Modeling vehicle collision instincts over road midblock using deep learning

Patil, Shubham; Raju, Narayana; Arkatkar, Shriniwas; Easa, Said

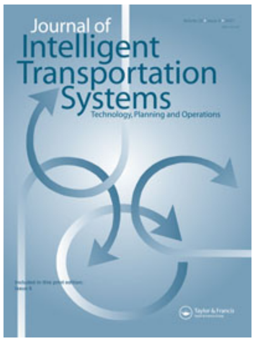
DOI
10.1080/15472450.2021.2014833

Publication date
2021

Document Version
Final published version

Published in
Journal of Intelligent Transportation Systems: technology, planning, and operations

Citation (APA)
Patil, S., Raju, N., Arkatkar, S., & Easa, S. (2021). Modeling vehicle collision instincts over road midblock using deep learning. *Journal of Intelligent Transportation Systems: technology, planning, and operations*, *27*(2), 257-271. https://doi.org/10.1080/15472450.2021.2014833



Journal of Intelligent Transportation Systems
Technology, Planning, and Operations

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gits20

Modeling vehicle collision instincts over road midblock using deep learning

Shubham Patil, Narayana Raju, Shriniwas S. Arkatkar & Said Easa

https://doi.org/10.1080/15472450.2021.2014833

Modeling vehicle collision instincts over road midblock using deep learning

Shubham Patil[a], Narayana Raju[b], Shriniwas S. Arkatkar[a], and Said Easa[c]

[a]Department of Civil Engineering, Sardar Vallabhbhai National Institute of Technology, Surat, Gujarat, India; [b]Transport & Planning Department, CiTG, Delft University of Technology (TU Delft), Delft, Netherlands; [c]Department of Civil Engineering, Ryerson University, Toronto, Canada

ABSTRACT

The present research aims to understand the safety over the midblock road sections and proposes a safety framework using the conventional Time to Collision (TTC) measure. In the present work, the safety framework underlines a supporting structure connecting the actions of the surrounding vehicles and assesses the collisions changes for a given subject vehicle. The Framework principally checks the likelihood of lateral overlap and the time gap between the subject vehicle and its surrounding vehicles. Later, for the trajectory data development, an automated trajectory data development tool is built with the help of image processing for generating the trajectory data from the study sections. In supporting the developed safety framework, the lateral movement of the vehicles is modeled precisely with the help of deep learning. Further, the conceptualized safety framework is tested with the developed trajectory data sets over the study sections. From the results, it is observed that, in mixed traffic, the collision points are over the entire geometry of the study section. In the case of homogeneous traffic, the collision instincts are clustered toward the median lanes. With the advancement of technology, trajectory data development can be a real-time exercise, and the safety framework can be implemented. By applying the study methodology, the critical spots over the road network can be flagged for better treatment and improve safety over the sections.

ARTICLE HISTORY
Received 17 March 2021
Revised 30 November 2021
Accepted 2 December 2021

KEYWORDS
Homogeneous traffic; mixed traffic; time-to-collision; traffic safety; trajectory data

Background

Monitoring traffic road safety is one of the critical challenges in traffic engineering practice. By quantifying the safety levels over the road sections, the sections can be treated well in advance, and future crashes can be limited. To understand traffic road safety and limit road crashes, researchers initially focused on historical crash data. With interpolation using geospatial tools, the critical crash locations over the road network are identified and treated to prevent future crashes. Using the Bayes model, Aguero-Valverde and Jovanis (2006) classified fatal and injury crashes for Pennsylvania, USA. Rowden et al. (2008) applied similar geospatial techniques to examine animal crashes in Australia. Recently, Schepers et al. (2017) used a similar framework and analyzed bicycle crashes in The Netherlands. Along with these, there are numerous other studies in this direction. However, the other methodologies highly depend on the historical crash data and fall under the reactive approach in limiting crashes.

In a parallel direction, researchers tend to relate driver behavior to understand the safety over the road sections. Vaiana et al. (2014) assessed the vehicles' acceleration profiles over the road sections, and safety was interpolated. By quantifying the vehicles' time gaps, Happee et al. (2017) estimated the vehicles' evasive maneuvers. Further, numerous other studies (Arvin et al., 2020; Ben-Bassat & Shinar, 2011; Raju, Kumar, et al., 2019; Sagberg et al., 2015; Taubman-Ben-Ari & Katz-Ben-Ami, 2012; Van Driel & Van Arem, 2010; P. Wang et al., 2010) analyzed safety over the study sections, with various surrogate safety metrics. Some of the surrogate metrics include Time To Collision (*TTC*) (Van Der Horst & Hogema, 1993), Deceleration Rate to Avoid Crash (*DRAC*) (Zheng et al., 2019), Time in TTC (*TIT*) (Vogel, 2003), Time Exposed in TTC (*TET*) (Behbahani et al., 2014), and Instantaneous Heading Time (*IHT*) (Raju et al., 2020), along with the mentioned measures, numerous surrogate safety measures are present in the literature. The preceding metrics demand high-quality

CONTACT Narayana Raju S.S.N.Raju@tudelft.nl Transport & Planning Department, CiTG, Delft University of Technology (TU Delft), 2600 AA Delft, Netherlands

data for the analysis. It can be noted that most of the above-mentioned surrogate safety metrics consider the leader-follower interactions and models the collision changes between the vehicles. The threat to the subject vehicle due to its lateral movement and the actions of surrounding vehicles is ignored.

Understanding driver behavior over the road sections demands quality data, where the driver responses can be quantified. In this direction, researchers used different types of data sources. Researchers used probe vehicles (Li et al., 2016; Nadimi et al., 2016; Schwarz, 2014), driving simulators (Engström et al., 2005; Meuleners & Fraser, 2015; Stavrinos et al., 2013; Yan et al., 2008), and probe vehicles embedded with video cameras and GPS (Das & Maurya, 2019; Ellison et al., 2015; Rogers et al., 1999; Strauss et al., 2015). Further, with the release of the Strategic Highway Research Program 2 (*SHRP*2) (Victor, 2016) datasets, researchers uncovered numerous driving instincts with respect to crashes, such as vehicle crash rate (Abdel-Aty et al., 2007; Blanco et al., 2016; Z. Chen et al., 2018; Hoseinzadeh et al., 2020), driving anger (Precht et al., 2017a), effect of secondary tasks while driving (Schneidereit et al., 2017), dynamic speed limits (Soriguera et al., 2013), and traffic violations (Precht et al., 2017b). On the other hand, considering the Next Generation Simulation (*NGSIM*) (US Department of Transportation and Federal Highway Administration, 2007) data as a base source, researchers tested the surrogate safety metrics to understand collision instincts, including intersection safety analysis (Cunto & Saccomanno, 2008), drivers risk efforts (Lyu et al., 2021; Przybyla et al., 2015; Wu et al., 2020), and effect of lane changes on safety (Pek et al., 2017).

Along with that, it highlighted the importance of trajectory data for quantifying driver behavior. Further, to further aid safety understanding, the Federal Highway Administration(*FHWA*) developed a Surrogate Safety Assessment Model (*SSAM*) (Gettman et al., 2008). Safety is related to the vehicles' time gaps at different driving instincts, such as following, lane change, and other gap acceptance behavior at intersections and roundabouts. Along with this, microsimulation packages are embedded as a safety tool.

From the driving behavior studies on NGSIM data, it is observed that monitoring the vehicles' movement plays a crucial role in assessing driver behavior. In achieving that, the vehicular positions must be traced over the section under consideration with a minimum update interval. The semi-automated image processing tools are reliable sources for trajectory development in the present context, particularly prevailing in developing countries like India. Given the advancements in image processing and artificial intelligence in the second half of the last decade, researchers identified the solution for automated trajectory development tools to enrich the concept of drive behavior analysis.

On the other hand, whereas the present surrogate studies can gauge the safety in the traffic streams, the existing surrogate methodologies have not found a place in the decision making. This can be attributed to the limited and unreliable historical data set and variation in results across the different surrogate safety measures. Most of the existing surrogate safety methodologies are heavily oriented toward the vehicles' longitudinal movement under midblock conditions. Whereas, along with longitudinal movement, the vehicles' lateral movement plays a considerable role in the overall driving behavior and the studies (Y. Chen et al., 2019; Deng et al., 2019; Qu et al., 2020; Raju et al., 2021) explicitly highlights the importance lateral behavior of the vehicles. However, the risk of the surrounding vehicles on the subject vehicle is mostly ignored. In this direction, toward developing a comprehensive surrogate safety measure, the vehicles' lateral movement must identify accurate collision instincts. Further, a real-time safety analysis can be a huge benefit in making timely decisions to improve driver compliance and the prevailing level of safety. But developing the trajectory data and processing the data is a challenging aspect in the present context, given the complexity and the stochasticity nature of the heterogeneous traffic.

Research methodology

To address the above-mentioned challenges in literature, the present study was carried in four stages, as shown in Figure 1. In Stage 1, based on the surrounding vehicles, a safety framework was developed using the conventional *TTC* measure. In Stage 2, an automated trajectory tool was programmed for developing the trajectory data over the study sections. In Stage 3, to support the developed safety framework the lateral movement of the vehicles are modeled with the help of deep learning models. Finally, in Stage 4, based on the developed safety framework, the collision instincts over the study sections were evaluated.

Safety framework

Recent studies on driving behavior have observed that vehicles' behavior, such as following and lateral

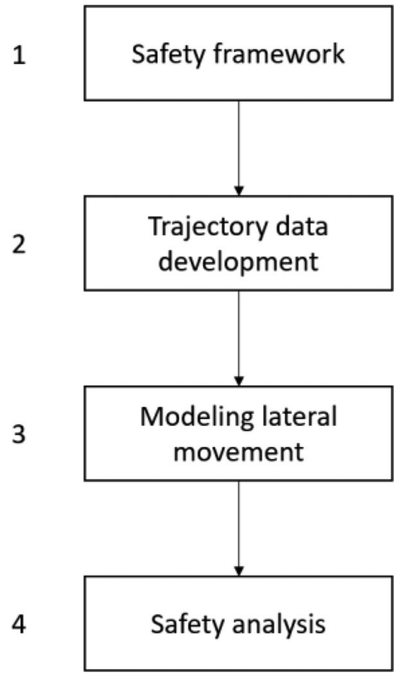


Figure 1. Research methodology adopted in the study.

movement of vehicles, is primarily influenced by its surrounding vehicles. On the other hand, in surrogate safety analysis, researchers heavily focused on leader-follower vehicle combinations in the traffic stream for assessing the collision changes. As a result, the collision threats due to the surrounding vehicles on a given subject vehicle are ignored. However, given the lateral movement of vehicles, there can be high chances for collisions from the surrounding vehicles.

Let n be the subject vehicle in the traffic stream. Given this, there can be a maximum of eight vehicles surrounding the vehicle n, as shown in Figure 2. On these lines, let s_1, s_2, …, s_8 be the surrounding vehicles for the subject vehicle n. In a general sense, with the current framework, the rear-end collision changes are given by its next leader vehicle. At the same time, it can be noted that the immediate leader can shift laterally, and some other vehicle can take its position in the next time frame, and the collision chances due to this lateral movement are not addressed in the present frameworks.

The surrounding vehicles for a given vehicle were classified into three vehicle categories: leader vehicles, adjacent vehicles, and follower vehicles. Further, with the leader and follower vehicles, the subject vehicle can have rear-ended collisions, whereas, with the adjacent vehicles, there can be sideswipe collisions.

Simultaneously, with a leader combination of vehicles, mostly the immediate leader having the lateral overlap with the subject vehicle can have the rear-end collision. Whereas with the left and right leader vehicles, the lateral movements of those vehicles play a role. For example, given the scenario, if the leader vehicle shifted laterally, the left leader vehicle tends to have lateral overlap with the subject vehicle, and the collision chances increase with the left leader. Similarly, given the lateral movement of the adjacent vehicles, the sideswipe collision chances vary. On the other hand, the follower vehicles can have similar instincts with the subject vehicle.

Further, to assess the collision instincts between the lead and trail vehicles, the TTC measure was adopted and was defined as

$$TTC(t) = \frac{X_{lead}(t) - X_{trail}(t) - l_{leader}}{V_{trail}(t) - V_{lead}(t)} \ \forall \ V_{trail}(t) > V_{lead}(t) \tag{1}$$

where:
$TTC(t)$ = time-to-collision at time t,
$X_{lead}(t)$ = longitudinal position of the leader vehicle at time t,
$X_{trail}(t)$ = longitudinal position of the trailing vehicle at time t,
l_{leader} = length of the leader vehicle in m,
$V_{trail}(t)$ = longitudinal speed of the trailing vehicle at time t, and
$V_{lead}(t)$ = longitudinal speed of the leader vehicle at time t,

Using the TTC measure, the collision chances for each surrounding vehicle to the subject vehicle was developed, as shown in Table 1. Then, the total number of collision instincts for a subject vehicle is given by

$$\begin{aligned} N = {} & P(L_{n-s2})*P\left(\frac{TTC_{limit}}{TTC_{n-s2}}\right) \\ & + P^c(L_{n-s2})* \ P(L_{n-s1})*P\left(\frac{TTC_{limit}}{TTC_{n-s1}}\right) \\ & + P^c(L_{n-s2})*P(L_{n-s3})*P\left(\frac{TTC_{limit}}{TTC_{n-s3}}\right) + \ P(L_{n-s4}) \\ & + \ P(L_{n-s5}) + P(L_{s7-n})*P\left(\frac{TTC_{limit}}{TTC_{s7-n}}\right) \\ & + P^c(L_{n-s7})* \ P(L_{n-s6})*P\left(\frac{TTC_{limit}}{TTC_{s6-n}}\right) \\ & + P^c(L_{n-s7})*P(L_{n-s8})*P\left(\frac{TTC_{limit}}{TTC_{s8-n}}\right) \end{aligned} \tag{2}$$

where:
N = number of collision instincts for the subject vehicle,

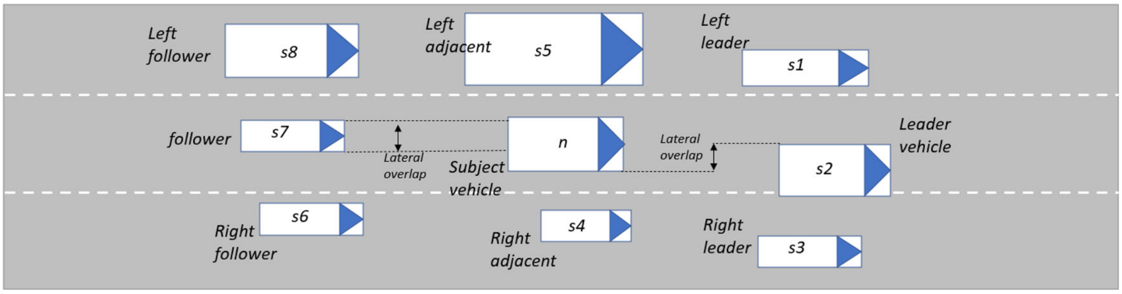


Figure 2. Surrounding vehicles classification over the subject vehicle.

Table 1. Collision probability of the surrounding vehicles.

Surrounding vehicle	Collision probability	Remark
Leader (s_2)	$P(L_{n-s2})\ P\left(\frac{TTC_{limit}}{TTC_{n-s2}}\right)$	Lateral overlap is present and TTC are in the critical limits
Left leader (s_1)	$P^c(L_{n-s2})\ P(L_{n-s1})\ P\left(\frac{TTC_{limit}}{TTC_{n-s1}}\right)$	Leader shifts laterally and other leaders have lateral overlaps
Right leader (s_3)	$P^c(L_{n-s2})\ P(L_{n-s3})\ P\left(\frac{TTC_{limit}}{TTC_{n-s3}}\right)$	
Left adjacent (s_5)	$P(L_{n-s5})$	Any lateral overlap can lead to sideswipe collision
Right adjacent (s_4)	$P(L_{n-s4})$	
Follower (s_7)	$P(L_{s7-n})\ P\left(\frac{TTC_{limit}}{TTC_{s7-n}}\right)$	Lateral overlap is present and TTC are in the critical limits
Right follower (s_6)	$P^c(L_{n-s7})\ P(L_{n-s6})\ P\left(\frac{TTC_{limit}}{TTC_{s6-n}}\right)$	Follower shifts laterally and other followers have lateral overlaps
Left follower (s_8)	$P^c(L_{n-s7})\ P(L_{n-s8})\ P\left(\frac{TTC_{limit}}{TTC_{s8-n}}\right)$	

$P(L_{n-si})$ = probability of lateral overlap between the n th and s_i vehicles, and

$P\left(\frac{TTC_{limit}}{TTC_{n-s_i}}\right)$ = probability that TTC between the n th and s_i vehicles is less than TTC_{limit}.

Data collection

Vehicular trajectory data can be a potent source in understanding the microscopic instincts between the vehicles. Therefore, in this study, three midblock study sections were selected for the analysis. The sections were selected in such a way that they capture varied traffic flow conditions. Section 1 is midblock in the western expressway India. In this section, traffic is mixed with five different vehicle categories, including Motorized Two Wheelers (MTW), Motorized Three Wheelers (MThW), Buses, Cars, and Trucks. Sections 2 and 3 are six-lane divided and four-lane divided highways with shoulders, respectively. Section 1 is a midblock road section from India, the driving behavior in section 1 is of non-lane based mixed traffic. On the other hand, the study sections 2 and 3 are motorways from Canada. The major proportion of the vehicles are cars in nature. Further, the vehicle tends to have good lane discipline. As a result, the driving behavior of the vehicles from the study sections is different, given the traffic conditions. Given this, those three study sections gave enough flexibility and variation in testing the developed safety framework. Unlike Section 1, the traffic in Sections 2 and 3 is homogenous, with a dominant car proportion. In developing the trajectory data from the study sections, video graphic data was used to track the vehicles. The snapshots of the study sections are presented in Figure 3.

In the present context, there are numerous semi-automated tools available for vehicular trajectory development. Computer mouse clicks can manually track a given vehicle over the road sections with their help. At the same time, employing those tools for trajectory data development will be laborious in developing large datasets. Considering this in the present study, it is planned to develop an image processing framework for automated trajectory data development with deep-learning architecture.

In the case of an automated trajectory data development, vehicle identification and tracking are the essential tasks, and vehicle identification is similar to object detection in an image frame. To achieve this, the Faster RCNN inception V2 model(Halawa et al., 2019) pre-trained on the Microsoft Coco dataset (Lin et al., 2015) is adopted to detect the vehicles in a given image frame.

It can be noted that five different vehicles are observed from the study sections, given these, five different objects (vehicle classes) have to be mapped. To do this, transfer learning methodology (Jain et al., 2019) is adopted. The pretrained faster RCNN is trained to classify different everyday objects from the Microsoft COCO dataset. But to build a classifier to

(a) Section 1 (b) Section 2 (c) Section 3

Figure 3. Snapshots of the study sections. (a) Section 1, (b) Section 2, and (c) Section 3.

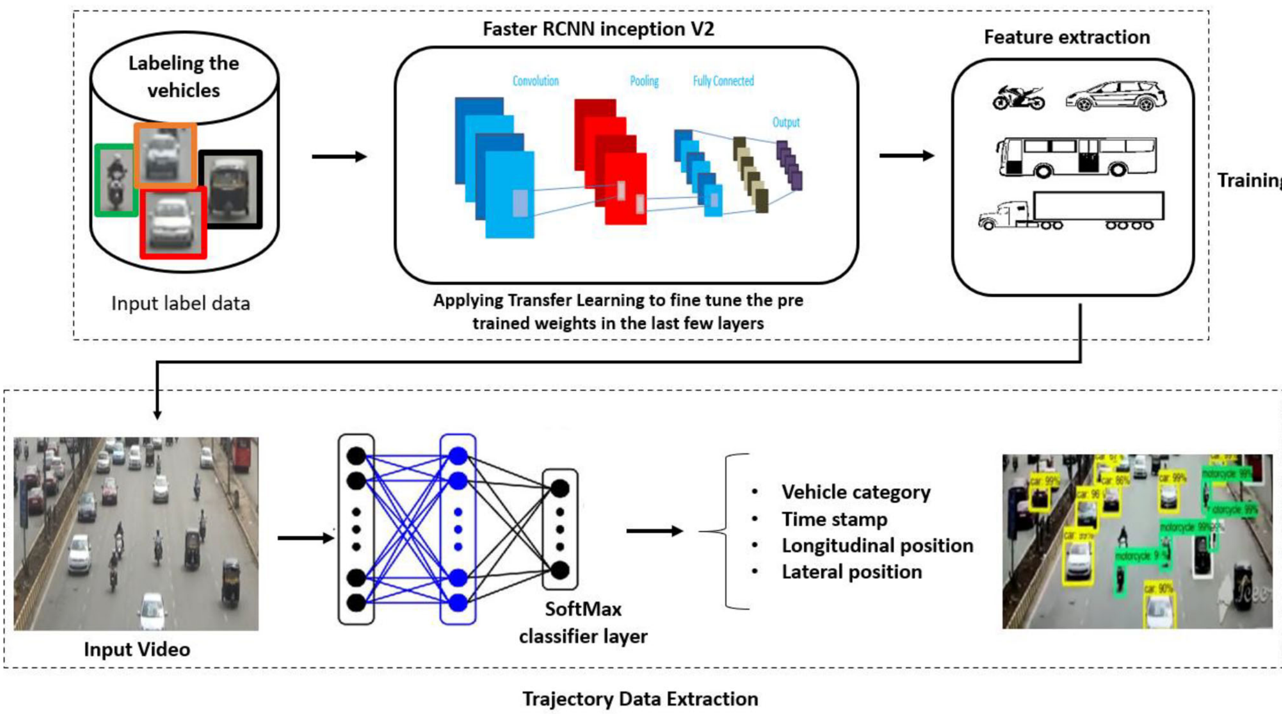


Figure 4. Automated trajectory data development tool.

perform a much more specific and narrower task, like identifying different types of vehicles over the study sections, for the vehicle images from the study section. On the trained RCNN with COCO dataset, transfer learning is applied to fine-tune the RCNN for identifying the vehicles with their category. The tensor flow module provides an API to train the customized objects (mainly vehicle types) with the pre-trained model RCNN model. The vehicles are labeled in the images extracted from the video files to generate the training data for each of the study sections. Later, those labeled images are loaded to the LABELING utility (Python Community, 2020) and converted to the XML files (Lalmas, 2011). Further, data in the XML files were used for training the Faster RCNN inception V2 model to record the vehicle features, as shown in Figure 4. Along with faster-RCNN, YOLO (Ren et al., 2020) is available for object detection. Both the faster-RCNN and YOLO are popular object detection approaches with an anchor-based framework. Faster-RCNN offers a local search in a given region of an image by applying the convolution principles. On the other hand, the YOLO follows the principles of looking once and finalizing the detection and classification outcomes. YOLO has difficulty detecting small and close objects due to only two anchor boxes in a grid predicting only one class of objects. It doesn't generalize well when objects in the image show rare aspects of ratio. Faster RCNN, on the other hand, do detect small objects well. However, it lags in

Table 2. Details of the trajectory data from the study sections.

Study section	Trap length (m)	No. of lanes	Road width (m)	Traffic composition (%) [a]	No. of vehicles tracked	Duration of trajectory data (minutes)
Section 1	120	5-lane	17.5	20, 29, 2, 45, 1, 3	1715	15
Section 2	150	3-lane	11.2	0, 0, 7, 84, 9, 0	1068	10
Section 3	150	2-lane	9.5	0, 0, 8, 82, 10, 0	211	10

[a]Traffic composition: MThW, MTW, buses, cars, trucks, LCV.

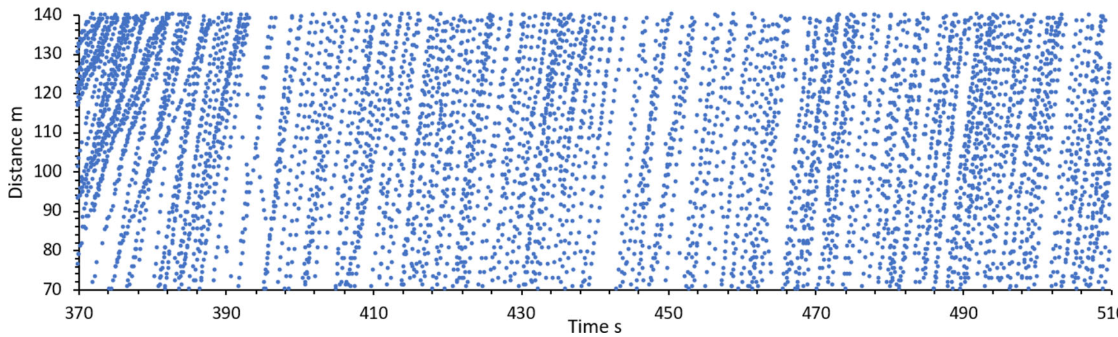


Figure 5. Sample time space plots from the study section 3.

comparison to YOLO for real-time detection with its two-step architecture. Considering this, in our present study, we adopted the Faster-RCNN rather than YOLO.

After recording the vehicles' features, the video files were given as inputs to the trained models. Initially, the video was divided into numerous image frames, and each subsequent frame is given as input to the model. Based on the trained Faster RCNN inception V2 model, the vehicular features were searched over the whole image frame. With the matching criteria, the objects were identified, and the vehicle category was assigned to each of the detected objects in a frame. Along with the vehicle category, other features such as time, centroidal longitudinal and lateral positions were recorded for the development of the trajectory data through Euclidean Object Tracking (Wang et al., 2005). The entire process is shown in Figure 4. The details of the trajectory data and the sample time-space plots are presented in Table 2 and Figure 5, respectively. The video footage recorded in the cameras is of perspective in nature. As a result, the object size varies with the distance in the image frames. On the other hand, the trajectory data should be orthogonal for any analysis. In achieving this, the recorded video files are initially processed, and the trajectory data is developed. In the next stage, trapezoidal transform is applied to project the trajectory data coordinates to the 3 D orthographic place. After the trapezoidal transform still, the axis of the trajectory data from the study sections is inclined. The concept of Cartesian coordinate rotation is applied to correct this and bring it into a planar form.

Further, to validate the automated tool's trajectory data, a semi-automated trajectory tool is selected to understand the authenticity of the data outcomes. This tool is adopted to obtain a better sample size of trajectory data by variation in vehicle class, traffic composition, and a combination of vehicular movements in longitudinal and lateral directions. Given the mixed traffic movement and vehicle classes, the traffic data on section 1 (western expressway) is selected for this purpose. Using a semi-automated trajectory tool, the trajectory data over the section is developed, in which the vehicle position was tracked using the pointer. In balancing the trajectory data accuracy with the manual efforts, the vehicles over the section are tracked with an update interval of 0.5 s. The detailed semi-automated methodology can be found in Raju et al. (2020). On these lines, the traffic composition is initially compared among the semi-automated (observed) with automated (outcome) datasets for every 5-minute interval each. Later, a *t*-test is performed to check the possible statistical difference, as shown in Table 3. From the results, it is observed that, in all instances, the null hypothesis is accepted at a 5% level of significance, demonstrating the good performance of the automated trajectory tool.

In addition to this, the instant longitudinal speed, longitudinal acceleration, lateral speed, and lateral acceleration distributions are also compared both from the semi-automated (empirical) and the automated (outcome) trajectory datasets. The probability function comparisons are shown in Figure 6, and the emerging values of the *t*-statistic are reported in Table 4. From the analysis, it is statically evident (5% level of significance) that the automated trajectory development tool can replicate the trajectory data very well when compared with the trajectory data developed using a semi-automated tool, as discussed earlier.

Table 3. Validation of vehicle composition data using *t*-test.

Time (minutes)		Traffic volume	Composition MThW	MTW	Bus	Car	Truck	*t*-statistic	*p*-value	Null hypothesis
0 – 5	Observed	551	19.1	23.8	2.4	40.6	14.1	0.185	0.99	Accept
	Outcome	547	17.5	25.7	1.9	43.2	11.7			
6 – 10	Observed	545	22.1	15	1.5	37	24.4	0.412	0.97	Accept
	Outcome	549	25.1	12.5	1.9	39	21.5			
11 – 15	Observed	593	12.7	23.2	0	51.8	12.3	0.193	0.99	Accept
	Outcome	598	17.2	17.6	0.6	44.2	20.4			

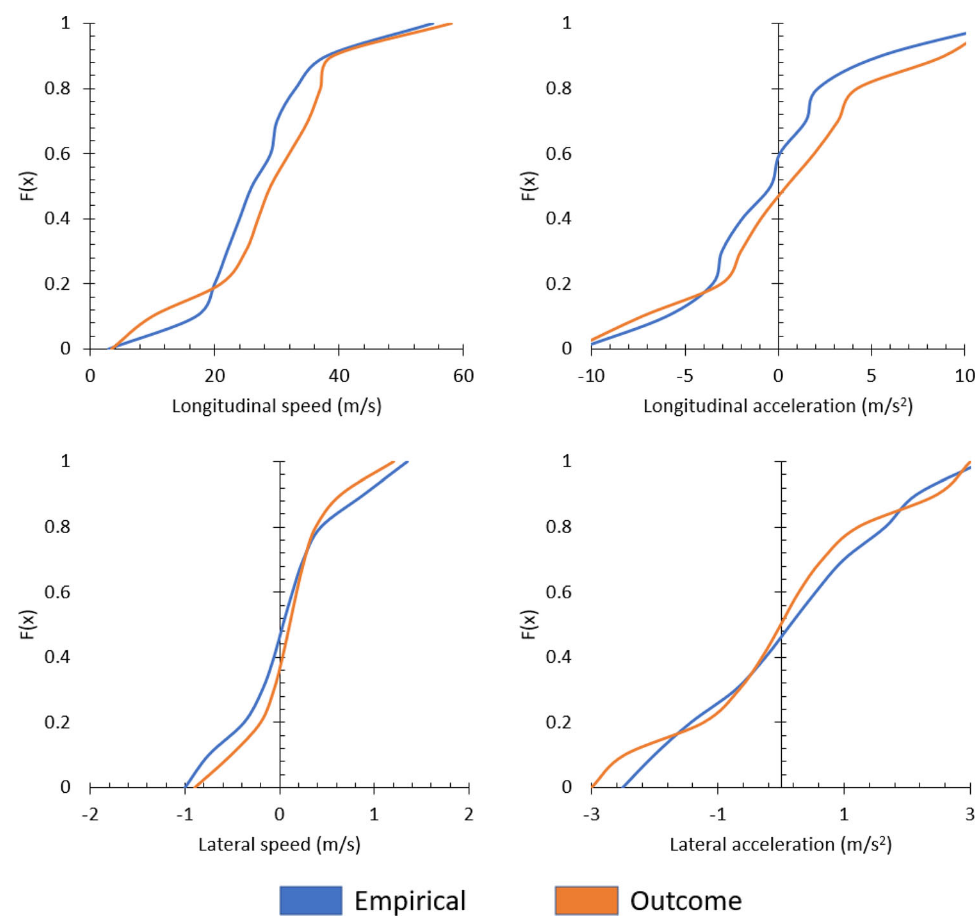


Figure 6. Comparison among the distributions of the parameters.

Table 4. Comparison of parameters using *t*-test.

Parameter	*t*-statistic	*p*-value	Degree of freedom	Null hypothesis
Longitudinal speed	−0.299	0.38	20	Accept
Longitudinal acceleration	−0.335	0.37	20	Accept
Lateral speed	−0.126	0.45	20	Accept
Lateral acceleration	0.217	0.41	20	Accept

Modeling lateral movement: deep learning

To achieve better results from the safety framework, the vehicles' lateral movement is precisely modeled. In overcoming this challenge, the literature on the lateral and lane changing behavior of vehicles is reviewed. It is identified that the vehicle's lateral movement depends a lot on the surrounding vehicles with numerous parameters from the traffic stream. Considering all these aspects in the present work, it is planned to model the vehicles' lateral movement with deep learning from the branch of artificial intelligence.

In deep learning, there will be three different layers: input layers, hidden layers, and output layers, as shown in Figure 7. Expressly, the input layers are

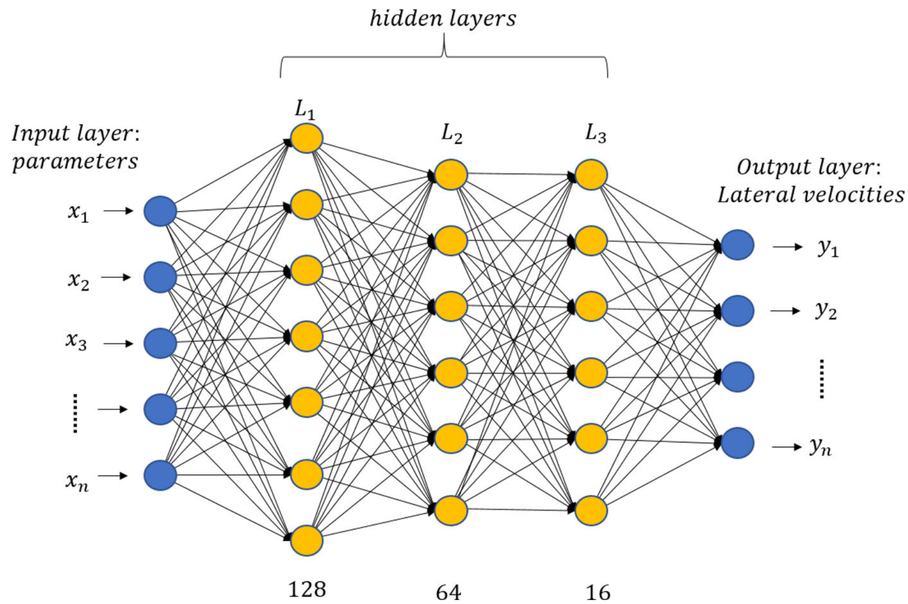


Figure 7. Deep learning architecture for modeling lateral movements.

provided with the input vectors as x_1, x_2, …, x_n. to map the outcomes in the output layer. Given this, the input data were filtered through a series of hidden layers. The hidden layers were sandwiched between the input and output layers. Typically, deep learning is developed based on the neuron's architecture in the human brain cells. This architecture is analogous to how electrical signals travel across the cells. Each subsequent layer of nodes is activated when it receives stimuli from its neighboring neurons. The accuracy of deep learning models predictions could be increased with the right amount of training data, as the brain also learns to identify things similarly.

At the same time, driving behavior is a function of numerous independent variables. With limited datasets, performing a correlation analysis will result in a smaller number of variables and can be highly local to the study. To limit this, in the present study and comprehensively model the lateral movement of the vehicles, authors depended on the literature for identifying the independent variables. Based on this, the authors identified the 22-variables and modeled the lateral movement of the vehicles. Further, based on the literature (Asaithambi & Joseph, 2018; Das & Maurya, 2019; Raju, Arkatkar, et al., 2019), the variables that can influence the vehicles' lateral movement are identified. On these lines, a total of 22-variables are identified, as shown in Table 5. By adopting the methodology of Raju, Arkatkar, et al. (2019), the surrounding vehicles that can influence the vehicles' lateral movement are labeled based on the position of the subject vehicle. For this purpose, authors marked a surrounding zone formed by the addition of 60 m distance in front (look-ahead) and 40 m distance behind (look-back) from the center of the subject vehicle, with a total longitudinal length of 100 m. A lateral distance of 5.5 m from the center position of a subject vehicle to the center position of the surrounding vehicle, including the total width of the subject vehicle, is considered over the entire road space. Based on this, the surrounding vehicle are label into eight categories based on their position from the subject vehicle. Further, based on the identified vehicles, all the other influencing variables in Table 5 are evaluated for a given vehicle at every instant of time. Later, all the 22-variables are computed for each vehicle at every recorded timestamp.

Further, to replicate the lateral movement of the vehicles, lateral speeds of the vehicles are computed. The lateral speed of the vehicles is considered as the dependent variable on the 22-influential variables. In training the deep learning models and testing their validity, the study sections' trajectory data is divided into two halves. With first used for model training and the second part for model testing and analysis. The deep learning models are developed with python programming (Rossum & Swallow, 2011) with the help of the Google TensorFlow (Google, 2020) programming framework. Later, the input variables and the output lateral speeds are used for training deep learning models with numerous combinations of hidden layers, neuron activation functions, and epochs. By applying a trial-and-error approach to increase the present case's accuracy, three hidden layers with 128,

Table 5. Influential variables on lateral movement of vehicles.

ID	Lateral Movement Variable	Description
lat_1	Leader presence	Presence of leader vehicle is taken, 0 is assigned when its absent and 1 is taken if this present.
lat_2	Leader vehicle category	Vehicle class of the leader vehicle.
lat_3	Subject vehicle category	Vehicle class of the subject vehicle.
lat_4	Relative speed with leader (m/s)	Relative speed (subject vehicle minus leader vehicle)
lat_5	Subject vehicle longitudinal speed (m/s)	Longitudinal speed of the subject vehicle
lat_6	Present lane	Present lane Id in the order as median side lane
lat_7	Left front vehicle	Presence of left front vehicle, 0 is assigned when its absent and 1 is taken if this present.
lat_8	Right front vehicle	Presence of right front vehicle, 0 is assigned when its absent and 1 is taken if this present.
lat_9	Left lateral clearance	Available lateral clearance in left side
lat_10	Right lateral clearance	Available lateral clearance in right side
lat_11	Left back vehicle speed (m/s)	Left back vehicle longitudinal speed
lat_12	Right back vehicle speed (m/s)	Right back vehicle longitudinal speed
lat_13	Left back vehicle acceleration (m/s^2)	Left back vehicle longitudinal acceleration
lat_14	Right back vehicle acceleration (m/s^2)	Right back vehicle longitudinal acceleration
lat_15	No. of surrounding vehicles	Number of vehicles in the surrounding vicinity of the subject vehicle.
lat_16	Lateral tilt with leader vehicle (m)	Lateral incline of the subject vehicle toward the leader vehicle in terms of lateral overlap (left side is taken negative value and right side as positive).
lat_17	Distance from left back vehicle (m)	Longitudinal distance from left back vehicle
lat_18	Distance from right back vehicle (m)	Longitudinal distance from right back vehicle
lat_19	Area occupancy the vehicles ahead of subject vehicle (m^2)	Area occupied by the vehicles in the frontal surrounding vicinity
lat_20	Subject vehicle longitudinal acceleration (m/s^2)	Instant longitudinal acceleration of the subject vehicle
lat_21	Left longitudinal gap(m)	Available left longitudinal gap for the subject vehicle
lat_22	Right longitudinal gap (m)	Available right longitudinal gap for the subject vehicle

Table 6. Details of the trained deep learning model [a].

Layer (Type)	No. of Nodes	Activation Function
L1 (dense)	128	ReLU
L2 (dense)	64	ReLU
L3 (dense)	16	SoftMax

[a] Model: Sequential, Epochs: 250, Input parameters: 22, Optimizer: RMSprop, Loss function: Mean square error, and Metrics: Mean absolute error.

64, and 16 nodes were adopted. In limiting the overfitting, 250 epochs were used for training. The architecture was based on sequential modeling (Goodfellow et al., 2016). Also, ReLU (Pattanayak, 2018) activation function is used for the layers other than the final SoftMax (Nelli & Nelli, 2018) layer. The details of the trained deep learning models are presented in Table 6.

In checking the trained deep learning models' validity, the remaining trajectory data is tested, and the lateral speeds are predicted. To verify the deep learning model performance, error metrics such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are computed. The prediction error, the difference between the observed and predicted lateral speeds are computed, and their distributions are presented in Figure 8. From the error distribution plots, it can be noted that most of the time in Sections 2 and 3, the prediction error is in the range of −1 m/s to 1 m/s. Whereas in the case of Section 1, due to mixed traffic and weak lane discipline, the vehicles' lateral movement is disordered compared to the other two study sections. Based on this, with the help of a deep learning framework, the vehicles' lateral movement in terms of lateral speeds is modeled with surrounding conditions.

In checking the validity of deep learning models, the remaining half of the trajectory data is used, and the lateral speeds of the vehicles are predicted. Given this, Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) are computed over the study sections considered in the present study, as reported in Table 7. From the results, it is identified that, in section 1, the error is observed to be marginally higher over the other two sections. This may be attributed to the variation in heterogeneous traffic conditions prevailing in India. Simultaneously, the RSME is obtained as less than 0.8 m/s, whereas MAPE is found to be less than 10 percent. This error is considered to be within acceptable limits (Gilliland, 2010; Ma & Qu, 2020), given the variation heterogeneous traffic conditions.

$$\mathrm{RMSE} = \sqrt{\frac{1}{\text{sample size}} \sum_{t=1}^{\text{sample size}} (\text{Estimated}_t - \text{Observed}_t)^2} \tag{3}$$

$$\mathrm{MAPE} = \frac{1}{\text{sample size}} \sum_{t=1}^{\text{sample size}} \frac{|\text{Estimated}_t - \text{Observed}_t|}{\text{Observed}_t} * 100 \tag{4}$$

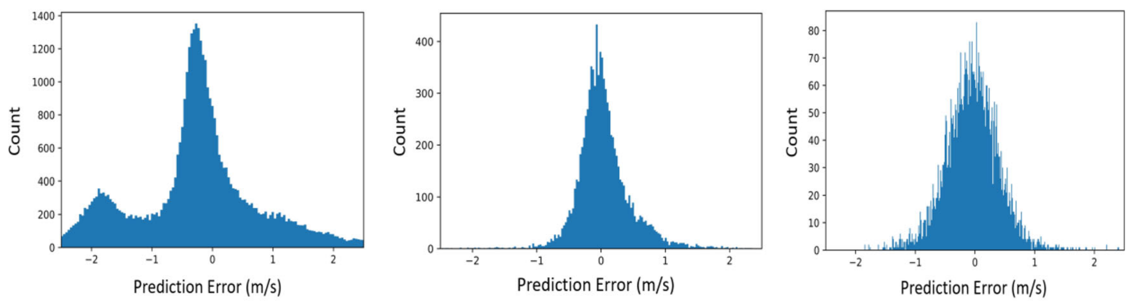


Figure 8. Comparison of prediction error of lateral speeds over the study sections. (a) Section 1, (b) Section 2, and (c) Section 3.

Table 7. Error comparison over the study sections.

Measure	Section 1	Section 2	Section 3
RMSE (m/s)	0.73	0.54	0.59
MAPE (%)	9.72	7.34	6.21

Results

In line with the study's objectives, the conceptualized safety framework, which includes the surrounding vehicle instincts, is tested. Initially, with the help of available trajectory data sets, the vehicles' lateral movement is predicted by the trained deep learning models in terms of lateral speeds for every time instant. It can be noted that the probability of lateral overlap between the vehicles will be the key in assessing the safety from the framework. With the help of predicted lateral speeds, the lateral positions in next time instants are predicted for all vehicles.

Let y_t be the lateral position of a given vehicle at t, and lat_{vel} be the predicted lateral speeds and T be the update time interval, the lateral position Y is given as follows.

$$Y(t+T) = y_t(t) + lat_{vel}{*}T \tag{4}$$

Further, the lateral overlap L, between the surrounding vehicle (s_i) and subject vehicle (n) of widths W at a given instant are computed as follows.

$$L_{n-s_i} = \begin{cases} |Y_n - Y_{s_i}| \ \forall \ |Y_n - Y_{s_i}| \leq \left(\dfrac{W_n + W_{s_i}}{2}\right) \\ \nexists \ L_{n-s_i} \ \forall \ |Y_n - Y_{s_i}| > \left(\dfrac{W_n + W_{s_i}}{2}\right) \end{cases} \tag{5}$$

The probability of lateral overlap is given as,

$$P(L_{n-s_i}) = \begin{cases} 1 \ \forall \ \exists L_{n-s_i} \\ 0 \ \forall \ \nexists L_{n-s_i} \end{cases} \tag{6}$$

Along with the lateral overlap between the vehicles, the TTC between the vehicles plays a vital role in understanding the collision instincts between the vehicles. As explained, the TTC is the time gap between the vehicles, the lesser time gap signifies more collision chances than a higher time gap. The TTC limit can be influenced by the change in traffic volume, composition, flow characteristics, vehicular properties, driver attentiveness, etc. However, in the present context, there is no such framework in identifying the limit. At the same time, this is one of the major limitations of the present existing surrogate safety measures in modeling the safety. In line with the literature (Li et al., 2016; Meng & Qu, 2012), in the present study a threshold value $TTC_{limit} = 2.5\ s$ was adopted to characterize collision and non-collision instincts. On these lines, the probability of collision instincts due to TTC is given by

$$P\left(\frac{TTC_{limit}}{TTC_{n-s_i}}\right) = \begin{cases} 1 \ \forall \ TTC_{n-s_i} < TTC_{limit} \\ 0 \ \forall \ TTC_{n-s_i} \geq TTC_{limit} \end{cases} \tag{7}$$

Based on the projected conditions, the probabilities of lateral overlaps and TTCs are estimated. Later, the number of collision instincts are evaluated for every vehicle over time, concerning its surrounding vehicles. The collision points were marked over the road space to understand the collision instincts better, as shown in Figure 9 for the three study sections.

From the safety analysis over the study sections, it is observed that the collision instincts are scattered over the road space in Section 1. Given the mixed traffic and weak lane discipline, the vehicles' movement is all over the road space, resulting in collision instincts. On the other hand, in Sections 2 and 3, the collision instincts are clustered toward the median side lanes. It is inferred that, in Sections 2 and 3, a proper lane wise movement of vehicles is observed. As a result, even collision instincts are more clustered over the lanes. The results revealed that the median side lane in Sections 2 and 3 is more prone to collisions than lanes. From the inspection, it is identified that, in the case of Section 1, most of the time,

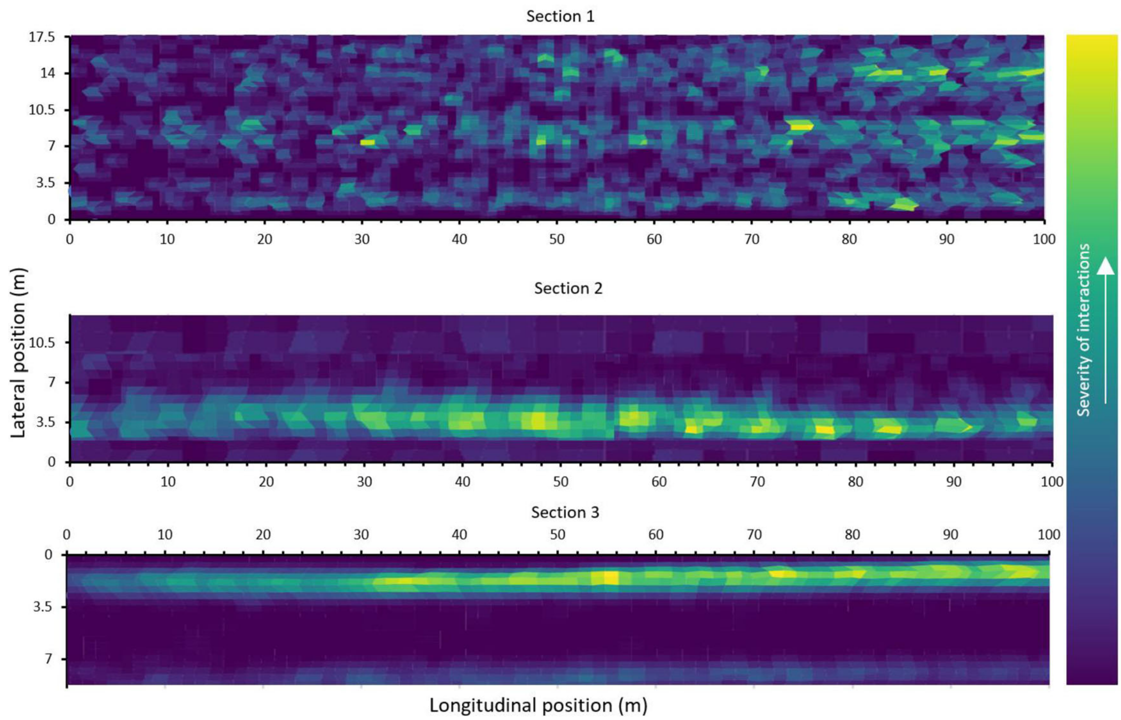


Figure 9. Collision instincts of vehicles over the geometry of the study sections.

Table 8. About the collision points.

		No. of collisions: vehicle category wise					
Study section	Total No. of collision instincts	MTW	MThW	Car	Bus	Truck	LCV
Section 1	42	19	11	7	3	2	–
Section 2	31	–	–	25	–	6	–
Section 3	23	–	–	13	–	10	–

smaller vehicles are more prone to collision instincts, which can be attributed to their size and better maneuverability. On the other hand, in Sections 2 and 3, given the significant proportion of cars, cars are more prone to collisions.

To understand the results better, the collision instincts are segregated based on the vehicle category over the study sections, as reported in Table 8. It can be noted that a total of 42 collision instincts are observed in the case of section 1; MTW and MThW are found to have a major proportion of collision instincts with 19 and 11, respectively. Given the non-lane-based movement in section 1, the collision instincts are scattered over the road space. On the other hand, in sections 2 and 3, around 31 and 23 collision instincts are observed (which are substantially smaller than the case of Section 1). Further, given the major proportion of cars, cars tend to have a major share of the collision instincts over both the study sections. Unlike in section 1, most of the collision points are clustered toward the median side lane (lane-based homogeneous traffic conditions) in both the study sections, which may be attributed to vehicles' faster movement (particularly cars).

Based on the safety framework, collision chances can be modeled and tested over different scenarios. Even though the safety analysis is carried with limited data sets in the present work, the safety framework can be extended to various study sections with different geometrics, such as curves, intersections, work zones, Etc. over different traffic volume conditions. The results would undoubtedly help in revealing the numerous implications of safety. The study sections can be treated for safe movement, well before any catastrophic mishaps. Considering the futuristic autonomous vehicles, these strategies will outline the researchers and practitioners in examining the safety of the study sections.

Conclusions

Surrogate safety measures are the key to assess the prevailing safety levels in traffic streams. At the same time, the surrogate safety metrics tend to have their limitations examining traffic safety. In this direction, through the present study, it is well brought out that the vehicles' lateral movement's character is not well incorporated in the existing surrogate metrics. At this backdrop, with the help of Artificial Intelligence (AI) based deep learning models, the present study also demonstrated the idea of involving the lateral movement of the vehicles in the safety analysis. In this context, the surrogate safety methodologies are highly dependent on the trajectory data for obtaining better insights. To this end, developing high-quality vehicle trajectory data demands a huge magnitude of effort given the variation in the involvement of multiple vehicle classes, their static and dynamic characteristics, and the non-lane based vehicular movements. The study presents a detailed framework for developing an automated trajectory development tool for the vehicles' real-time tracking in this direction. Based on this study, the following inferences are made:

1. Vehicular trajectory data have proven to be one of the potent sources in analyzing driving behavior. However, developing trajectory data can be a challenging task, given the complexity involved. Over time, numerous semi-automated tools have been developed, but they were laborious in developing trajectory datasets. Therefore, in the present study, an automated trajectory data tool was developed using advanced image processing logic and deep learning architecture. Based on this tool, the surveyed traffic videos were uploaded, and the trajectory data were developed with less human effort. Further, given the cost-effectiveness and the available programming libraries, the depicted trajectory data development framework will help researchers and practitioners working in the traffic engineering field for any driving behavior studies.
2. From the literature on midblock road sections, it is observed that most studies have focused mainly on the leader-follower vehicle combinations in understanding collision chances. At the same time, collision threats due to the surrounding vehicles have been ignored. In addressing this, the present study conceptualized a framework where the surrounding vehicles were included in modeling the collision chances and the conventional leader-follower interactions. The conventional TTC measure was used in the present study to sense the vehicular interactions. The framework can be extended with some other new advanced metrics in examining traffic safety.
3. In modeling the lateral movement of vehicles, lane-change models are adopted in the literature. Even the well-established simulation packages tend to adopt linear models for the lateral movement. Recent literature shows that vehicles' lateral movement, which is part of driving behavior, depends on various variables. In this direction, the present study uncovered the idea of deep learning in modeling the lateral movements. This idea can be explored and tested using different traffic flow modeling concepts.
4. From the safety analysis over the study sections, it is observed that in the case of the mixed traffic stream, the collision instincts were scattered over the road space. This is attributed to the weak lane discipline of vehicles and the high lateral maneuverability of smaller vehicles. On the other hand, in homogeneous traffic, the vehicles tended to have good lane discipline. As a result, the collision instincts were over the lanes, mostly on the median side lane.
5. Further, to understand the safety in a comprehensive manner, the safety analysis can be carried over different traffic flow conditions at different traffic compositions. Based on the results, the study sections can be treated, and the safety can be improved. In these lines, even different traffic management scenarios (i.e., bottleneck scenarios, ramp metering, lane segregation, etc.) can be tested for safety and the efficiency of the network elements.

Limitation and future scope

Along with the research findings, the present study has certain limitations, which should be considered in the work's future scope.

- In the present study, trajectory data is developed over three study sections (with varying traffic conditions; heterogeneous and homogeneous) for 10 minutes, each using a newly devised automated trajectory development tool. However, the present study framework must be tested over the study sections with even more different flow conditions with variation in vehicles' proportion. This can undoubtedly help in comprehensively assessing the safety levels over the study sections.
- The safety framework is only tested on the trajectory data from the study sections. However, the framework can be validated with the help of real collisions data from the field conditions. In this

direction, the video clippings of the vehicle collisions from traffic streams can be used to calibrate and validate the developed safety framework better. Though it may be difficult, with better availability of ground truth data, it should be aimed particularly under heterogeneous traffic conditions.

- Further, the collision instincts reported in the study is just an indirect representation of collisions. It is reported that, in some instants, the surrogate safety methodologies may overpredict the collision occurrence; this can act as a limitation in the present safety framework as well. Given this, reliable historical data can certainly provide more insights into the critical values of conflicts to decide the prevailing safety level. This will also supplement the effectiveness of traffic management or safety management measures toward improvement in safety and efficiency.
- The developed safety framework is tested with the conventional TTC metric. However, the safety framework can be tested with other available or newly developed metrics, and the results can be compared among the safety metrics with real crash data.
- In the present study, starting from developing the automated trajectory tool to processing the data, including training the deep learning models, the authors adopted various programming logics to conduct this research work. As a result, the involvement of many steps increases complexity while applying the present methodology to the roadway sections under consideration. Nevertheless, the study methodology can be sorted well with a simple Graphical User Interface (GUI) for the practical utilization of the study with a better visualization.
- In the present manuscript, the proposed safety framework is only tested. Further, the study can be extended, and comparative analysis can be conducted with the established safety measures with the proposed safety framework.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is supported by MITACS Global Research Award Canada, Application Ref IT16143.

ORCID

Shubham Patil http://orcid.org/0000-0002-1324-5678
Narayana Raju http://orcid.org/0000-0002-3561-5676
Shriniwas S. Arkatkar http://orcid.org/0000-0002-1804-9465
Said Easa http://orcid.org/0000-0003-0754-138X

References

Abdel-Aty, M., Pande, A., Lee, C., Gayah, V., & Santos, C. D. (2007). Crash risk assessment using intelligent transportation systems data and real-time intervention strategies to improve safety on freeways. *Journal of Intelligent Transportation Systems*, *11*(3), 107–120. https://doi.org/10.1080/15472450701410395

Aguero-Valverde, J., & Jovanis, P. P. (2006). Spatial analysis of fatal and injury crashes in Pennsylvania. *Accident Analysis and Prevention*, *38*(3), 618–625. https://doi.org/10.1016/j.aap.2005.12.006

Arvin, R., Khattak, A. J., Kamrani, M., & Rio-Torres, J. (2020). Safety evaluation of connected and automated vehicles in mixed traffic with conventional vehicles at intersections. *Journal of Intelligent Transportation Systems*, *25*(2), 170–187. https://doi.org/10.1080/15472450.2020.1834392

Asaithambi, G., & Joseph, J. (2018). Modeling duration of lateral shifts in mixed traffic conditions. *Journal of Transportation Engineering, Part A: Systems*, *144*(9), 04018055. https://doi.org/10.1061/JTEPBS.0000170

Behbahani, H., Nadimi, N., Alenoori, H., & Sayadi, M. (2014). Developing a new surrogate safety indicator based on motion equations. *PROMET – Traffic & Transportation*, *26*(5), 371–381. https://doi.org/10.7307/ptt.v26i5.1388

Ben-Bassat, T., & Shinar, D. (2011). Effect of shoulder width, guardrail and roadway geometry on driver perception and behavior. *Accident Analysis and Prevention*, *43*(6), 2142–2152. https://doi.org/10.1016/j.aap.2011.06.004

Blanco, M., Atwood, J., Russell, S., Trimble, T., McClafferty, J., & Perez, M. (2016). Automated vehicle crash rate comparison using naturalistic data. *Vtti*. https://doi.org/10.13140/RG.2.1.2336.1048

Chen, Y., Hu, C., & Wang, J. (2019). Human-centered trajectory tracking control for autonomous vehicles with driver cut-in behavior prediction. *IEEE Transactions on Vehicular Technology*, *68*(9), 8461–8471. https://doi.org/10.1109/TVT.2019.2927242

Chen, Z., Qin, X., & Shaon, M. R. R. (2018). Modeling lane-change-related crashes with lane-specific real-time traffic and weather data. *Journal of Intelligent Transportation Systems*, *22*(4), 291–300. https://doi.org/10.1080/15472450.2017.1309529

Cunto, F., & Saccomanno, F. F. (2008). Calibration and validation of simulated vehicle safety performance at signalized intersections. *Accident Analysis and Prevention*, *40*(3), 1171–1179. https://doi.org/10.1016/j.aap.2008.01.003

Das, S., & Maurya, A. K. (2019). Modeling maneuverability of motorized two-wheelers during filtering in urban roads. *Transportation Research Record: Journal of the*

Transportation Research Board, *2673*(5), 637–647. https://doi.org/10.1177/0361198119842818

Deng, Z., Chu, D., Wu, C., He, Y., & Cui, J. (2019). Curve safe speed model considering driving style based on driver behaviour questionnaire. *Transportation Research Part F: Traffic Psychology and Behaviour*, *65*, 536–547. https://doi.org/10.1016/j.trf.2018.02.007

Ellison, A. B., Greaves, S. P., & Bliemer, M. C. J. (2015). Driver behaviour profiles for road safety analysis. *Accident Analysis and Prevention*, *76*, 118–132. https://doi.org/10.1016/j.aap.2015.01.009

Engström, J., Johansson, E., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*(2), 97–120. https://doi.org/10.1016/j.trf.2005.04.012

Gettman, D., Pu, L., Sayed, T., & Shelby, S. (2008). Surrogate safety assessment model and validation: Final report. Publication No. FHWA-HRT-08-051.

Gilliland, M. (2010). Forecasting FAQs. The business forecasting deal: exposing myths, eliminating bad practices, providing practical solutions by Michael Gilliland Copyright © 2010 by SAS Institute, Inc., 193–246.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Sequence modeling: Recurrent and recursive nets. *Deep Learning*, 367–415.

Google. (2020). TensorFlow. An end-to-end open source machine learning platform. https://www.tensorflow.org/

Halawa, L. J., Wibowo, A., & Ernawan, F. (2019). *Face recognition using faster R-CNN with inception-V2 architecture for CCTV camera [Paper presentation]*. ICICOS 2019 – 3rd International Conference on Informatics and Computational Sciences: Accelerating Informatics and Computational Research for Smarter Society in the Era of Industry 4.0, Proceedings. https://doi.org/10.1109/ICICoS48119.2019.8982383

Happee, R., Gold, C., Radlmayr, J., Hergeth, S., & Bengler, K. (2017). Take-over performance in evasive manoeuvres. *Accident Analysis and Prevention*, *106*, 211–222. https://doi.org/10.1016/j.aap.2017.04.017

Hoseinzadeh, N., Arvin, R., Khattak, A. J., & Han, L. D. (2020). Integrating safety and mobility for pathfinding using big data generated by connected vehicles. *Journal of Intelligent Transportation Systems*, *24*(4), 404–420. https://doi.org/10.1080/15472450.2019.1699077

Jain, A., Mishra, A., Shukla, A., & Tiwari, R. (2019). A novel genetically optimized convolutional neural network for traffic sign recognition: A new benchmark on belgium and chinese traffic sign datasets. *Neural Processing Letters*, *50*(3), 3019–3043. https://doi.org/10.1007/s11063-019-09991-x

Lalmas, M. (2011). XML information retrieval. In *Understanding information retrieval systems: Management, types, and standards*. Auerbach Publications. https://doi.org/10.1201/b11499-29

Li, Y., Wang, H., Wang, W., Liu, S., & Xiang, Y. (2016). Reducing the risk of rear-end collisions with infrastructure-to-vehicle (I2V) integration of variable speed limit control and adaptive cruise control system. *Traffic Injury Prevention*, *17*(6), 597–603. https://doi.org/10.1080/15389588.2015.1121384

Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., & Dollár, P. (2015). *Microsoft COCO: Common objects in context [Paper presentation]. Proceedings of the IEEE computer society conference on computer vision and pattern recognition*. https://doi.org/10.1109/CVPR.2014.471

Lyu, N., Duan, Z., Ma, C., & Wu, C. (2021). Safety margins–A novel approach from risk homeostasis theory for evaluating the impact of advanced driver assistance systems on driving behavior in near-crash events. *Journal of Intelligent Transportation Systems*, *25*(1), 93–106. https://doi.org/10.1080/15472450.2020.1795846

Ma, L., & Qu, S. (2020). A sequence to sequence learning based car-following model for multi-step predictions considering reaction delay. *Transportation Research Part C: Emerging Technologies*, *120*, 102785. https://doi.org/10.1016/j.trc.2020.102785

Meng, Q., & Qu, X. (2012). Estimation of rear-end vehicle crash frequencies in urban road tunnels. *Accident Analysis and Prevention*, *48*, 254–263. https://doi.org/10.1016/j.aap.2012.01.025

Meuleners, L., & Fraser, M. (2015). A validation study of driving errors using a driving simulator. *Transportation Research Part F: Traffic Psychology and Behaviour*, *29*, 14–21. https://doi.org/10.1016/j.trf.2014.11.009

Nadimi, N., Behbahani, H., & Shahbazi, H. R. (2016). Calibration and validation of a new time-based surrogate safety measure using fuzzy inference system. *Journal of Traffic and Transportation Engineering (English Edition)*, *3*(1), 51–58. https://doi.org/10.1016/j.jtte.2015.09.004

Nelli, F., & Nelli, F. (2018). Deep learning with TensorFlow. In *Python Data Analytics*. https://doi.org/10.1007/978-1-4842-3913-1_9

Pattanayak, S. (2018). Pro deep learning with TensorFlow: a mathematical approach to advanced artificial intelligence in Python. In *Python Data Analytics*. Apress. https://doi.org/10.1007/978-1-4842-3913-1_9

Pek, C., Zahn, P., & Althoff, M. (2017). *Verifying the safety of lane change maneuvers of self-driving vehicles based on formalized traffic rules [Paper presentation]. IEEE Intelligent Vehicles Symposium, Proceedings*. https://doi.org/10.1109/IVS.2017.7995918

Precht, L., Keinath, A., & Krems, J. F. (2017a). Effects of driving anger on driver behavior – Results from naturalistic driving data. *Transportation Research Part F: Traffic Psychology and Behaviour*, *45*, 75–92. https://doi.org/10.1016/j.trf.2016.10.019

Precht, L., Keinath, A., & Krems, J. F. (2017b). Identifying the main factors contributing to driving errors and traffic violations – Results from naturalistic driving data. *Transportation Research Part F: Traffic Psychology and Behaviour*, *49*, 49–92. https://doi.org/10.1016/j.trf.2017.06.002

Przybyla, J., Taylor, J., Jupe, J., & Zhou, X. (2015). Estimating risk effects of driving distraction: A dynamic errorable car-following model. *Transportation Research Part C: Emerging Technologies*, *50*, 117–129. https://doi.org/10.1016/j.trc.2014.07.013

Python Community, D. (2020). *LABELIMG utility*. Project Description. https://pypi.org/project/labelImg/

Qu, D. Y., Han, L. W., Jia, Y. F., Zhou, J. C., & Hei, K. X. (2020). Vehicle behavior characteristics considering lateral offset and its full velocity difference expansion model. *Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)*, *50*(3), 963–969. https://doi.org/10.13229/j.cnki.jdxbgxb20190056

Raju, N., Arkatkar, S., Easa, S., & Joshi, G. (2021). Customizing the following behavior models to mimic the weak lane based mixed traffic conditions. *Transportmetrica B: Transport Dynamics*, 1–28. https://doi.org/10.1080/21680566.2021.1954562

Raju, N., Arkatkar, S., & Joshi, G. (2020). Effect of construction work zone on traffic stream parameters using vehicular trajectory data under mixed traffic conditions. *Journal of Transportation Engineering, Part A: Systems*, *146*(6), 05020002. https://doi.org/10.1061/JTEPBS.0000353

Raju, N., Kumar, P., Arkatkar, S., & Joshi, G. (2019). Determining risk-based safety thresholds through naturalistic driving patterns using trajectory data on expressways. *Safety Science*, *119*, 117–125. https://doi.org/10.1016/j.ssci.2019.01.017

Raju, N., Arkatkar, S., & Joshi, G. (2019). *Experiences of supervised machine learning to replicate driving behavior under traffic environment [Paper presentation]*. Transportation Research Board 98th Annual MeetingTransportation Research Board. https://trid.trb.org/view/1601033

Ren, P., Wang, L., Fang, W., Song, S., & Djahel, S. (2020). A novel squeeze YOLO-based real-time people counting approach. *International Journal of Bio-Inspired Computation*, *16*(2), 94. https://doi.org/10.1504/IJBIC.2020.109674

Rogers, S., Langley, P., & Wilson, C. (1999). *Mining GPS data to augment road models [Paper presentation]*. https://doi.org/10.1145/312129.312208

Rossum, G. V., & Swallow, U. (2011). Python (programming language). Tutorials Point (I) Pvt. Ltd.

Rowden, P., Steinhardt, D., & Sheehan, M. (2008). Road crashes involving animals in Australia. *Accident Analysis and Prevention*, *40*(6), 1865–1871. https://doi.org/10.1016/j.aap.2008.08.002

Sagberg, F., Selpi, Bianchi Piccinini, G. F., & Engström, J. (2015). A review of research on driving styles and road safety. *Human Factors*, *57*, 1248–1275. https://doi.org/10.1177/0018720815591313

Schepers, P., Twisk, D., Fishman, E., Fyhri, A., & Jensen, A. (2017). The Dutch road to a high level of cycling safety. *Safety Science*, *92*, 264–273. https://doi.org/10.1016/j.ssci.2015.06.005

Schneidereit, T., Petzoldt, T., Keinath, A., & Krems, J. F. (2017). Using SHRP 2 naturalistic driving data to assess drivers' speed choice while being engaged in different secondary tasks. *Journal of Safety Research*, *62*, 33–42. https://doi.org/10.1016/j.jsr.2017.04.004

Schwarz, C. (2014). On computing time-to-collision for automation scenarios. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 283–294. https://doi.org/10.1016/j.trf.2014.06.015

Soriguera, F., Torné, J. M., & Rosas, D. (2013). Assessment of dynamic speed limit management on metropolitan freeways. *Journal of Intelligent Transportation Systems*, *17*(1), 78–90. https://doi.org/10.1080/15472450.2012.719455

Stavrinos, D., Jones, J. L., Garner, A. A., Griffin, R., Franklin, C. A., Ball, D., Welburn, S. C., Ball, K. K., Sisiopiku, V. P., & Fine, P. R. (2013). Impact of distracted driving on safety and traffic flow. *Accident Analysis and Prevention*, *61*, 63–70. https://doi.org/10.1016/j.aap.2013.02.003

Strauss, J., Miranda-Moreno, L. F., & Morency, P. (2015). Mapping cyclist activity and injury risk in a network combining smartphone GPS data and bicycle counts. *Accident Analysis and Prevention*, *83*, 132–142. https://doi.org/10.1016/j.aap.2015.07.014

Taubman-Ben-Ari, O., & Katz-Ben-Ami, L. (2012). The contribution of family climate for road safety and social environment to the reported driving behavior of young drivers. *Accident Analysis and Prevention*, *47*, 1–10. https://doi.org/10.1016/j.aap.2012.01.003

US Department of Transportation and Federal Highway Administration. (2007). Next Generation Simulation (NGSIM). https://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm

Vaiana, R., Iuele, T., Astarita, V., Caruso, M. V., Tassitani, A., Zaffino, C., & Giofrè, V. P. (2014). Driving behavior and traffic safety: An acceleration-based safety evaluation procedure for smartphones. *Modern Applied Science*, *8*(1), 88. https://doi.org/10.5539/mas.v8n1p88

Van Der Horst, R., & Hogema, J. (1993). *Time-to-collision and collision avoidance systems [Paper presentation]*. 6th ICTCT Workshop—Safety Evaluation of Traffic Systems: Traffic Conflicts and Other Measures.

Van Driel, C. J. G., & Van Arem, B. (2010). The impact of a congestion assistant on traffic flow efficiency and safety in congested traffic caused by a lane drop. *Journal of Intelligent Transportation Systems*, *14*(4), 197–208. https://doi.org/10.1080/15472450.2010.516226

Victor, T. (2016). *Analysis of naturalistic driving study data: safer glances, driver inattention, and crash risk*. The National Academies Press. https://doi.org/10.17226/22297

Vogel, K. (2003). A comparison of headway and time to collision as safety indicators. *Accident Analysis and Prevention*, *35*(3), 427–433. https://doi.org/10.1016/S0001-4575(02)00022-2

Wang, P., Rau, P.-L P., & Salvendy, G. (2010). Youth and road safety. *Traffic Injury Prevention*, *11*(4), 425–432. https://doi.org/10.1080/15389581003754593

Wang, L., Zhang, Y., & Feng, J. (2005). On the Euclidean distance of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(8), 1334–1339. https://doi.org/10.1109/TPAMI.2005.165

Wu, Y., Abdel-Aty, M., Wang, L., & Rahman, M. S. (2020). Combined connected vehicles and variable speed limit strategies to reduce rear-end crash risk under fog conditions. *Journal of Intelligent Transportation Systems*, *24*(5), 494–513. https://doi.org/10.1080/15472450.2019.1634560

Yan, X., Abdel-Aty, M., Radwan, E., Wang, X., & Chilakapati, P. (2008). Validating a driving simulator using surrogate safety measures. *Accident; Analysis and Prevention*, *40*(1), 274–288. https://doi.org/10.1016/j.aap.2007.06.007

Zheng, L., Sayed, T., & Essa, M. (2019). Validating the bivariate extreme value modeling approach for road safety estimation with different traffic conflict indicators. *Accident Analysis and Prevention*, *123*, 314–323. https://doi.org/10.1016/j.aap.2018.12.007