak, A., 2017. Comparison of four statistical and machine learning methods for crash severity prediction. Accid. Anal. Prev. 108, 27–36.

Kaplan, S., Prato, C.G., 2012. Risk factors associated with bus accident severity in the United States: A generalized ordered logit model. J. Saf. Res. 43 (3), 171–180. https://doi.org/10.1016/j.jsr.2012.05.003.

Kim, H., Jang, J., Choi, Y., 2022. Spatial analysis of collision risk of child pedestrians–A case of urban elementary school districts in Busan, Korea. Travel Behav. Soc. 29, 165–175.

Lam, W.W.Y., Yao, S., Loo, B.P.Y., 2014. Pedestrian exposure measures: A time-space framework. Travel Behav. Soc. 1 (1), 22–30.

Lee, C., Abdel-Aty, M., 2005. Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. Accid. Anal. Prev. 37 (4), 775–786. https://doi.org/10.1016/j.aap.2005.03.019.

Liang, X., Meng, X., Zheng, L., 2021. Investigating conflict behaviours and characteristics in shared space for pedestrians, conventional bicycles and e-bikes. Accid. Anal. Prev. 158, 106167 https://doi.org/10.1016/j.aap.2021.106167.

Loo, B.P.Y., Hung, W.T., Lo, H.K., Wong, S.C., 2005. Road safety strategies: A comparative framework and case studies. Transport Rev. 25 (5), 613–639.

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transp. Res. A Policy Pract. 44 (5), 291–305.

Lou, Y., Zhang, C., Zheng, Y., Xie, X., Wang, W., Huang, Y., 2009. Map-matching for low-sampling-rate GPS trajectories. *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS '09*, 352. https://doi.org/10.1145/1653771.1653820.

McLeod, F., 2007. Estimating bus passenger waiting times from incomplete bus arrivals data. J. Oper. Res. Soc. 58, 1518–1525.

Mousavi, S.M., Osman, O.A., Lord, D., Dixon, K.K., Dadashova, B., 2021. Investigating the safety and operational benefits of mixed traffic environments with different automated vehicle market penetration rates in the proximity of a driveway on an urban arterial. Accid. Anal. Prev. 152, 105982.

Petrov, A., Evtyukov, S., 2020. Analysis of possibilities for achieving targets of Russian road safety strategy. Transportation Research Procedia 50, 518–527.

Planning Department, 2022, July 1. Hong Kong Planning Standards and Guidelines. https://www.pland.gov.hk/pland_en/tech_doc/hkpsg/full/index.htm#:~:text=The%20Hong%20Kong%20Planning%20Standards,various%20land%20uses%20and%20facilities.

Road Safety Council, 2017–2021. Road Safety Council Annual Report. https://www.roadsafety.gov.hk/en/information/annual_report.html.

RSC, 2021. *Road Safety Council Annual Report* [Government Report]. Road Safety Council. https://www.roadsafety.gov.hk/doc/annual2021/annual_report_2021_en.pdf.

Samerei, S.A., Aghabayk, K., Mohammadi, A., Shiwakoti, N., 2021a. Data mining approach to model bus crash severity in Australia. J. Saf. Res. 76, 73–82. https://doi.org/10.1016/j.jsr.2020.12.004.

Samerei, S.A., Aghabayk, K., Shiwakoti, N., Karimi, S., 2021b. Modelling bus-pedestrian crash severity in the state of Victoria, Australia. Int. J. Inj. Contr. Saf. Promot. 28 (2), 233–242. https://doi.org/10.1080/17457300.2021.1907597.

Sheykhfard, A., Haghighi, F., 2020. Assessment pedestrian crossing safety using vehicle-pedestrian interaction data through two different approaches: Fixed videography

(FV) vs In-Motion Videography (IMV). Accid. Anal. Prev. 144, 105661 https://doi.org/10.1016/j.aap.2020.105661.
Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinf. 8 (1), 1–21.
Su, J., Sze, N.N., 2022. Safety of walking trips accessing to public transportation: A Bayesian spatial model in Hong Kong. Travel Behav. Soc. 29, 125–135. https://doi.org/10.1016/j.tbs.2022.06.003.
Transport Department, 1983–2019. Road traffic accident statistics. https://www.td.gov.hk/en/road_safety/road_traffic_accident_statistics/index.html.
Transport Department, 2019. Road traffic accident statistics. https://www.td.gov.hk/en/road_safety/road_traffic_accident_statistics/2019/index.html.
Transport Department, 2020. Road traffic accident statistics. https://www.td.gov.hk/en/road_safety/road_traffic_accident_statistics/2020/index.html.
Truong, L., Somenahalli, S., 2011. Using GIS to Identify Pedestrian-Vehicle Crash Hot Spots and Unsafe Bus Stops. J. Public Transp. 14 (1), 99–114. https://doi.org/10.5038/2375-0901.14.1.6.
Wager, S., Athey, S., 2018. Estimation and inference of heterogeneous treatment effects using random forests. J. Am. Stat. Assoc. 113 (523), 1228–1242.
Wang, Z., Zheng, L., Liu, Y., Wang, S., 2020. Towards Real-Time Multi-Object Tracking. *The European Conference on Computer Vision (ECCV)*.
WHO, 2022, June 20. Road traffic injuries. *World Health Organization*. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.
Wojke, N., Bewley, A., Paulus, D., 2017. Simple online and realtime tracking with a deep association metric. IEEE International Conference on Image Processing (ICIP) 2017, 3645–3649.
Wong, S.C., Sze, N.N., Yip, H.F., Loo, B.P.Y., Hung, W.T., Lo, H.K., 2006. Association between setting quantified road safety targets and road fatality reduction. Accident Anal. Prev. 38 (5), 997–1005.
Yang, Y., Wang, K., Yuan, Z., Liu, D., 2022. Predicting freeway traffic crash severity using XGBoost-Bayesian network model with consideration of features interaction. J. Adv. Transp., 2022.
Yao, S., Loo, B.P.Y., 2012. Identification of Hazardous Road Locations for Pedestrians. Procedia Eng. 45, 815–823. https://doi.org/10.1016/j.proeng.2012.08.244.
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. Accid. Anal. Prev. 75, 320–332. https://doi.org/10.1016/j.aap.2014.12.005.
Zhang, Y., Qiao, Y., Fricker, J.D., 2020. Investigating pedestrian waiting time at semi-controlled crossing locations: Application of multi-state models for recurrent events analysis. Accid. Anal. Prev. 137, 105437 https://doi.org/10.1016/j.aap.2020.105437.
Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., Torralba, A., 2017. Scene parsing through ade20k dataset. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 633–641.