Vol 6, Issue 3, March 2026 www.ijesti.com E-ISSN: 2582-9734
International Journal of Engineering, Science, Technology and Innovation (IJESTI)

Application of Machine Learning and AI in Predicting and Managing Road Traffic Accidents

Purshottam Kumar
M. Tech. in Transportation Engineering, CBS Group of Institutions, Jhajjar, Haryana.

Shubender
A.P Civil Department, CBS Group of Institutions, Jhajjar, Haryana.

ABSTRACT

Road traffic accidents (RTAs) represent a significant public safety issue globally. Traditional methods for identifying accident-prone areas (black spots) have been limited, relying on historical data and statistical techniques. Recent advancements in Machine Learning (ML) and Artificial Intelligence (AI) are revolutionizing predictive traffic safety systems. AI-driven methods offer predictive analytics, processing large-scale data sets to identify patterns in traffic flow, environmental factors, and driver behavior. Models like LSTM and Random Forest demonstrate high accuracy in forecasting accident hotspots. The integration of IoT and real-time vehicle data further enhances predictive accuracy, enabling timely interventions for accident prevention.

Keywords: *Road Traffic Accidents, Machine Learning, Artificial Intelligence, Accident Prediction, Black Spots.*

I. INTRODUCTION

Road traffic accidents (RTAs) represent one of the most critical global public safety challenges, causing significant loss of life, economic damage, and social disruption. According to global transportation safety reports, millions of people are injured or killed annually due to road accidents, with developing countries bearing a disproportionate share of this burden. Traditional approaches to road safety, such as reactive interventions and generalized policy measures, have proven insufficient in effectively mitigating accident risks. Consequently, there is a growing emphasis on proactive, data-driven approaches that leverage advanced computational techniques for identifying accident-prone zones, commonly referred to as road accident black spots. A road accident black spot is typically defined as a location with a significantly higher frequency of accidents compared to surrounding areas. Identifying these locations accurately is crucial for targeted interventions such as infrastructure redesign, traffic regulation enforcement, and emergency response planning. Conventional methods for black spot identification often rely on historical accident frequency analysis and statistical techniques. However, these approaches are limited in their ability to capture complex, nonlinear relationships among multiple contributing factors such as traffic volume, weather conditions, road geometry, and driver behavior. In recent years, Machine Learning (ML) and Artificial Intelligence (AI) have emerged as transformative tools in transportation engineering, offering powerful capabilities for predictive modeling and pattern recognition. ML algorithms can process large-scale, heterogeneous datasets and uncover hidden patterns that traditional methods may overlook. As highlighted by Haq et al. (2026), AI-driven transportation systems significantly enhance traffic prediction accuracy, reduce accident frequency, and optimize traffic management through real-time analytics and adaptive decision-making. These advancements have enabled a shift from reactive safety management to proactive and preventive strategies. The integration of predictive analytics into transportation systems has further strengthened the capability to forecast traffic conditions and accident risks. Sharrab et al. (2025) demonstrated the effectiveness of models such as Facebook Prophet, ARIMA, and Long Short-Term Memory (LSTM) networks in predicting traffic flow patterns with high accuracy.

These models provide valuable insights into temporal variations in traffic behavior, enabling authorities to anticipate congestion and implement preventive measures. Such predictive frameworks are essential for identifying potential black spots before accidents occur. Similarly, Rehman et al. (2025) emphasized the importance of data-driven methodologies in predicting accident hotspots. Their study utilized machine learning algorithms including Random Forest, Support Vector Machines (SVM), and Neural Networks to analyze diverse datasets comprising traffic records, environmental factors, and road infrastructure characteristics. The findings indicated that ML models achieved high predictive accuracy and could be effectively used for optimizing patrol deployment and implementing targeted safety interventions.

The emergence of Internet of Vehicles (IoV) and Internet of Things (IoT) technologies has further enhanced the predictive capabilities of ML-based systems. Wei (2024) proposed a deep learning-based framework using Spatio-Temporal Conv-LSTM Autoencoders for real-time accident prediction, achieving an AUROC score of 0.94. This approach integrates real-time data from vehicles, roadside units, and environmental sensors, enabling dynamic risk assessment and timely intervention. In addition, edge computing has been introduced to address latency issues associated with cloud-based systems. Djazia et al. (2023) developed a hybrid cloud-edge architecture that combines deep reinforcement learning and edge-based deep learning models to improve accident prediction accuracy while reducing response time. This approach ensures faster decision-making, which is critical in real-time traffic management scenarios. Despite these technological advancements, understanding the underlying causes of accidents remains essential for effective black spot identification. Gheisari (2022) employed factor analysis and ranking techniques to identify key contributing factors to accidents, such as road conditions, traffic density, and environmental influences. Integrating such insights with ML models can enhance their interpretability and reliability. Machine learning applications in transportation are also supported by advancements in autonomous driving and intelligent systems. Muhammad et al. (2020) highlighted the role of deep learning in improving road safety through applications such as collision avoidance, pedestrian detection, and driver monitoring. These technologies contribute to reducing human error, which is a major cause of road accidents. Furthermore, predictive modeling techniques have been successfully applied to travel time estimation and traffic flow optimization. Qiu and Fan (2021) demonstrated that Random Forest models outperform other algorithms in predicting short-term travel times, thereby aiding in congestion management and route optimization. Efficient traffic flow reduces accident risks by minimizing sudden stops and erratic driving behavior. Deep learning advancements have also expanded the scope of ML applications in transportation systems. Pouyanfar et al. (2018) described deep learning as a powerful tool capable of handling complex data structures and extracting meaningful insights. However, challenges such as model interpretability, data quality, and computational complexity remain significant barriers to widespread adoption.

II. RESEARCH BACKGROUND

Haq et al., (2026) investigated the transformative role of Artificial Intelligence (AI) and Machine Learning (ML) in advancing smart transportation systems, particularly in improving road safety, optimizing traffic flow, and promoting environmental sustainability. The study employed a mixed-method research design, integrating quantitative data from transportation authorities and simulation outputs with qualitative insights from AI and traffic management experts. It was reported that AI-driven models significantly enhanced traffic prediction accuracy, reduced accident frequency, and optimized route management through real-time analytics and adaptive control mechanisms. The analysis also revealed that ML-based systems contributed to lower carbon emissions by facilitating fuel-efficient driving patterns and reducing idle times at intersections. Nevertheless, challenges such as data privacy concerns, lack of standardized regulatory frameworks, and limited public trust in AI-based decision-making were highlighted. The discussion suggested that sustainable implementation depended on collaborative governance, ethical design, and policy integration, while future research was encouraged to explore explainable AI, edge-based optimization, and multi-modal integration to improve system resilience and transparency.

Sharrab et al., (2025) investigated the use of predictive analytics for traffic management and accident prevention in smart cities by employing Facebook Prophet, Long Short-Term Memory (LSTM), and AutoRegressive Integrated Moving Average (ARIMA) models. They collected traffic data from road sensors to forecast future traffic patterns, aiming to enhance road safety and system efficiency. The study examined variations in traffic trends across different times of the day, weeks, and months, and demonstrated that the models effectively predicted traffic volumes, achieving accuracy rates of 86.96%, 83.16%, and 86.64% for Prophet, ARIMA, and LSTM, respectively. Additionally, the models were assessed using metrics such as MAPE, RMSE, and MAE, confirming their predictive reliability. The research highlighted that integrating predictive analytics could shift traffic systems from reactive to proactive management, optimizing flow, reducing congestion, and improving urban mobility. Challenges concerning data quality, real-time processing, and traffic complexity were also addressed, illustrating the promise of AI-driven solutions for infrastructure planning and traffic management.

Rehman et al. (2025) investigated the rising rate of road accidents in urban areas, highlighting it as a significant threat to both citizen safety and the efficiency of traffic systems. They noted that conventional reactive measures, such as post-incident reviews and general safety initiatives, were often insufficient in preventing crashes beforehand. To address this, the study proposed a proactive, data-driven approach utilizing machine learning (ML) techniques aimed at predicting accident-prone locations, or "hotspots." The researchers analyzed historical crash reports, environmental conditions like weather and lighting, temporal patterns, and structural features of road networks to understand their influence on crash occurrences. Publicly available datasets, including traffic incident logs, weather archives, and geospatial road data, were employed to train and validate ML models such as Random Forests, Support Vector Machines, and Neural Networks. The predictive performance of these models was assessed using accuracy, precision, and recall metrics, and the system was found to have potential for improving traffic management, optimizing patrol deployment, implementing targeted safety measures, and ultimately reducing urban accidents.

Wei (2024) investigated an Internet of Vehicles (IoV)-based Accident Prediction and Prevention System that employed Internet of Things (IoT) technologies to address road safety challenges caused by increased traffic volume and population growth. The study reported that IoV devices facilitated real-time data transmission and analysis to improve road safety and traffic efficiency. A multi-tier framework was developed to monitor vehicle and roadside unit (RSU) data, including road traffic conditions, vehicle status, weather, and other external factors. The research indicated that the Spatio-Temporal Conv-Long Short-Term Memory Autoencoder (STCLA) framework on a cloud-based control server processed and analyzed this data for accident prediction and prevention. Wei demonstrated that the proposed framework effectively integrated deep learning approaches for real-time accident management, with a year-long study in Hubei Province, China, revealing a significant improvement in predictive accuracy, achieving an AUROC score of 0.94, highlighting the system's potential for enhancing vehicular safety.

Djazia et al. (2023) investigated the impact of emerging technologies such as information and communications technology (ICT), artificial intelligence (AI), and the Internet of Things (IoT) on smart city development, highlighting the critical role of intelligent transportation systems (ITS). They emphasized that the application of machine learning (ML) in driver assistance systems had enhanced both safety and travel comfort. In their study, they proposed an intelligent driving system designed for predicting road accident risks, which extracted relevant information to alert drivers and prevent hazardous situations. They noted that existing Internet-of-vehicle (IOV) solutions depended heavily on cloud computing, which, despite its vast storage and processing capacity, was limited by connectivity and response time issues. To address these constraints, they introduced a vehicular edge computing (V.Edge.C) approach, leveraging local resources for faster processing. Their proposed ICEDAS framework integrated two complementary models, CLOUD_DRL and V.Edge_DL, which collectively improved the accuracy of crash prediction. The results reportedly demonstrated that the system was effective in reducing accidents and saving lives.

Gheisari (2022) examined accident black spots and the factors influencing accidents in the medium-sized city of Tirunelveli, India, by employing factor analysis. The study was reported to have drawn on previous literature to adopt geospatial techniques for identifying black spots and analyzing the contributing factors. The most influential factors were determined and ranked according to the frequency of accidents occurring in the identified black spot areas. Spearman's ranking system was applied to assess the correlations among the factors, while factor analysis was employed to group and highlight the key contributors to repetitive accidents. The study was suggested to provide insights for transportation planners to understand accident causation and to implement appropriate interventions aimed at reducing casualties during both the planning of new road construction and the management of existing road conditions.

Chy et al. (2021) investigated the challenges posed by traffic congestion on timely product distribution and highlighted the additional manpower required for traditional delivery processes. They proposed Delicar, a self-driving delivery vehicle designed to navigate roads autonomously while reporting its real-time geographical location to authorities via a mapping system. The study described how a camera module captured road images, which were transmitted to a computer through socket server programming. The images were processed by a pre-trained deep learning model to predict the steering angle, and this information was sent to a Raspberry Pi that controlled the L298 motor driver to direct the vehicle's movements. Power management was addressed using a 3-cell 12V LiPo battery and a buck converter supplying 5V to the Raspberry Pi. The steering angle model was developed following an Nvidia CNN architecture with nine layers, including five convolutional and three dense layers. Furthermore, GeoIP2 was employed to extract longitude and latitude from the system's IP address, and Folium was used to visualize the live location. The study emphasized that the system infrastructure was low-cost and easy to install.

Qiu and Fan (2021) investigated short-term travel time prediction (TTP) in metropolitan areas, emphasizing its significance for both travelers and traffic management due to increasing traffic volumes. They highlighted that accurately forecasting travel time could substantially aid vehicle routing and congestion mitigation, while noting that a primary challenge lay in developing and selecting the most appropriate prediction algorithm from available data. The study utilized travel time data collected from the Regional Integrated Transportation Information System (RITIS) and applied four machine learning algorithms—Decision Trees (DT), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Long Short-Term Memory neural networks (LSTM)—to predict travel times over short horizons of 15 to 60 minutes on selected freeway corridors. Various spatial and temporal factors affecting travel time were incorporated into model development. The results indicated that the Random Forest algorithm outperformed the other methods, demonstrating superior accuracy and stability in predictions.

Muhammad et al. (2020) reported that advances in information and signal processing technologies had significantly influenced autonomous driving (AD) by enhancing driving safety and reducing human effort through advanced artificial intelligence (AI) techniques. They observed that deep learning (DL) approaches had recently addressed several complex real-world problems, although their effectiveness in control processes for AD had not been thoroughly investigated. The survey emphasized the potential of DL architectures for reliable and efficient real-time performance and reviewed state-of-the-art strategies for safe AD, noting both their major achievements and limitations. It also discussed the applications of DL across the AD pipeline, including measurement, analysis, and execution, particularly in road, lane, vehicle, pedestrian, drowsiness detection, collision avoidance, and traffic sign recognition through sensing and vision-based methods. Furthermore, the authors evaluated the performance of various methods using different metrics, critically analyzing their strengths and weaknesses, and highlighted current challenges while providing recommendations for future research, offering a comprehensive reference for researchers entering Intelligent Transportation Systems.

Lin et al. (2019) proposed a deep learning–based assistive system aimed at enhancing the environmental perception of visually impaired (VI) individuals. The system was reported to consist of a wearable terminal equipped with an RGBD camera and an earphone, a high-performance processor dedicated to deep learning inferences, and a smartphone for touch-based interactions. They introduced a data-driven learning approach to predict safe and reliable walkable instructions using RGBD data and a semantic map. This map was further employed to facilitate VI users' understanding of their 3D surroundings and object layouts through well-designed touchscreen interfaces. Their study indicated that the learning-based obstacle avoidance method achieved notable performance on both indoor and outdoor datasets containing low-lying obstacles. In addition, user studies were conducted in diverse scenarios, which reportedly demonstrated significant improvements in the environmental perception experience of VI participants when using the proposed system.

Pouyanfar et al. (2018) discussed that the field of machine learning had entered a golden era as deep learning gradually emerged as the dominant approach. They reported that deep learning employed multiple layers to abstract data and construct computational models, with algorithms such as generative adversarial networks, convolutional neural networks, and model transfers significantly transforming information processing paradigms. They highlighted, however, that a comprehensive understanding of this rapidly evolving domain remained limited, and that deep learning was often treated as a "black-box," restricting fundamental development. The authors observed that it was frequently overestimated as a universal solution to all machine learning challenges, which they deemed inaccurate. Their review summarized both historical and contemporary state-of-the-art methods in visual, audio, and text processing, social network analysis, and natural language processing, and analyzed pivotal advances in deep learning applications. Furthermore, they noted that challenges such as unsupervised learning, black-box behavior, and online learning were addressed to suggest productive avenues for future research.

III. KEY FINDINGS FROM STUDY

Author & Year	Methodology	Data Used	Key Findings	Limitations
Haq et al. (2026)	Mixed-method AI & ML models	Traffic + simulation data	Improved prediction, reduced accidents	Privacy & trust issues
Sharrab et al. (2025)	Prophet, ARIMA, LSTM	Sensor-based traffic data	High prediction accuracy (~86%)	Data quality issues
Rehman et al. (2025)	RF, SVM, Neural Networks	Crash, weather, road data	Effective hotspot prediction	Complex data integration

Wei (2024)	STCLA Deep Learning	IoV & IoT real-time data	AUROC 0.94 accuracy	High infrastructure cost
Djazia et al. (2023)	Cloud + Edge ML models	IoV data	Reduced latency, improved accuracy	System complexity
Gheisari (2022)	Factor analysis	Accident records	Identified key risk factors	Limited predictive capability
Chy et al. (2021)	CNN (Nvidia model)	Camera & GPS data	Autonomous navigation success	Limited scalability
Qiu & Fan (2021)	RF, DT, XGBoost, LSTM	Travel time datasets	RF best performance	Data dependency
Muhammad et al. (2020)	DL survey	Multi-source data	Improved AD safety	Implementation challenges
Lin et al. (2019)	Deep learning system	RGBD data	Enhanced perception	Limited scope
Pouyanfar et al. (2018)	DL survey	Multi-domain datasets	Strong ML capabilities	Black-box nature

IV. CONCLUSION

The study of machine learning-based prediction of road accident black spots highlights a significant transition from traditional reactive approaches to proactive and intelligent traffic safety planning. The reviewed literature demonstrates that ML algorithms such as Random Forest, Support Vector Machines, Neural Networks, and deep learning models provide highly accurate predictions by analyzing complex relationships among multiple factors influencing road accidents. The integration of IoT, IoV, and edge computing technologies further enhances real-time prediction capabilities and system responsiveness. These advancements enable transportation authorities to identify high-risk locations effectively, optimize resource allocation, and implement targeted interventions, ultimately reducing accident frequency and severity. However, challenges such as data quality, model interpretability, infrastructure requirements, and privacy concerns must be addressed to ensure reliable and ethical implementation. Overall, ML-based approaches offer a robust framework for improving traffic safety and supporting sustainable urban mobility.

V. FUTURE SCOPE

- **Explainable AI (XAI):** Development of interpretable ML models to enhance transparency and trust in decision-making.
- **Real-Time Systems:** Integration of real-time traffic, weather, and vehicle data for dynamic black spot prediction.
- **Edge Computing:** Deployment of edge-based ML systems for faster processing and reduced latency.
- **Multi-Modal Data Integration:** Combining CCTV, satellite, sensor, and social media data for comprehensive analysis.
- **Smart City Integration:** Linking ML models with intelligent transportation systems (ITS) and urban planning frameworks.
- **Autonomous Vehicles:** Utilizing ML predictions to enhance safety in autonomous driving environments.

- **Policy and Regulation:** Establishing standardized frameworks for AI implementation in transportation systems.
- **Geospatial AI:** Advanced GIS-based modeling for spatial visualization and planning of accident-prone areas.
- **Sustainability Focus:** Incorporating environmental impact analysis alongside safety predictions.
- **Hybrid Models:** Combining statistical and deep learning approaches for improved accuracy and robustness.

REFERENCES

1. Haq, I. U., Ali, S., Shahani, S. A., Iftikhar, H., Ali, S., & Shakil, M. (2026). Artificial Intelligence and Machine Learning in Smart Transportation Systems: Improving Road Safety, Traffic Flow, and Environmental Sustainability. *Global Research Journal of Natural Science and Technology*.
2. Sharrab, Y. O., Irtahi, B. M., Eljinini, M. A. H., & Alsmadi, I. (2025). Integrating deep learning and statistical models for traffic prediction and accident prevention in smart cities. *Cluster Computing*, *28*(16), 1-23.
3. Rehman, A. U., Khan, A., Ahmad, T., Zeeshan, M., & Jan, M. A. (2025). Accident Hotspot Prediction and Prevention Using Machine Learning. Journal of Emerging Technology and Digital Transformation, 4(4), 82-92.
4. Wei, X. (2024). Enhancing road safety in internet of vehicles using deep learning approach for real-time accident prediction and prevention. *International Journal of Intelligent Networks*, *5*, 212-223.
5. Djazia, Z., Kazar, O., Harous, S., & Benharzallah, S. (2023). Collaborative Cloud-V. Edge System for Predicting Traffic Accident Risk Using Machine Learning Based IOV. *International Journal for Computers & Their Applications*, *30*(4).
6. Gheisari, M. (2022). Identifying Influencing Factors of Road Accidents in Emerging Road Accident Black spots.
7. Chy, M. K. A., Masum, A. K. M., Sayeed, K. A. M., & Uddin, M. Z. (2021). Delicar: A smart deep learning based self-driving product delivery car in perspective of Bangladesh. *Sensors*, *22*(1), 126.
8. Qiu, B., & Fan, W. (2021). Machine learning based short-term travel time prediction: Numerical results and comparative analyses. *Sustainability*, *13*(13), 7454.
9. Muhammad, K., Ullah, A., Lloret, J., Del Ser, J., & De Albuquerque, V. H. C. (2020). Deep learning for safe autonomous driving: Current challenges and future directions. *IEEE Transactions on Intelligent Transportation Systems*, *22*(7), 4316-4336.
10. Lin, Y., Wang, K., Yi, W., & Lian, S. (2019). Deep learning based wearable assistive system for visually impaired people. In *Proceedings of the IEEE/CVF international conference on computer vision workshops* (pp. 0-0).
11. Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M. P., ... & Iyengar, S. S. (2018). A survey on deep learning: Algorithms, techniques, and applications. *ACM computing surveys (CSUR)*, *51*(5), 1-36.