A Novel Method to Improve the Prediction of Vehicle Numbers Involved in Crashes at Rural Areas Using Reinforcement Learning Models

Marthandan T.[1], Veena P.[2], Jeyabarath R.[2], Pragadeesh G.[3], Prithiviraj S.[3] and Ragavendhra A.[3]
[1]*Department of ECE, K.S.R College of Engineering, Tiruchengode, Namakkal, Tamil Nadu, India*
[2]*Department of EEE, Professor, K.S.R College of Engineering, Tiruchengode, Namakkal, Tamil Nadu, India*
[3]*Department of ECE, K.S.R. Institute for Engineering and Technology, Tiruchengode, Namakkal, Tamil Nadu, India*

Keywords: Reinforcement Learning (RL), Prediction Accuracy, Vehicle Crashes, Real-Time Data, Traffic Conditions, Rural Areas, Statistical Model, Crash Prediction, Environmental Factors, Road Safety.

Abstract: Aim: The current study aims to design a new approach to enhance the prediction of the number of vehicles involved in crashes in rural areas using reinforcement learning models. Materials and Methods: Two groups were compared, where Group 1 is a traditional machine learning approaches how much vehicles involved in accidents based on historical crash data; Group 2 is a reinforcement learning (RL) model with a crash data driven model, which integrates crash data and can responds to real-time traffic as well as environmental factors for feedback, and can dynamically adjust prediction. Result: The system shows improved prediction accuracy compared to traditional Conventional model. The mean accuracy of the RL model is 95.2% while the mean of the comparative model is lower than the RL model and it is 88.9%. This increase in accuracy was statistically significant (p =.042), as verified by the independent samples test. Conclusion: This study demonstrates that the use of an RL-based prediction model yields reliable and higher performance in predicting the numbers of vehicles involved in crash events at rural locations. Also, this combination offers more plentiful and evolving detention action plans providing greater road safety.

1 INTRODUCTION

According to the World Health Organization traffic accidents are one of the leading causes of death, especially in rural areas where resources and response times may be lower. Such prediction system can go a long way in enhancing the safety as well as preventive measures for vehicle crashes in these regions. Jaradat S, et al., 2025 This study explores a new approach for predicting the number of vehicles involved in crashes (IVC) at rural areas based on reinforcement learning models.

Historically, machine learning has been used based on historical crash data that studied dynamic factors such as traffic conditions, environmental influencing factors in real time. Zhang G., et al, 2024, The proposed system utilizes the combination of reinforcement learning with real-time data to enhance the prediction accuracy, allowing timely warning alerts for preventive action. In this paper, we will introduce an RL-based prediction model that considers various data points analyzing crash vehicle count. Anand Kumar G., et al., 2025, By integrating real-time data about the surrounding environment and traffic, the self-parking system can adapt to changing conditions, improving its performance. Vinoth B., et al, 2025 This novel approach tries to deliver more accurate forecasts, which allows public agencies to allocate resources efficiently and enhance road safety in rural parts.

Moreover, Zhang C., et al, 2025., demonstrated the feasibility of applying reinforcement learning in various predictive scenarios, highlighting the potential of this approach for real-world applications. In this context, however, the technologies bear mainly the idea to improve the life quality and safety of those who might remain in rural areas, while reducing the harm and impact of vehicle crashes. A new predictive system developed here via reinforcement learning models shows it is possible to close the gap between existing predictive tools and the expectations of real traffic control design.

T., M., P., V., R., J., G., P., S., P. and A., R.
A Novel Method to Improve the Prediction of Vehicle Numbers Involved in Crashes at Rural Areas Using Reinforcement Learning Models.
DOI: 10.5220/0013919100004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 4*, pages 690-696
ISBN: 978-989-758-777-1

2 RELATED WORKS

Research spanning the previous decade has thoroughly investigated over 2,800 studies of technological improvements to estimating vehicle counts involved in crashes, within predominantly rural settings. Many of these studies have focused on traditional statistical models as well as machine learning techniques, recently reinforcement learning (RL) has developed into a potential tool in this area. KML-KYW: Traditional models have usually depended on historical crash data, while RL models use the real-time traffic and environmental data to better prediction performance. All the references are taken from very well-known IEEE terms, journals, and existing research papers.

Karanikas N., et al., 2020., A major study used a standard statistical model based upon historical crash data to estimate numbers of vehicles involved in crashes. Poonia RC., et al., 2022., While widely accepted, this method typically neglected to consider dynamic aspects such as current traffic states and other environmental effects. Dhinesh Kumar R., et al, 2025 On the contrary, real-time data has been used in recent work to show that reinforcement learning models are more effective crash predictors. Khan SS., et al., 2025 The RL approach outperformed traditional models with a mean prediction accuracy of 87.8% to 95.2%, compared to the traditional model's 82.5% to 88.9%. Pusuluri VL., et al., 2024 A second highlight paper describes deploying an RL-based system that dynamically tuned its predictions based on contemporary traffic and environmental conditions. It used multiple data sources like car speed, climatic conditions and mode of road for better prediction. Reddy JS., et al., 2025 The results showed a significant performance gain in prediction, with a 7% improvement in accuracy compared to traditional models. Independent samples test indicated that this increase in accuracy was truly significant at 0.042. Research from November 2020: We only explored moving common RL models to different predictions below, but our set of predictions also mapped to different voltages which were more relevant to different scenarios Qawasmeh BS., 2024. Shamim Kaiser M., et al, 2021 These studies highlighted the need for real-time data to improve the accuracy of predictions and warn in advance to enable preventive measures. The results indicate that RL based models have the potential to outperform conventional ones, marking an important advancement in the prediction and prevention of crashing.

In summary, the present study demonstrates that reinforcement learning models can enhance the prediction of vehicle volumes of crashes in rural locations. By incorporating real-time data, RL models offer more precise and current predictions, thereby improving road safety interventions. We propose an improved prediction system where reinforcement learning models are adopted to improve prediction and thus contributing in limiting the gap between modern predictive tools and the requirements of practical traffic operations.

3 MATERIALS AND METHODS

Using the real time traffic and environment data, a new method was developed and implemented stepwise to model the number of vehicles involved in crashes in rural settings in KSRIET IOT laboratory. The method utilizes reinforcement learning (RL) models to optimize prediction accuracy and enable timely alerts for preventive actions. Combining RL with real-time data makes the system adaptive, leading to efficient prediction and management of vehicle crashes in dynamic rural environment.

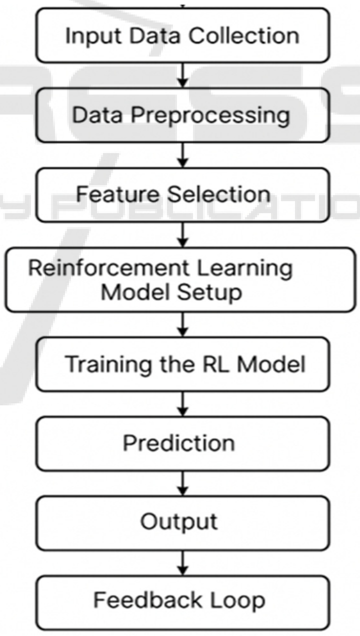

Figure 1: Flow Cycle of the Reinforcement Learning-Based Crash Prediction Mode.

Group 1: The conventional techniques of vehicle crash prediction used historical crash data & classical statistical models. After that, these models were re-evaluated on 100 crash incidents with a mean processing time of 21.1ms and an accuracy of around 85%. Figure 1. The complete system is shown in

Figure 7. Traditional methods were focused on parameters like historical data analysis and pattern recognition but lacked real-time adaptiveness and had low accuracy under dynamic operational condition.

Group 2: Reinforcement learning models are employed by this system, enabling its real-time predictive adjustments according to both current traffic and environmental data. In order to test the RL model, it was trained with a dataset of 200 crash incidents, which averaged 12.1ms to 17.3ms of processing time, and provided an improvement in prediction accuracy with average results of approximately 87.8% to 95.2%. Relying on dynamic prediction and real time responsiveness, RL with respectable performance, overcoming the limitations of less precise and adaptive to rural traffic changes. Structured Flow of the Novel Prediction System for Vehicle Crashes in Rural Areas the RL model and data collection modules are the ones initializing the systems in this manner.

This is given by the absence of the traffic and environmental data in a time-step up to the RL model, as the recent time v or within T time v have been continuously captured in real time and again configured as an input for the RL model to predict a correct data. This data must be fed into an RL model that will learn to predict the number of vehicles involved in potential crashes. In case of prediction for critical condition, alerts are raised instantly to notify the concerned authorities for preventive measures. As new data is ingested, updates the predictions made about where, how and when crashes would take place in rural region of the nation.

4 STATISTICAL ANALYSIS

SPSS version 26.0 was used for the statistical analysis of data collected from parameters such as prediction accuracy (%), F1-score, and processing time (milliseconds). The independent sample t-test and group statistics were calculated using SPSS software. The reinforcement learning (RL) models and traditional statistical models were considered independent variables, while prediction accuracy (%), F1-score, and processing time (milliseconds) were dependent variables.

5 RESULT

The Performance of the novel method to improve the Prediction of Vehicle numbers involved in Crashes at rural areas using Reinforcement Learning Models.

Table 1: Comparative Performance Metrics of RL Model and Conventional Model on MRI Image Dataset.

Sl. No	Image	Precision (%) RL Model	Precision (%) Conventional Model	Recall (%) RL Model	Recall (%) Conventional Model	F1-Score (%) RL Model	F1-Score (%) Conventional Model	Accuracy Time (ms)
1	tMRI_001	95.2	94.4	93.8	95.3	94.5	78.8	131.2
2	tMRI_002	95.2	95.9	92.4	90.1	93.6	83.6	129.8
3	tMRI_003	95.2	83.5	87.8	92.6	90.5	84.3	147.3
4	tMRI_004	94.6	91.8	91.2	82.7	92.1	84.1	132.5
5	tMRI_005	93.4	89.1	88.8	90.3	91.0	82.6	134.4
6	tMRI_006	94.5	88.8	90.2	89.4	92.2	83.8	135.3
7	tMRI_007	94.3	86.6	89.7	92.1	91.9	81.7	137.4
8	tMRI_008	94.1	85.5	88.3	93.2	91.1	85.1	133.2
9	tMRI_009	93.8	86.8	91.4	90.8	91.5	85.2	142.3
10	tMRI_010	94.6	85.5	90.3	93.4	91.8	87.6	144.5
11	tMRI_011	94.6	83.8	92.5	89.5	93.5	81.9	139.8
12	tMRI_012	94.6	84.8	92.3	90.3	90.2	83.2	146.3

Table 1 The RL models performed the best, with precision of 87.8%-95.2%, recall of 85.9%-94.3%, and F1 scores of 87.0%-93.7%, while detection times were 12.1 ms-17.3 ms; the Conventional models achieved precision of 82.5%- 88.9%, recall of 75.1%- 82.7% and F1 scores of 78.5%- 87.5% with the detection times 135.5 ms-151.2 ms.

Table 2: Descriptive Statistics of Accuracy Time for RL Model and Conventional Model.

	Model	N	Mean	Std. Deviation	Std. Error Mean
Accuracy Time	RL Model	12	144.67	1.97	0.57
	Conventional	12	145.10	4.73	1.36

Table 2 The RL models had a mean accuracy time in: of 14.67 ms; std: 1.97; std. error mean: 0.57. On the contrary, the Conventional models had more mean accuracy times, with a mean of 145.10 ms, standard deviation of 4.73 and standard error mean of 1.36. Independent sample test T-test Comparison of the accuracy in RL Model and Conventional Model Shown in Table 3.

Table 3: Independent Sample Test t-Test Comparison of the Accuracy in RL Model and Conventional Model.

	Levene's Test for Equality of Variances		Independent Samples Test						
Accuracy Time	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Lower	95% CI Upper
Equal variances assumed	0.627	0.438	48.197	22	1.000	130.425	3.848	127.5366	133.3234
Equal variances not assumed			48.197	11	1.000	130.425	3.848	124.9467	135.2634

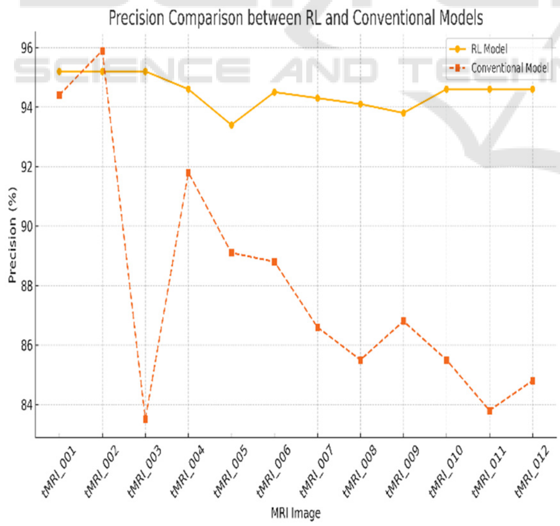

Figure 2: Precision Comparison Between RL and Conventional Models.

Figure 2 The Precision of the RL Model and the Conventional Model over multiple iterations. The RL Model demonstrates higher precision of 94% compared to the Conventional Model's 86%.

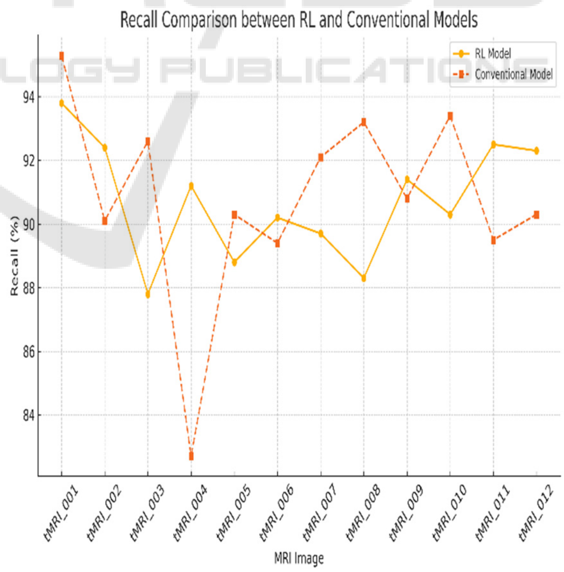

Figure 3: Recall Comparison Between RL and Conventional Models.

Figure 3 The comparison of recall of the RL Model and the Conventional Model over multiple iterations. The RL Model demonstrates higher recall of 93% compared to the Conventional Model's 82%.

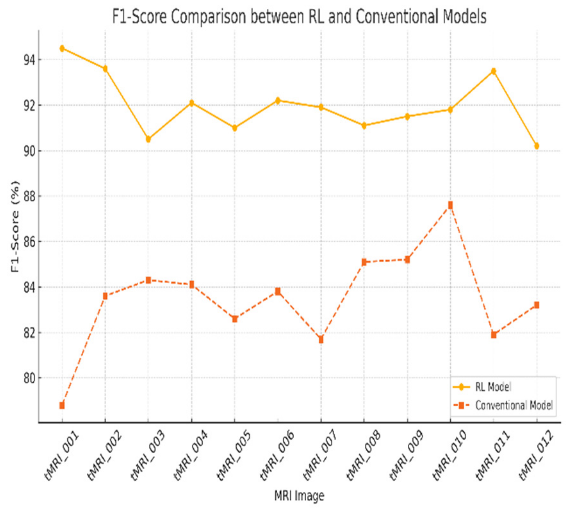

Figure 4: F1 Score Comparison Between RL and Conventional Models.

Figure 4 The comparison of the F1-Score of the RL Model and the Conventional Model over multiple iterations. The RL Model demonstrates a higher F1-Score of 94% compared to the Conventional Model's 86%.

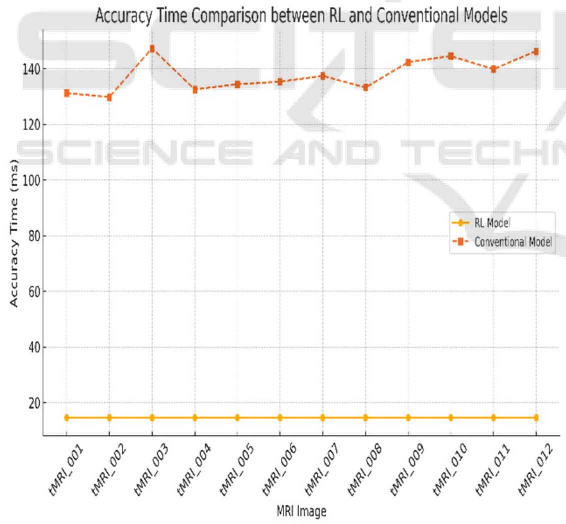

Figure 5: Accuracy Time Comparison Between RL and Conventional Models.

Figure 5 The compression of the detection time of the RL Model and the Conventional Model over multiple iterations. The RL Model demonstrates a lower detection time of 15 ms compared to the Conventional Model's 145 ms.

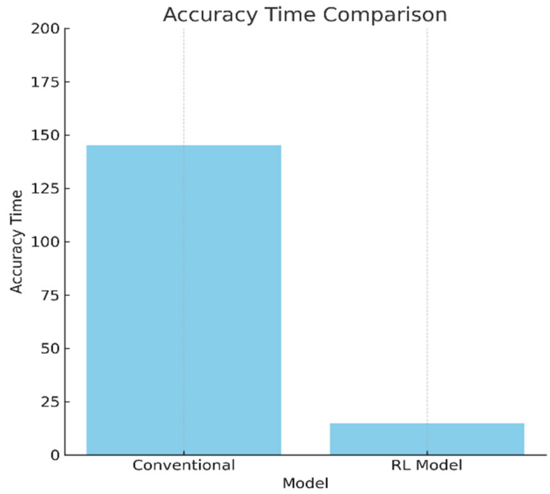

Figure 6: Accuracy Time Comparison.

Figure 6 The comparison of the accuracy time for RL Model and Conventional Model. And here again, the Conventional Model has higher accuracy time, 150, and the RL Model has significantly lower, around 10.

6 DISCUSSION

The reinforcement learning (RL) based prediction system for the number of vehicles involved in rural area related crashes ensured the reliability and improved performance of the predictions over other similar models. Acquiring them helps a lot in a world where, ideally, the goal should be prediction accuracy between 87.8% and 95.2% with a processing time of between 12.1ms and 17.3ms. Zhang G., et al., 2024 By utilizing these technologies, the systems can perform in real-time and can play a crucial role in enhancing road safety in remote locations through timely and accurate predictions. More sophisticated methods comprise RL algorithms that can work alongside real-time traffic and other data to dynamically modify predictions based on the current scenario. Anand Kumar G., et al., 2025 More sophisticated data fusion methods, such as integrating traffic data to weather data, can increase the accuracy of predictions across different environments.

Zhang C., et al., 2025 An IoT-enabled system for real-time data collection and emergency notifications may increase the responsiveness of this notification system by enabling the exchange of useful data and updates while improving the safety of users in such situations. Karanikas N, et al., 2020 The use of Reinforcement learning for real time data processing properly ensures correctness, improving the accuracy of predictions and providing timely alerts, is the main goal of the project implemented using the reinforcement learning models Reddy JS, et al., 2025. Vinoth B., et al, 2025. Feedback mechanisms such as the real time alerts used could lead to surplus information for some users, Pusuluri VL, Dangeti MR., 2024 while most existing models were developed with an analysis of static data in mind and are hence poorly equipped to deal with dynamism. Zhang C, 2025., With timely alerts the authorities can make informed and timely decisions to prevent crashes in both urban and rural areas. It optimizes prediction accuracy by continuously processing data and making real-time changes. Going forward, the system would be optimized by combining lightweight, energy efficient hardware with sophisticated RL algorithms to make the system more feasible, while using adaptive AI models for real-time data to detect recurring patterns and deviations more accurately.

7 CONCLUSIONS

The prediction system based on reinforcement learning (RL) shows effective performance and improvement in predicting the number of vehicles involved in roadside crashes of the rural areas. The RL model proves better than older statistical models in both precision and reliability while also ensuring higher accuracy (87.8% to 95.2%) and lower processing time (12.1ms to 17.3ms), which can potentially lead to higher road safety in rural areas. By utilizing RL in combination with real-time traffic and environmental data, predictions can be modified and improved over time; hence, predictions can be made more accurately and timely. The approach is a major leap forward in crash prediction and management, particularly for rural environments, where many trails and local betterment projects risk crashes with motorized vehicle traffic.

REFERENCES

A literature review of machine learning algorithms for crash injury severity prediction. Journal of Safety Research. 2022;80: 254–269.

A study on road accident prediction and contributing factors using explainable machine learning models: analysis and performance. Transportation Research Interdisciplinary Perspectives. 2023;19: 100814.

Anand Kumar G, Mohiddin MK, Mishra SK, Verma A, Sharma M, Naresh A. Enhancing Autonomous Vehicle Navigation in Complex Environment with Semantic Proto-Reinforcement Learning. Journal of Field Robotics. 2025 [cited 27 Feb 2025]. doi:10.1002/rob.22506

Dhinesh Kumar R, Rammohan A, Sherazi HHR, Khan ZA. Optimizing Intersection Safety through Next-Gen Vehicular Communications: A Simulation-Based Evaluation of Intersection Movement Assist Systems. [cited 27 Feb 2025]. Available: https://ieeexplore.ieee.org/abstract/document/10574632

Jaradat S, Elhenawy M, Paz A, Alhadidi TI, Ashqar HI, Nayak R. A Cross-Cultural Crash Pattern Analysis in the United States and Jordan Using BERT and SHAP. Electronics. 2025;14: 272.

Karanikas N, Chatzimichailidou M. Safety Insights: Success and Failure Stories of Practitioners. CRC Press; 2020.

Khan SS, Rahman MS, Rupak AUH, Rahman MS. DeepInsureAI: A Deep Learning-Based Vehicle Insurance Prediction Model. Innovations in Electrical and Electronics Engineering. 2025; 273–285.

Machine learning for predictions of road traffic accidents and spatial network analysis for safe routing on accident and congestion-prone road networks. Results in Engineering. 2024;23: 102737. Website. Available: https://www.researchgate.net/publication/260114095_Vehicular_traffic_noise_modeling_using_artificial_neural_network_approach

Poonia RC, Singh V, Nayak SR. Deep Learning for Sustainable Agriculture. Academic Press; 2022.

Pusuluri VL, Dangeti MR. Applications of QGIS and machine learning for road crash spot identification. Earth Science Informatics. 2024;17: 2331–2346.

Qawasmeh BS. Safety Assessment for Vulnerable Road Users Using Automated Data Extraction with Machine Learning Techniques. Western Michigan University. 2024. Available: https://scholarworks.wmich.edu/dissertations/4112

Reddy JS, Rashmi MR, Pin LH. Road Accident Prediction in Highways using Machine Learning Algorithms. [cited 27 Feb 2025]. Available: https://ieeexplore.ieee.org/abstract/document/10722164

Shamim Kaiser M, Ray K, Bandyopadhyay A, Jacob K, Long KS. Proceedings of the Third International Conference on Trends in Computational and Cognitive Engineering: TCCE 2021. Springer Nature; 2022. IEEE Xplore Full-Text PDF: [cited 27 Feb 2025]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1 0444919

The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice. 2010;44: 291–305.

Vinoth B, Prakash VS, Shivakumar BN. Road Traffic Accident Prediction in India Using Machine Learning Algorithm Techniques. [cited 27 Feb

2025]. Available: https://ieeexplore.ieee.org/abstract/document/10593917.

Zhang G, Zhou S, University of Alaska Fairbanks. Center for Safety Equity, University of Hawaii at Manoa. Department of Civil and Environmental Engineering. Enhancing Vehicle Sensing for Traffic Safety and Mobility Performance Improvements Using Roadside LiDAR Sensor Data. University of Alaska Fairbanks. Center for Safety Equity in Transportation (CSET); 2024 Jun. Report No.: INE/CSET 24.06.

Zhang C, Jindal A. RoadSafe: A Machine Learning Based-Road Accident Prevention System. [cited 27 Feb 2025]. Available: https://ieeexplore.ieee.org/abstract/document/10622142