Intelligent Analysis of Vehicle Accidents to Detect Road Safety: A Systematic Literature Review

Wei Sun

Computer Vision, Faculty of Computer Science and Information Technology
Universiti Putra Malaysia, Serdang, Malaysia
Email: gs66334@student.upm.edu.my

Lili Nurliyana Abdullah

Department of Multimedia, Faculty of Computer Science and Information Technology
Universiti Putra Malaysia, Serdang, Malaysia
Email: liyana@upm.edu.my
Corresponding Author

Fatimah binti Khalid

Department of Multimedia, Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
Email: fatimahk@upm.edu.my

Puteri Suhaiza binti Sulaiman

Department of Multimedia, Faculty of Computer Science and Information Technology
Universiti Putra Malaysia, Serdang, Malaysia
Email: psuhaiza@upm.edu.my

Abstract

Current road accident research focuses mainly on the key role and importance of Artificial Intelligence (AI) in road accident analysis and prevention. After reviewing the literature, this study found that AI had a wide range of potential applications. It can analyse traffic data more accurately and quickly through advanced machine learning and deep learning technologies, and identify accident risks and dangerous driving behaviours, thereby helping to predict and avoid accidents. Furthermore, the research provides insight into the different types of traffic accidents and the severity of injuries they cause, highlighting the importance of understanding these differences to improve road safety and help inform decision-making. The paper has attempted to develop a comprehensive and diverse road crash impact model by exploring some of the elements that influence road crashes, including human factors, vehicle factors and road environment factors. Finally, this paper identifies some innovations and future research directions for this study, including addressing imbalances and quality issues in collecting and processing data, improving the interpretability and transparency of injury severity expressions, and adopting a more comprehensive approach to analysing road crashes. These innovations will promote greater theoretical and practical progress in road crash research to improve road safety and reduce the damage caused by accidents.

Keywords—Artificial Intelligence (Ai), Traffic Accident, Accident Prevention, Deep Learning, Severity Of Injury, Multi-Factor Analysis

1. Introduction

Traffic accidents have become a major public health problem globally, causing a large number of deaths and injuries every year. According to the Global Status Report on Road Safety 2018 published by the World Health Organization (WHO), around 1.35 million people die in road crashes globally each year, with traffic injuries becoming the leading cause of death for people aged 5 to 29 years old reason for death (WHO, 2018). In addition, it is estimated that between 20 and 50 million people are injured in road crashes each year but survive and may suffer long-term adverse health effects (Organization et al., 2017). These statistics reveal the enormous global social and economic burden of road crashes, particularly in low and middle income countries.

As the global population increases and motorization levels rise, the number of traffic accidents is on a steady upward trend (WHO, 2018). Shockingly, more than 335,800 individuals within this age group lose their lives on the world's road- ways each year, with roughly 75 percent of them being men (Krug, 2012). Traffic accidents not only have a serious impact on victims and their families, but also place a heavy economic burden on communities and countries. The economic costs of traffic accidents include medical expenses, rehabilitation costs, lost labour and damage to roads and vehicles. To address this global problem, effective measures are needed to prevent traffic accidents and improve road safety, thereby reducing the negative impact of traffic accidents on people's lives.

Predicting and preventing traffic accidents is key to improving road safety. By analysing data on traffic accidents, it is possible to identify potential causes and risk factors for accidents. With the development of big data and computing power, it is now possible to use advanced data analytics to identify and predict traffic accidents in a more accurate and timely manner. To better understand and address the problem of traffic crashes, researchers and policy makers are increasingly relying on data analysis and advanced technologies. By collecting and analysing traffic accident data, we can better understand the causes of traffic accidents, identify high-risk areas, formulate effective traffic safety policies, and design safer transport

systems and infrastructure. In addition, intelligent analyses, and the application of advanced technologies, such as machine learning and artificial intelligence, offer new possibilities for traffic safety, enabling us to predict and prevent traffic accidents more accurately.

The aim of this paper is to provide an insight into how intelligent analytics can be applied to the processing of vehicle accident data to improve road safety through a systematic review of the literature. With the development of Machine Learning (ML), Artificial Intelligence (AI) and advanced statistical methods, intelligent analytics have opened up new possibilities for the analysis of traffic accident data. These techniques are capable of processing large amounts of data, identifying complex patterns, and providing real-time insights for traffic safety improvements. At the core of this review is an exploration of how smart analytics can improve road safety, particularly through the application of machine learning, AI and advanced statistical methods to the analysis of vehicle crash data, and their potential impact on crash detection and prevention.

In the further discussion of this study, the discussion was carried out according to six main chapters. Chapter 2 provided an in-depth discussion of the specific research methods and methodological framework used for this study, emphasizing the accuracy and reliability of the research. Chapter 3 provided a detailed analysis of traditional accident data analysis methods. The paper discussed the traditional processing and analysis methods of accident data and conduct an in-depth discussion of their advantages and disadvantages. Entering Chapter 4, this article focused on the practical application and potential value of artificial intelligence technology in accident injury level analysis. Chapter 5 explored in depth the various methods and techniques mentioned in previous chapters, and conducts detailed analysis and discussion in multiple dimensions. Finally, the main findings of the study were summarized and suggestions and prospects for future research directions are provided. Overall, this article aims to provide new insights and development directions for the field of accident injury level analysis through systematic analysis and discussion.

2. Methodological Approach
2.1 Strategies for literature search and selection

A clear literature search and selection strategy was developed to systematically review and understand how intelligent analysis can be applied to vehicle crash data to improve road safety. First, a set of keywords and phrases related to the topic were identified, such as 'traffic crash analysis', 'machine learning and road safety', 'application of artificial intelligence in traffic crash prediction', etc. Several databases and online resources, including PubMed, ScienceDirect, IEEE Xplore, and Google Scholar, were used for the literature search. The study also included a review of the top journals and conference proceedings related to traffic safety and data analytics. The search was primarily limited to literature published within the last five years.

During the initial stages of implementing the research project, a precise set of literature selection criteria was implemented with the aim of ensuring the relevance and quality of the research content. As table 1, this approach firstly ensured that the selected literature was directly relevant to the research topic vehicle crash analysis, road safety and its intelligent analytics applications and included in-depth exploration of vehicle crash data analysis, such as crash causation analysis and risk assessment. Second, we focus on literature published in the last decade to ensure that what is being studied is closely linked to current technologies, methods and issues, and that the research is real-time and relevant. Thirdly, we tend to select research published in journals or conferences with a strong reputation in the

field of traffic safety, intelligent systems or data analytics, with a view to relying on the rigorous academic review criteria of these platforms to ensure the high quality of the research cited. Finally, we emphasize the selection of studies that provide sufficient data and analysis to ensure that their results and recommendations are well supported, and pay attention to their data analysis methodology and conclusion extraction process to ensure that the studies are scientifically sound and accurate.

TABLE I. LITERATURE SEARCH AND EVALUATION STRATEGIES FOR INTELLIGENT ANALYSIS AP- PLIED TO VEHICLE CRASH DATA TO IMPROVE ROAD SAFETY.

Key points and steps	Notes and description
Literature search strategy	Keyword definition Literature search tools Top journals and conference papers Timeliness
Literature selection criteria	Direct correlation Data depth Journal/conference reputation Data and analytical Integrity Scientific of methods and conclusions
Evaluation and screening	Check documentation, select criteria Eliminate non-compliance with documents and quality requirements Preliminary assessment, content, methods, accuracy of conclusions
Synthesis and analysis	Read in detail and extract key information Analyse topics and identify research gaps Article review, providing re search overview

2.2 Data extraction and synthesis

Data extraction is a central component of the systematic review and is concerned with obtaining key information and data from selected literature. Through this process, we were able to gain a multi-dimensional understanding of how intelligent analyses can be applied to vehicle accident data processing and how they can contribute to improving road safety.

During the data extraction phase, we systematically focused on several key aspects to gain a deeper understanding of the content and value of each research article. Firstly, the authors of the articles and the year of publication were recorded, which helped us to understand the historical context of the research and the trends in the field of study. Next,

we delved into the research objectives and key findings of each article, which not only presented a clear picture of the focus of the research and the conclusions drawn, but also assisted us in assessing the contribution and impact of each study on the road safety field. We then analyse the data analysis methods and techniques used in each article, which helps us to understand the application and efficacy of different techniques in traffic accident analysis and provides us with new methods and techniques that may need to be considered in data analysis. Finally, exploring the practical applications and impacts of the research findings and understanding how the research findings have been applied in practice and the actual impacts they have had on road safety helps us to assess the usefulness and effectiveness of the different methods. Overall, our aim at this stage was to gain an in-depth understanding and documentation of all aspects of the research in order to accurately understand and assess the value and range of applications of each study.

TABLE II.   KEY READING POINTS FOR LITERATURE ON INTELLIGENT ANALYSIS IN TRAFFIC ACCIDENT DATA PROCESSING.

Stage	Precautions
Data Extraction	Author, year, background Goals, findings, impact Methods, techniques, effects Application, practical impact Understanding, value, scope
Data Synthesis	Data classification Theme analysis Cross-comparison Comprehensive Result presentation
Conclusion	Data extraction, intelligent analysis Application, impact Assessment, road safety

After data extraction, we performed data synthesis with the goal of understanding at a macro level how intelligent analyses can be applied to the processing of vehicle accident data and how these analyses can help improve road safety. Firstly, we classified and analysed the extracted data into themes, such as accident causation analysis, risk assessment, accident prediction and technology application, through thematic categorization and analyses, and conducted in-depth analyses of each theme in order to understand the strengths, weaknesses and scope of application of the different approaches and technologies. Secondly, in the cross-comparison and synthesis phase, we cross-referenced the methods and findings of different studies to identify commonalities and differences between them and synthesized the key findings of different studies to develop a comprehensive understanding of the application of intelligent analytics in traffic accident analysis. Finally, in the presentation and discussion phase of the results, we will demonstrate the application of intelligent analytics in traffic accident analysis based on the thematic categorization and analysis results. By discussing the application of different methods and techniques, we will explore their possible impact and value in improving road safety.

The process of data extraction and synthesis allows us to not only gain a comprehensive understanding of the use of intelligent analytics in the processing of traffic

accident data, but also a clear understanding of the effectiveness and impact of these applications. This provides us with valuable information to assess and recommend how intelligent analytics can be better utilized to improve road safety.

2.3. Evaluation guidelines for the methods

In order to comprehensively and objectively assess the efficacy of different intelligent analytics methods in crash data processing and their potential contribution to enhancing road safety, we have developed a set of clear and comprehensive assessment guidelines. These assessment guidelines include:

1) Accuracy - the methods are required to accurately identify and predict the risks and impacts of traffic crashes, with accurate prediction and identification providing a reliable basis for policy formulation; 2) Real-time - the focus is on whether the methods can provide real-time or near real-time traffic safety insights and early warnings, so that timely preventive measures can be taken to reduce the occurrence of accidents; 3) Interpretability - the ease of understanding and interpretability of the results, which should be clear and straightforward, and able to provide clear insights and recommendations for policy development and public understanding; 4) Scalability - assesses whether the method can handle large amounts of data and can be applied to different regions and conditions; 5) Privacy Protection - assesses whether the method considers data protection and individual privacy in data collection and analysis. Overall, these criteria ensure the reliability and efficacy of the method in identifying and predicting traffic accident risks, providing real- time insights, clarity of results, flexibility to cope with different data and scenarios, and privacy security during the analysis process.

By applying the above assessment guidelines, we can perform an in-depth comparison and evaluation of different intelligent analysis methods presented in the selected literature. This would not only help us understand the potential value of these methods in improving road safety, but also identify the strengths and limitations of each method, providing valuable insights and recommendations for future research and practice. At the same time, these guidelines also provide us with a framework to objectively, systematically and comprehensively evaluate intelligent analytics methods, which helps us to fully understand the effectiveness and application potential of these methods from multiple dimensions.

3. Traditional accident data analysis methods

3.1 Supervised

3.1.1 Naive bayes classifier

The paper's contribution is to provide a thorough mathematical explanation of the fundamentals of naive Bayes learning and the Bayes theorem (Berrar, 2018). A unique method was created by Narayan et al. to train a Naive Bayes classifier with privacy guarantee in a local environment in order to overcome the shortcomings of earlier schemes (Narayan, 2021). According to Xue et al., a field programmable gate array-based real-time hardware naive Bayes classifier (NBC) is an effective method (FPGA) (Xue et al., 2019).

3.1.2 Logistic regression

The study's goal was to compare the effectiveness of standardised and unstandardised estimates of the indirect effect and the proportion mediated using multiple regression, structural equation modelling, and the potential outcomes framework for mediation models with a binary outcome, according to Rijnhart et al. (Rijnhart et al., 2019). A secure logistic

regression training protocol and its implementation were published by Cock et al., along with a novel subprotocol for securely computing the activation function (De Cock et al., 2021). In a population- based cohort research involving Asian adults (n = 6,762), analyses were carried out by Nusinovici et al (Nusinovici et al., 2020).

3.1.3 Linear regression

The study by Kayes et al. sought to investigate the relationships between meteorological variables such as daily mean temperature (oC), relative humidity (percent), and rainfall (mm), as well as the concentration of specific air pollutants (SO2, CO, NOx, O3, PM2.5, and PM10) from January 2013 to December 2017 (Kayes et al., 2019). Demand forecasting in pharmaceutical supply chains: a case study was the main topic of Merkuryeva et al. Modern techniques and significant difficulties in demand forecasting for the pharmaceutical sector are explored. Numerous machine learning regression techniques have already been used, including neural networks, support vector regression, fuzzy logic, k closest neighbours' regression, multivariate adaptive regression spline, and random forest (Merkuryeva et al., 2019).

3.1.4 K-Nearest neighbor

K-nearest neighbourhood rough sets is a novel neighbourhood rough set approach Wang et al. presented (Wang et al., 2019). Gou et al. presented two locality limited representation-based k-nearest neighbour rules to address the problems and enhance the performance of KNN-based classification (Gou et al., 2019).

3.2 Unsupervised

3.2.1 K-Means

Sinaga et al. pioneered the unsupervised k-means (U-k- means) clustering technique, which automatically identifies the appropriate number of groups without needing setup or parameter selection. The k-means algorithm's inability to handle different data formats is one of its core issues (Sinaga and Yang, 2020). To address these issues, Ahmed et al. provide a well organized and comprehensive summary of the research done on the k-means method (Ahmed et al., 2020). A fair k-means objective and technique were developed by Ghadiri et al. to choose cluster centres with fair costs for various groups. This results in cluster assignments that are easily explained by a small number of feature values (Ghadiri et al., 2021).

3.2.2 Auto-Encoder

3D point-capsule networks were suggested by Zhao et al. as an auto-encoder for processing sparse 3D point clouds while maintaining the spatial arrangements of the input data (Zhao et al., 2019).The latent distribution of an auto-encoder was shaped into any sample prior distribution by Kolouri et al. using the sliced-Wasserstein distance (Kolouri et al., 2018). With few target training samples, He et al. examined deep transfer multi- wavelet auto-encoding for intelligent defect diagnostics of gear- boxes. Deep transfer multi-wavelet auto-encoder, an unique method for defect diagnostics of gearboxes with few training samples, is described (He et al., 2020).

3.3 Transfer learning

3.3.1 Sample based migration

Based on the first whole-season sample set from 2015 and all accessible archived Landsat 5 TM pictures on the Google Earth Engine cloud-based platform, Huang et al. created an automated training sample migration approach (Huang et al., 2020). With the aid of an X-ray computed tomography (X-CT) scanner, core flooding tests were carried out using an unique sandstone with several thin clay interlayers (Xu et al., 2021). Ullmann et al. devised a

significance criterion based on a statistical hypothesis test for the assessment of outside impacts on a sample in comparison to a reference sample (Ullmann et al., 2022).

3.3.2 Feature-based migration

Size dependency of grain boundary migration in metals under mechanical loading was examined by Zhou et al. In a wide size range, grain size dependences of GB migration in nanograined Ag, Cu, and Ni under tension were quantitatively examined (Zhou et al., 2019). A deep learning model and intrusion detection technologies are combined in Li et al's proposed IoT feature extraction and intrusion detection technique for intelligent cities (Li et al., 2019). A statistical hypothesis test-based significance criterion was developed by Ullmann et al. to evaluate external influences on a sample in comparison to a reference sample (Veeramuthu et al., 2022).

There are important research gaps in traffic accident safety research, and traffic safety data are highly imbalanced. This imbalance is manifested in the fact that the amount of data for some traffic accident categories far exceeds that of other categories, resulting in uneven distribution of categories in the data set. This imbalanced data distribution can negatively impact the results of modeling and analysis, as models tend to predict a higher number of classes and ignore a smaller number of classes.

Due to the imbalance of traffic accident safety data, existing technologies and methods have limitations in solving this problem. Imbalanced data can cause models to be biased in prediction and classification, potentially missing important features and trends in a minority of classes. Therefore, research is necessary to develop new methods and techniques to deal with imbalanced traffic safety data and ensure that the model can be accurate and robust for each category.

Solving this research problem will help improve the analysis and prediction accuracy of traffic accident safety data and provide a more reliable basis for improving traffic safety. This may involve new data balancing techniques, model tuning methods, or algorithm development specific to imbalanced data. By overcoming the challenges caused by imbalanced data, the occurrence and influencing factors of traffic accidents can be more comprehensively understood, thereby better preventing and reducing the occurrence of traffic accidents.

4. Artificial Intelligence In Accident Injury level analysis

4.1 Traffic safety

For the purpose of arranging the transit of sensitive goods, Semenov et al. developed an integrated approach to information analysis. An integrated method of traffic information analysis is presented in this article (Semenov et al., 2019). The purpose of the work is to conduct a critical assessment of the literature on various spatial approaches used by researchers to handle the dimension of space in its many manifestations in their research and analysis (Ziakopoulos and Yannis, 2020). Tao et al. investigated the probabilistic safety assessment approach for vehicle transportation of spent nuclear fuel. It is suggested to use the Probabilistic Safety Assessment (PSA) technique for SNF road transportation Tao et al. (2020). The purpose of the research was to discover prevalent road safety measures globally because there isn't a clear, complete analysis of countries' present and effective approaches to road safety in the world (Safarpour et al., 2020).

4.2 Type of traffic accidents

4.2.1 Fatal

In this study, Jeddah, a car-dependent city in Saudi Arabia, will be used as the study's case study to examine the regionally variable association between pedestrian deaths and other

associated urban environment characteristics (Aljoufie and Tiwari, 2021). The safety impact cannot be directly evaluated because CAVs are still in the development stage. Finland- based driver-managed passenger car accidents that resulted in fatalities were analysed qualitatively to determine the traits that CAVs must have to prevent these accidents (Utriainen and Po¨lla¨nen, 2021). The purpose of the research is to calculate the possible safety advantages of autonomous emergency steering (AES) and emergency lane keeping (ELK) systems with accurate and dependable lateral placement (Sternlund, 2021). In order to lower the number of pedestrian fatalities in traffic accidents, Yang et al. built the foundation for using optimum ma- chine learning approaches. There are few studies that include all the various factors that contribute to fatal collisions in older drivers and how these factors change over time (Yang et al., 2022).

4.2.2 Serious

The goal of the study was to examine how serious vascular injuries from traffic accidents affect the overall population, including both patients who pass away at the scene and those who survive and go to the hospital. To explain the injuries in incidents of automobile accidents that were reported to the Combined Military Hospital (CMH), Dhaka (Tjahjono et al., 2019). The accuracy of four models—radial basis function (RBF)-SVM, linear support vector machine (SVM), random forest (RF), and logistic regression (LR)—was examined in relation to mortality from traffic related injuries (Roshanfekr et al., 2020).

4.2.3 Light

Anambra State, South Eastern Nigeria, road traffic accident patterns and causes were described by Anebonam et al. (Anebonam et al., 2019). In a study conducted by Andrade et al., the number of victims of traffic accidents on federal highways in Brazil was compared before and after the decade of action for road safety. The number of fatalities, serious injuries, and minor injuries were calculated each month using the PraisWinsten technique (MPC) (Andrade and Antunes, 2019). Due of this, Qureshi et al used information from the Statewide Traffic Accident Records System to carry out an extensive analysis (Qureshi et al., 2020). Yuan et al. suggested an effective method and model for identifying imbalanced Chinese traffic accident texts based on severity since the number of minor traffic accidents on a daily basis is significantly larger than that of serious traffic accidents (Yuan and Wang, 2022).

In existing traffic accident research, although various factors and topics related to injury severity have been addressed, the interpretability and transparency of injury severity are still insufficient. Specifically, many studies focus on statistical and model analysis without fully explaining or clarifying the specific reasons or contextual situations behind certain results. This lack of depth and detail in explanations can make it difficult for policy- makers, traffic engineers, and the public to gain a deep understanding of the complexity and multifaceted contributing factors of accidents. Therefore, there is a significant research gap that requires more in-depth research into the contextual causes, situational and other potential factors of injury severity in traffic crashes to improve the interpretability and transparency of research. This deep insight can provide more precise and effective guidance for traffic safety interventions to better protect the safety of road users.

4.3 Factors of traffic accident

4.3.1 Individual factor

On a data set of road traffic accidents obtained from UK road traffic accidents of the year 2016, Kumeda et al. applied various machine learning classification algorithms and discussed here the six algorithms with the highest accuracy and best classification performances, such

as Fuzzy-FARCHD, Random Forest, Hierarchal LVQ, RBF Network (Radial Basis Function Network), Multilayer Perceptron, and Nave Bayes (Kumeda et al., 2019). With the aid of significant environmental elements of RTAs that can be utilised to develop the prediction model, the classification technique known as the Random Forest algorithm is used to find pertinent patterns and classify the kind of accident severity of various traffic accidents (Ramya et al., 2019).

4.3.2 Road condition

Given the circumstances of the road and the local environment, a Bayesian network model is created to forecast the severity of an accident, should one occur (Liu et al., 2019). The paper's contribution was to examine how various factors affect car accidents that occur at night using information gathered from a Chinese metropolis (Basavaraju et al., 2019). Using an ultrasonic sensor and an image processing technique, Sharma et al. devised an effective road surface monitoring system. One of the elements that could cause accidents is interference with the vehicle's steering due to uneven road surfaces (Sharma et al., 2020). The results were demonstrated by Roobini et al. by using extensive data mining techniques on the Finnish roadways (Roobini et al., 2020).

4.3.3 Weather factor

In order to investigate rearend collision avoidance behaviour in various foggy situations, Shangguan et al. proposed an approach that focused on changes in visibility and road alignment (Shangguan et al., 2020). The impact of inclement weather on the safety of road traffic was investigated by Fior et al. Because CAVs are still being developed, the safety impact cannot be directly evaluated (Fior and Cagliero, 2021). In order to evaluate the safety advantages of ACC in China, including the possible maximum impact and practical impact, Tan et al. used a national-level safety impact evaluation model (Tan et al., 2021). In traffic accident safety research, current focus is mainly on analysing local characteristics of accidents, such as communication between vehicles, individual behaviour, road conditions, and weather factors. However, research often neglects contextual information that incorporates the entire scene, a dimension that includes broader background and situational factors. This comprehensive contextual information can be critical in accurate classification and in-depth analysis of incidents.

Therefore, there is a significant research gap that requires more in-depth research on how to consider the contextual in- formation of the entire traffic accident scenario to improve the research precision and prediction accuracy of accident safety. This may involve integrating different data sources and information types, including communication between vehicles, individual behaviour, road conditions and weather factors, to create a more comprehensive analytical framework. Through this comprehensive approach, we can better understand the complexity of accidents and provide stronger evidence and strategies for traffic safety improvements. The solution to this research question will contribute to a more comprehensive and comprehensive understanding of the occurrence and influencing factors of traffic accidents, thereby providing more effective methods and strategies for preventing accidents and improving road safety.

5. Discussion and Analysis

In the previous four sections, various applications and practices of AI in traffic safety were explored through an in-depth literature review. In this section, these core findings are summarized and reviewed, and their implications for existing theory and practice are explored in depth.

5.1 Main findings

In describing the role of Artificial Intelligence (AI) in the field of traffic accident prevention and management, a dual role emerges. Primarily, accident prevention is a notable goal. The previous survey illustrates the ability of AI to predict and avoid traffic accidents by scrutinizing traffic dynamics, driver behaviour and environmental components. Using machine learning algorithms, large amounts of traffic data are dissected to identify risky behaviours and accident-prone conditions. In addition to prevention, AI makes a significant contribution to accident response and coordination. Automatically detecting the occurrence of an accident and then quickly forwarding critical information to the relevant entities can improve the efficiency of response while minimizing the subsequent harm caused by the accident.

In the area of accident type recognition, AI can accurately identify and classify different accident types, such as rearend and side-impact collisions, through analytical lenses trained on traffic video and sensing data. Meanwhile, AI paradigms generated from accident data and environmental indicators can provide predictions of potential injury severity, thus providing indispensable prognostic insights for emergency response and medical mediation.

The literature cited reveals a multitude of factors that influence accidents, most notably driver-centered factors. Numerous studies have highlighted the link between driver behaviour and state (e.g. fatigue, distraction) and accident occurrence, demonstrating how AI can mitigate these risks through real-time monitoring and timely alerts. Environmental variables also fall under the purview of AI, which analyses and reacts in real time to the impact of these factors (e.g., road conditions, weather) on traffic safety through automated traffic flow regulation or dissemination of safety recommendations. Furthermore, in terms of traffic mobility, AI enhances continuous vigilance and fine- tuning of traffic movements to ameliorate the risk of accidents associated with traffic congestion.

These pivotal discoveries underscore the variegated value of AI technology in championing traffic safety, encompassing, albeit not confined to, accident prevention, injury alleviation, accident response refinement, and real-time oversight and management of risk elements. Anchored on this bedrock, ensuing discussions will probe deeper into the theoretical and practical ramifications of these findings, propelling a more efficacious harnessing of this acumen and proficiency in forthcoming research and pragmatic endeavours.

5.2 Theoretical and Practical Significance

Studying the application of AI in traffic safety not only enriches the theoretical system in the field of scientific research, but also provides powerful tools and methods for practical operation and decision-making. The following section explores in detail the far-reaching significance of the research findings at the theoretical level and in practical applications.

The theoretical implications clearly indicated that the introduction of AI helped to expand the scope of existing research by providing researchers with a new perspective to understand the multidimensional aspects of traffic safety. It provides new theoretical frameworks and analytical methods in predictive modeling, risk assessment, and behavioural analysis, which in turn help to construct and optimize models. AI technology enables high-precision and real-time analyses of traffic accidents, which significantly contributes to the accuracy and responsiveness of theoretical models of traffic safety. In addition, by establishing new analysis dimensions, AI technology is able to process and analyse multi-source data, providing a more comprehensive and in-depth perspective for theoretical research, thus enriching and broadening the theoretical research horizons in the field of traffic safety.

The practical implications section highlights the value of the multifaceted application of AI technology in the field of traffic safety. Firstly, AI systems significantly improve the efficiency of traffic safety analysis by automating the collection and analysis of data, while the application of technologies such as deep learning further ensures the accuracy of the analysis. Second, the intelligent warning function of AI models can predict potential traffic risks and issue timely warnings, providing drivers and traffic management with valuable time to prevent possible traffic accidents. Meanwhile, AI's real-time traffic flow analysis and prediction capabilities can help dynamically adjust traffic management strategies, such as changing traffic signal cycles, to optimise traffic flow and reduce accidents. In terms of accident response and handling, AI is able to quickly identify and assess the accident scene after an accident occurs and notify the emergency response team in a timely manner, which effectively shortens the response time. By analysing data from the accident scene, AI can also provide scientific rescue advice to the rescue team to ensure the safety and effectiveness of rescue operations. Finally, AI's personalized risk assessment can provide accurate safety advice and early warning information for every driver, while based on AI's analysis results, traffic management authorities can optimize risk management strategies, such as adjusting traffic rules and setting up more traffic safety facilities, in order to achieve more scientific and efficient traffic safety management.

In short, AI plays a crucial role in the research and practice of traffic safety. It not only enriches the theoretical system of traffic safety and provides brand new research perspectives and methods; in practice, AI also greatly improves the efficiency and accuracy of analyses, enhances the ability of accident prevention and response, and provides powerful support for modern traffic safety management. In future research and application, further exploring and utilizing the potential of AI in traffic safety will be a direction of interest.

5.3 Limitations Of The Methodology
In the course of carrying out an in-depth study and research, it is important to recognize that the methodology of this study has some limitations in a number of areas.

5.3.1 Limitations Of Data Acquisition And Processing
The high imbalance of traffic safety data and limitations in data acquisition and processing negatively affect the effectiveness of deep learning models. During the data acquisition phase, fatal accidents are less frequent compared to minor accidents, and this imbalance may lead to a tendency for the model to recognize most classes, while ignoring more critical and rare events. In addition, there are issues with the integrity and quality of the data. For example, data collected may be missing or incomplete due to equipment failures or data transmission problems, which can affect the accuracy of analyses. Varying quality of data from different sources or regions, such as inconsistencies in precision, format and standards, may further affect model performance. Finally, the issue of data representativeness should not be overlooked, as in some cases the data obtained may not be fully representative of the target study population or phenomenon, and there is a problem of sample bias, which may lead to limited generalization ability of the model.

5.3.2 The need for interpretability and transparency of the level of injury
In current traffic accident research, although a wide range of factors and themes related to injury severity have been explored, there is still a lack of interpretability and transparency regarding injury severity. Specifically, many studies have focused primarily on statistical and

modelling analyses without adequately explaining or elucidating the specific causes or contextual scenarios behind certain results, and this lack of depth and detail may make it difficult for policy makers, traffic engineers and the public to gain a deeper understanding of the complexity and multifaceted influencing factors of crashes.

5.3.3 Deepening Traffic Accident Safety Research

In traffic accident safety research, the focus of current studies is mainly on analysing local features of accidents, such as intervehicle communication, individual behaviour, road conditions and weather factors. However, studies usually omit contextual information that integrates the whole scenario, which includes a wider range of contextual and situational factors. Such integrated contextual information may play a crucial role in the accurate classification and in-depth analysis of accidents.

6. Conclusion

In the process of this paper, we have analysed in depth the role and importance of Artificial Intelligence in the field of Traffic Accident Analysis and Prevention, thoroughly examined different types of traffic accidents and the severity of the injuries they produce, as well as assessed the multiple factors that influence traffic accidents.

6.1 The importance of artificial intelligence

We have focused on the role played by artificial intelligence in traffic accident analysis, highlighting its significant benefits in data analysis, accident prediction, simulation and decision support. Through the use of advanced machine learning and deep learning techniques, AI is not only able to analyse accident data more accurately and quickly, identifying accident risks and dangerous driving behaviours, but can also simulate traffic flows under different conditions to some extent to predict potentially dangerous situations.

6.2 Analysis of accident types and injury severity

In exploring the association between accident type and injury severity, we have analysed in depth the injuries in a variety of accident scenarios. Different accident types, such as side impacts, frontal collisions and pedestrian collisions, have different injury characteristics and hazards, which are of guiding value when providing decision support for improving road safety.

6.3 Multiple factors affecting traffic accidents

By exploring a variety of factors (including human factors, vehicle factors, road environment factors, etc.), we have attempted to construct a comprehensive and multivariate traffic accident impact model. By analysing the relationship between these factors and traffic accidents, we realize that the challenge of improving traffic safety lies in the integrated management and coordination of multiple factors to achieve multifaceted and systematic risk management.

6.4 Innovative points and future directions of the study

Combining existing theories and studies, this study presents some new observations and findings in certain aspects. While carefully navigating the intricacies of traffic safety data and accident research, this study candidly acknowledges its methodological limitations, paving the way for future investigations to enhance data collection, improve model transparency, and deepen the research context. A notable limitation lies in the imbalance and quality of data collection and processing, particularly of traffic safety data, which can mask the effectiveness of deep learning models, especially in identifying serious but rare fatal accidents. Additionally, current approaches encounter challenges related to data integrity, source variability, and representational adequacy, which affect model performance and generalization capabilities. Another key reflection revolves around the clear need for enhanced interpretability and transparency in the articulation of injury severity, with future research likely to endeavour to elucidate the specific reasons and contextual underpinnings of the recorded data, thereby providing stakeholders with a more nuanced understanding of the influencing factors. Finally, embarking on a more holistic approach to traffic crash safety research requires moving beyond a primary focus on localized crash characteristics and conducting integrated analyses that encompass a wider range of contextual and situational factors, thereby enhancing accurate analytical results. Traffic accident classification and in-depth analysis. Going forward, these innovations will lead to advances in model robustness, interpretability and wider contextualization in traffic accident research, maintaining theoretical richness and relevance in future investigations and interventions.

Declaration of competing interest

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

Acknowledgment

References

Ahmed, M., Seraj, R., Islam, S.M.S., 2020. The k-means algorithm: A comprehensive survey and performance evaluation. Electronics 9, 1295.

Aljoufie, M., Tiwari, A., 2021. Modeling road safety in car-dependent cities: Case of jeddah city, saudi arabia. Sustainability 13, 1816.

Andrade, F.R.d., Antunes, J.L.F., 2019. Tendêncía do núímero de vícítimas em acidentes de trâ̂nsito nas rodovias federais brasileiras antes e depois da déícada de açã̃o pela seguranç̧a no trâ̂nsito. Cadernos de Saú́de Púíblica 35.

Anebonam, U., Okoli, C., Ossai, P., Ilesanmi, O., Nguku, P., Nsubuga, P., Abubakar, A., Oyemakinde, A., 2019. Trends in road traffic accidents in anambra state, south eastern nigeria: need for targeted sensitization on safe roads. The Pan African Medical Journal 32.

Basavaraju, A., Du, J., Zhou, F., Ji, J., 2019. A machine learning approach to road surface anomaly assessment using smartphone sensors. IEEE Sensors Journal 20, 2635–2647.

Berrar, D., 2018. Bayes' theorem and naive bayes classifier. Encyclopedia of bioinformatics

and computational biology: ABC of bioinformatics 403, 412. De Cock, M., Dowsley, R., Nascimento, A.C., Railsback, D., Shen, J., Todoki,

A., 2021. High performance logistic regression for privacy-preserving genome analysis. BMC Medical Genomics 14, 1–18.

Fior, J., Cagliero, L., 2021. Estimating the incidence of adverse weather effects on road traffic safety using time series embeddings, in: 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), IEEE. pp. 402–407.

Ghadiri, M., Samadi, S., Vempala, S., 2021. Socially fair k-means clustering, in: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, pp. 438–448.

Gou, J., Qiu, W., Yi, Z., Xu, Y., Mao, Q., Zhan, Y., 2019. A local mean representation-based k-nearest neighbor classifier. ACM Transactions on Intelligent Systems and Technology (TIST) 10, 1–25.

He, Z., Shao, H., Wang, P., Lin, J.J., Cheng, J., Yang, Y., 2020. Deep transfer multi-wavelet auto-encoder for intelligent fault diagnosis of gearbox with few target training samples. Knowledge-Based Systems 191, 105313.

Huang, H., Wang, J., Liu, C., Liang, L., Li, C., Gong, P., 2020. The migration of training samples towards dynamic global land cover mapping. ISPRS Journal of Photogrammetry and Remote Sensing 161, 27–36.

Kayes, I., Shahriar, S.A., Hasan, K., Akhter, M., Kabir, M., Salam, M., 2019. The relationships between meteorological parameters and air pollutants in an urban environment. Global Journal of Environmental Science and Management 5, 265–278.

Kolouri, S., Pope, P.E., Martin, C.E., Rohde, G.K., 2018. Sliced wasserstein auto-encoders, in: International Conference on Learning Representations.

Krug, E., 2012. Decade of action for road safety 2011–2020. Injury 43, 6–7. Kumeda, B., Zhang, F., Zhou, F., Hussain, S., Almasri, A., Assefa, M., 2019.

Classification of road traffic accident data using machine learning algo- rithms, in: 2019 IEEE 11th international conference on communication software and networks (ICCSN), IEEE. pp. 682–687.

Li, D., Deng, L., Lee, M., Wang, H., 2019. Iot data feature extraction and intrusion detection system for smart cities based on deep migration learning. International journal of information management 49, 533–545.

Liu, J., Li, J., Wang, K., Zhao, J., Cong, H., He, P., 2019. Exploring factors affecting the severity of night-time vehicle accidents under low illumination conditions. Advances in Mechanical Engineering 11, 1687814019840940.

Merkuryeva, G., Valberga, A., Smirnov, A., 2019. Demand forecasting in pharmaceutical supply chains: A case study. Procedia Computer Science 149, 3–10.

Narayan, Y., 2021. Comparative analysis of svm and naive bayes classifier for the semg signal classification. Materials Today: Proceedings 37, 3241–3245. Nusinovici, S., Tham, Y.C., Yan, M.Y.C., Ting, D.S.W., Li, J., Sabanayagam, C., Wong, T.Y., Cheng, C.Y., 2020. Logistic regression was as good as machine learning for predicting major chronic diseases. Journal of clinical epidemiology 122, 56–69.

Organization, W.H., et al., 2017. Save lives: a road safety technical package . Qureshi, A.I., Huang, W., Khan, S., Lobanova, I., Siddiq, F., Gomez, C.R.,

Suri, M.F.K., 2020. Mandated societal lockdown and road traffic accidents. Accident Analysis & Prevention 146, 105747.

Ramya, S., Reshma, S., Manogna, V.D., Saroja, Y.S., Gandhi, G.S., 2019. Accident severity prediction using data mining methods. International Journal of Scientific Research in

Computer Science, Engineering and Information Technology 5.

Rijnhart, J.J., Twisk, J.W., Eekhout, I., Heymans, M.W., 2019. Comparison of logistic-regression based methods for simple mediation analysis with a dichotomous outcome variable. BMC medical research methodology 19, 1–10.

Roobini, M., Mulakalapally, S., Mungamuri, N., Lakshmi, M., Ponraj, A., Deepa, D., 2020. Car accident detection and notification system using smart- phone. Journal of Computational and Theoretical Nanoscience 17, 3389– 3393.

Roshanfekr, P., Khodaie-Ardakani, M.R., Sajjadi, H., Ardakani, H.M.A., 2020. Income-related inequality in traffic accident health outcomes (injury, dis- ability and mortality): evidence from the nationwide survey in iran. Iranian journal of public health 49, 718.

Safarpour, H., Khorasani Zavareh, D., Mohammadi, R., 2020. The common road safety approaches: A scoping review and thematic analysis. Chinese journal of traumatology 23, 113–121.

Semenov, I., Filina Dawidowicz, L., Trojanowski, P., 2019. Integrated approach to information analysis for planning the transport of sensitive cargo. Archives of Transport 51.

Shangguan, Q., Fu, T., Liu, S., 2020. Investigating rear-end collision avoidance behavior under varied foggy weather conditions: A study using advanced driving simulator and survival analysis. Accident Analysis & Prevention 139, 105499.

Sharma, S.K., Phan, H., Lee, J., 2020. An application study on road surface monitoring using dtw based image processing and ultrasonic sensors. Applied Sciences 10, 4490.

Sinaga, K.P., Yang, M.S., 2020. Unsupervised k-means clustering algorithm.
IEEE access 8, 80716–80727.

Sternlund, S., 2021. The safety potential of enhanced lateral vehicle positioning. Traffic injury prevention 22, 139–146.

Tan, H., Zhao, F., Liu, Z., 2021. Impact of adaptive cruise control (acc) system on fatality and injury reduction in china. Traffic injury prevention 22, 307– 312.

Tao, L., Wang, J., Long, P., Wang, D., Wang, F., Zhou, B., Chen, J., 2020. Probabilistic safety assessment method for spent nuclear fuel road transportation. Annals of Nuclear Energy 137, 107043.

Tjahjono, T., Kusuma, A., Pratiwi, Y.Y., Purnomo, R.Y., 2019. Identification determinant variables of the injury severity crashes at road railway level crossing in indonesia. Transportation research procedia 37, 211–218.

Ullmann, S., Nordsiek, S., Lowke, D., 2022. Significance or scatter—statistical evaluation of rapid chloride migration test results for sprayable cement mortar. Materials 15, 6050.

Utriainen, R., Pöllänen, M., 2021. The needed features of connected and au- tomated vehicles to prevent passenger car crashes caused by driving errors. Future transportation 1, 370–386.

Veeramuthu, A., Meenakshi, S., Mathivanan, G., Kotecha, K., Saini, J.R., Vijayakumar, V., Subramaniyaswamy, V., 2022. Mribrain tumor image classification using a combined feature and image-based classifier. Frontiers in Psychology 13, 848784.

Wang, C., Shi, Y., Fan, X., Shao, M., 2019. Attribute reduction based on k-nearest neighborhood rough sets. International Journal of Approximate Reasoning 106, 18–31.

WHO, V., 2018. Global status report on road safety 2018. World Health Organization .

Xu, L., Li, Q., Myers, M., White, C., Tan, Y., 2021. Experimental and numerical investigation of supercritical co2 migration in sandstone with multiple clay interlayers. International Journal of Greenhouse Gas Control 104, 103194.

Xue, Q., Zhu, Y., Wang, J., 2019. Joint distribution estimation and naïve bayes classification

under local differential privacy. IEEE transactions on emerging topics in computing 9, 2053–2063.

Yang, L., Aghaabbasi, M., Ali, M., Jan, A., Bouallegue, B., Javed, M.F., Salem, N.M., 2022. Comparative analysis of the optimized knn, svm, and ensemble dt models using bayesian optimization for predicting pedestrian fatalities: an advance towards realizing the sustainable safety of pedestrians. Sustainability 14, 10467.

Yuan, S., Wang, Q., 2022. Imbalanced traffic accident text classification based on bert-rcnn, in: Journal of Physics: Conference Series, IOP Publishing. p. 012003.

Zhao, Y., Birdal, T., Deng, H., Tombari, F., 2019. 3d point capsule networks, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 1009–1018.

Zhou, X., Li, X., Lu, K., 2019. Size dependence of grain boundary migration in metals under mechanical loading. Physical review letters 122, 126101.

Ziakopoulos, A., Yannis, G., 2020. A review of spatial approaches in road safety. Accident Analysis & Prevention 135.