Applied Computer Systems
ISSN 2255-8691 (online)
ISSN 2255-8683 (print)
2025, vol. 30, no. 1, pp. 122–132
https://doi.org/10.2478/acss-2025-0014
https://content.sciendo.com

PathGuard: Dynamic Large Vehicle Detection and Real-time Alerts on Narrow Roads Using Mobile Sensors

Sukhitha T. Sandunwala[1], B. M. Thosini Kumarika[2*]
[1,2]*Department of Statistics & Computer Science, University of Kelaniya, Kelaniya, Sri Lanka*

Abstract **– On a narrow road, an accident is hard to avoid even for a responsible driver. If vehicles are stuck in traffic, driving on a single lane is worrying and takes time. For small vehicles, narrow roads pose unique challenges, especially in identifying large vehicles, hence reducing the likelihood of an accident. The study discovers these issues and presents how an inventive Intelligent Transportation System (ITS) has been developed as a worldwide phenomenon that aims at enhancing safety on narrow roads by integrating with mobile sensors. Smartphones are used by almost everyone today because their prices have gone down. The study examines the effectiveness of different machine learning models for the task of classifying vehicle type using (accelerometer, and gyroscope) sensors. The results reveal that the Random Forest model is the most effective having a mean accuracy rate of 99.78 %. Moreover, the trained Random Forest Model has been combined with an originally developed unique warning algorithm that integrates geofencing methods for drawing polygons around narrow roads and location data from smartphones. To summarise, this study adds to the development of safety systems in transport and offers useful ideas for developing and implementing real-time safety applications for narrow roads.**

Keywords **– Accelerometer, gyroscope, intelligent transportation system, machine learning, narrow roads.**

I. Introduction

In today's rapidly developing world, mobility has become a major aspect of human life, which plays a huge role in the economy, society, and human well-being. However, this increased mobility has also created challenges, especially in traffic management and ensuring road safety [1], [2]. The coexistence of large and small vehicles on narrow roadways has become a notable disruption in effective traffic management.

The differences in vehicle types increase the difficulty of navigating through limited spaces, leading to traffic jams that slow down the smooth flow of traffic. This not only leads to time wastage but also increases the risk of accidents and problems in traffic management. Accidents on narrow roads can have considerable impacts, such as human casualties, economic losses, and prolonged interruptions to traffic flow. The risks of navigating through these restricted spaces are compounded by the limited awareness of drivers about the real-time nature of the road, particularly when large and small vehicles simultaneously traverse narrow roads. This lack of awareness is the main reason for increased safety risks [1]. This also creates an urgent need for an innovative solution that completely overcomes these challenges.

The idea of smart roads, together with advanced technologies, provides huge support for overcoming these challenges. In fact, smart road systems regularly depend on integrated sensors and computer vision to monitor and manage traffic conditions [3]–[5].

However, the broad execution of such smart road systems has been interrupted by the prohibitive costs of installing and maintaining allocated sensor networks. Considering this aspect, taking advantage of existing technologies, such as smartphone sensors, seems to be a more open and cost-effective approach. This study that creates an intelligent transportation system making use of smartphone accelerometer, gyroscope, and global positioning system (GPS) information aims to handle the challenges presented by narrow roads.

The main goal of this study is to form an intelligent transportation system that utilises mobile phone sensors and machine learning to identify vehicles and issue warnings inside the circumstances of narrow roadways. By combining sensor technology and advanced algorithms, the research intends to improve road safety by delivering real-time information to drivers.

- Developing Algorithms for Large Vehicle Detection
  - To develop and execute machine learning algorithms capable of recognising between large and small vehicles using data obtained from mobile phone sensors.
- Integrating the Detection System with Warning Mechanisms
  - To develop a robust system architecture that consistently incorporates the vehicle detection algorithms with real-time warning mechanisms.
  - To establish boundaries around narrow roads and deploy launching mechanisms to issue timely warnings to drivers when large vehicles exist in narrow roads.

[*] Corresponding author's e-mail: thosini@kln.ac.lk
Article received 2025-08-09; accepted 2025-10-16

The proposed system is significant for several reasons: the usage of ubiquitous mobile phone sensors and the system's applicability to a wide and varied user base stand top on the list. This accessibility guarantees that drivers of all vehicle types and economic backgrounds can access and get the advantage of the real-time warnings and safety features provided by the intelligent transportation system. Also, the usage of commonly used mobile phone sensors, specifically the accelerometer and gyroscope, presents a cost-effective solution compared to traditional methods. In both developed and developing nations, this affordability enables the integration of the intelligent transportation system guaranteeing that even areas with constrained resources can take the advantage of improved road safety measures.

II. Literature Review

The section provides a comprehensive review of recent research endeavours that address related aspects of the study. Although many studies have applied Random Forest models to vehicle classification, our study addresses a unique objective by classifying vehicles based on their physical size for road safety applications, rather than classifying the mode of transportation [6]–[14]. Most related research focuses on identifying whether an individual is in a car, on a bike, walking, or using other transportation modes. On the other hand, most research [4] on determining a vehicle's physical size is based on computer vision and costly infrastructure. In contrast, our research leverages smartphone sensors, which are both cost-effective and ubiquitous, to determine a vehicle's physical size. This approach offers a novel, economically feasible alternative to traditional methods of vehicle classification, setting our study apart from others in the field. Therefore, the investigation is mainly conducted in two domains. The first is about transportation mode detections, and the other is about intelligent transportation systems (ITS) that use smartphone sensors and warn smartphone users.

A. Vehicle Recognition Using Sensors

When discussing transportation mode detections, as shown in Fig. 1, previous research has primarily utilised GPS sensors, cell tower signals, barometers, magnetometers, inclinometers, gyroscopes, accelerometers, etc.

As the research primarily focuses on detecting large and small vehicles using accelerometers and gyroscopes, the studies incorporating these technologies for vehicle recognition will be mainly examined. Most of the previous studies, such as [6], have employed LSTM (Long Short Term Memory) networks for determining one of eight different modes of locomotion from testing datasets containing GPS, Wi-Fi, and cell tower signal data. They found GPS location to be the most crucial data, particularly when combined with feature engineering to calculate the direction and speed of the mobile device. These new features were integrated into the dataset and fed into an LSTM-based model. Additionally, feature engineering to derive direction proves significant for our research, albeit their system utilised more sensors than ours.

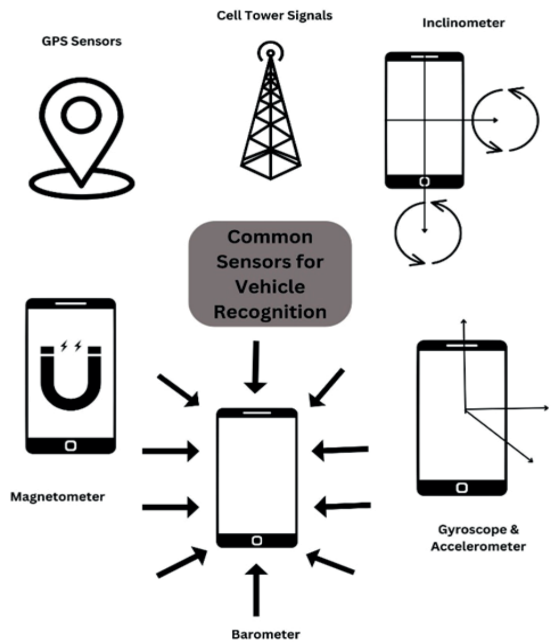

Fig. 1. Common sensors for vehicle recognition.

Furthermore, in [7], a residual and LSTM recurrent network-based transportation mode detection algorithm was proposed, leveraging multiple lightweight sensors (accelerometer, gyroscope, magnetometer, and barometer) integrated into commodity smartphones. Feature representation learning was separately adopted on multiple pre-processed sensor data using deep residual and LSTM networks, enhancing identification accuracy, and supporting one or more sensors. Lastly, [8] introduced a hybrid deep learning model, Deep Convolutional Bidirectional-LSTM (DCBL), which combines convolutional and bidirectional LSTM layers. This model is trained directly on raw sensor data to predict transportation modes, eliminating the need for handcrafted feature extraction methods. In addition to LSTM, several studies have employed the Random Forest algorithm. [9] classified walking, biking, car, bus, and rail using GPS and accelerometer data, achieving an overall accuracy of 96.8 % with the Random Forest classification model. Similarly, [10] achieved slightly higher accuracy using a Random Forest classifier. They utilised two datasets, achieving different accuracies for each set. The first dataset, collected by the research, attained an accuracy of 98.33 %, while the second dataset, obtained from HTC company in New Taipei, Taiwan, achieved 97.33 % accuracy. Additionally, this research identified driving events of motorcyclists, which was not directly relevant to our study. In [11], Random Forest was used to classify smartphone data among six modes, achieving an overall accuracy of 99.96 %, with no mode less than 99.8 % for data collected at a frequency of 10 Hz. The study observed a decrease in accuracy with decreased data frequency, but also noted a decrease in computation time. Variables such as data frequency, moving window size, and proportion of data used for training were optimised to improve results. Although this study utilised GPS, accelerometer, and gyroscope sensors, our research focuses solely on accelerometer and gyroscope sensor data to identify vehicle types, reducing computation time for the algorithms.

In [12], data collection involved capturing GPS, accelerometer, and gyroscope sensor data while subjects engaged in activities such as walking, running, biking, and traveling by bus or car. The application ran continuously for over 8 hours, tagging sensor data at 12-second intervals, resulting in 2500 patterns. Utilising the logistic regression method, the study achieved an impressive accuracy of 99.6 %.

Additionally, studies such as [13] and [14] employed deep learning techniques for classifying transportation modes. In [13], the researchers proposed a method to identify which vehicle the subject was using based on recorded accelerometer and gyroscope sensor data embedded in a smart device. In other words, this research implemented a machine learning model, specifically Artificial Neural Network, to classify vehicles using the sensor data.

This study primarily focuses on the dataset collected in [14]. However, the dataset was enhanced by incorporating sensor data from some more vehicle types using the Vieyra Software app. In this dataset, accelerometer and gyroscope data were collected from a bicycle, bus, and Toyota Camry using a Google Pixel phone, as well as from a bicycle, train, and Honda Insight using a Samsung Galaxy phone. Notably, the Honda Insight was driven in various conditions, including rain, traffic, and clear highway roads, for comparison purposes. Unlike [14], our study does not include train data. [14] utilised accelerometer and gyroscope data, commonly found in smartphones, to detect vehicle classes such as car, bus, train, or bike. After employing several machine learning algorithms, the highest classification accuracy was achieved by the designed 1D CNN model.

While previous studies have primarily focused on classifying vehicles according to different transportation methods, our research endeavours to classify vehicles into two distinct classes: large and small. Most systems that classify vehicles into these two categories rely on computer vision techniques. Examples of such research include [4] and [5].

B. Intelligent Transportation Systems (ITS)

The integration of vehicle recognition using smartphone sensors contributes to the advancement of Intelligent Transportation Systems (ITS). The evolution of safety measures from seat belts to traffic lights and red-light cameras showcases a continuous effort to enhance road safety. Our proposed system aligns with this evolution, utilising smartphone sensors to contribute to the growing field of Artificial Intelligence (AI) in transportation.

According to [15], for road safety and collision avoidance, three main methods are utilised: traditional methods based on traffic rules and traffic facilities, sensor-based methods, and AI-based methods. When considering AI-based methods, various AI techniques, such as neural networks, hidden Markov models (HMMs), and reinforcement learning, are applied in communication and sensing technologies for collision avoidance (CA) systems. However, even with AI, these systems cannot yet achieve 100 % collision avoidance. For real-time risk estimation, [16] proposes a framework for estimating traffic risks using advanced models and high-resolution autonomous vehicle (AV) sensor data. It introduces a 2D high-fidelity time-to-collision (TTC) indicator, extreme value theory (EVT), and hierarchical Bayesian structure models to capture real-time near-miss events and predict crash risks across diverse traffic scenarios, improving upon traditional crash-based methods. For real-time traffic management and optimisation, [17] presents a novel AI-driven fog computing framework. This framework leverages real-time data from IoT devices and sensors to predict traffic patterns, optimise traffic signals, and manage congestion dynamically, utilising advanced algorithms such as genetic algorithms, particle swarm optimisation, and ant colony optimisation for enhanced traffic flow and incident response. To enhance road safety in real-time, [18] introduces a highly accurate (99.75 %) real-time detection system for distracted driving, utilising the YOLOv8 model on Nvidia Jetson Nano with efficient processing at 70 milliseconds per frame. By integrating GPS/GSM modules, the system enables rapid classification of driver behaviours, aiming to reduce distraction-related accidents effectively. The proposed system can also be used with autonomous vehicles for path planning. In [19], AI techniques for autonomous vehicle path planning are explored, emphasising real-time adaptability and efficiency in dynamic environments. It highlights methods such as reinforcement learning, deep learning, and optimisation algorithms to improve safety, decision-making, and navigation performance.

Being more specific about AI-based systems used in ITS, in [20] these systems are highlighted as critical components of modern traffic management, enhancing safety, efficiency, and sustainability in transportation. ITS integrates technologies such as IoT, AI, and real-time data analytics to optimise traffic flow, reduce congestion, and improve decision-making. Key applications of ITS include adaptive traffic signal control, dynamic route guidance, and real-time monitoring of traffic conditions. The use of AI in ITS is particularly significant, enabling traffic pattern prediction, incident management, and enhanced communication between vehicles (V2V) and infrastructure (V2I). By leveraging machine learning and optimisation algorithms, ITS aims to create a connected and responsive transportation network that minimises environmental impact and supports smart city initiatives. Rear-end collision prediction is crucial for improving traffic safety in smart cities, but traditional methods often face challenges in feature extraction and prediction accuracy. A study [21] proposes a Rear-end Collision Prediction Mechanism (RCPM) using a convolutional neural network. This mechanism, through dataset smoothing and expansion, surpasses existing methods like Honda, Berkeley, and MLP-based algorithms in prediction performance. Considering ITS leveraging computer vision and AI techniques, another study [22] applies these technologies to extract vehicle speeds and analyse Automated Guided Vehicle (AGV) trajectories from port surveillance videos. By integrating a feature-enhanced descriptor, Kalman filter, Hungarian algorithm, and perspective projection theory, this framework achieves high accuracy in kinematic data extraction, demonstrating minimal measurement error and strong correlation with real-world scenarios, thereby aiding efficient port traffic management.

Previous studies on warning systems, as demonstrated in [23]–[25], underscore the potential of smartphone sensors in detecting and reporting emergencies, accidents, and incidents. Our research builds upon these concepts by focusing on the unique challenges presented by narrow roads. Unlike [4], which utilises a traffic management system employing RGB cameras for narrow roads, our system provides a cost-effective alternative with global applicability, leveraging ubiquitous smartphones for data collection and analysis.

In conclusion, our study not only advances vehicle recognition using smartphone sensors but also establishes itself as an accessible and cost-effective solution within the domain of intelligent transportation systems.

III. Research Data

The research primarily deals with sensor data collected by smartphones. A modern smartphone is a multipurpose device commonly used for day-to-day tasks by human beings. Using smartphones for various fields is a common scenario these days. In many such applications, smartphone sensors are the main source of data. The reason for using mobile phones for such applications is because they are commonly used and cost-effective. These sensors are integrated directly into the hardware of the phone and provide valuable data about the device's surroundings and the user's interactions with it. Figure 2 represents the coordinate system used by accelerometer and gyroscope sensors and some other sensors integrated into modern smartphones.

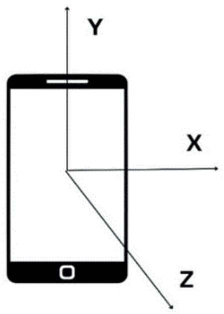

Fig. 2. Coordinate system used by accelerometer and gyroscope.

When the phone is held in the upright position with the screen facing the user, the Z-axis points to the outside of the screen, the X-axis is horizontal and points to the right, and the Y-axis is vertical and points up [26].

A. Accelerometer and Gyroscope Data for ML Models

Primarily, a dataset [27] available in Kaggle and some other data from an open-source Android application (Vieyra Software) [28] were collected for the training and testing purposes of the ML models used for vehicle type classification. The primary features included accelerometer readings along the x, y, and z directions and gyroscope readings around the x, y, and z axes. This data was collected using a freely available app, Physics Toolbox, developed by Vieyra Software. It includes data from cars, buses, bicycles, and trains; however, we excluded train data as it was not relevant to this study. Since motorbikes are a widely used mode of transportation in Sri Lanka and other developing countries, we collected additional data from motorbikes using the Physics Toolbox app. The research goal was not solely to classify vehicles accurately but also to integrate ML-based classification with other technologies to create a safety-oriented navigation application for users. After combining both collected and available data, we selected 2800 data points labelled as "large" and another 2800 data points labelled as "small".

Figures 3 and 4 present the basic line plots for different values corresponding to features ax, ay, az, wx, wy, wz (where x, y, z represent x y z dimensions, respectively, in Fig. 2) of accelerometer (a) and gyroscope (w) after concatenating the dataset to two different classes as small and large. By examining these plots and values, it becomes evident that it is challenging to distinguish between these classes based solely on visual inspection. Therefore, in this study, machine learning was decided to be used for classification.

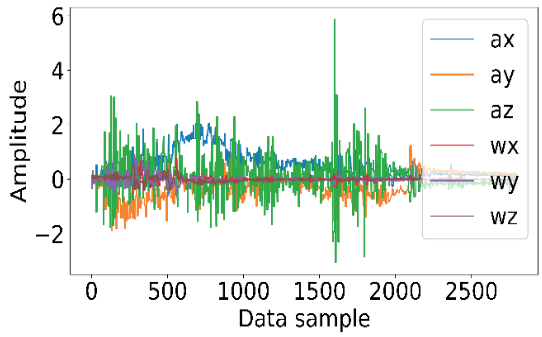

Fig. 3. Basic line plot for the large vehicle data.

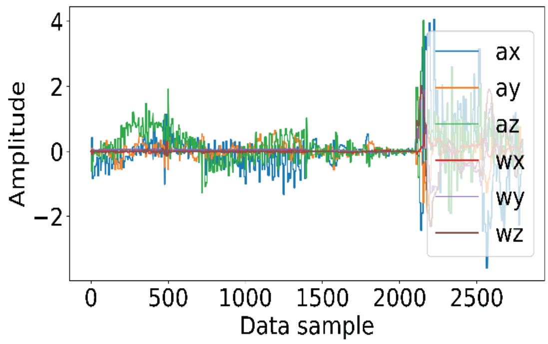

Fig. 4. Basic line plot for the small vehicle data.

B. Geospatial Data

To achieve the goal of determining whether a vehicle is inside a narrow road, latitude and longitude data of the vehicle's current location are needed. Without using a GPS tracking device, a modern smartphone with a GPS sensor can be used for this purpose.

Utilising a smartphone as a GPS tracker allows real-time location monitoring, route tracking, and geo-fencing. By leveraging the device's built-in GPS receiver and appropriate software applications, users can accurately track and manage assets, vehicles, or individuals for various purposes such as navigation, safety, and asset management. For developing the prototype, we collected GPS data from a narrow road with approximate width of 2.5 metres in Sri Lanka. We chose a road where vehicles could easily travel in only one direction; if two small vehicles met in opposite directions, they could manage with difficulty, but when a large vehicle arrived from the opposite direction, it caused traffic congestion or even accidents. Such roads are common in developing countries, and there are often no safety measures in place. These roads were selected because current navigation systems provided route options only based on traffic rules for given vehicles. However, roads with approximate widths of 2.5 metres present significant navigation challenges, as they are often suggested for both large and small vehicles – an issue that no existing system fully addresses. Using geofencing, we checked whether smartphone sensor readings matched the map data for these roads. The geofencing method used confirmed that GPS data from a smartphone could reliably indicate when the device was on the road.

IV. Research Methodology

The overall methodology followed throughout the study can be divided into three main parts: (1) the vehicle type recognition using sensor data and the use of machine learning models in the process; (2) the identification, extraction of narrow roads and the use of geofencing technologies to determine vehicle presence in narrow roads; (3) the development and design of the warning algorithm and how alert messages are issued to drivers. This systematic approach ensures a comprehensive investigation into vehicle detection and safety enhancement on narrow roads.

A. Vehicle Type Classification Using Sensor Data

Machine learning models were trained using the accelerometer and gyroscope sensor data under the x, y, and z dimensions. Initially, models such as Logistic Regression, Naïve Bayes, Support Vector Machine, Decision Tree, K Nearest Neighbour and Random Forest were utilised. These six models were trained by splitting the dataset into 20 % for testing and 80 % for training. Out of all machine learning models, the random forest model achieved the highest accuracy. However, recognising the accuracy alone may not suffice for developmental purposes, a k-fold validation was conducted. Following the k-fold validation, mean accuracies and mean area under the curve values were calculated, reaffirming the superiority of the random forest model for classification purposes. Subsequent validation with the confusion matrix further solidified this decision, leading to the selection of the trained Random Forest model for deployment in the production environment of the intelligent transportation system.

B. Identification and Extraction of Narrow Roads

Identifying and extracting narrow roads posed a bit of a challenge during the research process. We selected roads with an approximate width of 2.5 m. In Sri Lanka, where roads with such conditions are common, identifying these narrow roads is relatively straightforward. For this study, we selected narrow roads by querying in Overpass Turbo [29] with the width criteria, as shown in Fig. 5. Additionally, road data can easily be retrieved into our code using the Overpass API.

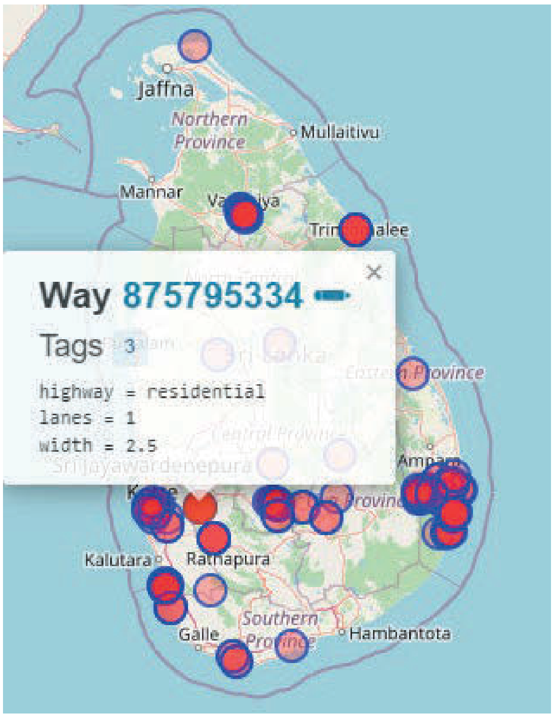

Fig. 5. Identifying roads with approx. width of 2.5 m.

The ensuing step involved determining whether a GPS location extracted from a smartphone lies within the narrow road. To accomplish this, the concept of geofencing was employed.

Geo fencing: This is a location-based service used to define a virtual perimeter around a geographical area, typically identified by GPS coordinates. The most common method involves drawing a simple circle or a more complex polygon around the geographical area and obtaining its coordinates. Geofencing enables triggering alerts when a device enters a defined virtual perimeter.

The next challenge involved drawing polygons around narrow roads and extracting their polygon GPS coordinates. For this purpose, the online tool called geojson.io was utilised [30], [31]. These polygons define boundaries around selected geographic areas. As polygons are drawn, geojson.io captures the latitude and longitude coordinates of each vertex, stores them as an ordered list, and converts them into a GeoJSON object, a standard format for representing geographic data. Finally, this data is used to identify whether a GPS coordinate derived from the smartphone lies inside the narrow road.

C. Design of Algorithm for Warning

The algorithm for warning was developed incorporating several python libraries [32], [33]. Also, as it was difficult to collect, demonstrate and validate the algorithm in real time, pre-collected data in the csv file format was used. In the algorithm, the test called point inclusion test, which uses points in polygon algorithms, was utilised. This involves checking whether the GPS point's coordinates fall within the boundaries of the defined polygon (Geofenced narrow road). First, the algorithm filters the times when there is a large vehicle in the geofenced narrow road. Then, all the vehicles in the system at that time are selected, and from them, vehicles in the narrow road are extracted. These vehicles receive the warning. This method ensures that if a large vehicle is alone on the road, it will be warned. Similarly, if one or more small vehicles and a large vehicle are in the specified narrow road, they will all be warned. However, if there are only small vehicles in the narrow road, no warnings will be issued. Considering these scenarios and in order to avoid distracting the driver with unnecessary warnings, a customized set of warning messages was created under categories A, B, C, D.

- A: "Warning: Vehicle has entered a narrow road."
- B: "Caution: Vehicle is in a narrow road. No large vehicles detected."
- C: "All clear: Vehicle is not in a narrow road."
- D: "Danger: Vehicle is in a narrow road with large vehicle(s) present."

Warning message A is customized only for large vehicles. Whenever a large vehicle enters a narrow road, this message will be issued. When a small vehicle enters a narrow road, there are two types of warning messages, namely B and D. In this case, when there are no large vehicles in the narrow road, a caution-type message B will be passed, considering that although the vehicle is in a narrow road, as there are no large vehicles, the driver can traverse safely while minding that he/she is in a narrow road. Danger-type message D will be issued if there is at least one large vehicle in the narrow road. Both large and small vehicles will receive message C when they are not in a narrow road. Although there is the possibility to add more messages, these critical cases were only selected so that the driver would not be distracted unnecessarily.

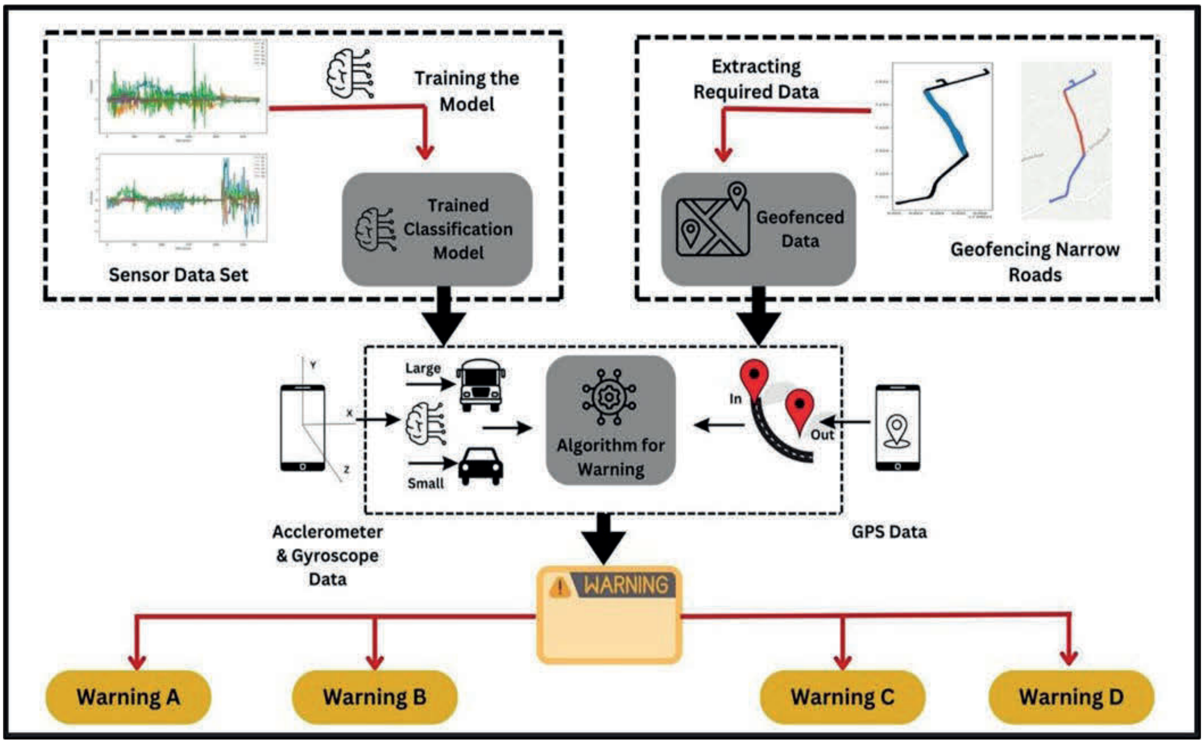

Fig. 6. High-level description of the research method.

The diagram above (Fig. 6) portrays a high-level description of how the warning algorithm was utilised with the research. Including all these parts, the final prototype with featuring more tailored warning messages was developed, and a dataset collected from the Sri Lankan Road context was used to evaluate the warnings of the warning algorithm.

As this research is currently at the prototype stage, we have used smartphone sensor data provided by testers with their informed consent, ensuring them that the data is solely for research purposes and the development of a system aimed at public well-being. For the final product, we plan to implement a comprehensive privacy policy to inform users and obtain their agreement. This policy will follow standard privacy management practices and trusted agreements, ensuring that both the service provider and any third parties access, store, and use location information with anonymisation (use generated id) and encryption.

V. Results

The section presents the outcomes of the experiments conducted to assess the effectiveness of the machine learning models utilised in detecting vehicle types, extracting narrow roads, and implementing the warning system designed to issue timely alerts to users navigating narrow roads. Through rigorous experimentation and analysis, key insights into the performance of these components were obtained, shedding light on their capabilities and limitations. The section highlights the accuracy rates, precision, recall, and other relevant metrics used to evaluate the performance of the ML models in accurately classifying vehicle types and identifying narrow road segments. Additionally, it provides insights into the effectiveness of the warning system in alerting users to potential hazards, thereby contributing to enhanced road safety on narrow roads. Overall, the section serves as a comprehensive evaluation of the proposed methodologies and technologies, providing valuable insights for future research and development in the field of intelligent transportation systems.

A. Results of Vehicle Type Classification Using Sensor Data

Analysis of the machine learning models revealed that the Random Forest model outperformed other models, achieving a highest mean accuracy of 99.78 %. This accuracy and other validation techniques depict the ability to use the model for large real-time datasets and capture intricate patterns in vehicle classification. As shown in Table I, ML models like KNN and decision tree also exhibited high accuracies, but they showed slightly lower accuracies and AUC compared to Random Forest Classifier. SVM, on the other hand, demonstrated moderate performance, indicating its limitations in capturing non-linear relationships within the data.

TABLE I

MEAN ACCURACIES AND AUC VALUES OF THE SELECTED ML MODELS

ML Model	Mean Accuracy	Mean AUC
Logistic Regression	73.839 ± 0.743 %	82.424 ± 0.531 %
Naïve Bayes	61.25 ± 1.259 %	83.456 ± 1.682 %
SVM	90.554 ± 1.242 %	95.67 ± 0.757 %
Random Forest	99.786 ± 0.071 %	99.997 ± 0.002 %
Decision Tree	99.5 ± 0.23 %	99.52 ± 0.249 %
KNN	98.661 ± 0.53 %	99.853 ± 0.079 %

Our study used sensor data, which is often noisy and subject to variability. Since the Random Forest model outputs the most frequent prediction from multiple decision trees, it helps reduce the impact of noisy data, making it superior to a single decision tree. Additionally, sensor data often exhibits complex, non-linear relationships due to the physics of motion and orientation changes. Random Forests are well-suited to model such non-linear interactions, as each decision tree captures a different pattern or subset of these relationships. This adaptability is particularly beneficial when working with multi-axis sensor data. In contrast, models like Logistic Regression and Naïve Bayes struggle with these complexities due to assumptions of linearity and feature independence, which do not hold true for sensor data. While effective in some cases, SVM can be computationally expensive and sensitive to noisy data, requiring careful kernel tuning. KNN is highly sensitive to feature scaling, making it less suited for this data. Unlike some models, such as neural networks, Random Forests do not require extensive scaling or normalization of the sensor data, simplifying pre-processing. This explains the outstanding performance of Random Forest compared to other models.

The study also explored advanced deep learning models, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, for vehicle classification. CNNs processed reshaped sensor data, leveraging spatial hierarchies with optimized hyper-parameters like filters, kernel sizes, dropout rates, and learning rates using Keras Tuner. The best CNN model achieved a test accuracy of 96.79 %. LSTM networks modelled temporal dependencies in sequential data, with tuneable units, dropout rates, and learning rates, achieving a test accuracy of 99.29 % after optimisation. While these methods demonstrated high accuracy, lightweight computation and reduced inference time make shallow machine learning models like Random Forest more practical in scenarios like vehicle size classification using sensor data.

B. Results of Identification and Extraction of Narrow Roads

In the initial stages of the research, a predefined set of Sri Lankan narrow roads, classified as residential roads in OpenStreetMap (OSM), served as the basis for experiments related to narrow roads. Using the online tool GeoJson.io, polygons were manually drawn around these narrow roads to establish geofences. To evaluate whether the polygon data and smartphone GPS data synchronize correctly to identify whether the vehicle is inside the narrow road, a test dataset was used. This dataset was manually annotated and used as ground truth. After determining whether the test GPS data lays within the narrow road, it was compared with the ground truth, and the confusion matrix was derived (Fig. 7). Using the confusion matrix, metrics in Table II and overall accuracy of 0.96 were calculated, showing how accurate the above-mentioned procedure is in extracting narrow roads.

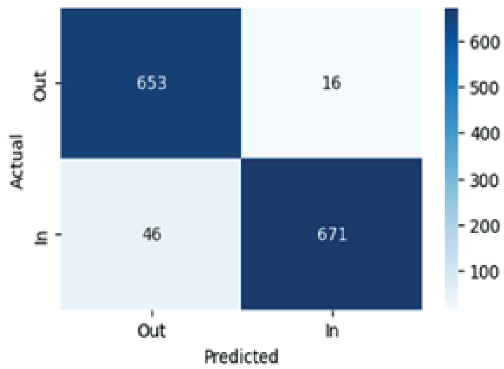

Fig. 7. Confusion matrix for the classification of geofence in or out.

TABLE II

CLASSIFICATION REPORT FOR THE CLASSIFICATION OF GEOFENCE IN OR OUT

	Precision	Recall	F1-Score
Geofence In	0.93	0.98	0.95
Geofence Out	0.98	0.94	0.96

C. Results of the Warning Algorithm

The warning system demonstrated its effectiveness in mitigating the risk of collision and road congestion on narrow roads. To assess its performance and validate the system, a test dataset collected from vehicles on Sri Lankan roads was used. First, the dataset was manually annotated and used as the ground truth for comparison. Then, the developed simulation was run using the test data, and the results were recorded under warning class labels corresponding to A, B, C, and D. Finally, each detection from the simulation was compared with the corresponding instance in the ground truth and categorised as TP, FP, TN, or FN, and the metrics were calculated as shown below. In Fig. 8, rows represent the actual class labels, columns represent the predicted class labels, and the diagonal elements show the number of correct predictions for each class. Table III represents the metrics calculated using the confusion matrix. These values and the overall accuracy of 0.95 calculated using the confusion matrix indicate that the simulation accurately identifies and issues warnings to vehicles navigating these hazardous narrow roads.

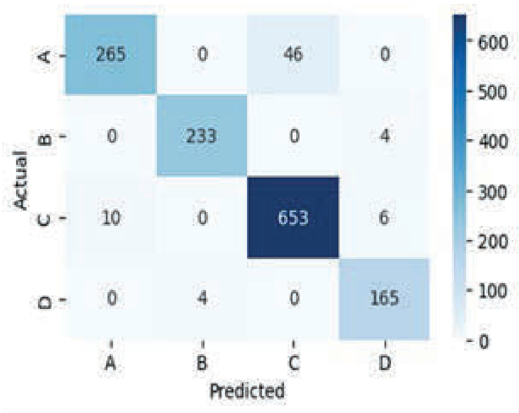

Fig. 8. Confusion matrix of the algorithm for warning.

TABLE III
CLASSIFICATION REPORT OF THE ALGORITHM FOR WARNING

	Precision	Recall	F1-Score
A: Warning	0.96	0.85	0.90
B: Caution	0.98	0.98	0.98
C: All clear	0.93	0.98	0.95
D: Danger	0.94	0.98	0.96

Figures 9–12 depict different scenarios, all of which showcase the effective functioning of the warning algorithm.

- Upon entry of a large vehicle into a narrow road, the warning message “A” will be issued.

Fig. 9. Warning message A.

- When a small vehicle navigates through a narrow road and no large vehicles are present, a caution message “B” will be transmitted indicating the absence of large vehicles on the narrow road.

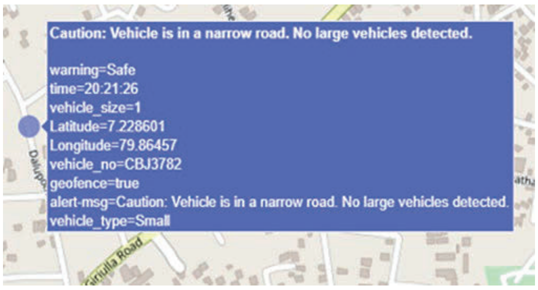

Fig. 10. Caution message B.

- Upon vehicles exiting the confines of a narrow road, the warning system issues an all-clear message “C”.

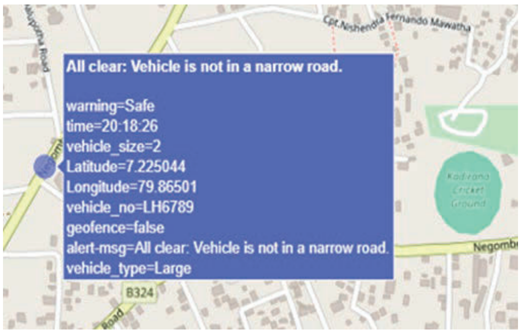

Fig. 11. All-clear message C.

- In scenarios where both a small vehicle and at least one large vehicle are present within the confines of a narrow road, the warning system prioritizes alerting the small vehicle with a danger message “D”, indicating the presence of a large vehicle on the road.

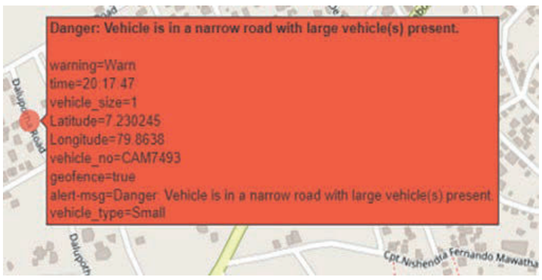

Fig. 12. Danger message D.

D. Results of User Study

To evaluate the usability and effectiveness of the warning messages, we conducted a user study involving a selected number of drivers with varying levels of experience. For each warning message scenario (A, B, C, and D), we presented the outputs of the prototype for simulated instances and provided participants with a questionnaire featuring a Likert scale (1–5). Initially, participants were briefed about the purpose of the study and the procedure. Each warning case was displayed, and participants' responses were recorded on paper. To minimise bias, the four scenarios were presented in random order. For each scenario, we calculated the average scores for clarity, relevance, actionability, and urgency. The accompanying bar graph (Fig. 13) illustrates these scores for each message. Based on the results, we concluded that while all messages generally performed well, Scenario C required improvements in clarity and urgency.

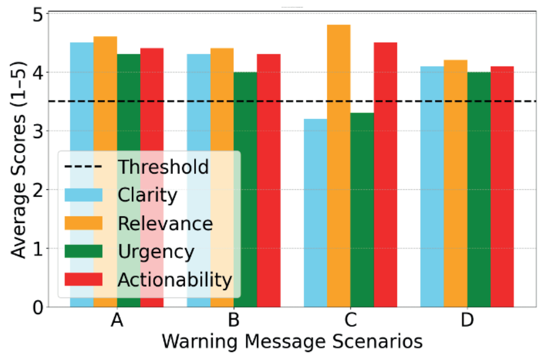

Fig. 13. Clarity, relevance, urgency, and actionability of warning messages.

E. Results of Real-World Experiments

The system is currently in the prototype stage. For real-world experiments, as shown in Figs. 14–16, we collected data from different vehicles simultaneously in real traffic conditions and used this data to enable the system to issue warnings accurately. In these experiments, the classification of whether a vehicle was on a specified narrow road was highly accurate, utilising the geofencing technologies explained in the previous section. The GPS data from the phone and the map data, received as a GeoJSON object, aligned well in our experiments. However, the model occasionally failed to accurately identify the vehicle size in some instances. Since we performed classification for every individual sensor reading, the model sometimes misclassified the vehicle type. To address this, we plan to base predictions on multiple readings and classify the vehicle using the class with the highest predicted probability as the final classification.

Fig. 14. Conducting real-world experiments.

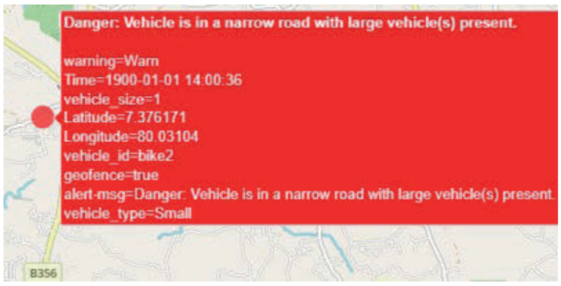

Fig. 15. Warning message to a specific vehicle at 14.00.36.

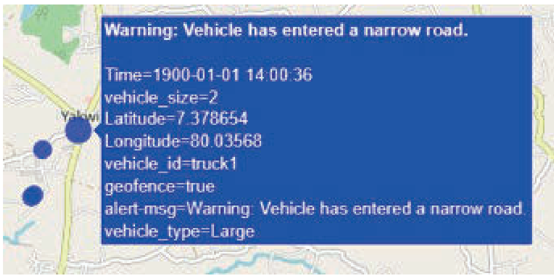

Fig. 16. Vehicles on the narrow road at 14.00.36.

VI. Conclusion

In summary, this study has illuminated the effectiveness of the created intelligent transportation system in improving road safety, particularly on narrow roads. Through the integration of machine learning models, the system has illustrated its ability to classify the vehicle type accurately, whether it is large or small. By incorporating these models into the warning system, vehicle drivers can receive real-time alerts about potential risks, improving their situational awareness. Moreover, the system has illustrated its capability to identify narrow roads by leveraging existing technologies such as geofencing and map data.

The limitations of the developed ITS prototype have several key areas that warrant consideration for future enhancements and refinements. First, while the prototype effectively utilised accelerometer and gyroscope data for vehicle type detection, it would have benefited from the inclusion of a more diverse range of large vehicle types, such as heavy vehicles and long vehicles, in the sensor dataset. Incorporating data from various vehicle types would have enhanced the accuracy and robustness of the detection model, particularly in scenarios involving different types of large vehicles. Moreover, the development of the prototype was constrained to a selected set of narrow roads, limiting its generalizability and applicability to broader transportation networks. Overcoming these limitations through continued refinement and development efforts can pave the way for a more comprehensive and effective ITS solution capable of addressing the complex challenges of modern transportation systems.

Moving forward, the refinement and improvement of the developed ITS into a fully functional IoT system incorporating cloud technologies can significantly improve its effectiveness and applicability in real-world scenarios. Furthermore, future research could focus on utilising OpenStreetMap and incorporating APIs to automatically extract all narrow roads according to predefined criteria. This approach would enable the ITS to operate across a broader network of narrow roads, improving its utility and impact. By embracing these future directions, researchers can advance the ITS from a prototype to a comprehensive and scalable solution capable of addressing the complex challenges of modern transportation systems.

VII. Abbreviations

API: Application Programming Interface
AUC: Area Under the Curve
CNN: Convolutional Neural Network
CSV: Comma-Separated Values
GPS: Global Positioning System
IOT: Internet of Things
ITS: Intelligent Transportation System
KNN: K-Nearest Neighbours
LSTM: Long Short-Term Memory
ML: Machine Learning
OSM: OpenStreetMap
SVM: Support Vector Machine

VIII. Declarations

Conflicts of interest: The authors declare that they have no conflict of interest.

References

[1] D. S. Samaraweera, "Road traffic accidents in Sri Lanka," Weekly Epidemiological Report, vol. 51, no. 6, Feb. 2024. https://www.epid.gov.lk/storage/post/pdfs/en_65fc60581a048_Vol_51_no_06-english.pdf

[2] "Global status report on road safety 2023." [Online]. Available: https://assets.bbhub.io/dotorg/sites/64/2023/12/WHO-Global-status-report-on-road-safety-2023.pdf

[3] L. Sun, H. Zhao, H. Tu, and Y. Tian, "The smart road: Practice and concept," *Engineering*, vol. 4, no. 4, pp. 436–437, Aug. 2018. https://doi.org/10.1016/j.eng.2018.07.014

[4] D. Triantafyllou, N. Kotoulas, S. Krinidis, D. Ioannidis, and D. Tzovaras, "Large vehicle recognition and classification for traffic management and flow optimization in narrow roads," in *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, San Francisco, CA, USA, Aug. 2017, pp. 1–4. https://doi.org/10.1109/UIC-ATC.2017.8397670

[5] J.-S. Chou and C.-H. Liu, "Automated sensing system for real-time recognition of trucks in river dredging areas using computer vision and convolutional deep learning," *Sensors*, vol. 21, no. 2, Jan. 2021, Art. no. 555. https://doi.org/10.3390/s21020555

[6] G. Dogan, J. D. Sturdivant, S. Ari, and E. Kurpiewski, "Locomotion-transportation recognition via LSTM and GPS derived feature engineering from cell phone data," in *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*, USA, Sep. 2021, pp. 359–362. https://doi.org/10.1145/3460418.3479379

[7] C. Wang, H. Luo, F. Zhao, and Y. Qin, "Combining residual and LSTM recurrent networks for transportation mode detection using multimodal sensors integrated in smartphones," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 9, pp. 5473–5485, Sep. 2021. https://doi.org/10.1109/TITS.2020.2987598

[8] J. V. Jeyakumar, E. S. Lee, Z. Xia, S. S. Sandha, N. Tausik, and M. Srivastava, "Deep convolutional bidirectional LSTM based transportation mode recognition," in *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, Singapore, Oct. 2018, pp. 1606–1615. https://doi.org/10.1145/3267305.3267529

[9] B. Martin, V. Addona, J. Wolfson, G. Adomavicius, and Y. Fan, "Methods for real-time prediction of the mode of travel using smartphone-based GPS and accelerometer data," *Sensors*, vol. 17, no. 9, Sep. 2017, Art. no. 2058. https://doi.org/10.3390/s17092058

[10] D.-N. Lu, D.-N. Nguyen, T.-H. Nguyen, and H.-N. Nguyen, "Vehicle mode and driving activity detection based on analyzing sensor data of smartphones," *Sensors*, vol. 18, no. 4, Mar. 2018, Art. no. 1036. https://doi.org/10.3390/s18041036

[11] M. Shafique and E. Hato, "Travel mode detection with varying smartphone data collection frequencies," *Sensors*, vol. 16, no. 5, May 2016, Art. no. 716. https://doi.org/10.3390/s16050716

[12] S. Ballı and E. A. Sağbaş, "Diagnosis of transportation modes on mobile phone using logistic regression classification," *IET Software*, vol. 12, no. 2, pp. 142–151, Apr. 2018. https://doi.org/10.1049/iet-sen.2017.0035

[13] T. S. Pias, D. Eisenberg, and M. A. Islam, "Vehicle recognition via sensor data from smart devices," in *2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE)*, Yunlin, Taiwan, Oct. 2019, pp. 96–99. https://doi.org/10.1109/ECICE47484.2019.8942799

[14] T. S. Pias, D. Eisenberg, and J. Fresneda Fernandez, "Accuracy improvement of vehicle recognition by using smart device sensors," *Sensors*, vol. 22, no. 12, Jun. 2022, Art. no. 4397. https://doi.org/10.3390/s22124397

[15] Y. Fu, C. Li, F. R. Yu, T. H. Luan, and Y. Zhang, "A survey of driving safety with sensing, vehicular communications, and artificial intelligence-based collision avoidance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6142–6163, Jul. 2022. https://doi.org/10.1109/TITS.2021.3083927

[16] M. Anis, S. Li, S. R. Geedipally, Y. Zhou, and D. Lord, "Real-time risk estimation for active road safety: Leveraging Waymo AV sensor data with hierarchical Bayesian extreme value models," *arXiv: arXiv:2407.16832*, Oct. 2024. https://doi.org/10.48550/arXiv.2407.16832

[17] K. Meduri, G. S. Nadella, H. Gonaygunta, and S. S. Meduri, "Developing a fog computing-based AI framework for real-time traffic management and optimization," *International Journal of Sustainable Development*, vol 5, no 4, May 2023. https://www.researchgate.net/publication/391019625_Developing_a_Fog_Computing-based_AI_Framework_for_Real-time_Traffic_Management_and_Optimization

[18] O. N. Neamah, T. A. Almohamad, and R. Bayir, "Enhancing road safety: Real-time distracted driver detection using Nvidia Jetson Nano and YOLOv8," in *2024 Zooming Innovation in Consumer Technologies Conference (ZINC)*, Novi Sad, Serbia, May 2024, pp. 194–198. https://doi.org/10.1109/ZINC61849.2024.10579437

[19] J. Singh "AI-driven path planning in autonomous vehicles: Algorithms for safe and efficient navigation in dynamic environments," *Journal of AI-Assisted Scientific Discovery*, vol. 4, no. 1, pp. 48–88, 2024.

[20] L. S. Iyer, "AI enabled applications towards intelligent transportation," *Transportation Engineering*, vol. 5, Sep. 2021, Art. no. 100083. https://doi.org/10.1016/j.treng.2021.100083

[21] X. Wang, J. Liu, T. Qiu, C. Mu, C. Chen, and P. Zhou, "A real-time collision prediction mechanism with deep learning for intelligent transportation system," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 9497–9508, Sep. 2020. https://doi.org/10.1109/TVT.2020.3003933

[22] X. Chen, Z. Wang, Q. Hua, W.-L. Shang, Q. Luo, and K. Yu, "AI-empowered speed extraction via port-like videos for vehicular trajectory analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 4, pp. 4541–4552, Apr. 2023. https://doi.org/10.1109/TITS.2022.3167650

[23] F. Bhatti, M. A. Shah, C. Maple, and S. U. Islam, "A novel Internet of Things-Enabled accident detection and reporting system for smart city environments," *Sensors*, vol. 19, no. 9, May 2019, Art. no. 2071. https://doi.org/10.3390/s19092071

[24] A. Khan, F. Bibi, M. Dilshad, S. Ahmed, Z. Ullah, and H. Ali, "Accident detection and smart rescue system using Android smartphone with real-time location tracking," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 6, pp. 341–355, 2018. https://doi.org/10.14569/IJACSA.2018.090648

[25] M. S. Roobini, S. Mulakalapally, N. Mungamuri, M. Lakshmi, A. Ponraj, and D. Deepa, "Car accident detection and notification system using smartphone," *Journal of Computational and Theoretical Nanoscience*, vol. 17, no. 8, pp. 3389–3393, Aug. 2020. https://doi.org/10.1166/jctn.2020.9192

[26] Z. Ma, Y. Qiao, B. Lee, and E. Fallon, "Experimental evaluation of mobile phone sensors," in *24th IET Irish Signals and Systems Conference (ISSC 2013)*, Letterkenny, Ireland, 2013, pp. 49–49. https://doi.org/10.1049/ic.2013.0047
[27] "Vehicle recognition dataset." Accessed: Jun. 08, 2024. [Online]. Available: https://www.kaggle.com/datasets/tanmoypias/vehicle-recognition-dataset
[28] "Vieyra software," vieyra-software. Accessed: Mar. 17, 2024. [Online]. Available: https://www.vieyrasoftware.net
[29] "Overpass turbo." Accessed: Nov. 30, 2024. [Online]. Available: https://overpass-turbo.eu/
[30] "geojson.io." [Online]. Available: https://geojson.io/#new&map=2/0/20
[31] J. Dougherty and I. Ilyankou, "Draw and edit with GeoJson.io | hands-on data visualization." Accessed: Mar. 17, 2024. [Online]. Available: https://handsondataviz.org/geojsonio.html
[32] "Documentation – GeoPandas documentation." Accessed: Mar. 17, 2024. [Online]. Available: https://geopandas.org/en/stable/docs.html
[33] "plotly.express: high-level interface for data visualization – 5.20.0 documentation." Accessed: Mar. 17, 2024. [Online]. Available: https://plotly.com/python-api-reference/plotly.express.html

Sukhitha T. Sandunwala received the B. Sc. Honours degree in Computer Science at the University of Kelaniya, Sri Lanka in 2024 with First Class Honours, and he was awarded the Gold Medal for the best performance in computer science from the University of Kelaniya. Currently he is working as a Lecturer in Computer Science at the University of Kelaniya. His research interests include mobile computing, artificial intelligence and machine learning.
E-mail: stsan241@kln.ac.lk

B. M. Thosini Kumarika received the PhD and the M. Sc. degrees in Computer Science from Georgia State University, Atlanta, the USA in 2021 and the B. Sc. Honours degree in Computer Science from the University of Kelaniya, Sri Lanka in 2014. She is currently a Senior Lecturer in Computer Science at the University of Kelaniya. She was awarded Molecular Basis of Disease fellowship by Georgia State University where she contributed as a research fellow. She received the Bachelor of Science with First Class Honours, and she was awarded the Gold Medal for the best performance in computer science in 2014 from the University of Kelaniya. Her research works have been published in many SCI journals and conference proceedings. Her current research interests include deep graph learning, artificial intelligence in healthcare, machine learning, and deep learning.
E-mail: thosini@kln.ac.lk
ORCID iD: https://orcid.org/0000-0002-0964-7826