Sigma Journal of Engineering and Natural Sciences, Vol. 42, No. 6, pp. XX-XX, December, 2024
Yildiz Technical University Press, Istanbul, Turkey

Enhancing road safety through blackspots mitigative measures - a review

Prasad MANOJ[1,*], Kotagi B PUNITH[2], Kerekoppa Chandrashekar MANJUNATH[2]

ABSTRACT

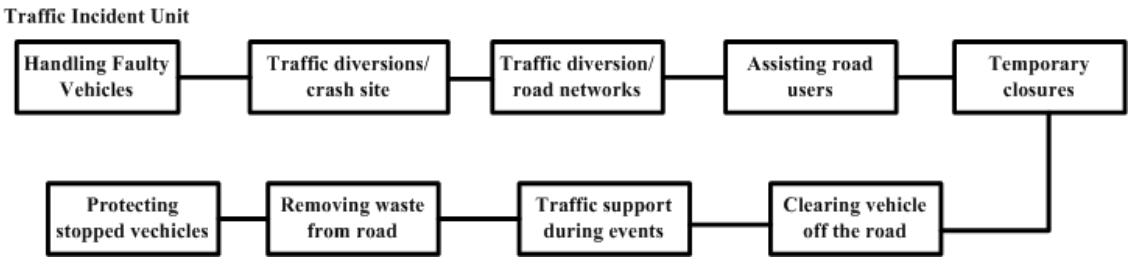
Precise identification of black spots can ensure the timely adoption of remedial measures to mitigate road accidents. The review of various trends of approaches towards identifying the black spots by studying the evolved-up techniques towards a pre-emptive investigation of crash events is the focus of the paper. This study contributes to briefing all the open research issues that existing researchers have merely explored. The paper also discusses addressing such open-end problems that offer more edge to the precise analysis of black spots, followed by a possible plan towards solving these issues to mitigate blackspots.

Keywords: *Road Safety Management, Black Spots, Analysis of Accidents, Crash Events, Predictive Analysis.*

INTRODUCTION

Safety is a critical requirement of any land transportation system for any nation with an increasing rate of crash events including severe fatalities. Majority of the road accidents occur due to over-speeding at the National Highways and Freeways. [1]. Such fatal accidents not only cost lives but also cost extreme economic losses. Currently, the standards of the management techniques for traffic incidents have multiple objectives such as; i) offering a faster response to the emergency incident associated with road crash, ii) management of the causalities and its connected problems iii) safeguarding and vacating the location of the accident iv) providing clearance to the traffic at a faster pace (Figure 1). In order to achieve this, the initial process is to manage the defective vehicle, followed by a diversion of traffic or the location of the crash event. Another important traffic incident unit is to help the road users, followed by the temporary closure of the lane system where the accident has happened. The next process is to clean all the debris and clear the vehicle with the help of support staff during the post-crash event. This eventually mean that there are various operation forms after the accident event, which results in economic losses. With a consistent effort from the government and varied sources, there is a continuous evolution towards leveraging the road safety process. In this regard, the term blackspot plays a critical role while defining the road-safety process.

A black spot is a location that has recorded maximum accidents owing to diversified reasons. [2, 3]. Some of the frequent reasons are hairpin bends, no traffic lights, sharp corners of the road, concealed oncoming vehicles, fast roads with a hidden junction, low-quality traffic signs, etc. [4]. Apart from this, various other factors also result in road crash events such as i) physical and mental state of the driver, ii) overloading state of the vehicle, iii) intoxication, iv) over-speeding, etc.



This paper was recommended for publication in revised form by Regional Editor Ahmet Selim Dalkilic
[1] Department of Civil Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka India
[2] Department of Civil Engineering, The National Institute of Engineering, Mysuru, Karnataka India
** E-mail address: manoj.p@vvce.ac.in*
Orcid id: https://orcid.org/0000-0002-7185-5992 Prasad Manoj, 0000-0001-5504-0534 Kotagi B. Punith, 0000-0002-0482-5910 Kerekoppa Chandrashekar Manjunath
Manuscript Received 02 March 2023, Revised 04 April 2023, Accepted 13 May 2023

Figure 1. Management of Traffic Incident

At present, the study of the road-safety is carried out over a mechanism that can collect accident data in order to perform the following operations such as i) Road Safety Audit, ii) Road Safety Impact Assessment, iii) Black Spot Management, iv) Network Safety, v) Road safety Inspection. [5, 6]. Although all these operations are equally important, black spot management is treated as more critical and challenging to be solved, and hence, the majority of the work towards road safety emphasizes black spot management approaches. In such an approach, the emphasis is on identifying the black spots with the highest priority and offering a remedial solution. From the viewpoint of management of road infrastructure, there are two types of strategies in practice, i.e., preventive and reactive strategy. The preventive strategies are Road Safety Impact Assessment, Road Safety Audit, and Road Safety Inspection, while the reactive strategies are Black Spot Management and Network Safety Management with discrete applicability on existing roads and novel schemes (Figure 2). The existing research work is directed towards these strategies, while more exploration towards exploring the effectiveness of such studies has been further investigated to evolve up with new and robust safety measures consistently.

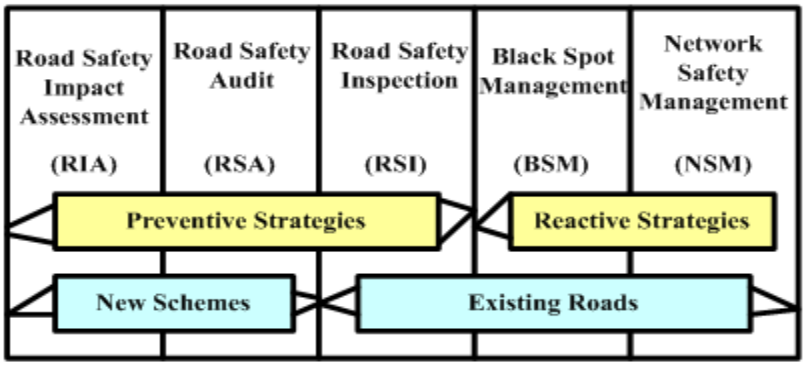


Figure 2. Classifications of Strategies of Road Safety

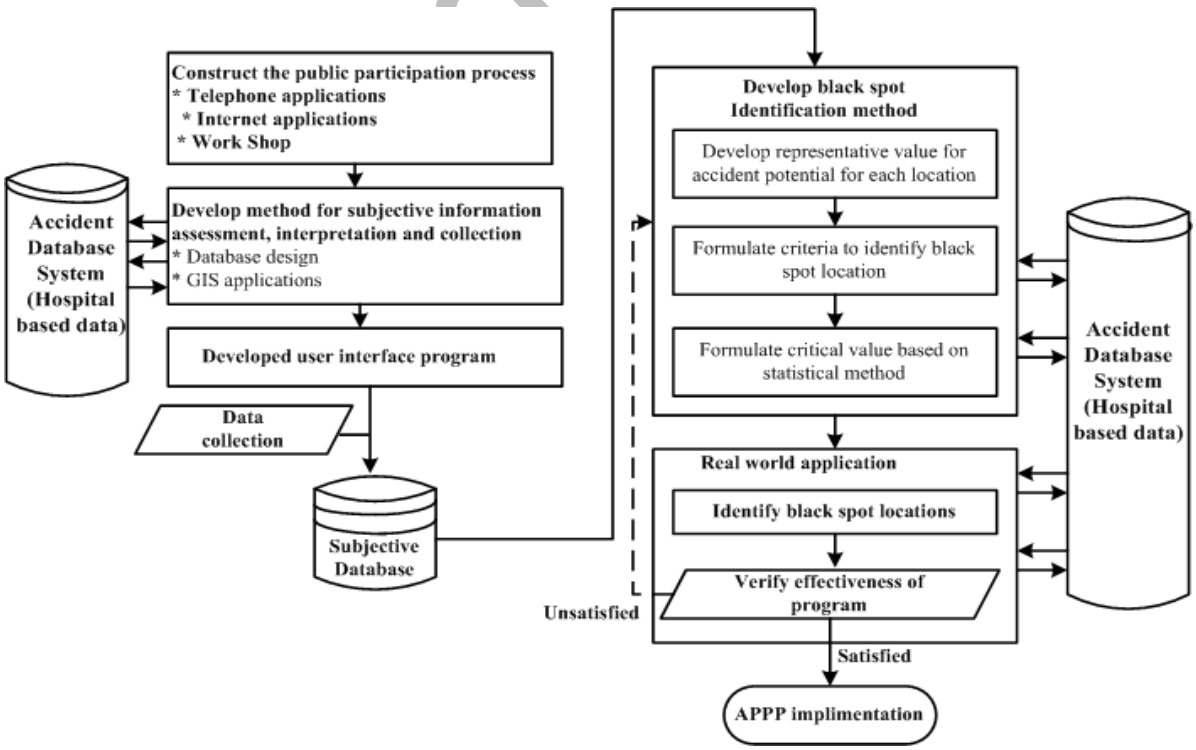


Figure 3. Conventional Process of Blackspot Identification

The conventional process of blackspot identification (Figure 3) is all about using an accident database system followed by constructing a process for the subjective assessment of information. The primary information can be qualitatively collected publicly. Finally, a model for identifying blackspot detection is formulated, which could use various methods to confirm the accurate identification of the potential blackspot.

This paper discusses existing methods of blackspot detection by various researchers with an agenda to extract open-end research problems associated with the topic to make constructive suggestions towards possible solutions for traffic safety issues.

The organization of article is as follows: Section 2 discusses existing approaches of road accident studies, followed by Section 3 which highlights the open research issues and Section 4 highlights the discussions and future scope of research.

EXİSTİNG APPROACHES OF ROAD ACCİDENT STUDİES

At present, there are massive archives of the research work towards road and traffic safety over multiple domains, multiple problems, and multiple solutions. It is found that there are mainly 3 types of standard methods for black spot analysis, i.e., discussion-based, observational-based, and modelling-based (Figure 4). All these approaches have their uniqueness and setbacks which are discussed in the following section.

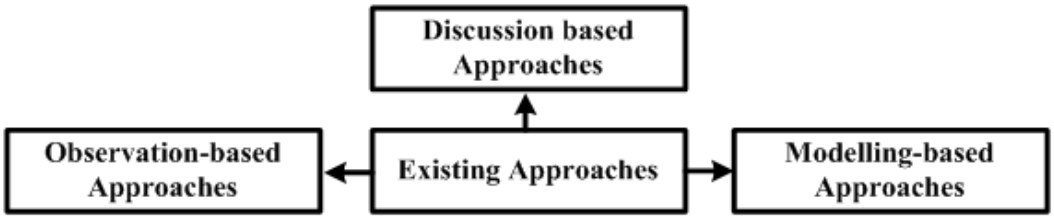


Figure 4. Classification of Existing Approach

DİSCUSSİON-BASED APPROACHES

At present, various studies have offered a theoretical baseline of the approaches associated with analysing the road accidents. According to researchers, there are various ranges of factors that are responsible for causing road accidents. The discussion-based approaches highlights the already existing research works with respect to data and concept. However, the researchers choose to illustrate it further to extract various causes and factors that could be treated as a latent attribute towards crash events.

Shinar [7] has carried out a study to establish connectivity between accidents with different reasons, e.g., such as safety rule and protocol implementation. The authors discussed that there are various other reasons such as human reason (texting, speaking over the phone while driving, distraction, pre-occupied thoughts), environmental reason (sign missed, usage of non-standard road safety sign, glare from the sun, sign view obstructed), and vehicle reason (loss of brake fluid, exhausted brake pads, lack of disc brakes). This study concludes that associating human error or any other mechanical failures with the accident's predominant causes is usually not true and is found to be highly biased.

In another recent investigation, that road infrastructure was another indicator of vehicle crash events. Different flaws associated with infrastructure have been considered for the investigation such as features of the cross-section of road, properties of alignment, deficiencies of the road surface, junctions, etc. [8]. However, different studies have explored different attributes resulting in road accidents. This investigation reports the adoption of expert opinion to extract data from police officials and the public. [9]. Quantitative analysis in this study shows that flaws in enforcement of law are the primary reason for effective reporting policies.

Sleep factor is one of the prominent cases resulting in road accidents by latent drivers. A discrete set of discussion group has investigated the relationship between the sleep factor and road accident. A simulation-based analysis has been carried out considering the road's test environment with high and low traffic situations over

different times of a day. [10]. The study reveals that the driver's exhaustion condition for long drives significantly affects the vehicle's driving behaviour; however, there is no impact on the sleep factor's road environment. The analysis of the study was carried out considering the data collected by real-participants (drivers). A similar mode of research work was also carried out by Alvaro *et al.* [11], where it was discussed that if the drivers are updated with precautionary-based information associated with driving over odd situations, then the frequency of accident events can be reduced. However, the outcome of the study has been offered using a smaller number of participants.

Certain interesting studies state about diversified factors associated with road safety. A unique case study discussed by Ando *et al*. [12] has correlated crime events with road accident considering Japan's case study. Statistical analysis was considered over the city's data showed that there are significant effects of crimes over road accidents. Savolainen *et al*. [13] has carried out statistical analysis to review the criticality of the accidents.

In all the discussion mentioned above, data connected with an accident are the primary source of investigation. The investigation carried out by Lu *et al*. [14] reveals that current approaches are frequently associated flaws in analyzing and modelling the crash data. The study reports that irrespective of progressive improvement towards predictive modelling applied over such crash-related data, much of the forecasting operation's effectiveness was missing. Additionally, it was also discussed that there is a necessity for data with more detailed information about the vehicle's mechanical operation along with black-box system-based data for crash reports.

According to various researchers, e.g., Lu *et al*. [15] and Ekapun *et al*. [16], artificial intelligence and machine learning plays an essential role in analyzing data associated with civil works, e.g., road management and crash analysis. Approaches associated with machine learning offer the more predictive and exclusive operation of classification that could provide more edge to the analysis of road accidents. The discussion carried out by Ekapun *et al*. [16] offers an exciting finding where heavy makeup acts as an impediment to eyesight while driving. Like this, many other models offer comprehensive implementation of deep learning for various forms of transportation improvements.

Underwood *et al*. [17] presented an analysis model that can perform both numerical as well as hypothetical investigation of road accidents.According to study carried out by Gutierrez *et al*. [18], best results are obtained when two or more machine learning techniques are combined and Silva *et al*. [19] founded that Machine Learning technique is promising and road –environmental problems are the most used in modelling approaches.

Table 1. Summary of Discussion-Based Approach

References	Objective	Essential findings
Shinar [7]	Exploring non-conventional reason for road accident	Diversified factors, apart from human/technical errors.
Papadimitriou *et al.* [8]	Impact of road infrastructure on crash events	The study resulted in a ranking of eleven risk factor
Ahlstrom *et al*. [10]	Factors affecting the sleep of driver	Road condition does not affect sleep directly Night driving significantly affects sleep
Alvaro *et al.* [11]	Impact of imparting safety information on the accident	Study outcomes not completely validated for effectiveness
Ando *et al*. [12]	Impact of City crimes	Length of road, public facility number, can overcome accidents
Savolainen *et al*. [13]	Investigating criticality of accidents	The existing system is diversified and lacks benchmarks
Lu *et al*. [14]	Analysis of statistical approaches	Study limited to the scope of crash data
Ekapun *et al.* [16]	Impact of eye status on accident	Effective recognition using deep learning
Gutierrez *et al*. [18]	Analysis & forecasting of road accidents	Best results are obtained when two or more machine learning techniques are combined
Silva *et al*. [19]	Crash Predictive Model	Machine Learning technique is promising and road – environmental problems are the most used in modelling approaches.

OBSERVATİON-BASED APPROACHES

Performing an analysis of the existing system will offer insight into the strengths and weaknesses of analyzing accident events over roads. The observation-based studies deal with the discussion of the existing approaches toward accident analysis, where the researchers apply their analysis technique over the existing data or perform a more in-depth investigation of existing approaches of analysis.

Existing approaches are found to apply various evaluation schemes over available crash data for observing the outcomes followed by an analysis of it. However, while doing so, most of the existing schemes have been considering the stability factor connected with the parameters of the model for analyzing crash data. However, the discussion presented by Mannering [20] has concluded that the factors responsible for road accidents have no significant change associated with temporal factor and is also considered to be continued owing to different behavioral factors. Such forms of temporal stability factor are yet undisclosed in any research work. Apart from this, evaluations of the adoption of existing methods are also questionable for their success rate. It is because analysis of road accidents should focus on multiple objectives that could be synchronized and implemented for resisting accidents. [21]. This study discusses the methods used for assessing the performance of human factors integrated with robust safety technology usage, followed by a trustworthy organization.

A road accident is studied concerning different types of vehicles, but it is also studied concerning pedestrians. It has been studied that the pedestrian makes mistakes in judgment while crossing the road, which is influenced by different external factors. [22]. Study shows that narrow road offers confusing decision making by the pedestrian for crossing based on the remaining time. Most interesting findings in a most recent study by Intini *et al.* [23] showed that familiarity of the route impacts the crash event's occurrence. Considering the Norwegian highway's case study, the researchers have investigated possible connectivity between crashes and route familiarity using a logistic regression model. The study outcome shows that there are significant factors that affect the driver with route familiarity to encounter a crash are i) fluctuation on summer traffic, ii) small driveways, iii) winter/autumn season, iv) speed limits, v) travel goal. Similarly, for the driver with non-route familiarity are affected by i) maximum density of traffic, ii) summer, iii) head-on collision, iv) young drivers, and v) involvement of heavy vehicles. Some other impacting factors are distraction and lethal driving behavior.

Existing research has also been initiated towards exploring crashing occurring at midblock of roads. [24]. The study discusses a multivariate binomial framework towards investigating to come up with some non-conventional outcome. The observation carried out using the statistical approach shows that road accidents are not significantly caused due to lane density, the number of lanes, and the segment's length. Amazingly, the study shows that even a higher speed limit has no significant impact on road crashes than the low-speed limit. Apart from this, even the conventional reasons for accidents (e.g., parking on the street, shoulders, the width of the lane, one-way traffic, etc.) are proven to be less significant in this study.

There was a study where analysis of accidents connected with light vehicles has been discussed. [25]. The study has also discussed that the driver's age factor has a significant impact on the driving quality where younger drivers are claimed to have lesser faults in contrast to aged drivers. Considering the case study of Singapore, the authors have used binary logit model [26] over crash data to find that maximum accident has been witnessed with the older driver during festive seasons as well as different road types (singe/multiple lane/wet surfaces, etc.). Study towards light moving vehicles like bicycles has also been investigated for its possibilities of crashes over the road. [27]. The study findings highlight that the usage of cycles increases over the shared road that significantly causes an accident due to the economic viewpoint.

Another unique study has been carried out by Wang *et al*. [28], where the authors have discussed the prediction-based approach to analyze the crash, even in real-time. The study shows that parameters associated with the trip generation and socio-demographic aspects significantly impact the road accident. The authors have used a

support vector machine and Bayesian Logistic Regression to forecast the reasons for the accident. The study outcome also says that the number of vehicles, the vehicle's speed, and the proportion of production contribute to the accident.

A similar age factor was also considered as the core attribute in the recent study by Gargoum *et al*. [29]. The study discussed that with the increasing rise of the aged driver, it is required to offer facilitation over the road to assist the aged drivers from fault-free driving. For this purpose, the authors have introduced few parameters, e.g., Available Sight Distance (ASD) and Stopping Sight Distance (SSD), that are distance-based parameters connecting with stopping the vehicle in the event of identifying a crash. The authors have also introduced multiple SSD forms based on the (weak/good) drivers' ability considering the case study of Canadian road. The study outcome shows that accidents are more likely to happen over road regions that do not offer enough information about SSD, and aged drivers are mostly a part of it.

Another observation-based approach was carried out by Mardula *et al*. [30], where the analysis was carried out to assess the potential of existing studies considering India's case study; however, it was purely a survey-based study. Researchers have found that various information can be derived from the accident data that could offer more insights into the causes of the accident. The study found that the truck's involvement, e.g., the truck was always studied frequently, but the researchers e.g., Chen *et al.* [31] have investigated that number of such heavy vehicles is also a factor influencing accident. The study outcome shows that roads filled with snow and traffic density are the prime indicators of accidents. The study connected with the heavy vehicle is also carried out by Hickman *et al*. [32] with a secondary emphasis on the crash dataset. The study says that crash data is only beneficial if they offer comprehensive information about the behavioral aspect of the driver and driving sequences. The study advocates the usage of crash data characterized by continuous data and baseline event of driving and events with safety and critical attributes.

It has been seen that road safety is being associated with the adherence to standard design factors too. An existing study has been carried out to investigate the impact of roadways' design factor on road safety. [33]. Considering the dataset of Indiana highways, the study has used statistical analysis to find that existing design standards practices contribute to accidents. The study also found that a similar process could be used for evaluating occurrences of accidents for roadways with and without adherence to design standards.

Table 2. Summary of Observation-based Approaches

References	Objective	Technique	Findings
Mannering [20]	Temporal stability of data	Qualitative	An open problem
Petkov [21]	Quantified context, the human factor	quantitative	The integrated analysis offers more outcomes.
Zhuang *et al*. [22]	Pedestrian crossing	Statistical	Remaining time and decided to crossroads
Intini *et al.* [23]	Familiarity of routes	Logistic regression model	Drivers with both familiar and non-familiarity of routes encounter accidents due to different reasons.
Liu *et al.* [24]	Effects of Midblock	Multivariate binomial model	Non-biased findings improve road safety
Chin *et al*. [25]	Facts affecting old drivers	Binary logit model	singe/multiple lane/wet surfaces
Wang *et al*. [28]	Effects of Socio-demographic	Support Vector Machine, Bayesian, Logistic Network	Vehicle speed, number, production
Gargoum *et al*. [29]	Distance-based stopping vehicle	Experimental	Road with lesser SSD is highly accident-prone, aged drivers.
Chen *et al.* [31]	Number of heavy vehicles	Qualitative	Snowy road, light traffic
Hickman *et al*. [32]	Significance of dataset	Risk analysis	A dataset with normative, critical, and

			safety offer a better behavioral aspect of driving
Malyshkina *et al*. [33]	Compliance to standard roadways design	Statistical	No significant impact process of design on road safety.

MODELLİNG-BASED APPROACHES

In such an approach, the researchers attempt to construct a representation of objects associated with the crash investigation problems for road safety. These approaches mainly focus on essential issues for crash investigation by filtering all the non-essential information and criteria. Apart from this, the adoption of modelling approaches also assists researchers to idealize their concept towards road safety considering various parameters. Another beneficial part of the modeling-based approach is its capability to adapt to various changes according to the researchers' constraint.

There are various modelling aspects introduced by the researchers towards improving traffic safety by addressing critical issues associated with identifying essential indicators leading to crash. Bener *et al*. [34] developed a hybrid model that uses path analysis and a hidden class to explore all the essential factors for the crash. Experiments and analyses have been carried out using various datasets that play a vital role in investigating road safety. However, if the data itself has a problem, it may pose an impediment during the investigation. This problem has been addressed by Squillante *et al*. [35] by introducing a multivariate model integrated with the Bayesian network's learning approach. Apart from this, it was explored that various event associated with the physical environmental condition also affects the road safety. A model is constructed using a methodological approach to assess the criticality of the crash events. [36].

A specialized case of Euclidean space, i.e., the fractal theory, is reported to be considered for developing a model establishing a relationship between traffic congestion and vehicle crash events. [37]. It was also found that the adoption of a modelling approach using latent classes over data related to traffic speed can be used for better classification of the problems connected with the congestion factor. The model contributes to finding the black spot for the given speed data. Apart from space and time, various other factors are also proven to be accountable for accidents. It was discussed in the existing literature that a static and dynamic factor also contributes towards crash-investigation. [38]. The authors have used a logit framework for over the model's random attributes (i.e., static and dynamic) connected with the accident.

The climatic condition and natural factors have also been considered for the cause of accidents. Various natural calamities lead to varied intensity of rainfall and water depth level. [39]. The authors have used structural equation modelling based on natural entities (rain, water depth), human factors, traffic, road, environment, etc. The contribution of this model was to assess the criticality of road accidents caused due to such factors.

A unique study was carried out towards such investigation, which claims the positioning of the red-light image capturing device. [40]. This study has used a state-choice model that can evaluate the proper positioning of such a camera. Adopting machine learning and regression modeling was also found to offer a space-time gap connection with the accident. The authors have used a support vector machine and neural network considering various attributes, e.g., class of violation, the severity of the crash, condition of weather, condition of the road surface, etc. The study outcome shows that the integrated learning approach offers better goodness of fit than the linear model. A similar form of investigation using a support vector machine was carried out by Xu *et al*. [41], where wavelets have been used to denoise the data. The model looks for improving the rate of detection of traffic incidents.

The proliferation of the predictive approach, as well as machine learning, was also seen in existing approach using diversified techniques [24, 42-44] have used a Bayesian modelling approach where multiple models have been developed using time and spatial attributes, while the goodness of fit is used to assess the model performance. The outcome shows the superiority of the linear model in offering better forecasting performance. Another recent study of predictive approach was reported [45], where the auto-regression model has been used with a moving average for constructing the model. The study finds that a regression model with explanatory variables offers better performance than the conventional regression operation. Existing studies have also found that the degree of congestion also affects

crash events on the road. The work carried out by Park *et al*. [46] has constructed a neural network-based model along with the Bayesian approach of learning. The author has contributed to developing an approach for rule extraction to improve the predictive values' accuracy further. The study contributes to finding that a sudden change in traffic congestion leads to accidents.

Daylight Saving Time has a significant impact on the rate of accidents, which was analysed using a comprehensive model constructed by Robb *et al*. [47], where the rate of accident is considered as the essential predictor of the date of the accident. The study reflected some interesting outcomes that states that Daylight Saving Time is not much linked with road crash events; however, they are specifically high over certain particular days in a week.

However, the foundation of the adoption of machine learning approaches for constructing a predictive model is already discussed by Kargah *et al*. [48]. The predictive approach was also developed by Dhamaniya [49], where statistical methods have been utilized considering the case study of west India. According to this modelling, a regression is performed in order to forecast the accident frequency. A case study of Indian roads were also considered [50], where an epidemiological study has been carried out by constructing a regression model of Poisson-lognormal mixture. The model is constructed to establish connectivity between the travel distance and the road's fatal accidents using various core attributes, e.g., highway length, fuel consumption, population ratio, etc. However, linear modeling's beneficial factor is questionable using the statistical approach. [51]. Shafabakhsh *et al*. [52] and Cela *et al*. [53] developed a model using GIS and spacial analytical techniques as a management system for accident analysis and to establish a relationship between accident rate and road condition, road surface, road characteristics, main cause and time. Wang *et al*. [54] developed Hierarchical Bayesian logistic model and stated that Seven variables viz., road type, compatibility difference, age, vehicle type, drunk driving, driving unregistered vehicle, and driving year are significant and gender was not significant. The conventional tree-based approach was to be re-modeled towards the additive Poisson Regression method of multiple forms with an agenda to investigate the crash events with the pedestrian specific to certain geographic locations. Gu *et al*. [55] presented a framework for predicting intersection crash frequency using connected vehicle data and geographical random forest, demonstrating improved accuracy compared to traditional models. The study contributes to traffic safety by offering an effective approach to analyze crash occurrences at intersections and provides insights for targeted interventions. Ahmed *et al*. [56] explores the use of explainable machine learning models for road accident prediction and analyzes their performance, contributing to understanding the factors influencing accidents and aiding in the development of road safety measures. The study evaluates the accuracy and reliability of these models in predicting accidents and provides interpretable insights into the contributing factors, offering valuable contributions to the field of transportation research.

The comparison results indicate that Random Forest is the top-performing classifier, achieving an accuracy of 81.45%, precision of 81.68%, recall of 81.42%, and an F1-Score of 81.04%.

The modeling-based approach has proven to offer an effective solution to solving the problems associated with the blackspot identification with varied uniqueness, mainly focusing on accuracy achieved computationally.

Table 3: Summary of Modelling-based Approaches

Authors	Objective	Technique	Findings
Chand *et al.* [37]	Development of Predictive Model	Hurst Model Metric	Qualities of the motorist, Curvature, Pavement Quality, and Weather.
Cheng *et al.* [57]	Accident Analysis	Spatiotemporal Models	Reorganization of hot sport, Site ranking
Dhamaniya [49]	Crash Analysis	Accident Calculation Models	Mixed traffic forms
Ding [51]	Crash Detection	A machine learning approach	Road safety program, Policy proposals
Fountas *et al*. [38]	Accident Analysis	Logit Framework	High-risk location

Goel [50]	Preemptive model development	Poisson-lognormal Mixture Regression Model	Cycling, Walking, and Interestingly
Ihueze [45]	Crash Detection	ARIMAX Model	Decrease road crashes in Nigeria
Kargah *et al*. [48]	Crash Detection	Machine Learning	An illustrated perforation prediction models
Lee *et al*. [39]	Accident Analysis	Structure Equation Modelling	Road factor, Human factor, Water depth factor, Environment
Egbendewe *et al*. [40]	Development of Predictive model	Stated Choice Model	Enhance the red light camera program
Park *et al*. [46]	Crash Detection	Bayesian learning approach	Alert drivers
Shafabakhsh *et al*. [52]	Urban Accident Analysis	GIS – based spatial stastical method	GIS can be used as a management system for accident analysis
Cela *et al*. [53]	Accident Analysis	GIS and Spacial Analytical Techniques	Relationship between Accident rate and road condition, road surface, road characteristics, main cause and time.
Wang *et al*. [54]	Driver injury severity analysis of crashes	Hierarchical Bayesian logistic model	Seven variables viz., road type, compatibility difference, age, vehicle type, drunk driving, driving unregistered vehicle, and driving year are significant and gender was not significant
Cheng *et al.* [57]	Road Accident Prediction	Machine Learning Approach	The Random Forest algorithm gave the highest accuracy of 80.78 % when the accuracies of the Machine Learning models were compared.
Gu *et al*. [55]	Predicting intersection crash frequency	Geographical random forest method	Rear-end crashes are more likely to happen at intersections connecting minor roads compared to major road
Ahmed *et al*. [56]	Predict road accident Injury Severity	Machine Learning Approach	The comparison results show that Random Forest is the best classifier with 81.45% accuracy, 81.68% precision, 81.42% recall, and 81.04% of F1-Score

In India, the number of road accidents and fatalities is alarmingly high, with over 400 deaths per day on average. To address this issue, researchers have adopted various modelling-based approaches to identify the key factors associated with road accidents and develop effective strategies for improving road safety.

One of the main factors contributing to road accidents in India is the lack of infrastructure and inadequate road design. Researchers have used modelling approaches to assess the impact of road characteristics, such as width, curvature, and gradient, on accident rates. In addition, they have also studied the impact of traffic flow, pedestrian behaviour, and vehicle speed on road safety.

Another important factor contributing to road accidents in India is driver behavior. Researchers have used statistical models to analyze the impact of factors such as age, gender, education, and experience on driver behaviour and accident rates. They have also examined the role of factors such as fatigue, distraction, and alcohol consumption in contributing to road accidents.

In addition to these factors, environmental conditions such as weather, visibility, and road surface quality have also been found to play a significant role in road accidents in India. Researchers have developed models to assess the impact of these factors on accident rates and identify effective strategies for improving road safety under different environmental conditions.

Overall, the adoption of modelling-based approaches has proven to be an effective tool for identifying the key factors contributing to road accidents in India and developing effective strategies for improving road safety. These approaches can help policymakers and stakeholders make informed decisions and allocate resources more effectively to reduce the number of road accidents and fatalities in India.

OPEN RESEARCH ISSUES

The existing methods towards Analysis and Prediction of Accident Models have various associated advantages concerning various individual applications. However, it is also associated with limitations too. There is no doubt that there is increasing adoption of technological advancement in the traffic management process too. A closer look into the existing system will show that all the studies towards blackspot identification have been carried out towards using a specific form of case studies of a certain location. This restricts the workability of the existing models. Simultaneously, all the models are developed in a different manner, which also restricts the general applicability of the existing system. Therefore, some of the unaddressed issues are as follows:

- Usage of stale data: Almost all the research work has been carried out considering certain datasets, which is considered base input towards all the models. The study considers data collected from IRS P-6 LISS III data, FIR data from the nearest police station, or public work department. However, such a dataset is quite older than 1 year and undergoes less updating process even if repairing work has recently been done. Lack of communication between a police station and the public work department is another prominent cause of this stale data. Therefore, modelling carried out over such a dataset could form outliers in any computational model. With an increase of time, the blackspot's severity could further vary, which is never considered in any modelling. Hence, such experiments are quite restricted to offline analysis and never for online analysis of blackspot. Offline analysis means performing analysis with accident data that has already been collected, and its outcome is not meant for immediate usage. However, its outcome may be useful for investigation in the future if required. Online analysis means performing a predictive operation to obtain the target road's instantaneous outcome considering it has processed, analyzed, and trained data from the beginning. Online analysis is highly essential for users compared to offline analysis from user experience performance and uprising of new technological navigational devices.
- Outdated procedure for analysis of Non-Adoption of Current Technologies: The massive amount of accident data must be migrated to cloud-based services. Hence, when different forms of data are hosted over cloud-based services, there are different complexities of big data management [58], which gives complications when an analysis is performed. At present, there is no such discussion of such inevitable adoption of cloud-based services for road-safety measures. In order to understand the importance of real-time data on blackspot identification, Figure. 5 represents a pictorial depiction. Consider that a car driving through Lane-C targets reach its destination which can be taken using both Lane-A and Lane-B. Lane-A is at a higher distance while Lane-B is at a reduced distance from the car. Assume that the overall distance taken by lane-B is time consuming compared to lane-A. Hence, in normal road conditions, the vehicle will always choose to drive by Lane-A. Also, assume that Blackspot-1 is the older information that is already known to the car. In that condition, the driver decides to drive the car towards either lane-A or lane-B. However, assume of Blackspot-2 is a newly reported accident area officially reported in police FIR but not updated in a distributed road safety service system that also assists in navigation. Hence, stale information (i.e., driver possess older information of blackspot-1 but does not possess new information of blackspot-2). In such a case, the vehicle will either meet the accident or divert it back to Lane-C once it reaches near Blackspot-2. Therefore, the presence of real-time information (it is usually for accident events within a day or a couple of days) could significantly assist in reducing the inconvenience and resist the possibility of a new accident. It requires faster registration of accident information and uploading in a distributed navigational system for the driver to know about it. Unfortunately, there are no such devices or services that exist at present in this regard.

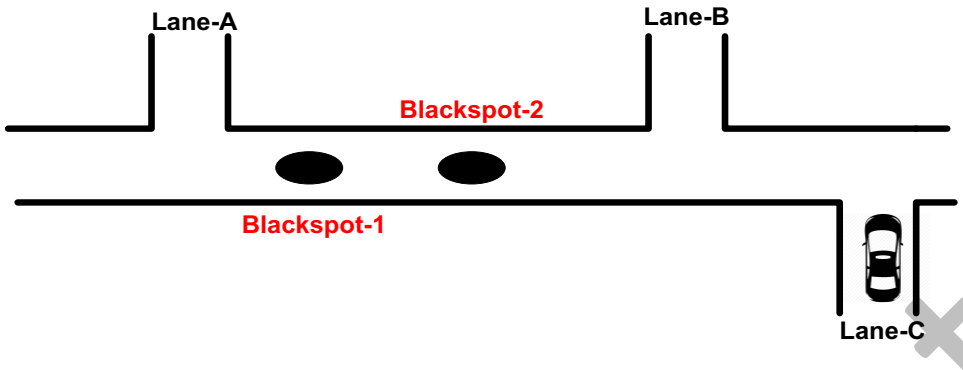


Figure 5. Pictorial representation of the need for real-time blackspot data

The data sources can be classified according to i) Government data ii) Open data iii) Onboard equipment iv) Measurement technologies and iv) Social media and best results are obtained when two or more machine learning techniques are combined. [18].

This section briefs about the open research issues in the line of various potential problems of Road Accident data analysis and the development of predictive models.

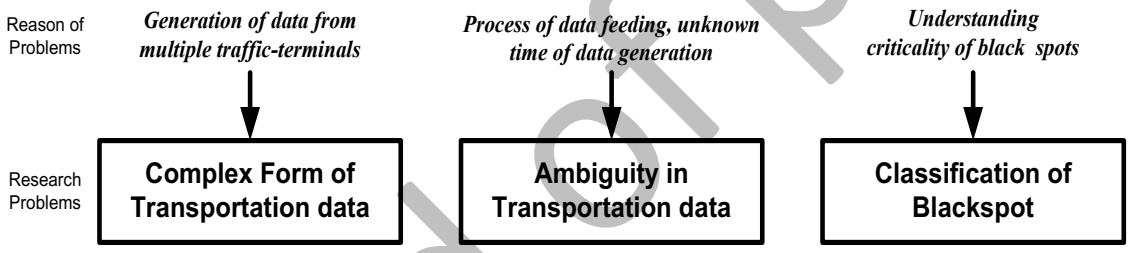


Figure 6. Identification of Research Problems

The identified problems in the existing literature are as follows:

- The majority of existing studies [45-47, 57] on black spots have not studied real-time data complexity in detail. The data arriving from the transportation system (RTO/Police Station/PRAMC) is highly complicated compared to other traffic-related data forms. Such problems are not addressed in the existing literature.
- Not many research efforts have been made to process the data before using them to analyze and develop models. The data processing will provide opportunities for analyzing complex data.

Hence, the problem statement of the proposed study can be stated as - "It is a challenging task to collect complex real-time data from the selected black spots and to process the data methodically to eliminate the error in the data for analyzing road accidents using appropriate techniques and to develop suitable models for predicting the road accidents and to suggest mitigative measures for Indian conditions.

DİSCUSSİONS & FURTHER SCOPE OF RESEARCH

Identification of the black spot is quite critical for ensuring road safety and the land transportation system. This paper has discussed various approaches towards the identification of the blackspot. At present, there are many discussion-related papers on traffic safety; however, this paper differs from others by contributing to some essential findings that will help the researchers make more concrete decisions towards road safety. The learning outcomes of the proposed paper are as follows:

- It was found that there are certain sincere attempts towards modelling based studies where the accuracy of predictive value towards blackspot identification was identified.
- None of the studies were found to adopt technological advancement of data processing or storage through the cloud or equivalent large network-based resources.
- Cloud and Internet-of-services are increasingly penetrating in almost all sectors offering better automation and supportability of real-time processing.
- IoT has some potential analytics that can offer better identification of blackspot. However, existing approaches and real-time practices are yet not ready for such an adaptation.

Therefore, the future direction of the work will be to design an architectural and predictive model for accident analysis for assessing the possible occurrences of black spots for highly complex transportation data. The initial emphasis should be on developing a novel and straightforward computational model for converting complex transportation data to highly-structured data. This will be followed by analyzing the transformed road accident data associated with the selected black spot and addressing the uncertainty problem related to it. Finally, a novel predictive model will be constructed for precise accident analysis and formulate a useful inference for mitigating accidents on black spots and to validate the same.

A closer look into the future work being presented above will show that they are mainly targeted to handle the issues related to blackspot management. The first implementation is completely focused on addressing the complex form of accident data of black spot that is also highly heterogeneous, which is one of the most critical preliminary stages of normalizing accidental events. The second implementation of the study emphasizes data analysis by addressing the ambiguity of black spot detection, which will solve real-time modeling of continually growing data and make a significant decision out of it. The third implementation is focused on incorporating a cost-effective intellectual model to carry out the precise classification of the black spots for the given accident data. This study will formulate a reliable and cost-effective prediction model for extracting more valuable features in accident data. Therefore, the proposed study's core aim is - for a given stream of incoming real-time accident data of black spot, the system could perform a continual assessment of decision and predict if the decisions are correct or error-prone towards resisting accidents. The proposed system's anticipated outcome is to offer fast processing of the complex stream of accident data to perform highly precise accident analysis. The outcome is also expected to offer information about geographical and geometric approaches and techniques to be engineered for mitigating accidents on black-spots.

CONFLICTS OF INTEREST

" The authors do not have any conflict of interest with other entities or researchers".

Reference

[1] Road Transport Accidents. (2019). *https://www.nhp.gov.in/*. Available: https://ncrb.gov.in/sites/default/files/Chapter-1A-Traffic-Accidents_2019.pdf
[2] J. Lefler, *The Black Spot*: Xlibris, 2017.
[3] G. Gartner and H. Huang, *Progress in Location-Based Services 2016*: Springer, 2017.
[4] X. ZOU, W. L. YUE, and C. LI, "Road Traffic Accident Prediction using an SCGM (1, 1) c-Markov Model," *Asian Transport Studies,* vol. 5, pp. 191-205, 2018.
[5] B. Strnad, "Road Safety Audit and Road Safety Inspection on the TEM network," 2018.
[6] W. Kustra, K. Jamroz, and M. Budzynski, "Safety PL–a support tool for road safety impact assessment," *Transportation research procedia,* vol. 14, pp. 3456-3465, 2016.
[7] D. Shinar, "Crash causes, countermeasures, and safety policy implications," *Accident Analysis & Prevention,* vol. 125, pp. 224-231, 2019.

[8] E. Papadimitriou, A. Filtness, A. Theofilatos, A. Ziakopoulos, C. Quigley, and G. Yannis, "Review and ranking of crash risk factors related to the road infrastructure," *Accident Analysis & Prevention,* vol. 125, pp. 85-97, 2019.

[9] J. J. Rolison, S. Regev, S. Moutari, and A. Feeney, "What are the factors that contribute to road accidents? An assessment of law enforcement views, ordinary drivers' opinions, and road accident records," *Accident Analysis & Prevention,* vol. 115, pp. 11-24, 2018.

[10] C. Ahlström, A. Anund, C. Fors, and T. Åkerstedt, "Effects of the road environment on the development of driver sleepiness in young male drivers," *Accident Analysis & Prevention,* vol. 112, pp. 127-134, 2018.

[11] P. K. Alvaro, N. M. Burnett, G. A. Kennedy, W. Y. X. Min, M. McMahon, M. Barnes*, et al.*, "Driver education: Enhancing knowledge of sleep, fatigue and risky behaviour to improve decision making in young drivers," *Accident Analysis & Prevention,* vol. 112, pp. 77-83, 2018.

[12] R. Ando, K. Higuchi, and Y. Mimura, "Data analysis on traffic accident and urban crime: a case study in Toyota City," *International journal of transportation science and technology,* vol. 7, pp. 103-113, 2018.

[13] P. T. Savolainen, F. L. Mannering, D. Lord, and M. A. Quddus, "The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives," *Accident Analysis & Prevention,* vol. 43, pp. 1666-1676, 2011.

[14] H. Lu, S. Huang, Y. Li, and Y. Yang, "Panel data analysis via variable selection and subject clustering," *Data Mining for Service,* pp. 61-76, 2014.

[15] P. Lu, S. Chen, and Y. Zheng, "Artificial intelligence in civil engineering, mathematical problems in engineering," *F. Kang, Hindawi Publishing Corporation,* vol. 2012, 2012.

[16] P. Ekapun and T. Y. Pang, "Design and performance analysis of an electromagnetic tricycle operated in an airport," *Procedia Engineering,* vol. 99, pp. 1330-1338, 2015.

[17] P. Underwood and P. Waterson, "Accident analysis models and methods: guidance for safety professionals," *Loughborough University,* 2013.

[18] C. Gutierrez-Osorio and C. Pedraza, "Modern data sources and techniques for analysis and forecast of road accidents: A review," *Journal of traffic and transportation engineering (English edition),* vol. 7, pp. 432-446, 2020.

[19] P. B. Silva, M. Andrade, and S. Ferreira, "Machine learning applied to road safety modeling: A systematic literature review," *Journal of traffic and transportation engineering (English edition),* vol. 7, pp. 775-790, 2020.

[20] F. Mannering, "Temporal instability and the analysis of highway accident data," *Analytic methods in accident research,* vol. 17, pp. 1-13, 2018.

[21] G. Petkov, "Symptom-based context quantification for dynamic accident analysis," *Safety science,* vol. 121, pp. 666-678, 2020.

[22] X. Zhuang and C. Wu, "Display of required crossing speed improves pedestrian judgment of crossing possibility at clearance phase," *Accident Analysis & Prevention,* vol. 112, pp. 15-20, 2018.

[23] P. Intini, N. Berloco, P. Colonna, V. Ranieri, and E. Ryeng, "Exploring the relationships between drivers' familiarity and two-lane rural road accidents. A multi-level study," *Accident Analysis & Prevention,* vol. 111, pp. 280-296, 2018.

[24] C. Liu, M. Zhao, W. Li, and A. Sharma, "Multivariate random parameters zero-inflated negative binomial regression for analyzing urban midblock crashes," *Analytic methods in accident research,* vol. 17, pp. 32-46, 2018.

[25] H. C. Chin and M. Zhou, "A study of at-fault older drivers in light-vehicle crashes in Singapore," *Accident Analysis & Prevention,* vol. 112, pp. 50-55, 2018.

[26] Z. Ma, C. Shao, H. Yue, and S. Ma, "Analysis of the logistic model for accident severity on urban road environment," in *2009 IEEE Intelligent Vehicles Symposium*, 2009, pp. 983-987.

[27] Y. Li, L. Xing, W. Wang, M. Liang, and H. Wang, "Evaluating the impact of Mobike on automobile-involved bicycle crashes at the road network level," *Accident Analysis & Prevention,* vol. 112, pp. 69-76, 2018.

[28] Y. Wang, D. Zhang, Y. Liu, B. Dai, and L. H. Lee, "Enhancing transportation systems via deep learning: A survey," *Transportation research part C: emerging technologies,* vol. 99, pp. 144-163, 2019.

[29] S. A. Gargoum, M. H. Tawfeek, K. El-Basyouny, and J. C. Koch, "Available sight distance on existing highways: Meeting stopping sight distance requirements of an aging population," *Accident Analysis & Prevention,* vol. 112, pp. 56-68, 2018.

[30] G. Mridula, "Traffic Accident Analysis and Mitigation Measures at Kariyad (NH-544), Ernakulam, Kerala," 2016.

[31] F. Chen and S. Chen, "Injury severities of truck drivers in single-and multi-vehicle accidents on rural highways," *Accident Analysis & Prevention,* vol. 43, pp. 1677-1688, 2011.

[32] J. S. Hickman, R. J. Hanowski, and J. Bocanegra, "A synthetic approach to compare the large truck crash causation study and naturalistic driving data," *Accident Analysis & Prevention,* vol. 112, pp. 11-14, 2018.

[33] N. V. Malyshkina and F. L. Mannering, "Empirical assessment of the impact of highway design exceptions on the frequency and severity of vehicle accidents," *Accident Analysis & Prevention,* vol. 42, pp. 131-139, 2010.

[34] A. Bener, T. Özkan, and T. Lajunen, "The driver behaviour questionnaire in Arab gulf countries: Qatar and United Arab Emirates," *Accident Analysis & Prevention,* vol. 40, pp. 1411-1417, 2008.

[35] R. Squillante Jr, D. J. S. Fo, N. Maruyama, F. Junqueira, L. A. Moscato, F. Y. Nakamoto*, et al.*, "Modeling accident scenarios from databases with missing data: A probabilistic approach for safety-related systems design," *Safety science,* vol. 104, pp. 119-134, 2018.

[36] J. L. Sullivan, D. C. Novak, L. Aultman-Hall, and D. M. Scott, "Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: A link-based capacity-reduction approach," *Transportation Research Part A: Policy and Practice,* vol. 44, pp. 323-336, 2010.

[37] S. Chand and V. V. Dixit, "Application of Fractal theory for crash rate prediction: Insights from random parameters and latent class tobit models," *Accident Analysis & Prevention,* vol. 112, pp. 30-38, 2018.

[38] G. Fountas, M. T. Sarwar, P. C. Anastasopoulos, A. Blatt, and K. Majka, "Analysis of stationary and dynamic factors affecting highway accident occurrence: a dynamic correlated grouped random parameters binary logit approach," *Accident Analysis & Prevention,* vol. 113, pp. 330-340, 2018.

[39] J. Lee, J. Chae, T. Yoon, and H. Yang, "Traffic accident severity analysis with rain-related factors using structural equation modeling–A case study of Seoul City," *Accident Analysis & Prevention,* vol. 112, pp. 1-10, 2018.

[40] A. Egbendewe-Mondzozo, L. M. Higgins, and W. D. Shaw, "Red-light cameras at intersections: Estimating preferences using a stated choice model," *Transportation research part A: policy and practice,* vol. 44, pp. 281-290, 2010.

[41] T. Xu, L. Cheng, and Z. Chen, "Traffic Incident Detection Algorithm Based on Wavelet Support Vector Machine," in *ICCTP 2011: Towards Sustainable Transportation Systems*, ed, 2011, pp. 541-554.

[42] Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang, "A hybrid deep learning based traffic flow prediction method and its understanding," *Transportation Research Part C: Emerging Technologies,* vol. 90, pp. 166-180, 2018.

[43] J. Zhang, V. Kwigizile, and J.-S. Oh, "Automated hazardous action classification using natural language processing and machine-learning techniques," in *CICTP 2016*, ed, 2016, pp. 1579-1590.

[44] Z. Zhang, Q. He, J. Gao, and M. Ni, "A deep learning approach for detecting traffic accidents from social media data," *Transportation research part C: emerging technologies,* vol. 86, pp. 580-596, 2018.

[45] C. C. Ihueze and U. O. Onwurah, "Road traffic accidents prediction modelling: An analysis of Anambra State, Nigeria," *Accident analysis & prevention,* vol. 112, pp. 21-29, 2018.

[46] H. Park, A. Haghani, S. Samuel, and M. A. Knodler, "Real-time prediction and avoidance of secondary crashes under unexpected traffic congestion," *Accident Analysis & Prevention,* vol. 112, pp. 39-49, 2018.

[47] D. Robb and T. Barnes, "Accident rates and the impact of daylight saving time transitions," *Accident Analysis & Prevention,* vol. 111, pp. 193-201, 2018.

[48] N. Kargah-Ostadi, "Comparison of machine learning techniques for developing performance prediction models," in *Computing in Civil and Building Engineering (2014)*, ed, 2014, pp. 1222-1229.

[49] A. Dhamaniya, "Development of accident prediction model under mixed traffic conditions: a case study," in *Urban Public Transportation Systems 2013*, ed, 2013, pp. 124-135.

[50] R. Goel, "Modelling of road traffic fatalities in India," *Accident Analysis & Prevention,* vol. 112, pp. 105-115, 2018.

[51] C. Ding, P. Chen, and J. Jiao, "Non-linear effects of the built environment on automobile-involved pedestrian crash frequency: A machine learning approach," *Accident Analysis & Prevention,* vol. 112, pp. 116-126, 2018.

[52] G. A. Shafabakhsh, A. Famili, and M. S. Bahadori, "GIS-based spatial analysis of urban traffic accidents: Case study in Mashhad, Iran," *Journal of traffic and transportation engineering (English edition),* vol. 4, pp. 290-299, 2017.

[53] L. Çela, S. Shiode, and K. Lipovac, "Integrating GIS and spatial analytical techniques in an analysis of road traffic accidents in Serbia," *International Journal for Traffic and Transport Engineering,* vol. 3, pp. 1-15, 2013.

[54] L. Wang, R. Li, C. Wang, and Z. Liu, "Driver injury severity analysis of crashes in a western China's rural mountainous county: Taking crash compatibility difference into consideration," *Journal of traffic and transportation engineering (English edition),* vol. 8, pp. 703-714, 2021.

[55] Y. Gu, D. Liu, R. Arvin, A. J. Khattak, and L. D. Han, "Predicting intersection crash frequency using connected vehicle data: A framework for geographical random forest," *Accident Analysis & Prevention,* vol. 179, p. 106880, 2023.

[56] S. Ahmed, M. A. Hossain, S. K. Ray, M. M. I. Bhuiyan, and S. R. Sabuj, "A study on road accident prediction and contributing factors using explainable machine learning models: analysis and performance," *Transportation Research Interdisciplinary Perspectives,* vol. 19, p. 100814, 2023.

[57] W. Cheng, G. S. Gill, Y. Zhang, and Z. Cao, "Bayesian spatiotemporal crash frequency models with mixture components for space-time interactions," *Accident Analysis & Prevention,* vol. 112, pp. 84-93, 2018.

[58] K. Stylianou, L. Dimitriou, and M. Abdel-Aty, "Big data and road safety: A comprehensive review," *Mobility Patterns, Big Data and Transport Analytics,* pp. 297-343, 2019.