DriveCare – Automated and Sustainable Points Manipulation System for Drivers and Traffic Police Officers

Lakshika Medhavini Wickramasinghe
Department of Computer Science and Software Engineering
Sri Lanka Institute of Information Technology
Sri Lanka

Pasindu Bhasura
Department of Computer Science and Software Engineering
Sri Lanka Institute of Information Technology
Sri Lanka

Ishanka Pramidu Lakshan
Department of Computer Science and Software Engineering
Sri Lanka Institute of Information Technology
Sri Lanka

Charith Widanapathirana
Department of Computer Science and Software Engineering
Sri Lanka Institute of Information Technology
Sri Lanka

Nelum Amarasena
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe
Sri Lanka

ABSTRACT
Adhering to traffic rules is an important aspect to follow while traveling on the road. Currently, there are no such applications addressing the issues that are faced by drivers and traffic police officers. In Sri Lanka, lots of traffic law violations are made by drivers daily, further, some traffic officers misuse the law and take advantage of victims. Introducing "DriveCare", a point-based monitoring system that has a mobile application and a web application to build connectivity between traffic police officers, court, police department, and drivers, in Sri Lanka. According to the actions done by the police officers or drivers, their overall points get increased or decreased which leads to either a cancelation of the driving license if the guilty party is the driver, or a transfer if the subject of the action is the police officer.

Analyzing the criticality of an accident, managing police officer transfers, reporting false allegations against drivers, and predicting accidents and early warnings are the features of this system. Images and videos are processed with image processing. OpenCV (Open-Source Computer Vision Library) is used to analyze images and videos. Natural Language Processing (NLP) is used to analyze and identify text-based datasets.

Keywords
Natural Language Processing (NLP), OpenCV (Open-Source Computer Vision Library), Image Processing, XGBoost, Random Forest, Artificial neural network (ANN), Support Vector Machines (SVMs).

1. INTRODUCTION
Road safety is an important aspect and a critical issue today in Sri Lanka. According to statistics on road accidents, there were 22319 accidents in the previous year (2021). Not only that, 2414 fatal accidents, 8070 minor accidents, 6401 critical accidents, and 5434 damages happened within that year. Preventing road accidents is the main goal of road safety. Maintain road safety, the police department plays a major role. When analyzing the previous records of incidents, the connectivity between the traffic police department and the drivers was low.

In this research paper, following four main components will be discussed.

- Analyze the situation of a particular accident using the images of vehicles in an accident and send emergency messages to parties.
- Identifying police officers who need to be transferred and where to transfer.
- Traffic Violation Reporting when a traffic violation happened.
- Accident prediction and early warning before a possible accident occurs.

Below diagram illustrates how the above mentioned four components are interconnected in the system.

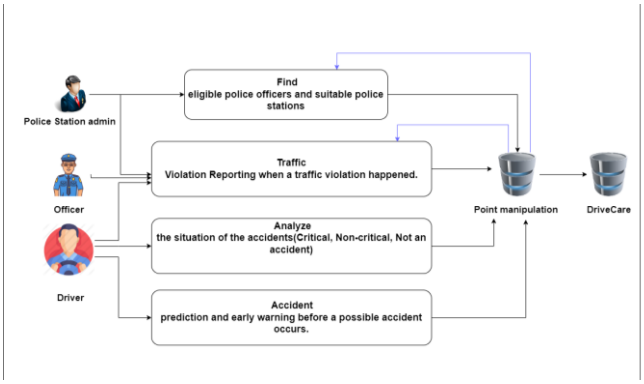

Figure 1: Connectivity of four components

The main intent of the first component is to verify whether damage happened or not and if it happened to verify whether it is a critical accident or a non-critical accident or not accident. Most of the existing systems only focused on giving alerts to police stations, hospitals, and relations using mobile applications. The main drawback of those existing systems available in the market today is those systems do not have any kind of mechanism to verify the criticality of the accident vehicle using images of that vehicle. When an accident situation happened on the road, by

considering the criticality of the accident, the system can notify related parties; the nearest active police officer, and relative, given in the system, about the accident situation. The alert message contains the current location of the vehicle and the status of the accident.

According to the second component, the issue with the current police officer transferring process in Sri Lanka is that it's slow, inefficient, and time incentive. Also, it's a manual task that requires human handling and sometimes it gets unnecessary third-party involvements as well. So, this component is built to find police officers who can be transferred and suitable police stations to which they can be transferred. When a station admin logs in to the system, they can see the list of officers that need to be transferred. For each officer, eligible police stations can be seen as well. Using this information, police stations can make the final decision on police officer transfer.

Traffic violations in Sri Lanka are increasing day by day and police officers are misusing the law in issuing fines. There is still no mechanism to report false allegations against the driver and take decisions on the offenses of traffic police officers and drivers. For decades, fines have been issued only in written form and no new techniques have been used. Due to such situations, there is a problem with making efficient and correct decisions for the police department as well as the court. This subsection of the proposed point base system proposes a new technology system for police officers and drivers, the police department, and the court to address these issues. A mobile application has been created to make it easier for police officers to enter fine data and track traffic violations committed by drivers. In the same way, a method is proposed here to make it easier for the driver to report the abuse of authority by the police officers through the mobile device with evidence videos. Here, attention has been focused on two types of violations. Related to issuing fines for lane crossing and traffic light violations. The accuracy of the mistake is also obtained through the dash camera video and using that data the points of the driver and the police officers are calculated in the point base system. A web-based software has been created to make this data easy to use for decision-making by the police department and court.

When considering day-to-day accidents in Sri Lanka, there are various factors involved with that, such as speeding, driving under the influence of alcohol, driver distractions, the nature of the road surface, and vehicle defects. Based on police statistics in, the most affecting factors to three-wheeler-related accidents are speeding, reckless driving, poor condition of roads, lack of road safety awareness and vehicle defects. This component mainly focuses on traffic accidents caused by careless driving by three-wheeler taxi drivers. Because three-wheeler is the second most registered vehicle in Sri Lanka among developing countries in South Asia [1]. Therefore, early prediction of such accidents can reduce many possible accidents. There are currently some proposed systems that can predict accidents, but most of them have considered the association with the human factor. Furthermore, there is currently no way to evaluate drivers based on performance. To bridge that gap, this accident prediction system is capable of analyzing vehicle speed, real-time location, weather conditions and traffic conditions. A machine learning model trained using the accident dataset as well as the speed limit on the road and, using the above factors, detects whether the driver has committed any violations. When a violation is detected, the system displays an early warning message on the driver's phone as well as a warning alert. Also, the driver is evaluated on the basis of points earned for his performance during the journey.

2. RELATED WORKS

Past research on "Design and development of a context-aware decision support system for real-time accident handling in logistics" describes the design and development of a context-aware fleet management system to optimize the vehicles during an accident. It is for real time vehicle positioning and eSeal enables RFID technology. A CFMS prototype was used for that [2]. An analysis of challenges when creating emergency management is described in "Addressing challenges for future strategic level emergency management" [3]

Research on "Emergency Button – a Telco 2.0 application in the e-health environment" uses emergency buttons, when elder people during their driving. It was built as a smartphone function. For the location tracker, it used a GPS-based location system. But it didn't get to the nearest location and distance to send messages. It is a simple emergency location system built on a cellular network infrastructure [4].

Research on "Landmark-based shortest path detection by using A* Algorithm and Haversine Formula" [5] is focused on the route between two applications inside a smart city environment. Google Map is used to get the exact route with landmarks and rode connecting them.

Recent research and trends focus on developing or improving several existing image-processing methods to capture road violations. [6] They focus on the types of violations on the road and detect traffic violations. Application examples for these techniques include object and image recognition [7]. Optical character recognition (OCR) methods are the most used solutions to plate recognition problems. The optical flow approach works well for detecting moving objects. It detects moving objects based on the relationship between the structure and motion of the objects and the time-domain change of pixel intensity in a sequence of images.

The main focus of intelligent detection systems is to improve systems that manage interactions between road components such as private or public transport, pedestrians, and existing traffic systems. Real-time systems and plate character recognition, vehicle plate location systems can arrest violators immediately after traffic violations. [8] But existing systems only focus on driver mistakes. In this proposed system, the video evidence submitted by the driver focuses on police officers and drivers both parties based on the accuracy obtained on the image processing technology.

According to an existing literature study, taxi drivers are overworked and stressed due to long shifts which directly affects their road safety. Drowsy driving, aggressive driving, mental states, and poor working environments can cause road accidents [9]. Although, drivers who's having a high number of daily taxi journeys and annual driving distances, as well as drivers driving at speeds above 50 km/h, are more likely to be involved in road accidents. Moreover, there are certain associations between aggressive violations, general violations, and accident involvement [10, 1].

The study based on drivers' wellness data has considered 4 stages different stages to produce an optimized model, starting with data collection, then prioritizing those factors, and after that optimizing the model using an artificial neural network (ANN) based on the collected factors. Finally, choose an enhanced model [9].

When considering a real-time warning it's a major part of this process. Hence, that result should be presented with accurate information at the right moment. In the present study on the accident early warning system, a video recognition-based method is used, which includes an accident detection device and an

accident early warning device. When the accident detection device detects an accident, it immediately activates the spike light alarm at the location of the first accident. According to the established vehicle driver state discrimination model, then the state, voice warning will be activated if the traffic is in a dangerous collision; At the same time, the upper area of danger uses variable information boards to issue hierarchical warnings according to the level of danger [11].

3. PROPOSED WORK

Since the system is based on points, both drivers and police officers are given initial points to start with. For drivers 100 points are given because their points can only be decreased. However, police officers are given 80 points because it can be either increased or decreased. If it is a violation five points will be deducted. This is applicable to both users. But if it's not a violation five points will be added to police officers. If the police officer's total points get equal to 100 rest of the points will be added as extras to motivate the officers.

Data collection and implementation methodologies will be discussed in these areas.

3.1 Analyze the Situation of the Accidents(Critical, Non-critical, Not an Accident)

When an accident happens, any user can submit images of the vehicle through the app. Images are analyzed and inform the nearest police officers of a critical or non-critical accident.

Since this component is completely based on images that will be uploaded by the users of the vehicle images, data collection also contains images of the accidents. To predict the accident is critical using images, there are three types of image data sets collected.

- Images of vehicles
- Images of critical accidents
- Images of non-critical accidents

Those images of accidents are mainly collected from local accident reporters like Accidents 1st, and Accident Prevention 1st Facebook pages. Apart from those two pages, images are collected from several global news websites. The main source that has been used to collect images of vehicles was official websites of vehicle manufacturers like Honda, Toyota, Mitsubishi, Mazda, Nissan, Ford, etc. Apart from that Pinterest and Google are used as the main sources of collecting images of previous-generation vehicles.

The main criteria that have been used to decide whether the image of the accident is a critical accident image or a non-critical accident image are the news reports that has been published by global news websites and local accident reporters. When analyzing news reports published with images of the accident, if the report mentioned that one or more persons that were in the vehicle had serious damages or died, that accident was categorized as a critical accident. If the report mentioned that those people had minor damages, those types of the image of accidents were categorized as non-critical accidents.

For selecting and analyzing the images, CNN (Convolutional Neural Network) is used. It is a way of implementing deep learning and image understanding. Deep learning recognizes objects in an image using CNN. It's a kind of network architecture to use for understanding the image and processing the pixel data. After capturing the image, there are two things to do first.

1. Image Normalization – Normally the pixel intensity of an image can vary from 0 – 255 so the model cannot be converged. Within that scale, images cannot be processed. Images need to decrease the pixel intensity to 0, 1 range or

-1, 1 range. After submitting one image, using image normalization the pixel intensity of that image can be decreased. It can be done in two ways as min-max scale and standard normalization. Min, max scale normalized the image up to 0, 1 and it is done by dividing the pixel count into 255. Standard normalization will be covered the intensity to -1, 1 range by decreasing the image pixel count to 127.5 and dividing by 255.

2. Data Augmentation – Data augmentation is used to process the images and improve the model performance [12] It is also used to verify the images. To demonstrate the augmentation rotation, cutout, and grayscale can be used. [13] Rotate is a safe operation. To process the model correctly, the images must be the correct way. Also, augmentation can be used to expand datasets and solve overfitting issues. [14]

DriveCare emergency alert message system mainly consists of a mobile application. Drivers and both police officers use this.

Driver – Drivers need to register to the system via the app. They need to provide their mobile number. After login to the system, the driver needs to add the details of the relative along with the mobile number. Those contact details are considered as the emergency contact information of the user. When an accident happens to the drivers, they can click the SOS button in the app. When the driver clicks the SOS button, an emergency SMS will be sent to the nearest traffic police officer and to the relative. When an accident happens on the road, any user who is logged in to the app can inform about the accident by uploading images of the vehicle which faced the accident. It is better to upload 4 images of the four sides of the vehicle, so it will be easy to analyze the situation of the accident. Then the system will analyze the situation of the accident as critical or non-critical. When the user uploads 4 images and at least one image is classified as a critical image, the situation is classified as critical. When it is a critical accident, the SMS will be sent to the nearest police officer, relative, and to the nearest hospital. When it is a non–critical accident, the SMS will be sent to the nearest police officer and relative. The images are uploaded to the AWS S3 bucket as storage.

Police Officer – When notifying the police officer, the system finds the nearest active police officer. To find the nearest police officer, the system considers the distance between the location of the police officer and the location of the accident happened. The app tracks the live locations of both users and to get the live location of the officers, they need to log in to the app when they are on duty. The system finds the nearest police officer by considering the accident location and the officer's location by using Google Distance Matrix API and then sends the SMS with the live location and criticality of the accident.

They need to log out of the app when their duty period is over. All the login officers will be counted as active officers. Police officers' duty locations are periodically checked and updated when the location is changed.

Live locations of police officers and accident places are tracked using Global Positioning System (GPS). Twilio is used as a third-party messaging service and is integrated into the system to send SMS.

3.2 Find Eligible Police Officers and Suitable Police Stations

When a police station admin logs in to the system, they can see a list of police officers who are eligible to get transferred. If a

specific officer is clicked, the most suitable police stations to which they can be transferred can be seen.

This component is based on text-based data, so no images or videos are used to implement it. When it comes to data sets, two data sets are used to train the models. The first one is for finding eligible police officers. The second set is to select suitable police stations.

For the first data set, points of police officers, tenure duration, and transfer requested status are taken into the consideration. For the second data set, position availability and previously worked or not are given priority.

Since police stations and officer details are highly confidential data, those data can't be obtained for any reason. So, those two data sets are created manually in order to train the models. As mentioned above, two data sets are created with 1500 records each to train the models.

When it comes to finding police officers to transfer, the points of each officer will be considered first. If it's above 80 or below 40, the police officer is eligible to transfer. And if their tenure years also vary between 3-5 years it also boosts the chance to get transferred plus their willingness to get transferred is also considered. To find suitable stations, the previous work status of that station and position availability are considered.

Three machine learning techniques have been used to train three models and the most performed one has been selected which is Random Forest. Three machine learning techniques that have been used are XGBoost, Random Forest, and Artificial Neural Networks (ANN).

To prepare text data for machine learning models, it must be converted into numerical features using techniques such as bag-of-words, term frequency-inverse document frequency, or word embeddings. This is also known as data pre-processing. In order to maximize the accuracy and efficiency of the models, this data pre-processing is essential. After converting the text into numerical features, the data can be scaled, divided into training and validation sets, and used to train machine learning models like XGBoost, Random Forest, or ANN. To improve the model's accuracy and generalization performance, hyperparameters such as the number of trees, depth of trees, number of hidden layers, activation functions, learning rate, batch size, and regularization strength have been fine-tuned.

XGBoost:- XGBoost (eXtreme Gradient Boosting) is a popular supervised learning algorithm, particularly for classification and regression tasks. It is a gradient-boosting method that makes use of decision trees as base models. XGBoost works by adding decision trees to a model iteratively and adjusting their parameters to minimize residual errors from the previous iteration. The algorithm is well-known for its efficiency, scalability, and high predictive accuracy, and it has been successfully used in a variety of applications, including credit risk modeling, fraud detection, and customer churn prediction.

Random forest:- Random Forest is another popular supervised learning algorithm that uses decision trees as base models. Unlike XGBoost, Random Forest constructs an ensemble of decision trees by selecting a subset of features at each node of the tree at random and then aggregating the predictions of the individual trees to make the final prediction. This procedure aids in the reduction of overfitting and improves the model's generalization performance. Random Forest is widely used in classification and regression tasks, and it has been used in a variety of domains including finance, healthcare, and natural language processing.

ANN:- ANN is made up of several layers of interconnected nodes (neurons) that process and transmit data through the network. Each neuron performs a simple computation on its input and sends the outcome to the next layer. Backpropagation allows ANN to learn complex patterns in data by adjusting the strength of connections between neurons. The ANN algorithm is a versatile algorithm that can be used for a variety of tasks including image recognition, speech recognition, and natural language processing. Training ANN, on the other hand, necessitates a large amount of data and computing resources and can be computationally demanding.

After getting police officers and their respective possible police stations to transfer, the police station admin can see the station list with approximate distance details. This distance data is fetched from Google map API by giving station and officer's residence location details. Using these details, police station admins can make the final call to decide which station is the best for each police officer and continue with the transferring process.

3.3 Traffic Violation Reporting when a traffic violation happened

When a police officer issues a fine, that data needs to be entered into the data through the mobile app. After following up on a report of a violation by a police officer through the mobile app, the app has been updated to allow drivers to view the specific violations they have committed recorded in their accounts. If the driver has dash camera video evidence related to misuse of authority by police officers, a facility has been provided here to access the relevant violation and upload it as evidence. Here, attention has been focused on two types of violations. It deals with issuing fines for lane crossing and traffic light violations. Python language and Open CV (Open-Source Computer Vision Library) were used to determine the probability of lane crossing violations or Traffic light violations. Based on the accuracy of the violation, the police officer's point and the driver's score are determined.

When a police officer correctly reports a violation, who gets five points. The system has been set up in such a way that marks are given based on the accuracy of the witness video uploaded by the driver in case a police officer abuses the law and issues a fine. Accordingly, if there is no violation according to the analyzed video, in cases where the accuracy is more than 60%, five points will be deducted from the police officer's score after being reviewed by the admin panel. Driver's marks will not be affected. Furthermore, in cases where the accuracy of the witness video uploaded by the driver is less than 60%, five points will be deducted from the driver's score.

Table 1: Police Officers and Driver Point Deduction

	If police officer report right	False Allegation from police officer	Driver Violation
Lane Crossing Violation	+5	-5	-5
Traffic light Violation	+5	-5	-5

To detect lane-crossing uploaded dash camera video, a method has been developed using OpenCV and Kalman filters. First, pre-processed the video by converting it to grayscale, applying Gaussian blur and canny edge detection to extract edges. Next, a region of interest (ROI) is defined focusing on the area of the

video where the lane lines were expected to appear. The Hough Transform was then used to detect the lane lines in the ROI, resulting in a set of line segments corresponding to the lane lines. To track the lane lines from frame to frame, Kalman filter was implemented, which predicts the state at each time step and updates the state based on new measurements of the lane lines.

A buffer area is defined around the lane lines that indicate when a vehicle has crossed. To detect lane crossings, the vehicle's position in the current frame is compared with its position in the previous frame. If the vehicle position was outside the buffer area in the previous frame and inside it in the current frame, a lane crossing will be detected. To assess the accuracy of this lane-crossing detection algorithm, the detected lane-crossing video is compared with a ground truth dataset containing actual instances of lane crossings. To assess the accuracy of this approach, precision, recall and F1 score are calculated.

To detect traffic light violations in dash camera recorded videos, a methodology has been developed using OpenCV and YOLO (You Only Look Once) object detection algorithm. Firstly, pre-processed the video by resizing the frames and applying normalization. Next, loaded the pre-trained YOLOv4 model using OpenCV's DNN (Deep Neural Network) module and ran it on each frame of the video to detect traffic lights and vehicles. Then analyzed the output of the YOLOv4 model to determine the location and state of the traffic lights and the location and movement of the vehicles. Based on the condition of the traffic lights and the movement of the vehicles, the traffic signal determines whether there is a violation.

3.4 Accident Prediction and Early Warning before a Possible Accident Occurs

When the driver logs into the mobile app, he can navigate to the DriveCare map. Then after specifying the destination the driver can start his daily journey. While driving, the map detects the vehicle's real-time location and speed. Apart from that, the traffic condition of the road is also obtained through this, and the weather data related to the area where the road passes is also obtained.

The system is integrated with various API services to collect real-time data. There, The Global Positioning System (GPS) is used to identify the real-time locations of the three-wheeler, where the longitude and latitude data of the vehicle is obtained. The GPS system is also used to get the speed of the vehicle. Used Google Map API to get real-time traffic status and integrate DriveCare map functionality. And finally, the AccuWeather API is used to retrieve weather-related information for the area the vehicle is travelling through.

In terms of data collection, this component is based on text-based data. The sources of road accident data in Sri Lanka are the police and motor vehicle insurance companies. Although the police have a centralized system, it is not currently in use, hence a manual data entry system is used. Though data is available through car insurance companies, it is difficult for third parties to obtain sensitive data. Therefore, data related to road accidents had to be created manually. It includes three-wheeler-related road accident data in Sri Lanka, along with accident location (latitude, longitude), city, date and time of the accident, speed, traffic conditions, weather condition, and road surface.

The obtained data is then cleaned and pre-processed, in order to remove any outliers, missing values and irrelevant features. The dataset contains both categorical and numerical data, but the model can only feed numeric values, so labelEncoder is used to convert all data with categorical values into a numeric value, such as city, traffic conditions, weather conditions, and road surface. After converting the data labels to a numeric value, all labels must be scaled between 0 and 1 when they are in different ranges or units. Then feature engineering involves analyzing data and extracting relevant features such as latitude & longitude, weather conditions, road surface, and traffic conditions. This is an important step because it helps create a robust model that can accurately predict accidents.

In this system, two machine learning (ML) algorithms had to be used to train the models for this accident prediction. The model that shows high accuracy based on the accuracy obtained is used for prediction. They are XGBoost, and Support Vector Machines (SVMs).

XGBoost: is an enhanced implementation of gradient boosting algorithms designed to be more efficient, scalable, and accurate. Moreover, it enhanced the performance of the weaker learner. This algorithm is categorized under supervised learning. This can use used for regression and classification tasks. It uses a variety of techniques in order to improve the performance of the model, including regularization, parallel computing, and gradient descent optimization.

Support Vector Machines (SVMs): This algorithm is also one of the popular classification algorithms. Which is also a supervised learning algorithm. SVMs are mostly useful for solving problems where the data are not linearly independent, and have many applications in various fields, including text classification, image classification, and handwriting recognition.

The above factors are referred to in the well-trained machine learning model, where the prediction is made based on the speed of the vehicle, the traffic condition of the road, weather condition, and the nature of the road surface. Furthermore, the above factors are analyzed in real-time by a well-trained ML model. Then, if the probability of the resulting model is greater than 0.5, the possibility of an accident on the road is high. Also, if the probability is minimal, it can be ascertained that the possibility of an accident is minimal. When there is a higher probability of accident the system notifies the driver by displaying an early warning message. This early warning message is displayed in three levels based on the nature of the traffic violation: 40kmph, 45kmph, and 50kmph. A warning message as well as a warning alarm is trigger when the vehicle speed is exceeded at a speed of 50kmph.When comes to the final stage, If the driver commits more than 3 violations of these rules in a day, through the point base system, a point will be deducted from his account. According to the Gazette of the People's Socialist Republic of Sri Lanka, the speed limit of a three-wheeler is 40 km per hour.

Table 2: Driver point deduction

	Driver volition
Over speed	-1

4. RESULTS AND CONCLUSION

In Sri Lanka, there was no system to connect drivers and police officers. Because of this reason, this system has been introduced to make this happen according to the Sri Lankan road regulation. Both police officers and drivers are given some points to start with and this will be manipulated according to their actions. Since the system analyze, predict, and provide recommendations based on certain actions, methodologies such as machine learning, image processing, and deep learning have been used to implement.

Deep learning is used to analyze big data. Convolutional Neural

Network (CNN) is a type of artificial neural network of deep learning. It can detect patterns of images. Using that images and objects can be identified and recognized whether the vehicle is critically damaged or not. For image processing issues, CNN can be introduced to analyze the images within an emergency period. As a messaging service, Twilio has been used to inform the about the accident to the relevant parties. Since the performance was high and the size is small, "Xception" is used as the image classification model. 0.91 of accuracy was obtained as the test accuracy with Xception.

For text-based datasets, XGBoost, Random Forest, and ANN can all be used, but the best option depends on the task and the data. XGBoost and Random Forest excel at sentiment analysis and text classification, while ANN excels at natural language processing. Random Forest is a popular choice for text-based datasets due to its high performance and ability to handle high-dimensional and sparse data. Random Forest can handle nonlinear relationships between features and reduce overfitting by constructing an ensemble of decision trees with randomly selected features at each node. Furthermore, Random Forest can generate feature *importance* scores, which can be used to interpret the results and identify the most important features in the data. Overall, Random Forest can be a powerful algorithm for achieving high performance and accurate predictions when working with text-based datasets. After testing the model with test data, 0.89 of test accuracy (89%) was achieved with the Random Forest model so it has been chosen as the final algorithm.

After obtaining videos from drivers for lane crossing and traffic violations, image processing techniques were used to analyze them and OpenCV and Kalman filter methods were used for lane violations. Similarly, OpenCV and YOLO object detection algorithms were used for traffic light violations from driver upload videos. This technology is great for showing lane-crossing locations and their accuracy. Also, it was possible to use technical methods more effectively to show exactly if traffic light violations have occurred. The F1 score is calculated as 0 to 0.8. This methodology provides an effective solution for lane-crossing detection in dash camera videos.

Three wheelers, also known as "tuk-tuks", are a common mode of transport in Sri Lanka, but are involved in a significant number of road accidents. In order to reduce such accidents as much as possible, the government should focus on strengthening the existing legal regulations, providing good training to drivers, informing drivers and passengers about safe driving practices and the risk of careless driving, and use of new technological tools.

When an accident occurs, it is impossible to make a conclusion based on it at once. Because it can be affected by various factors. Mainly, human factors, data availability, and some political and social factors will influence this. By identifying such factors, it is easy to identify the causes of the accident and improve the safety practices that can be taken.

The proposed solution is used a text-based dataset to train a machine learning (ML) model. XGBoost, Support Vector Machines (SVMs) were used as machine learning models. First, divide the obtained data into training and test sets. There, the training set was used to train the XGBoost and SVM models. After that hyperscale tuning was done to optimize the performance of the model and the test set was used to evaluate the performance of the model. Hyperscale tuning can be done by increasing the amount of training data, using performance-enhancing algorithms, and changing the model architecture. Finally, the XGBoost algorithm was selected as the model with the highest efficiency and accuracy. Thus, in predicting road accidents, if the probability of an accident is more than 0.5, it is considered an accident and if it is less, it is considered as no accident involved. Overall, 1.0 was obtained as test accuracy (100%) using XGBoost and 0.95 test accuracy was obtained for SVM. Since XGBoost achieved the highest test accuracy, it is used as a prediction algorithm.

5. REFERENCES

[1] I. U. A. S D R Perera, "The false safety of three-wheelers: A study in the Sri Lankan Context," International Journal of Engineering and Technical Research (IJETR), 2021.

[2] T. L. Y. W. M. L. P. C. Y. C. E.W.T. Ngai, "Design and development of a context- aware decision support system for real time accident handling in logistics," 2011.

[3] C. O. S. C. Karyn Bosomworth, "Addressing challenges for future strategic- level emergency management: reframing, networking, and capacity- building," 2016.

[4] K. L. J. L. Andrzej Podziewski, "Emergency Button – a Telco 2.0 application in the e-health," 2012.

[5] M. M. K. N. Prof. Nitin R.Chopde, "Landmark based shortest path detection by using A* Algorithm and Haversine Formula," 2013.

[6] W. K. M. Katanyoo Klubsuwan, Traffic Violation Detection Using Multiple Trajectories Evaluation of Vehicles, 15 April 2013.

[7] X. Wang, B. Z. J. L. Li-Min Meng and K.-L. Du, A video-based traffic violation detection system, 22 December 2013.

[8] A. R. F. Q. R. A. B. Aaron Christian P. Uy, A. B. Alexander Abad and Edwin Sybingco, Automated traffic violation apprehension system using genetic algorithm and artificial neural network.

[9] J. H. K. ,. J. P. ,. C. O. a. G. L. Seolyoung Lee, "Deep-Learning-Based Prediction of High-Risk Taxi Drivers Using Wellness Data," International Journal of Environmental Research and Public Health, 2020.

[10] A. S. M. Z. T. M. M. Javadreza Vahedi, Aberrant Driving Behaviour, Risk Involvement, and Their Related Factors Among Taxi Drivers, 2018.

[11] J. Liu, Highway Secondary Accident Warning System Based on Video Detection, 2020.

[12] T. M. K. C. Shorten, "A survey on image data augmentation," 2019.

[13] G. W. T. T. DeVries, "Improved regularization of convolutional neaural networks with cutout," 2017.

[14] C. S. T. L. C. R. A. Maier, "A gentle introduction to deep learning in medical image processing," 2019.

[15] A. R. F. Q. R. A. B. Aaron Christian P. Uy, A. B. Alexander Abad and E. Sybingco, Automated traffic violation apprehension system using genetic algorithm and artificial neural network.

[16] A. (. M. D. T. Amir Bahador Parsa, Real-time accident detection: Coping with imbalanced data, 2019.