JOURNAL OF ARTIFICIAL INTELLIGENCE AND ROBOTICS
Oct-Dec 2024, VOL. 1, ISSUE 4, pp. 27-34
https://doi.org/10.61577/jaiar.2024.100020

RESEAPRO JOURNALS

ORIGINAL ARTICLE

OPEN ACCESS

Integrating AI algorithms and GIS technology in traffic accident prediction: Towards safer roads through advanced modeling

Khanh Giang Le

Department of Civil Engineering, University of Transport and Communications, Hanoi, Vietnam

ABSTRACT

This paper explores the integration of Artificial Intelligence (AI) algorithms and Geographic Information System (GIS) technology for predicting and preventing traffic accidents. Traditional statistical methods are insufficient for capturing the complex factors influencing accidents, while AI and GIS allow for enhanced predictive accuracy and proactive safety interventions. The integration of AI and GIS enables the analysis of diverse datasets—historical crash data, environmental conditions, and road characteristics—through advanced algorithms, enhancing accident prediction and offering data-driven insights. Key benefits include precise data collection, real-time processing, and high mapping accuracy, which enable transportation agencies to identify high-risk areas, implement targeted safety measures, and improve emergency response. The paper presents a framework for AI-GIS accident prediction, involving stages like data collection, spatial analysis, model development, and monitoring. It details data preparation, emphasizing data cleaning, GIS feature extraction, and temporal-spatial aggregation to improve prediction accuracy. Geospatial models (like spatial regression) and Machine Learning (ML) algorithms (e.g., Random Forest and Optimized Support Vector Machines (SVM)) are applied to identify accident hotspots and assess risks effectively. The study, set in the United Kingdom, demonstrates AI-GIS's potential in enhancing traffic safety, though challenges such as technology costs and data quality persist.

KEYWORDS

Traffic Accident Prediction; Artificial Intelligence (AI); Machine Learning (ML); Geographic Information System (GIS); Road Safety

ARTICLE HISTORY

Received 06 November 2024; Revised 27 November 2024; Accepted 03 December 2024

Introduction

According to the World Health Organization's (WHO) 2023, there were about 1.19 million deaths and 50 million injured every year, mostly in low- and middle-income countries [1]. Traffic accidents pose serious risks to public safety and disrupt the efficiency of road networks. Integrating machine learning with advanced data analysis has transformed the way we predict and prevent these incidents. By harnessing powerful algorithms and neural networks, transportation agencies can analyze vast datasets to detect potential risk factors and implement proactive safety measures. This approach not only improves traffic safety but also helps reduce accident rates and the associated economic costs.

Developing advanced accident prediction models leverages deep learning alongside comprehensive traffic accident analysis. These models analyze complex data inputs, such as historical crash reports, environmental conditions, and roadway characteristics, to generate accurate forecasts. By utilizing artificial intelligence, the accuracy of these models increases, offering valuable insights into accident patterns and trends. This data-driven strategy empowers transportation authorities to implement targeted interventions, ultimately enhancing road safety and reducing the likelihood of traffic accidents.

Conventional traffic accident research often depends on statistical analysis, which has limited predictive accuracy and fails to fully capture the multiple factors that influence accidents [2]. These traditional methods fall short of addressing the inherent spatial complexity of traffic accidents. The success of any accident prevention strategy largely depends on the quality and accuracy of collected data and the appropriateness of the analysis methods applied.

Furthermore, the limited adoption of modern traffic management technologies often leads to congested and hazardous road conditions. Traditional approaches struggle to keep up with the complexities of expanding road networks and the growing volume of vehicles, underscoring the need for a more data-driven approach to studying accident patterns and pinpointing risk factors.

The structure of this paper is as follows: Section 2 provides a review of previous studies. Section 3 describes the methodology and data utilized in the study. Section 4 presents the results and includes a discussion. Lastly, Section 5 outlines the conclusions, discusses limitations, and proposes directions for future research.

Literature Review

Combining AI, ML, and GIS presents a powerful solution to deal with the drawbacks of traditional traffic analysis methods. These advanced technologies enhance predictive accuracy and enable proactive safety measures. By utilizing AI and ML models to anticipate and address accident risks, this approach offers a more precise and timely alternative to conventional methods, promising substantial improvements in public safety and traffic management [3-5].

*Correspondence: Dr. Khanh Giang Le, Department of Civil Engineering, University of Transport and Communications, Hanoi, Vietnam, e-mail: khanhgiang298@gmail.com

GIS technology is pivotal for conducting effective safety analyses, as it integrates diverse data sources—such as traffic volumes, road geometry, pavement conditions, and weather data—into a unified analytical framework. This comprehensive integration allows for in-depth, root-cause analysis of traffic incidents [6-10]. Each technology has its own strengths; when GIS and AI are combined, their power is maximized, especially in the field of transportation in general and traffic accident analysis in particular, as big data is fully utilized. Key benefits of AI and GIS integration include:

1. Enhanced data collection and management: GIS technology enables precise field data collection and seamless data management across systems, ensuring data accuracy and reliability.
2. Advanced analytical capabilities: AI-driven mapping can generate detailed vector layers for essential infrastructure, including crosswalks, sidewalks, and bike lanes, providing valuable insights for safety planning.
3. Real-time data processing: Modern systems allow for real-time data updates, supporting timely and data-informed decision-making.
4. Greater accuracy: AI-based solutions achieve a high level of mapping precision, comparable to skilled manual GIS processing, thereby boosting overall data quality and relevance.

Integrating AI and GIS thus empowers transportation authorities to predict, manage, and mitigate traffic risks more effectively, contributing to safer and more efficient road systems, including:

1. Accident hotspot identification: By identifying high-crash-density areas, agencies can focus on accident-prone locations, fostering safer driving environment.
2. Proactive safety interventions: Pinpointing high-risk zones enables the implementation of targeted measures to mitigate accident risks.
3. Enhanced emergency response: Accurate severity predictions support optimal emergency response, helping reduce fatalities and injuries.
4. Optimized traffic management: Advanced models contribute to improved traffic flow control and better transport planning.
5. Data-driven decision-making: The interpretability of these models empowers transportation planners to prioritize interventions that will have the greatest safety impact.

In reality, several studies have applied GIS and ML integration in traffic accident analysis, but most have focused only on utilizing clustering and association rule mining techniques to identify key factors causing accidents [11]. The study applied ML models to predict the severity of traffic accidents through R Studio and ArcGIS [12]. Although challenges such as technology costs, data quality, and infrastructure deployment remain to be addressed, the potential of integrating AI and GIS to enhance traffic safety is highly promising. This combination is expected to drive meaningful advancements in traffic safety management, contributing to accident reduction and the protection of human lives. Therefore, this study provides a methodological contribution by integrating GIS and AI technologies for traffic accident prediction and analysis.

Methodology and Data

Methodology

The integration of AI and GIS has revolutionized the approach to predicting and preventing traffic accidents. An effective AI-GIS accident prediction system comprises several key components that work in tandem to provide accurate and actionable insights. These components include a data collection module, a geospatial analysis engine, a machine learning core, and a visualization interface (Figure 1).

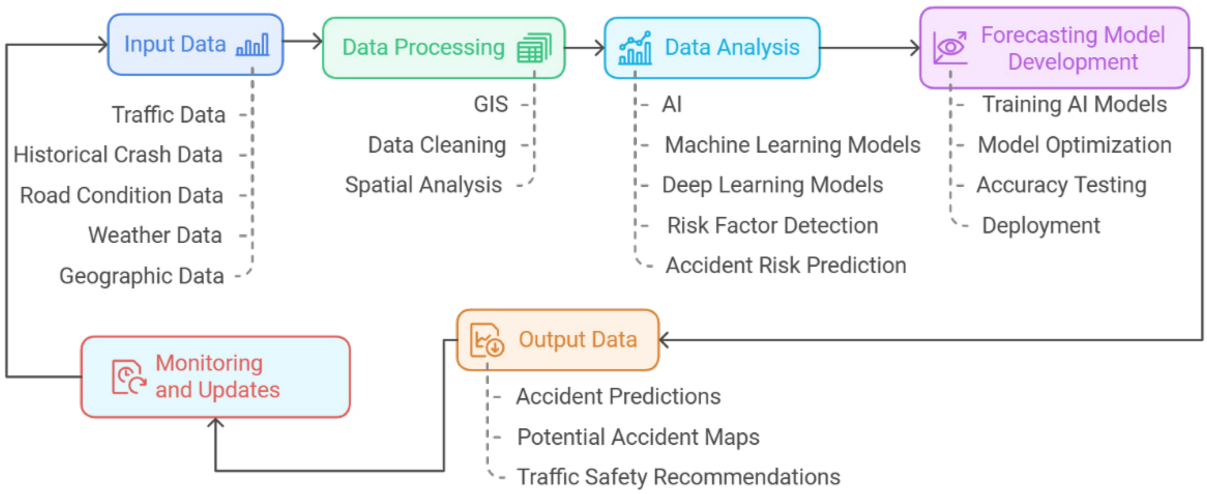

Figure 1. Diagram of the traffic accident prediction analysis model based on AI and GIS.

https://doi.org/10.61577/jaiar.2024.100020

Figure 1 illustrates the methodological framework for traffic accident prediction and safety enhancement. The process starts with input data acquisition, which includes various sources such as traffic, historical crashes, road conditions, weather, and geographic data.

Next, data processing is performed, where GIS is utilized for data cleaning and spatial analysis. This leads to data analysis, where AI, ML, and deep learning models detect risk factors and predict accident risks. Following analysis, the forecasting model development stage involves training AI models, optimizing, testing accuracy, and deploying them.

The output data generated includes accident predictions, accident risk maps, and traffic safety recommendations. Finally, a monitoring and updates process ensures continuous refinement, improving model performance and adapting to new data.

This methodological flow supports a comprehensive approach to analyzing and forecasting traffic risks and enhancing safety recommendations based on data-driven insights.

Although AI and GIS have been widely applied in traffic accident prediction analysis, they have mainly been used independently, with limited research on integrating these two technologies. Each technology has its own strengths, and when combined, they can form a highly effective tool. This integration represents a significant contribution to this paper.

Spatial regression models

Spatial regression models have become increasingly important in transportation safety analysis because they effectively address spatial correlations between neighboring areas. These models have been utilized in several areas of traffic safety research, including evaluating signalized intersections, assessing how traffic congestion affects safety, and analyzing accident mechanisms under varying levels of congestion.

There are two main types of spatial regression models widely used [14]:

1. Spatial Lag Model (SLM): This model is ideal when the dependent variable is influenced directly by the dependent variable values in neighboring regions. It assumes that spatial autoregressive processes affect only the dependent variable.
2. Spatial Error Model (SEM): This model is applicable when both the dependent variable and surrounding values are influenced by unobserved spatial characteristics, rather than directly by neighboring dependent variable values. It assumes spatial autoregressive processes are present only in the error term.

An advanced model, the Time-Fixed Effect Error Model (T-FEEM), has shown promising potential for traffic accident analysis. This model accounts for both spatial correlation among neighboring Traffic Analysis Zones (TAZs) and temporal correlation across time. In comparative studies, the T-FEEM has outperformed models such as ordinary least squares, SLM, and SEM, demonstrated by lower values in both the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) [14].

Advanced machine learning algorithms

Random Forests for Accident Prediction: Random forest algorithms have seen notable growth in use for processing traffic accident data in recent years. This ensemble learning technique integrates multiple decision trees to create a highly accurate and reliable predictive model. Its adaptability has led to widespread application across fields such as medicine, meteorology, and statistics.

In traffic accident prediction, random forests have shown exceptional performance, particularly in managing high-dimensional data and maintaining strong classification accuracy even when data points are missing. This robustness makes them especially suitable for the complex and sometimes incomplete datasets typical in accident analysis.

A major benefit of random forests is their capacity to evaluate feature importance, helping researchers pinpoint the most critical factors influencing accident severity. For example, studies have found that newly considered features like accident location, accident type, road segment details, and driving speed are often more influential than traditional factors like time of day or vehicle age [15].

Support Vector Machines (SVMs) in Risk Assessment: SVMs have proven to be highly effective in risk assessment for predicting traffic accidents. Their performance can be further enhanced by applying various heuristic optimization algorithms [16].

SVM is a classification algorithm that aims to find the optimal hyperplane to separate data classes. Here are the relevant formulas for SVM:

Separating Hyperplane: A hyperplane can be defined as:

$$w\,.x+b=0 \tag{1}$$

where w is the weight vector, x is the feature vector, b is the bias term.

SVM Objective Function: The goal is to optimize the margin between classes and the hyperplane. The objective function is:

$$min\frac{1}{2}\parallel w\parallel^{2} \tag{2}$$

Heuristic optimization algorithms for SVM

Genetic Algorithm (GA): Fitness Function: Measures the “goodness” of each solution; for SVM, a classification error function can be used.

Crossover and mutation operators are used to create new generations of solutions from existing ones to find the optimal solution [17].

Sparrow Search Algorithm (SSA): SSA simulates the foraging behavior of sparrows. The formula for updating a sparrow’s position is [18]:

$$x_i^{t+1} = x_i^t * e^{\left(-\frac{i}{max\_iter}\right)} \tag{3}$$

where xi is position of sparrow i; t: iteration number; max_iter: maximum number of iterations.

Gray Wolf Optimizer (GWO): GWO simulates the leadership hierarchy and hunting strategy of gray wolves [19]. Position update:

https://doi.org/10.61577/jaiar.2024.100020

RESEAPRO JOURNALS

$$X(t+1) = x^{*} - A \cdot D \quad (4)$$

where A and D are vectors based on the distance between the wolf and prey

Heuristic algorithms like GA, SSA, and GWO can enhance SVM performance by identifying optimal parameters, improving the accuracy and effectiveness of traffic accident prediction models.

Study area and data

The study area is the United Kingdom (UK). The UK covers a total area of about 244,376 km², with 242,741 km² being land. Its population stands at 67,837,434. The country has a radial road network comprising 46,904 km of main roads, 3,497 km of motorways, and 344,000 km of paved roads. As of 2022, there were 40.8 million licensed vehicles in the UK.

In 2022, the UK recorded 1,711 road fatalities, a 2% drop from 2019, and 29,742 cases of death or serious injury, down 3%. Total casualties of all severities were 135,480, a 12% decrease. Vehicle travel returned to pre-COVID levels, with 328 billion miles traveled, but fatalities per billion miles rose by 2% to 5% (Figure 2).

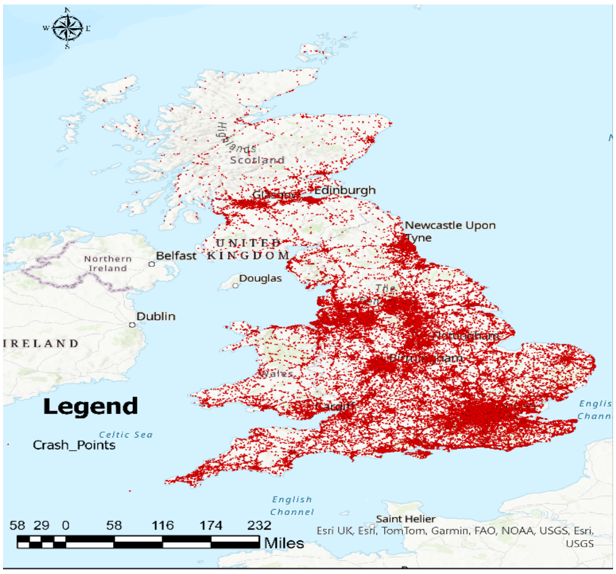

Figure 2. The locations of traffic crashes in the UK in 2022 displayed on the GIS map.

The traffic accident data used for analysis was collected from January to December 2022, with a total of 106,005 accidents, including full spatial information (such as accident location, as shown in (Figure 2) and attributes (such as date, time, cause, vehicle type, accident type, etc., as shown in Table 1), all provided via the website. Additionally, the road network data was sourced from the OpenStreetMap website, including both spatial and attribute data of the road network.

Table 1. The sample of study data.

accident index	acciden t year	accident_ reference	location_ easting_one	location northing_cost	longitude	latitude	police force	accident severity	number of vehicles	number of casualties	date	day_of week	time
2.02E+12	2022	10352073	525199	177928	-0.198224	51.486454	1	3	2	1	05-01-22	4	16:40
2.02E+12	2022	10352573	546214	179866	0.105042	51.49883	1	3	2	1	01-01-22	7	1:17
2.02E+12	2022	50352575	551119	174789	0.173482	51.451924	1	3	2	1	01-01-22	7	1:15
2.02E+12	2022	10352578	528889	192230	-0.139873	51.614152	1	3	2	2	01-01-22	7	2:24
2.02E+12	2022	10352580	539773	190404	0.016495	51.595151	1	3	4	3	01-01-22	7	2:30
2.02E+12	2022	10352588	543159	181263	0.065636	51.512146	1	2	1	5	01-01-22	7	2:55
2.02E+12	2022	10352501	535435	168744	-0.0544	51.401567	1	3	2	1	01-01-22	7	4:20
2.02E+12	2022	10352594	546736	189194	0.116442	51.58251	1	3	4	4	01-01-22	7	2:19
2.02E+12	2022	10352596	530501	178263	0.007763	51.486-112	1	3	1	1	01-01-22	7	5:05
2.02E+12	2022	10352600	534597	190727	-0.05806	51.599313	1	2	1	1	01-01-22	7	2:15
2.02E+12	2022	10352601	532694	183413	-0.088278	51.536038	1	3	1	1	01-01-22	7	2:10
2.02E+12	2022	10352605	534094	182549	-0.068434	51.525643	1	2	1	1	01-01-22	7	1:13
2.02E+12	2022	10352606	546303	179840	0.106283	51,498,574	1	2	2	3	01-01-22	7	3:51
2.02E+12	2022	10352610	508764	174030	-0.425995	51.454835	1	2	1	2	01-01-22	7	5:45
2.02E+12	2022	10352611	509533	182582	-0.422338	51.531533	1	3	3	4	01-01-22	7	7:40
2.02E+12	2022	10352612	528473	189140	-0.147037	51.58648	1	3	2	1	01-01-22	7	10:45
2.02E+12	2022	10352613	533304	178605	-0.081305	51.490683	1	3	2	1	01-01-22	7	6:40
2.02E+12	2022	10352615	529403	177887	-0.137725	51.48514	1	3	2	1	01-01-22	7	4:49
2.02E+12	2022	10352616	528111	177157	-0.012675	51.476523	1	3	4	1	01-01-22	7	9:05

https://doi.org/10.61577/jaiar.2024.100020

Data preparation and enrichment

An efficient AI-GIS accident prediction system is built on the foundation of data preparation and enrichment. To guarantee the accuracy, consistency, and applicability of the data utilized in the model, this critical stage entails a number of tasks.

Data Cleaning and Validation: The initial phase of data preparation involves cleaning and validating the data. This entails the removal or modification of data that is incorrect, incomplete, irrelevant, duplicated, or improperly formatted. Data cleaning can be a time-intensive process, often consuming up to 45% of a data scientist's time. It includes correcting spelling and syntax errors, standardizing datasets, rectifying issues like empty fields, and identifying duplicate entries.

To maintain data quality, a systematic approach is required, which involves:

1. Eliminating duplicate observations and irrelevant data
2. Filtering out unwanted outliers
3. Correcting structural errors, such as inconsistent naming conventions or improper capitalization
4. Addressing any missing data
5. Validating the cleaned dataset

For machine learning applications, a dataset should ideally contain at least 1,000 rows and 5 columns, with the first column designated as an identifier. The data should be consolidated into a single file or table, with minimal missing values and no personally identifiable information.

Feature Extraction from GIS: GIS is essential for extracting significant features for accident prediction models. This involves generating new metrics such as distance, proximity, and density, which can be fed into machine learning algorithms. Techniques like spatial clustering and spatial interpolation are commonly used to create these features.

To effectively integrate geospatial data, a strong framework is necessary to unify information from various sources, including GPS, satellite imagery, and social media. GIS tools excel in this area, allowing for the consolidation and alignment of disparate datasets into a comprehensive geospatial database.

Key static features obtained from GIS data include:

1. Road geometry (e.g., curvature)
2. Speed limits
3. Population density
4. Road orientation
5. Sinuosity
6. Traffic signs

Temporal and Spatial Aggregation: Temporal and spatial aggregation are critical for preparing data for accident prediction models. Temporal aggregation involves grouping data into specific time frames, such as hourly, daily, weekly, or monthly intervals, enabling the model to identify time-dependent patterns in accident occurrences. In contrast, spatial aggregation groups data based on geographic areas, such as countries, cities, or states, which helps identify spatial patterns and accident hotspots.

Incorporating both spatial and temporal dimensions is vital for models reliant on dynamic location data. Temporal-spatial models that consider changes over time and space offer a more comprehensive context for analysis. These models can incorporate dynamic features such as:

1. Weather conditions
2. Traffic volumes
3. Solar geometry
4. Time variables (hour, month, day)

To manage the extensive geospatial datasets needed for accident prediction, efficient data management strategies are essential. This may involve using optimized storage solutions and cloud-based platforms. Additionally, employing spatial indexing techniques, like R-trees or Quadtrees, can enhance the speed of spatial queries and minimize computational overhead. By implementing these data preparation and enrichment steps, organizations can establish a solid foundation for their AI-GIS accident prediction models, leading to more accurate and reliable outcomes.

Results and Discussions

The table 2 compares the performance of the standard SVM model and its optimized versions using heuristic algorithms, including GA-SVM, SSA-SVM, and GWO-SVM. Each model demonstrates significant differences in effectiveness, evaluated through metrics such as Accuracy, F1 Score, Precision, Recall, and AUC.

The standard SVM model exhibits the lowest performance among all models. Specifically, its accuracy is only 0.76, with an F1 Score and Precision of 0.66, indicating weak classification capability, especially in balancing Precision and Recall. The AUC for standard SVM is 0.74, reflecting average performance in distinguishing between data classes.

In contrast, the GA-SVM model, optimized with the Genetic Algorithm, shows a significant improvement over the standard SVM. Both accuracy and F1 Score increase to 0.83, while AUC reaches 0.87, demonstrating better classification performance. This marks a major improvement, highlighting the effectiveness of applying heuristic algorithms.

The GWO-SVM model, optimized using the Grey Wolf Optimizer, also achieves strong performance with an accuracy and F1 Score of 0.84. Its Precision is 0.93, close to SSA-SVM, but Recall is lower at 0.83. The AUC of GWO-SVM is 0.88, better than GA-SVM but slightly behind SSA-SVM.

Finally, the SSA-SVM model, which employs the Sparrow Search Algorithm, is the best-performing model in the table. With accuracy and F1 Score both at 0.87 and an outstanding Precision of 0.94, SSA-SVM showcases superior classification capability, particularly in minimizing false positives. Its AUC of 0.89, the highest among all models, indicates excellent ability to differentiate between classes.

Table 2. Comparison of prediction results between AI algorithms.

Model	SVM	GA-SVM	GWO-SVM	SSA-SVM
Accuracy	0.76	0.83	0.84	0.87
F1 Score	0.66	0.83	0.84	0.87
Precision	0.66	0.83	0.93	0.94
Recall	0.76	0.83	0.83	0.86
AUC	0.74	0.87	0.88	0.89

https://doi.org/10.61577/jaiar.2024.100020

These findings highlight the potential of optimized SVM models to improve the accuracy and reliability of traffic accident risk assessments.

In addition, the results indicate that integrating AI algorithms with GIS significantly enhances the ability to pinpoint high-risk areas. By leveraging spatial data and machine learning, the system generated detailed accident risk maps, which provide actionable insights for transportation agencies. These insights enable targeted safety measures, such as enhanced signage, speed regulation in high-risk zones, and improved road infrastructure. In addition, dangerous locations are also classified into different levels of danger from low to very high shown as in Figure 3. This will help traffic managers easily manage and promptly have effective solutions.

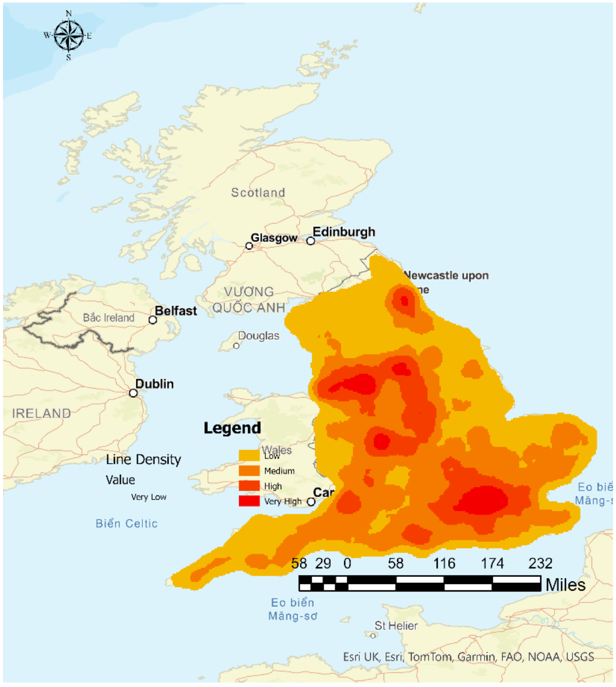

Figure 3. The dangerous accident locations.

Improved prediction accuracy has real-world safety implications, including:

1. Accident hotspot identification: By identifying areas with high accident densities, transportation agencies can prioritize these locations for safety improvements.
2. Proactive safety interventions: Accurately predicting high-risk zones enables targeted interventions such as:
   - Adjusting speed limits
   - Improving road design
   - Enhancing visibility
   - Implementing traffic calming measures
3. Enhanced emergency response: Predicting accident severity can optimize resource allocation for emergency response teams, reducing fatalities and injury severity. This could involve:
   - Dispatching appropriate resources
   - Planning efficient routes
   - Preparing hospitals for trauma cases
4. Informed urban planning: Reliable models guide policymakers in designing safer transport systems, considering traffic volume, road geometry, and population density.
5. Optimized traffic management: These insights can be used to improve traffic flow control and inform better transportation planning. This could lead to:
   - Reduced congestion
   - Smoother traffic flow
   - Minimized travel delays
6. Data-driven decision-making: AI-GIS models offer insights into the factors contributing to accidents, allowing transportation planners to prioritize interventions with the greatest impact on safety.

These advancements translate into measurable safety enhancements, such as a reduction in accident rates and associated costs, ultimately saving lives and fostering safer communities.

Conclusions

The standard SVM model performs the poorest among all models, underscoring the need for optimization techniques to improve its effectiveness. SSA-SVM stands out as the best-performing model, suitable for tasks requiring high accuracy and the ability to minimize false positives. GA-SVM and GWO-SVM also demonstrate significant improvements over the standard SVM but fall short compared to SSA-SVM overall. These optimized models are highly effective for complex classification tasks that demand high precision.

The combination of AI and GIS technologies is transforming accident prediction and prevention. By merging advanced data analysis with spatial insights, this integration enhances the accuracy of identifying high-risk zones, which supports more effective, targeted interventions and improved road safety. Real-world applications demonstrate significant benefits, from optimizing urban traffic patterns to improving emergency response times.

Future Directions

Looking forward, advancements in AI-GIS accident prediction models show great potential for further reducing traffic incidents and saving lives. As these technologies progress, even more adaptive systems can be developed to respond to real-time changes. However, fully leveraging these advancements requires ongoing collaboration among technology developers, urban planners, and policymakers to ensure that these solutions are implemented both effectively and responsibly.

Limitations and Mitigation Strategies

While the study presents promising results, it is important to acknowledge certain limitations:

1. Data quality and availability: The accuracy of predictions relies heavily on the quality of input data. Issues like missing data, errors, and inconsistencies can impact the model's performance. Mitigation strategies include:
   - Implementing rigorous data cleaning and validation procedures
   - Developing data quality control measures

https://doi.org/10.61577/jaiar.2024.100020

- Investing in data collection infrastructure to improve accuracy and completeness

2. Computational complexity: The integration of GIS and advanced AI algorithms requires significant computational resources, potentially limiting scalability. Mitigation strategies include: Optimizing algorithms for efficiency and employing distributed computing frameworks can reduce computational burdens.
3. Generalizability of results: The findings are based on data from a specific geographic region (UK), which may limit their applicability to other contexts. Mitigation strategies include: Expanding the dataset to include diverse regions and conducting localized calibration can improve the generalizability of the model.

Disclosure Statement

No potential conflict of interest was reported by the author.

References

1. Global Status Report on Road Safety 2023, Available at: https://www.who.int/publications/i/item/9789240086517.
2. Ahmed S, Hossain MA, Ray SK, Bhuiyan MM, Sabuj SR. A study on road accident prediction and contributing factors using explainable machine learning models: analysis and performance. Transp Res Interdiscip Persp. 2023;19:100814. https://doi.org/10.1016/j.trip.2023.100814
3. Chai AB, Lau BT, Tee MK, McCarthy C. Enhancing road safety with machine learning: Current advances and future directions in accident prediction using non-visual data. Eng Appl Artif Intell. 2024;137:109086. https://doi.org/10.1016/j.engappai.2024.109086
4. Lan HH, Hoc TQ. Applying a two-step cluster algorithm in traffic accident data analysis. Tạp chí Khoa học Giao thông vận tải. 2024;75(4):1673-1687. https://doi.org/10.47869/tcsj.75.4.16
5. Le KG. Enhancing traffic safety: A comprehensive approach through real-time data and intelligent transportation systems. Sci J Silesian Univ Technol Ser Transp. Series Transport/Zeszyty Naukowe Politechniki Slaskiej. Seria Transport. 2024;122. https://doi.org/10.20858/sjsutst.2024.122.8
6. Semira M, Alkhereibi AH, Abulibdeh A, Jawarneh RN, Balakrishnan P. GIS-based spatiotemporal analysis for road traffic crashes; in support of sustainable transportation planning. Transp Res Interdiscip Persp. 2023;20:100836. https://doi.org/10.1016/j.trip.2023.100836
7. Le KG, Liu P, Lin LT. Determining the road traffic accident hotspots using GIS-based temporal-spatial statistical analytic techniques in Hanoi, Vietnam. Geo-spat Inf Sci. 2020;23(2):153-164. https://doi.org/10.1080/10095020.2019.1683437
8. Le KG, Liu P, Lin LT. Traffic accident hotspot identification by integrating kernel density estimation and spatial autocorrelation analysis: a case study. Int J Crashworthiness. 2022;27(2):543-553. https://doi.org/10.1080/13588265.2020.1826800
9. Le KG, Tran QH. The sustainable development of railway system in Vietnam by GIS-based Technologies. InE3S Web of Conferences 2021;310;03003. EDP Sciences. https://doi.org/10.1051/e3sconf/202131003003
10. Le KG, Liu P, Lin LT, Student P. Road traffic accident black spot determination by using kernel density estimation algorithm and cluster statistical significant evaluation. In FIG Working Week 2019.
11. Le KG, Tran QH, Do VM. Urban traffic accident features investigation to improve urban transportation infrastructure sustainability by integrating GIS and data mining techniques. Sustainability. 2023;16(1):107. https://doi.org/10.3390/su16010107
12. Al-Mistarehi BW, Alomari AH, Imam R, Mashaqba M. Using machine learning models to forecast severity level of traffic crashes by R Studio and ArcGIS. Front Built Environ. 2022;8:860805. https://doi.org/10.3389/fbuil.2022.860805
13. Bilaşco Ş, Man TC. GIS-Based spatial analysis model for assessing impact and cumulative risk in road traffic accidents via Analytic Hierarchy Process (AHP)—Case Study: Romania. Appl Sci. 2024;14(6):2643. https://doi.org/10.3390/app14062643
14. Wang W, Yuan Z, Yang Y, Yang X, Liu Y. Factors influencing traffic accident frequencies on urban roads: A spatial panel time-fixed effects error model. PLoS one. 2019;14(4):e0214539. https://doi.org/10.1371/journal.pone.0214539
15. Yang J, Han S, Chen Y. Prediction of traffic accident severity based on random forest. J Adv Transport. 2023;2023(1):7641472. https://doi.org/10.1155/2023/7641472
16. Li X, Lord D, Zhang Y, Xie Y. Predicting motor vehicle crashes using support vector machine models. Accid Anal Prev. 2008;40(4):1611-1618. https://doi.org/10.1016/j.aap.2008.04.010
17. Konak A, Coit DW, Smith AE. Multi-objective optimization using genetic algorithms: A tutorial. Reliab Eng Syst Saf. 2006;91(9):992-1007. https://doi.org/10.1016/j.ress.2005.11.018
18. Gharehchopogh FS, Namazi M, Ebrahimi L, Abdollahzadeh B. Advances in sparrow search algorithm: a comprehensive survey. Arch Comput Methods Eng. 2023;30(1):427-455. https://doi.org/10.1007/s11831-022-09804-w
19. Hou Y, Gao H, Wang Z, Du C. Improved grey wolf optimization algorithm and application. Sensors. 2022;22(10):3810. https://doi.org/10.3390/s22103810

https://doi.org/10.61577/jaiar.2024.100020