INTERNATIONAL JOURNAL OF AUTOMOTIVE AND MECHANICAL ENGINEERING
ISSN: 2229-8649 e-ISSN: 2180-1606
VOL. 20, ISSUE 2, 2023, 10422 – 10440
DOI: https://doi.org/10.15282/ ijame.20.2.2023.08.0806

REVIEW ARTICLE

Review of Computational Techniques for Modelling Eco-Safe Driving Behavior

Neetika Jain and Sangeeta Mittal*

Department of Computer Science & Engineering and Information Technology, Jaypee Institute of Information Technologies, Noida, Uttar Pradesh-201309, India

ABSTRACT - Driving is a complex task involving the perception of the driving event, planning response, and action. Safe driving ensures the vehicle's and its passengers' safety, whereas economical driving brings down fuel consumption. Eventually, eco-safe driving that ensures economical as well as safe driving is the best bet. This review paper provides a systematic comprehensive analysis across cross-cutting dimensions such as data collection mechanisms, features affecting eco-safe driving, computational models for driving behavior analysis, driver motivational approaches towards eco-safe driving, exploiting research gaps and opportunities for further research in this domain. Driving behavior along with environmental context, including weather information, road conditions, traffic flow and time of travel, represent the most effective factors for doing eco-safe driving analysis. 82% of reviewed papers recommended OBD as a reliable data collection source, along with supplementary information from body sensors and cameras. The K-Mean clustering is an effective driving profiling technique clubbed with dimensionality reduction techniques based on Random Forest regressor, PCA and autoencoders. Deep learning and ensemble learning-based safety approaches utilizing Recurrent Convolutional Networks (RCN), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) and Decision Tree (DT) have achieved impressive accuracies surpassing 99%, followed by Neural Networks (NN), Support Vector Machines (SVM) and Random Forest (RF) with accuracy ranging from 91% to 96%. Long Short-Term Memory (LSTM) yielded superior Area Under Curve (AUC of 0.836) for fuel prediction, in comparison to Support Vector Machines (SVM) and Neural Networks (NN). Gated Recurrent Unit (GRU) represents fine-grained accurate fuel-prediction methods with accuracy comparable to Long Short-Term Memory (LSTM). Prominent research gaps identified during this study are the lack of a comprehensive approach covering all aspects related to safety, fuel economy, the scope of improvement in real-time driving risk assessment at appropriate granularity level, missing effective and engaging driving feedback, dealing with uncertain and incomplete driving events, driver's personal traits affecting driving safety and fuel-economy. The review will help in establishing the readiness of automation of driving analysis for reinforcement of eco-safe driving for various kinds of vehicles plug-in hybrid vehicles, hybrid electric vehicles, electric vehicles, and self-driving cars.

ARTICLE HISTORY

Received	:	22nd Sept 2022
Revised	:	09th June 2023
Accepted	:	24th June 2023
Published	:	04th July 2023

KEYWORDS

Crash prediction,
Driving behaviour,
Driving risk,
Onboard telematics,
Machine learning

1.0 INTRODUCTION

The automotive industry is revolutionized with innovation in communication technologies. Nowadays, sensors in connected vehicles generate massive driving data in real-time, which can be potentially analyzed and processed to generate meaningful inferences or advice for drivers in real-time. The real-time advice may help the driver avoid dangerous driving, maintain fuel economy, or assist the driver in maneuvering [1]. Driving behavior can be adapted towards safe and economical driving via various mechanisms such as issuing alerts, visual and audio feedback, and using gaming and incentives-based schemes. Advanced Driver-Assistance Systems (ADAS) provide a real-time human-machine interface to assist drivers with vehicle safety and eco- driving and warn the driver if the driver deviates from safe or fuel-economical driving behavior [2]. Driving safety and fuel economy are critical aspects of driving [3]. The driving practices that help to reduce fuel economy while adhering to safety norms are collectively referred to as eco-safe driving. Capturing driving behavior in real-time is critical to identifying dangerous driving and improving fuel efficiency. Disastrous situations like crashes can be avoided by controlling driving behavior. The driver profiling information along with their driving behavior risk attributes, can be utilized to assist the driver in car-following, lane-changing, and steering control scenarios [4-5]. The driver's physiological features (eye blink, head movement, facial expression, pulse rate, electrocardiogram (ECG), electroencephalogram (EEG), Electrooculography (EOG)) help in detecting driver's states, namely fatigue, sleepiness, drunken state, or distracted state [3] [6-7]. The appropriate advice can be issued to alert the driver after identification of the driver's state. The techniques mentioned help in driver profiling based on the usage of social characteristics such as age, cognitive abilities, route preferences, and trip characteristics and customize the driving assistance according to the driver's profile [1] [8-10]. The other approaches focus on improving fuel efficiency by controlling driving behavior [11-12]. These approaches helped in predicting fuel consumption taking driving style, driving

***CORRESPONDING AUTHOR | S. Mittal** | sangeeta.mittal@jiit.ac.in

behavior and contextual environment variables as a parameter via machine learning models, deep learning models and custom models. Incentive-based or gamification mechanisms maintain continuous user engagement and facilitate eco-safe driving. Incentive-based or gamification mechanisms maintain continuous user engagement and facilitate eco-safe driving.

In this work, existing approaches concerning the correlation of driving behavior, socio- economic characteristics or effect of congestion, and road type with safe driving and eco-driving have been compiled and reviewed. The paper is organized as follows. Section 2 details the services, technologies and applications associated with connected cars. Section 3 evaluates the various parameters such as driving behavior, social, contextual, and physiological parameters for driving behavior modelling in the existing state-of-the-art. Section 4 elaborates on state-of-the-art computational models for vehicle crash avoidance, driver profiling, driving distraction/fatigue/sleep detection, or avoidance techniques. Section 5 mentions computational models for fuel prediction based on driving behavior attributes, driving assistance for fuel-economical driving, and optimizing vehicle powertrain model to save fuel consumption. Section 6 elucidates approaches based on eco-safe driving and driving assistance or motivational approaches. Section 7 mentioned about limitations and research gaps of the existing state of the art. Section 8 concludes the review of state- of-the-art.

2.0 RELATED WORKS

This part provides a brief overview of applications, services and technologies that can connect a connected car to other devices via the Internet of Things (IoT) based on various connectivity mechanisms such as Bluetooth, cellular network, and satellite connectivity [12-13]. These can provide multiple applications such as infotainment, remote-diagnostics, convenience, parking assistance, roadside assistance and advanced driving assistance as shown in Figure 1.

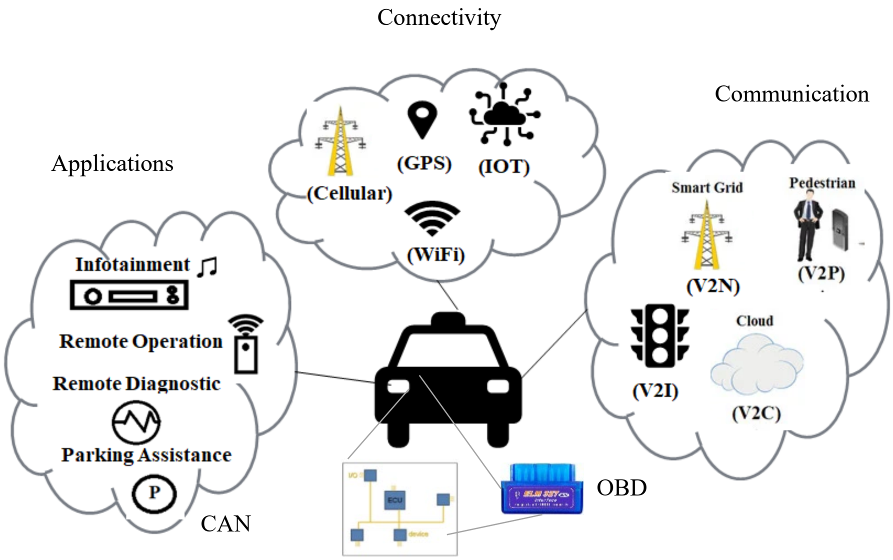

Figure 1. Connected car architecture

Connectivity Mechanisms - These cars can receive satellite data from radio stations to provide infotainment services. Bluetooth and Wi-Fi connectivity enable users to connect their mobile to the car's head unit and do simple activities such as attending phone calls, connecting to music apps, and navigation. Most of the connected cars also come pre-fitted with onboard devices and a built-in embedded sim (e-sim) for cellular communication. The onboard device (OBD) is connected to Electronic Control Units (ECU) that control the powertrain, chassis, steering, and vehicular functions via Control Area Network (CAN) [12]. The built-in e-sim helps to send data gathered via OBD, vehicle sensors and ECUs to cloud-based backend IoT applications.

Applications - Some convenience features enabled via connected cars include remotely executing the car's functions such as start, stop, lock, unlock, switching AC ON. Communication Mechanisms - Connected cars are enabled via different technologies such as vision/camera systems, onboard sensors, vehicle data networks and V2X communication where X may represent infrastructure, network, cloud, pedestrian, or vehicle. Depending on the underlying application, vehicle-to-cloud (V2C), vehicle-to-infrastructure(V2I), vehicle-to-network (V2N), and vehicle-to-pedestrian (V2P) communication takes place [13]. V2X communication is based on either IEEE 802.11p (5.9 GHz) or LTE-based communication mechanisms. IEEE 802.11 communication has also become a popular and standardized method for broadcast or multicast messages sent by applications meant for traffic regularization, cooperative routing, and communication with roadside infrastructure units. Communication of vehicle data to the real world happens via different kinds of interfaces based on V2X systems.

Telematics in the automotive domain is a technology that enables communication from the sensor to vehicle and vehicle to infrastructure via wireless or GPS-based communication mechanisms.

Table 1. Different types of sensors involved in connected car

Serial number	Type of Sensor	Description
S-1	Global Navigation Satellite System (GNSS)	Captures accurate latitude/longitude of the vehicle via Global position System (GPS) satellite
S-2	Wheel odometry sensors	Tracks wheel's speed and distance travelled by the vehicle is reflected in the odometer reading.
S-3	Throttle pressure sensor	Monitors the throttle opening rate and its relative position. It is located on spindle/shaft to monitor the position of the throttle.
S-4	IMU - Gyroscope	Measures the angular velocity in degrees per second.
S-5	IMU - Accelerometer	Measures vehicle acceleration displacement vector in x, y, and z directions
S-6	IMU - Magnetometer	Measures magnetic field strength on each axis to detect heading
S-9	Fuel level sensor	Measures the level of fuel in tank by monitoring the movement of float.

There are multiple types of sensors used in connected cars to gather various types of data related to location, direction of movement, lateral/angular speed, lateral/angular acceleration or deceleration, and throttle paddle statistics, as mentioned in Table 1. Controller Area Network (CAN) is a high-speed bus protocol to broadcast all the traffic to all nodes on a given bus. OBD (Onboard Diagnostics) is a device that connects to the car's OBD-II connector and gets access to values from Electronic Control Units (ECUs) and different sensors such as Global Position System (GPS), General Packet Radio Service (GPRS) via Controller Area Network (CAN). CAN bus helps to broadcast sensor-measured values for the vehicle's location, speed, engine RPM, fuel level to ECUs and OBD logger. Initially, the use of OBD devices was limited to the vehicle's health diagnosis and reporting any malfunction indicator. OBDs retrieve raw data gathered by electronic control units by connecting to CAN. The following are different in-vehicle sensors that may be installed in a connected vehicle; and can be referred to in Figure 2.

i. GNSS provides accurate vehicle location via GPS satellite. The accuracy of location mentioned by the GNSS system may get affected by signals being obstructed by monumental buildings and mountains. If the GPS signal is unavailable, then the previous location, acceleration, and direction value are utilized to compute the estimated current vehicle's location.
ii. Wheel odometry sensors track the wheel's speed, and the distance travelled by vehicle is reflected in an odometer reading.
iii. Throttle pressure sensor measures the throttle opening rate and its relative position. The position of the accelerometer pedal is sent to the engine control unit for controlling the power supplied to it.
iv. Inertial Measurement Unit (IMU) - accelerometer, gyroscope, magnetometer combines multiple outputs from accelerometer, gyroscope and magnetometer and tracks the movement and angular orientation of the vehicle. The accelerometer embedded in IMU, measures vehicle acceleration displacement vector in x, y, and z directions. The gyroscope measures angular velocity in degrees per second. The magnetometer measures magnetic field strength on each axis and helps to detect heading.
v. Fuel level sensor measures fuel-level in the fuel tank by combining the value received from fuel level sensor with GPS values and helps to calibrate fuel consumed within each sampling time.

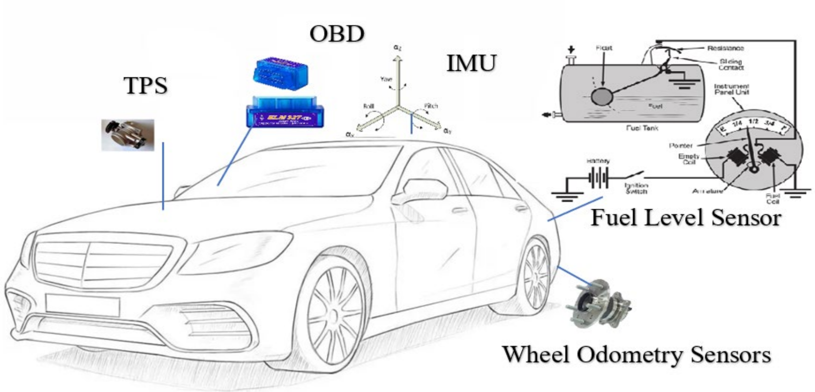

Figure 2. In-vehicle sensor placement

Electric and hybrid vehicles are gaining popularity due to several advantages over traditional internal combustion engine (ICE) vehicles. These include environmental benefits with lower emissions, improved energy efficiency, reduced dependence on fossil fuels, potential cost savings, technological advancements, and government incentives [24],[36]. An electric vehicle (EV) relies on electric motors or traction motors for propulsion, utilizing regenerative braking to capture and utilize lost kinetic energy during braking. Through regenerative braking, the EV's electric motor acts as a generator, absorbing the vehicle's motion energy and converting it into electrical energy during deceleration or when the vehicle slows down. This energy is then stored in the vehicle's batteries. In contrast, mechanical braking in traditional vehicles converts motion energy into heat and friction. The implementation of regenerative braking in EVs improves energy efficiency and contributes to an extended driving range by harnessing and utilizing energy that would otherwise be wasted during braking.

Hybrid Electric Vehicles (HEVs) combine both gasoline/diesel engine and electric batteries for power. Similar to EVs, HEVs also utilize regenerative braking to capture and store energy during braking. However, HEVs cannot be charged directly from external battery charging devices. On the other hand, Plug-in Hybrid Vehicles (PHEVs) can be connected to an electric grid to recharge their batteries, similar to EVs. This allows PHEVs to draw power from the grid to charge their batteries in addition to utilizing regenerative braking.

3.0 DATA PREPARATION TECHNIQUES FOR MODELLING ECO-SAFE DRIVING

Various steps of data processing of vehicle data can be briefly explained in Figure 3. The vehicle data for GPS, accelerometer, velocity, and fuel level is sensed via different sensors. Values received from vehicles are broadcasted by CAN to various ECUs and vice-versa. Information received is used to detect driving behavior events, and contextual conditions. Contextually mapped information is used to develop various computational models for safe and fuel-economical driving. The first step for modelling driving behavior includes suitable parameter extraction from data. Extracted parameters are then taken as input to computation models for real-time driving modelling. In this section, parameters selection and modelling techniques proposed for this domain have been reviewed.

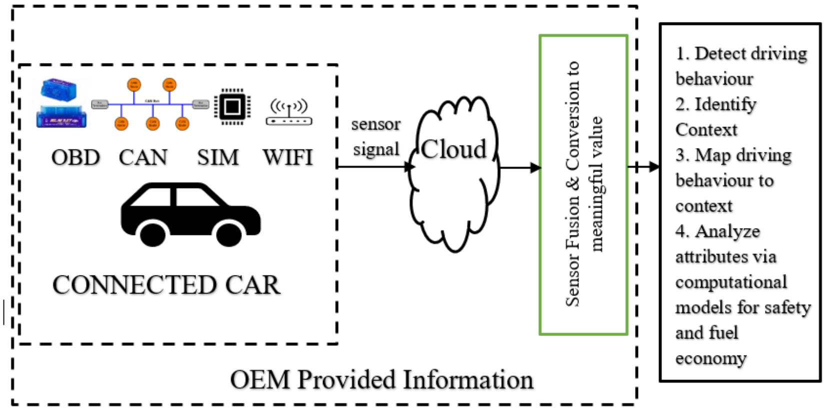

Figure 3. Process flow for data processing

3.1 Parameters used for Modelling Fuel-Economical and Safe Driving Behavior

Driving is an abstract phenomenon. Capturing driving behavior includes the extraction of suitable attributes from raw driving data. Extracted parameters are then utilized for developing computational models for real-time driving modelling. Table 2 summarizes various input attributes and factors considered in the literature for the abstraction of safe and eco-driving. Attributes in various publications can be categorized into different types. Vehicle attributes help to consider vehicle-specific attributes such as engine rpm and torque that affect power and fuel consumption. Socio-economical attributes consider multiple parameters such as age, cognitive skills, gender, and economic state [17],[32]. Few approaches discussed the driving behavior of old drivers, their cognitive skills, and route familiarity [10],[17]. One of the approaches suggested that socio-economic characteristics affect driving behavior during off-peak hours [17]. These approaches offered older adults feedback mechanisms to help them improve their driving abilities and decisions. Few approaches used trip attributes such as frequency, length, and duration of trips for driver characteristics profiling. These approaches helped to reflect the driver personality or preferences for travelling.

Many approaches considered environmental contextual attributes such as road type (road surface type, number of lanes, road curvature, road gradient, presence of speed bumps and portholes, road conditions as wet, dry or snowy) weather conditions (clear, sunny, foggy, rainy, snow, windy thunderstorm), time of travel (day, night) and road congestion to refine the computational model [4],[7-8],[10-11],[15]. Factors such as weather conditions, road surface conditions (e.g., wet, icy, or uneven surfaces), visibility, traffic congestion, and the presence of pedestrians or obstacles can significantly impact the safe speed and acceleration limits on the road. For example, driving at a certain speed on a dry and straight road may be considered safe, but the same speed could be dangerous on a wet or icy road due to reduced traction. Adhering to local traffic laws and regulations is also essential for safe driving practices in different road conditions. Speed limits are established based on various factors, including road design, traffic flow, and the surrounding environment, to ensure the safety of all road users. Similarly, rapid acceleration in heavy traffic or crowded areas may pose a higher risk of accidents compared to open highways with minimal traffic.

Table 2. Study of attributes for the abstraction of safe and eco-driving.

		Vehicle attributes		Socio-economical attributes	Trip attributes		Environmental context				Driving behavior					Physiological attributes	
Reference	Purpose of model - Safe/eco driving	Fuel consumed	Engine speed or torque	Age/gender/cognitive skills	Trip length	# of trips	Time of travel	Road information	Traffic flow	Weather information	Harsh/frequent paddle	Driving speed	Idling time	Cruising time	Turn behavior	Eye blink/closure	ECG/EEG/EOG
[1]	SAFE	√	√								√	√					
[2]	ECO	√									√	√	√		√		
[4]	SAFE								√		√	√			√		√
[6]	SAFE			√						√	√	√	√		√		√
[7]	SAFE								√		√	√			√	√	√
[8]	SAFE							√			√	√	√		√		
[9]	SAFE				√	√	√				√	√			√		
[10]	SAFE			√	√	√	√	√		√	√	√					
[11]	ECO	√	√		√			√	√		√	√	√				
[12]	ECO	√	√								√		√				
[14]	ECO	√	√			√					√	√	√				
[17]	SAFE			√	√	√	√	√			√				√		
[19]	SAFE						√	√	√		√	√			√		√
[20]	SAFE						√	√								√	√
[21]	ECO	√							√		√	√	√	√			
[22]	SAFE		√								√	√	√		√		
[24]	ECO	√	√								√	√	√				
[25]	ECO	√										√					
[26]	SAFE						√	√								√	√
[27]	SAFE										√		√				
[27]	SAFE										√		√		√		
[28]	SAFE										√	√	√	√	√		
[30]	ECO	√	√								√	√	√		√		
[31]	ECO	√							√		√	√		√			
[32]	ECO	√		√	√	√	√		√		√	√		√			√
[33]	SAFE										√	√	√		√		√
[34]	ECO	√	√					√			√	√	√				
[35]	ECO	√	√		√	√	√	√	√		√	√	√				
[36]	ECO	√									√	√	√	√			
[37]	ECO	√	√								√	√	√		√		
[38]	ECO	√						√			√	√	√				
[39]	ECO	√									√	√	√				
[40]	ECO	√	√								√	√	√				
[41]	ECO	√	√								√	√	√		√		
[43]	SAFE						√	√			√				√		√
[44]	SAFE								√		√	√					

Weather conditions can have a significant impact on safe driving [6],[22],[47],[18]. Different weather conditions present unique challenges and hazards that drivers need to be aware of and adapt to, such as maintaining a safe distance from other vehicles, being attentive, and anticipating potential hazard:

i. Rain; rainfall reduces road traction, making the road surface slippery. Reduced visibility and the need for windshield wipers also add to the challenges. It is important to drive at a reduced speed, increase the following distance, and use headlights in the rain to improve visibility.
ii. Snow and ice; snow and ice create extremely slippery road surfaces. Traction is greatly reduced, leading to longer stopping distances and difficulties in maintaining control. It is also crucial to leave ample headway space between vehicles and brake gently to prevent skidding.
iii. Fog; fog reduces visibility and makes it challenging to see other vehicles, pedestrians, or road signs. Drivers should use fog lights or low beam headlights, maintain a safe distance from the vehicle ahead, and drive at reduced speeds to have enough time to react to hazards.

Most of the approaches used negative driving behavior attributes, namely harsh acceleration/deceleration, overspeed, engine idling, sharp turn behavior, and steering angle as parameters. Cruising has been identified as recommended driving to improve fuel economy [31-32]. Driving assistance approaches consider multiple attributes such as lane-keeping, longitudinal/lateral lead distance from the preceding vehicle, time headway and relative speed between vehicles. Driver's response time and the mechanical capability of vehicle dictate recommended lead distance and vehicle speed. The use of physiological attributes is being used to detect sleep, mind-wandering, or driving stress, fatigue, nausea, headache [32-33],[43]. The driver is alerted after detecting the driver's abnormal driving behavior or driving state.

Vehicle data for GPS, accelerometer, velocity, and fuel level is sensed via different sensors. Values received from vehicles are broadcasted by CAN to various ECUs and vice-versa. OBD-II gathers information from the OBD-II port and transmits the same to the cloud via a cellular connection or Wi-Fi connection. Raw sensor signals are converted to actual attribute values by applying sensor fusion and pre-processing techniques by OEM. The processed attributes received from OEM are analyzed to extract attributes corresponding to driving behavior, and environmental context and mapped together. Contextually mapped attributes are processed via different computational models to predict fuel consumption, assess driving risk, identify anomalies, and provide driver assistance to plan further driving actions. Table 3 elaborates on various features extracted from the datasets and their mapping to respective sensor sources.

Table 3. Mapping of features to their source of sensor

Serial number	Feature name	Source sensor
1. Trips by driver	Distance covered between two subsequent data samples	GNSS, wheel odometry, TPS, APS IMU
	Duration between two subsequent data samples	Timestamp
	Speed of vehicle between two subsequent data samples	GNSS, wheel odometry, TPS, APS
2. Vehicle parameters	Mileage observed between two subsequent data samples	Fuel level sensor
3. Environmental context parameters	Road information for a particular sample	GIS, Camera
	Weather conditions for a particular sample	Weather database from local weather authorities
	Traffic congestion for a particular sample	Heremap API
	Travel time for a particular sample	Timestamp difference
4. Driving behavior parameters	Distance and duration for harsh acceleration or deceleration	Throttle pressure sensor, IMU
	Overspeed distance and duration	Wheel odometry, Here map API
	Sharp turn distance and duration	Angular velocity using steering sensor and gyroscope

Different attributes mentioned above affect driving safety and fuel economy. The quality of gathered attributes also depends on the data collection mechanism. Section 3.2 discusses different mechanisms for data collection.

3.2 Mechanisms for Capturing Raw Driving Data

Real-time driving information is retrieved via drives on test routes or routes under naturalistic conditions. Test drives on known routes by real drivers are referred to as Real-time Test Driving (RTD). Data collection under naturalistic conditions is labelled as Real-Time Naturalistic Driving (RND). RND helps in capturing real-time naturalistic driving behavior with natural road and traffic maneuvering without being obtrusive. Real-time data collection for driving attributes can be done through in-vehicle sensors on OBD devices, smartphone-based sensors, cameras, or body sensors to monitor driving, contextual, vehicular, trip, and physiological features. As shown in Figure 4, driving simulators have a steering wheel, an engine simulator with control over acceleration or deceleration paddle, a camera to study the driver's facial features or eye movement, body sensors to study psychological attributes, and a driving console to visualize the trajectory [22].

The simulator enables driving scene visualization and simulates traffic. The driver can visualize the trajectory via a monitor display. The use of the simulator may cause simulator sickness to participants and may result in fatigue, headache, dizziness, blurred vision, or nausea, directly affecting the participant's attention span and response time. Furthermore, the results based on the simulator cannot be equivalent to those resulting under naturalistic conditions representing real-time road manoeuvres, traffic, and weather conditions. Other types of simulators include feeding standard driving cycles to chassis dynamometers to study various vehicle-related parameters. The driving simulator can be fed with captured driving cycles generated on test drives referred to as simulator-based test drive (STD) [23],[25]. Real-time driving data fed to the simulator is labelled as simulator-based naturalistic driving (SND) and has been utilized in [32], [36-37].

Figure 4. An example of a driving simulator [48]

Real-time data collection under naturalistic conditions (RND) is an ideal way to get unbiased data for applying machine learning based computational model. Real-time test drive data (RTD) provides a mechanism to standardize test drive routes for easy comparison of different algorithms. Simulator-based naturalistic driving (SND) and simulator-based test drive (STD) provide a mechanism to simulate real-time driving conditions under a simulated environment. The following are various kinds of data collection modes, their sources, and other statistics of data used in existing studies (as given in Table 4).

i. *Data Collection via OBD Sensors* helps to gather telematics data from various sensors and Controller Area Network (CAN) embedded in vehicles. OBD captures the vehicle's location, speed, engine RPM and fuel level via OBD-II port of the car.
ii. *Data collection via smartphones* is done using in-built smartphone sensors such as accelerometers, magnetometers, gyroscopes, and GPS [36],[43]. Though smartphones are easier and fast, smartphone-based data may be erroneous due to frequent orientation changes, low sampling rate, filtering mechanism, positioning, battery life, and internet connectivity.
iii. *Data Collection via camera or body sensors* helps to capture physiological features such as eye movement, facial expression, head movement, ECG, EEG, EOG, so that the driver's state of mind can be determined [30],[36].
iv. *Data collection via OBD* is more reliable than data collected via smartphone sensors as it is firmly fitted within vehicle. Due to this, it does not lead to noisy data and allows data to be captured seamlessly. Data collection via camera or body sensors might be intrusive to drivers due to privacy concerns.

Table 4. Driving data collection mechanisms

Reference	Purpose of study	Driving data	# of drivers	Source of data	Availability of dataset
Real-Time Naturalistic Driving (RND)					
Abdennour et al., 2021 [1]	SAFE	46 km (round-trip) of 23 hours duration between Korea University and SANGAM World Cup Stadium for multiple trips	10	OBD	Publicly available
Marzet et al., 2021 [11]	ECO	Normal driving - 3,573 kms Eco- Driving - 3190 kms	11	OBD, smartphone sensors	Madrid Spain dataset, not available publicly
Baetz et al., 2020 [14]	ECO	Real-time driving data of 495 drivers	495	OBD	Not available
Yao et al., 2020 [21]	ECO	15 days driving data of 20 drivers	20	OBD, smartphone sensors	Available on request
Liu et al., 2020 [30]	SAFE	485 video clips of the driving stop event with causes of traffic light (stop4light), pedestrian (stop4ped), stop sign (stop4sign), and congestion (stop4cong).	NA	Camera, body sensors	Honda Research Institute driving dataset, available on request
Osman et al., 2019 [22]	SAFE	200 driving events for 373 drivers, divided as 95 drivers calling, 96 drivers texting, 84 drivers engaging in conversation with the adjacent passenger, and 98 drivers not engaged in any secondary task.	373	OBD	SHRP2 NDS database, available on request

Payyanada et al., 2019 [17]	SAFE	Driving data of drivers aged 65 years and older	29	Smartphone sensors	Not available
Bian et al., 2018 [9]	SAFE	198 driving observations containing 8 behavior-related features and 1 risk-level label.	198	OBD	China Insurance Company data – not available
Yin et al., 2017 [15]	SAFE	Three rounds, each lasting 10 min, total of 1,800 driving moments (1500 training/84 test)	1	OBD	Not available
Yu et al., 2017 [28]	SAFE	6 months of driving data	20	Smartphone sensors	Not available
Real-time Test Drive (RTD)					
Abukhalil et al., 2020 [12]	ECO	Test drive of 3 sedan vehicles, each on 66 km hilly route for 40 minutes	3	OBD	Not available
Hoffman et al., 2019 [38]	ECO	Test trip data of 7332 driving cycles on 21 routes.	331	OBD	Not available
Pekkanen et al., 2018 [4]	SAFE	Test driving data for car-following driving scenarios, 40 Real Car + 37 Simulator	40	OBD	Publicly a vailable https://zenodo.org/record/1341081#.Ye-crPjhU2w
Ferreira et al., 2017 [8]	SAFE	4 trips of duration 13 mins on test drives	2	Smartphone Sensors	Not available
Hsu et al., 2017 [40]	ECO	Total 30 routes of 45 km on various road types	1	OBD	Not available
Simulator-based Naturalistic Driving (SND)					
Ozkan et al., 2021 [25]	ECO	45 minutes of data segmented into three parts (NGSIM - US Highway 101 Dataset)	9	OBD	NGSIM US-101 dataset, a vailable
Guo et al., 2021 [24]	ECO	Past standard drive cycles fed to a simulator for urban and highway	NA	OBD	Not available
Zhang et al., 2020 [21]	ECO	Driving behavior, demographic attributes of 66 drivers	66	OBD	Available on request
Gadde et al., 2019 [37]	ECO	6000 km driving data of urban and highway driving in 20 shifts	NA	OBD	Not available
Branislav et al., 2019 [34]	ECO	Real-time driving cycles (WLTP) with vehicle driving time of 600 seconds.	NA	OBD	Not available
Barua et al., 2018 [26]	SAFE	540 drives on 6 routes	30	OBD	Not available
Simulator-based Test Drive (STD)					
Filippos et al., 2021 [31]	ECO	Data of 4156 trips taken by 100 drivers.	100	OBD	Proprietary to OSeven Telematics, London, UK
Mubarak et al., 2020 [35]	ECO	1100 km drive by 2 drivers on a fixed highway route	2	OBD	Not available
Ping et al., 2019 [2]	ECO	30 passenger cars driven by 202 drivers on a predefined route	202	OBD	Experimental data, not available
Lasocki et al., 2019 [39]	ECO	12 driving cycles across - Vehicle 1, Vehicle 2	2	OBD	https://www.anl.gov/, Proprietary
Darji et al., 2018 [6]	SAFE	45 mins of driving data for day and night on urban and highway roads	21	OBD	Included within manuscript
Jacobé et al., 2018 [23]	SAFE	driving data of 110 minutes	20	Camera, body sensors	Not available
Kumada et al., 2018 [27]	SAFE	25mins of car-following driving	40	OBD	Publicly available
Hu et al., 2017 [33]	SAFE	36 vehicle test data samples provided by Ford with 1080 km per sample	NA	OBD	Not available

3.3 Feature Engineering

Feature engineering transforms and maps raw data to a form that provides meaningful interpretation and makes it simple for processing by machine learning models. Some of the advantages of feature engineering are listed below:

i. feature engineering helps in reduction of dimensionality and reduces overfitting,
ii. models using processed features are less complex and faster to train,
iii. better features result in more accurate models,
iv. it also helps in noise removal or reduction.

Feature engineering can be classified broadly as Filter method, Wrapper method, and Embedded method. Filter methods help to select features independent of machine learning algorithm by applying statistical tests that correlate feature with output. There are different ways of correlation methods, such as Pearson correlation, linear discriminant analysis, analysis of variance (ANOVA), and chi-square for continuous or categorical features.

i. Pearson correlation method measures linear dependence between two continuous variables, and its value ranges from -1 to 1,
ii. linear discriminant analysis helps to find the linear combination of categorical features that helps in segregation of different classes,
iii. ANOVA correlates categorical independent features with one continuous dependent feature,
iv. chi-square evaluates correlation among categorical features based on their frequency distribution.

Correlation and scatter diagrams were used to correlate driving behavior parameters with fuel consumption in [14],[18],[31]. Few approaches used Pearson correlation coefficient (PCC) to determine correlation of vehicles and driving behavior parameters with fuel consumption [2],[21]. Other approaches used PCC to determine the relationship between driving behavior and driving risk parameters [8],[12]. Though the filter method acts as a simple pre-processing step, but is unable to remove multicollinearity. Wrapper method identifies the subset of features and evaluates their efficacy for training a machine learning model. Based on the iterative evaluation, a final subset of features is selected for a computational model. There are three methods for iterative feature selection.

i. Forward selection iteratively adds the best features one by one until the addition of any more features does not lead to performance improvement.
ii. Backward elimination starts with all features and iteratively removes the least significant feature until the performance of model no longer improves by eliminating more features.
iii. Recursive feature elimination repeatedly creates different models by eliminating different features and ranks different features based on the model's performance. Best-ranked features are shortlisted as part of this technique.

Stepwise forward feature selection is used to shortlist features out of 60 features based on the threshold significance level of 0.005 in [6]. Exploring all possible combinations iteratively makes this approach computationally intensive. Embedded methods combine the advantages of filter and wrapper methods to identify the best features. They also apply regularization techniques to reduce overfitting. Some examples of these methods include LASSO regression, RIDGE regression, and Random Forest Regressor. Random Forest Regressor was used to shortlisted features for the car-following model based on relative importance of feature to predict speed of the preceding car [25]. Random Forest Regressor is an ensemble technique that helps to identify relative feature importance by averaging values of different decision trees and thereby producing more generalized result. Filter methods use statistical methods to shortlist features, whereas wrapper method use cross validation on actual ML model. Due to this reason, filter methods are faster in comparison to wrapper methods as they need not train ML model. On the other hand, wrapper methods are computationally intensive. Wrapper methods may be prone to model overfitting as these are trained on ML model. Embedded methods balance the advantages of filter and wrapper methods and avoid model overfitting by applying the regularization technique.

Few dimensionality reduction techniques used principal component analysis (PCA) to shortlist important principal components [1],[22]. One of the approaches applied a fast Fourier transform along with PCA to reduce 15 principal components for driver identification based on his driving behavior [1]. The other approach used PCA to find out 10 principal components that helped to classify driers' engagement in secondary tasks [22]. PCA-based techniques reduce dimensionality by linear transformation of features. PCA-based techniques are not suitable for feature compression in the case of non-linear data. Autoencoder is an effective dimensionality reduction technique for non-linear data that helps to retain only the most salient features due to its inherent architecture [46].

4.0 COMPUTATIONAL MODELS FOR SAFE DRIVING

Driving risk assessment is the first step towards crash avoidance and safe driving. Aspects affecting safe driving include driver's state detection related to fatigue, sleep, distraction, and identification of harsh and frequent negative driving behavior. Adapting driving according to the environmental context and improving the driver's personal driving style can go a long way in the direction of safe driving. In the existing literature, data from connected cars have been utilized to propose ML-based computational models as detailed in Table 5.

Table 5. Computational models for safe driving

Reference	Purpose of study	Computational model for analytics	Evaluation metrics
Shallow Machine Learning based approaches			
Darji et al., 2018 [6]	To find about driver's hazardous state using driving style, physiological state, and vehicle movement statistics	Logistic regression, SVM, ensemble boosted DT	Accuracy rate alert vs drowsy - 98.8%, cell phone use - 82.3%, dense vs light traffic -91.4%, snowy vs clear - 71.5%
Osman et al., 2019 [22]	Detect secondary task involvement, and find the type of secondary task (calling, texting, talking, driving)	Hierarchical classification DT – Level 1 RF – Level 2	Accuracy rate
Bian et al., 2018 [9]	To develop vehicle insurance, pricing model using driving behavior	SVM, NN, bagging	Accuracy rate, Kappa score, MAE, RMSE

Gwakl et al., 2018 [7]	Detection of driver's drowsiness	Logistic regression, SVM, KNN, RF.	Accuracy rate, p recision rate, and recall rate
Barua et al., 2018 [26]	To detect the state of driver sleepiness based on contextual and physiological information	KNN, SVM, RF, and case-based reasoning	Accuracy rate, precision rate, recall rate
Kumada et al., 2018 [27]	To detect mind wandering for driver	SVM, KNN, decision tree, ensemble learning, Naïve Bayes	Accuracy rate, precision rate, recall rate, Kappa, Friedman
Deep learning based approaches			
Abdennour et al., 2021 [1]	Detect driving signature via pattern identification	Neural network	Accuracy rate, precision rate, recall rate, F1 score, Kappa score, AUC score
Jacobé et al., 2018 [23]	Prediction and detection of driver drowsiness on simulated road types and traffic flows	Adaptive artificial neural networks (Ad-ANN)	RMSE
Ferreira et al., 2017 [8]	Derivation of driving behavior attributes and driver profiling	BN, ANN – Multi-Layer Perceptron, RF, SVM	Area Under Curve (AUC)
Yu et al., 2017 [28]	Derive six types of fine-grained driving behavior using acceleration/orientation patterns	SVM, NN.	Accuracy rate, precision rate, recall rate, false positive rate
Liu et al., 2020 [30]	Predict driving behavior by learning human causal reasons	3DResnet-TRB	Average precision
Hu et al., 2017 [33]	Detect three typical abnormal driving conditions, fatigue/drunk, recklessness, and use of phone while driving	Cerebellar Model Articulation Controller (CMAC) – NN	Coefficient of variation (CV)
Miscellaneous approaches			
Payyanadan et al., 2019 [17]	Study the effect of age, cognitive abilities, and route familiarity for driving behavior profiling	Bayes conditional probability, generalized linear mixed-effects regression model	Standard error, t- value, confidence interval
Pekkanen et al., 2018 [4]	To design a car-following model taking driver's attention span, event's uncertainty, perception response, acceleration control as inputs	Cognitive modelling and state transition model	Spearman correlation – p-value
Yin et al., 2018 [15]	Detection of dangerous driving behavior using driver, vehicle, and lane attributes	Fuzzy PSO model	MAE, MSE

4.1 Shallow Machine Learning-based Approaches

ML-based approaches applied for modelling safety-related aspects have been discussed in this section. Few approaches applied machine learning classification algorithms to detect driver's distraction due to secondary tasks [22],[27]. Osman et al. proposed a bi-level hierarchical classification methodology based on Decision Tree (DT) and Random Forest (RF) [22]. Decision trees were recommended for detection of distraction at the first level, whereas random forest was found to be optimal for identifying secondary tasks. In [27], the author compared various supervised ML algorithms such as support vector machines, k-nearest neighbour classifier, ensemble learning, Naïve Bayes, decision tree and recommended building a participant-specific mind wandering detection model.

Multiple approaches detected driver's sleepy state or drowsiness by analyzing driving behavior [15],[26]. To detect reasons for unsafe driving, Darzi et al. studied parameters related to driver's style, physiological characteristics, weather conditions and vehicle movement statistics on the simulator, WYOSIM [6]. The reason for dangerous driving was either due to intrinsic factors(sleep deprivation) or extrinsic factors(adverse weather conditions). As per author, weather conditions such as snowy, foggy can affect visibility and road friction differently hence affecting driving safety. The DT came out to be most accurate while identifying alert vs drowsy state, whereas Logistic Regression helped in classification such as phone usage and driving on a congested route. SVM could accurately identify weather conditions such as snowy vs clear. Authors in [15] compared logistic regression, SVM, k-nearest neighbour (KNN), and RF to detect driver's drowsiness and RF achieved the maximum accuracy based on vehicle data, driving data, road type, weather information and traffic flow. In [26], Support Vector Machine(SVM) was recommended for binary classification (with 93% accuracy) as follows.

i. Driver profiling for safe driving - Machine learning models based on SVM, Bayesian approaches, and neural networks for driver profiling using smartphone sensors have been discussed by Bian et al. [9]. Knoefel et al. suggested a naturalistic driving framework that focused on the fact that age-related cognitive disorder affects driving ability and choices made by the driver [10]. The study also included how drivers adapted their behavior as per weather or seasons and road conditions(wet, ice, snow).
ii. Change in driver behavior due to adverse weather conditions – Das et al. applied a logistic regression model to study the effect of adverse weather conditions on drivers' lane changing behavior under naturalistic conditions

[47]. Drivers demonstrated poorer lane-keeping abilities under low visibility conditions of adverse weather, making standard deviation of lane position (SDLP) 1.37 times as in normal scenarios.

4.2 Deep Learning-Based Approaches

Abdennour et al. proposed a deep learning Residual Convolutional network (DL-RCN) for driver identification and profiling [1]. Liu et al. suggested the Temporal Reasoning Block model(TRB) based on Convolutional Neural Networks(CNN) that could predict accurate driving behavior by learning human causal reasons [30]. A neural network-based model to detect three types of abnormal driving behavior, namely reckless driving, use of phone while driving, and fatigue/drunk driving has been proposed in an approach by Hu et al. [33]. This model uses real-world test data and analyses it using a numerical model and considers three driving style classifications under city and highways roads; it does not consider traffic congestion and travel time. It considered ‘stopping’ a critical unsafe driving behavior considering various causes such as traffic lights, pedestrians, and congestion. TRB with 3DResNet demonstrated the highest accuracy of 86.3% as per numerical and visual evaluations. Ferreira et al. applied driver profiling and derived a driving safety score by applying machine learning to driver behavior attributes [8].

Few deep learning approaches helped in the identification of abnormal driving behavior [12],[28]. Jacob et al. proposed an adaptive artificial neural networks (Ad-ANN) approach to detect drivers’ drowsiness under various road types and traffic flow simulations [23]. RMSE was used to evaluate the Ad-ANN Model for accurate measurement. Bagging-based learning approach was more accurate than other approaches such as Naïve Bayes, SVM and NN for fine-grained driving pattern identification such as fast U-turns, weaving, sideslipping, turning with a wide radius, swerving, and sudden braking [28]. Though this approach warns users of fine-grained negative driving behavior, it does not predict driving crashes depending on negative driving behavior.

4.3 Miscellaneous Approaches

Traditional car insurance schemes assess risks only based on the type of car, mileage and usage and completely ignore how the driver is driving the car [9]. It was suggested to use Pay As You Drive (PAYD) over usage-based insurance (UBI). Behavior-centric insurance pricing based on driver behavior profiling has been proposed using attributes such as total mileage, driving duration in hours per month for night or weekday driving, average speed, overspeed, acceleration, deceleration, and sharp turn. But while assessing the driving data approach, however did not consider contextual information such as weather, road, and traffic conditions.

In Payyanadan et al., the driving behavior of old drivers affecting their choice of route depending on route familiarity and traffic congestion [17]. As per this study, familiarity with routes leads to decreased driver’s attention span. Due to this, the driver tends to over-speed, resulting in vehicle crashes, especially in adverse conditions such as congestion and at intersections. For collecting training data, drivers maintained a manual trip diary to record their observations regarding various trips taken by drivers. Diary-based data collection can be error-prone, especially for older drivers with cognitive issues. Moreover, the author considered driving risk based only on route characteristics such as distance and manoeuvres such as left and right turns. Studying the effect of route familiarity on risky driving events is limited to older drivers using simulated driving to provide navigational assistance to them.

In Yin et al. a fuzzy particle swarm optimization model was suggested that took vehicle and road type as inputs to estimate the intensity of dangerous driving [15]. Using SenseFleet approach, the driving score is calculated based on harsh acceleration/deceleration, over-speeding, and steering action within the context of current weather and time of day using smartphone-based sensors [22]. This paper does not consider road conditions and traffic congestion in the driving context. Moreover, route familiarity and driving habits include taking long trips/short speeds, vehicle speed, taking familiar routes or avoiding highway routes and linking it with driving behavior. Pekkanen et al. suggested a state-transition model for the car-following scenario and computed a safe distance according to the difference in acceleration of leading and following vehicles [4]. It is challenging to estimate the acceleration of leading and following vehicles in real-time under naturalistic driving conditions. The deceleration-based surrogate safety measure (DSSM) estimated the leading vehicle’s speed and maintained a safe distance from the leading vehicle [44]. A threshold for safe distance is decided based on the leading vehicle speed, the driving pattern for acceleration, and vehicle’s mechanical capability. It used the Next Generation Simulation (NGSIM) trajectory data. Chen et al. studied the effects of different weather conditions on traffic flow for a car-following model [18]. The study generated driving behavior data under different weather conditions using a driving simulator and fed this to traffic simulator. Observations indicated that different weather conditions have a significant impact on driving behavior and road capacity. Here are the key findings:

i. Heavy rain and fog; under heavy rain and fog, the average speed of vehicles decreased by approximately 7.6% to 27.5%, and there is an observed reduction in road capacity of around 11.1% to 20.5%. Reduced visibility and the need for cautious driving contribute to slower speeds during these weather conditions.

ii. Snowy weather; in snowy weather, the average speed reduction is even more significant, ranging from 19.2% to 45.6%. The road capacity reduction is also higher (43.7~71.1%) under snowy weather. This is attributed to both reduced visibility and the necessity of driving at slower speeds due to slippery road conditions. Drivers need to maintain larger headways to avoid collisions, resulting in lower traffic density.

These observations highlighted the impact of adverse weather conditions on driving patterns, including decreased average speeds, lower traffic density, and reduced road capacity. Drivers tend to adjust their behavior to ensure safety under challenging weather conditions, leading to slower and more cautious driving. It is crucial for drivers to be aware of these effects and adjust their driving accordingly to maintain safety on the road.

4.4 Safe Driving Approaches – Observations and Discussion

Few approaches focused only on driver profiling instead of driving risk analysis [1],[8],[17],[28],[30],[33]. DL approaches based on Residual Convolutional network (RCN) outperformed ML approaches such as SVM, and DT on accuracy evaluation in the same problem domain. RF, SVM, and DT have been popular approaches for driver's physiological state detection for stress and fatigue. On the other hand, DL methods such as Ad-ANN and ensemble techniques such as bagging, Naïve Bayes, boosted DT are also becoming prevalent nowadays for abnormal driving behavior and pattern identification.

Figure 5 shows comparative accuracy evaluation for various approaches based on ML and deep learning as claimed by the authors in their work. DL approaches based on RCN, CNN and LSTM achieved accuracy of more than 99%. Ensemble learning boosted DT algorithm also performed close to 99% accuracy, whereas NN achieved 96%. SVM achieved 93%-95% accuracy, whereas RF achieved an accuracy of 91% to 92%. Deep learning-based approaches in general outperform ML approaches as they can learn useful representations of input data and automatically extract features. ML approaches need to be retrained every time features are changed.

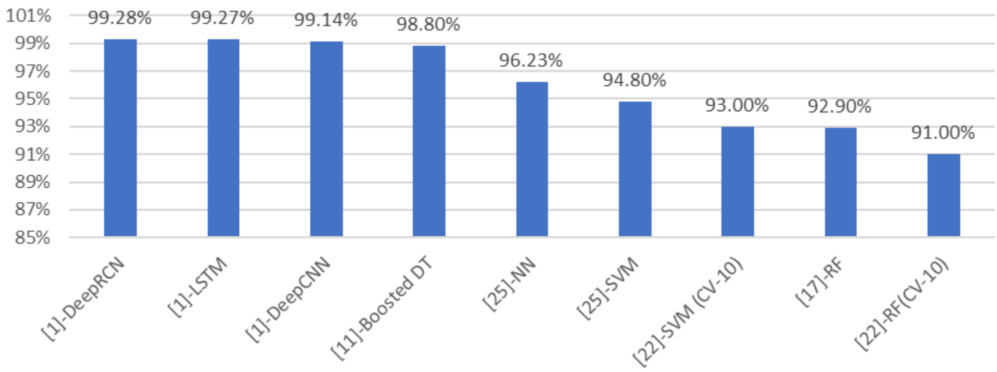

Figure 5. Accuracy evaluation for classification approaches

Some proposed methods that monitor drivers' physiological conditions via body sensors during driving are obtrusive in nature, and people may not refrain from doing so [7],[18],[23-25]. Driver profiling approaches are too coarse-grained and broadly classify drivers into three categories [8-10],[30]. Coarse-grained classification does not help in monitoring real-time driving risk. Some works are constrained and monitor driving risk for only car-following scenarios and help to assist drivers only in speed correction [4],[44]. Approaches in [9],[15],[42-44] help to assess dangerous driving behavior, but these approaches consider just one or two environmental conditions. There are others that are limited to driving pattern identification but do not correlate driving patterns to driving risk [28],[33]. Most of these approaches do broad-level driving risk classification but do not take any lead for crash prediction. A comprehensive approach that considers driving behavior, the driver's personality style, and all environmental conditions to assess driving risk at appropriate granularity are lacking.

5.0 COMPUTATIONAL MODELS FOR FUEL-ECONOMICAL DRIVING

Eco-driving is a set of driving practices that reduce fuel consumption, such as maintaining steady-state speed and avoiding harsh acceleration or deceleration. Eco-driving helps in resource conservation of limited fuel resources and reduces the resulting vehicular emission of carbon and air pollutants. In this section, different shallow machine learning and deep learning approaches were explored and addressed for fuel consumption prediction or driver profiling based on their driving practices and fuel consumption.

Table 6. Computational models for eco-driving

Reference	Purpose of study	Computational model for analytics	Evaluation metrics
Shallow Machine Learning based approaches			
Marzet, 2021 [11]	To study the reduction of fuel consumption post-eco-driving training	VSP model, K-Means Clustering	t-statistics comparison for pre and post-training
Abukhalil, 2020 [12]	To predict fuel consumption using Engine RPM and Throttle position as predictor variables	SVM, Polynomial Regression Model	RMSE, R-squared
Filippos, 2021 [31]	To study the effect of driving behavior profiles on fuel-pollutant reduction	K Means Clustering	t-value

Mubarak, 2020 [35]	To predict engine power and fuel consumption by analyzing driving behavior	Classification Model	Precision rate
Rios-Torres, 2019 [36]	To predict fuel economy using driving behavior using regression	Linear regression model	R Square, TI, VIF, RMSE, MAPE
Branislav, 2019 [34]	To predict fuel consumption using Engine RPM and Throttle position as predictor variables	3D regression model.	R-square value, error rate
Gadde, 2019 [37]	To frame the fuel prediction model using Driving behavior data	Machine learning-based OLS Linear model.	The root-mean-square error (RMSE)
Hoffman, 2019 [38]	To define regression model to predict fuel economy using driving behavior	Non-linear regression	Average relative ranking change
Hsu, 2017 [40]	To model ODI index to measure and improve fuel efficiency for a trip	Model tree based on M5' algorithm	MAPE, relative error, MAE
Deep learning based approaches			
Ping, 2019 [2]	Apply long-term historical data to learn short-term data for the prediction of fuel consumption	Long Short-Term Memory (LSTM) based model	Accuracy rate, area under curve
Guo, 2021 [24]	To optimize fuel consumption for Plug-In Hybrid vehicle	C/GMRES for fuel and battery optimization	Cost reduction percentage
Ozkan, 2021 [25]	To minimize fuel consumption using the speed profile of the preceding vehicle.	Gated recurrent unit non-linear model predictive control (GRU-NMPC)	Percentage efficiency gain, RMSE
Feedback based approaches			
Payyanadan et al., 2019 [17]	Study the effect of age, cognitive abilities, route familiarity for driving behavior profiling	Bayes conditional probability, Generalized linear mixed-effects regression model	Standard error, t-value, confidence interval
Pekkanen et al., 2018 [4]	To design a car-following model taking driver's attention span, event uncertainty, perception response, and acceleration control as inputs	Cognitive modelling and state transition model	Spearman correlation – p-value
Yin et al., 2018 [15]	Detection of dangerous driving behavior using driver, vehicle, and lane attributes	Fuzzy PSO model	MAE, MSE

5.1 Shallow machine learning-based approaches

Driver profiling for eco-driving - Many approaches established a correlation between fuel consumption and driving behavior for driver profiling[11],[31],[35]. Boggio-Marzet applied k-means clustering on driving data and analyzed the impact of traffic congestion, road type, and driving behavior on fuel consumption [11]. Adamidis et al. applied k-means clustering to categorize driving profiles into three different categories from an eco-driving point of view [31]. Mubarak and Al-Samari estimated power requirements of vehicle based on the vehicle's longitudinal movement and derived fuel consumption on an advanced vehicle simulator [35]. The speed profile for the luxury vehicle "Chrysler 300" had a smooth acceleration profile leading to a reduction of 12.8% in fuel consumption, unlike speed profile for Dodge Charger model that had a great acceleration and speed variance.

Rios-Torres et al. predicted the fuel consumption as per their driving style, vehicle dynamics, and powertrain and suggested fuel reduction techniques using linear regression [36]. The study examined three distinct driving styles, normal, calm, and volatile, in both urban and highway driving scenarios. The aim was to compare the fuel consumption of conventional vehicles with that of Hybrid Electric Vehicles (HEVs). The findings revealed that employing effective power distribution optimization strategies can lead to significant improvements in fuel consumption. In urban driving conditions, fuel consumption could be enhanced by up to 12%, while in highway driving, a 4% improvement was observed. These results highlight the potential benefits of optimizing power distribution strategies to achieve better fuel efficiency in both urban and highway driving settings. Gadde et al. developed the OLS Linear model to predict fuel economy using engine RPM and average speed; principal components were selected if their variance was within 95% [37]. As a result, engine RPM and average speed were identified as major principal components. The limitation of this study is that the effect of environmental conditions is not identified. Hoffman et al. evaluated the impact of route, payload, and driver behavior on fuel economy [38]. It suggested compensating effects of route inclination and payload on fuel consumption. The variation in fuel economy significantly decreased after compensating effects of route inclination and payload. Sarkan et al. developed a 3D regression model for fuel-consumption prediction using throttle position and engine speed and demonstrated stronger dependence of engine speed and throttle position on engine RPM [34]. The engine idle time was removed before applying the 3D regression model, and the model could achieve better accuracy as compared to linear regression model by manually revising the model by excel based tools. Manual formulation of model on specific dataset may not lead to generalized effective model development for other datasets.

Abukhalil et al. studied the correlation of fuel consumption with engine speed and throttle position using SVM and found it to be more effective as compared to regression-based methods or neural networks-based methods [12]. The approach referred to by Yao et al. compared fuel consumption prediction models such as support vector regression (SVR), backpropagation (BP) neural network, and random forest [21]. They recommended the Random forest (RF) model for fuel consumption prediction as it demonstrated the highest accuracy among all other approaches. Hsu et al. proposed a weka-based M5' model tree to measure the Overall Driving Effectiveness (ODE) index as the measure of efficiency for eco-driving [40]. Model tree exhibited comparable performance to the artificial neural network, but it required expert knowledge to be able to document all possible rules efficiently.

5.2 Deep Learning-Based Approaches

Ping et al. suggested a macroscopic model that relates driving behavior with fuel economy and a subscopic model related to driving behavior and environment [2]. LSTM approach when compared to SVM and NN gave better AUC (area under the curve) results. Guo et al. proposed Continuation/Generalized Minimal Residual Algorithm (C/GMRES) combined predicted velocity and state-of-charge (SOC) reference generator for optimal use of battery and fuel for the plug-in hybrid vehicle [24]. Expected engine power demand is estimated while minimizing the cost of fuel consumption, battery charging and battery degradation in real-time. Ozkan et al. used Gated Recurrent Unit (GRU) network to estimate the speed profile of the preceding vehicle and applied non-linear model predictive control (NMPC) [25]. GRU- NMPC gave better fuel efficiency gain than the constant distance/time headway approach. The LSTM approach by Ping et al. [2] did help to figure out instantaneous fuel consumption but the approach was not extended to provide feedback for reporting real-time driving behavior resulting in excessive fuel consumption.

5.3 Miscellaneous Approaches

Bätz et al. framed a Comprehensive Factor Model (CFM) to study the effect of providing eco-feedback on driving and indicated that eco-feedback helped reduce hard acceleration [14]. Hsu et al. used ODI as a measure of eco-driving and it provided quantitative feedback to drivers for reviewing their long-term driving behavior. Similarly, another approach used a driving style indicator as a measure of driving aggressiveness and indicated a strong correlation with fuel economy [40]. The above methods provided feedback for long-term driving aspects and lack providing real-time feedback.

5.4 Fuel-Economical Driving Approaches – Observations and Discussion

The k-means clustering is a preferred approach for driver profiling along with some custom advanced vehicle simulator based model [35]. Although k-means clustering is simple to implement and scales easily according to dataset size, outliers and varying density data cannot be effectively clustered. As the number of dimensions increases, k-means becomes less effective in categorising underlying data. It must be combined with other dimensionality reduction techniques, such as PCA. Profiling-based methods are coarse-grained and not very useful for providing real-time feedback.

Most fuel prediction approaches are based on linear regression, and polynomial regression [12],[32],[34]. The drawback of these approaches is that they assumed a linear relationship between predictor and response variable fuel consumption. Custom approaches [39-41] based on CFM, DSI, Model Tree, and fuzzy rule model also correlated driving behavior with fuel economy, but these require expert knowledge base for building an effective model. Few other approaches are specific to vehicle-specific powertrain models based on the vehicle's longitudinal or lateral movement and involve a lot of computational effort [30],[35]. Other approaches, such as SVR, and RF are better than linear regression approaches due to their non-linear transformation. The disadvantage of the above shallow machine learning approaches is that they support either linear or require labelled data, but ML algorithms are computationally intensive for temporal data processing.

Deep learning approaches GRU-NMPC, C/GMRES and LSTM are suitable for processing historical time series data to learn and predict fuel economy due to underlying structure and learning mechanism [24-25]. Though these approaches are better than regression-based approaches for univariate time series data processing, they are not effective for correlating multi-variate temporal input sequence with fuel economy.

6.0 MOTIVATING DRIVERS FOR ECO-SAFE DRIVING

Safe driving is of utmost important, and it includes driving practices that do not include the application of harsh or frequent negative driving behavior and maintaining proper lead distance with other vehicles. On the other hand, eco-driving practices include driving styles such as maintaining steady speed, avoiding harsh brakes, and avoiding overspeed. An efficient driver must take care of both aspects of safety and fuel economy. Computational approaches discussed in Sections 2.4 and 2.5 either discuss safe or fuel-economical computational approaches. The approaches considering both critical driving aspects together need to be enhanced. Challenges in maintaining eco-safe driving include drivers' unawareness about good driving practices, reactive or excessive feedback mechanisms, and inadequate trade-off of driving behavior to balance safety and fuel economy.

Drivers can be engaged towards eco-safe driving via various ways such as training, and real-time driving visual/audio feedback [10],[15],[42]. Knoefel et al. offered older adults feedback mechanisms to help them improve their driving abilities and decisions [10]. Yin et al. issued warnings and feedback to a driver regarding driving safety [15]. Feedback based approaches are reactive approaches post-observation of any negative driving behavior. Table 7 summarizes proactive eco-safe driving approaches for driving engagement. Bian et al. classified drivers into a driving category based on their driving and framed insurance pricing premium based on driving category [9]. State-transition-based models or fuzzy Particle Swarm Optimisation(PSO) models do attempt to model the driver's perception, response, and psychological aspects, but these approaches are more intuitive and are not based on facts[4],[28]. The car-following model suggested by Pekkanen et al. adjusts its speed according to a safe distance from leading vehicles [4]. It is tough to judge the driving actions of the leading vehicle and estimate the leading vehicle's speed. Driving does involve interaction with fellow drivers that need to be taken care of along with self-driving aspects. Information about fellow drivers is not available in real scenarios, drivers estimate the aggressiveness and actions of fellow drivers before planning responses.

Game theory helps to control the driver's response proactively by evaluating the benefits of choosing any outcome as a strategic response against the actions of fellow drivers. Few game theory based approaches helped drivers proactively in lane changing, interaction with cyclists, and trajectory routing[19-20]. Michieli et al. proposed a Bayesian gaming model and provided controlled feedback to assist the driver during their interaction with cyclists [19]. Meanwhile, Li et al. defined a 2-level controller for predicting the optimal trajectory and required driving manoeuvring under different traffic conditions [20]. Gaming approaches for driver-to-driver interaction where information about other driver's aggressiveness and actions is not known; need to be further enhanced.

Table 7. Comparative study of proactive eco-safe approaches

Reference	Purpose of study	Computational model for analytics	Evaluation metrics
Bian et al., 2018 [9]	To develop Vehicle Insurance Pricing model using driving behavior	SVM, NN, bagging	Accuracy rate, kappa score, MAE, RMSE
Yin et al., 2018 [15]	To detect dangerous behavior based on driver, vehicle, and lane attributes.	A fuzzy PSO model	Dangerous driving intensity (DDI)
Pekkanen et al., 2018 [4]	To design a car-following model taking driver's attention span, event uncertainty, perception response, and acceleration control as inputs	Cognitive modelling and state transition model	Spearman correlation – p-value
Michieli et al., 2018 [19]	To model for interaction between a cyclist and a vehicle	Simultaneous Bayesian game with the concept of Nash Equilibria (NE)	number of accidents Payoff

Drivers can be engaged towards eco-safe driving via various ways such as training, real-time driving visual/audio feedback, and haptic acceleration paddle. Bellotti et al. suggested different pervasive games for drivers and passengers [16]. Michieli et al. proposed Nash Equilibria(NE) based Bayesian gaming approach to assist the driver and develop an autonomous model by providing controlled feedback inside vehicles [19]. Li et al. defined a 2-level controller for predicting the path and required driver's action corresponding to various traffic conditions [20]. Knoefel et al. offered older adults feedback mechanisms to help them improve their driving abilities and decisions [10]. Yin et al. ensured driving safety by issuing warnings and feedback to the driver[15]. Lin et al. proposed a Social Vehicle Route Selection (SVRS) algorithm based on the gaming strategy of Nash equilibrium that helped in reducing traffic congestion [29]. Magaña and Organero framed a gamification tool to generate a relative score and leader- board ranking [41]. As per Jamson et al., haptic force or stiffness feedback performed better than visual feedback under high traffic congestion [3]. Ji et al. developed a game-based control framework to assist driver steering control via feedback [45]. It had been observed that training or immediate feedback sometimes might not result in long-term driving behavior re-enforcement. Motivational incentive-based schemes or gaming-based approaches help continuously engage drivers and motivate them towards positive driving behavior retention.

7.0 DISCUSSION AND RESEARCH GAPS

Data collection is an important step for developing computational models for eco-safe driving. Approaches mentioned in the literature fall under one of the below categories:

i. Smartphone sensors include using built-in GPS, magnetometer, gyroscope, and accelerometer sensors. However, smartphone-based data may be erroneous due to frequent orientation changes and may get interrupted due to calls or low battery.
ii. Data collection via cameras for capturing the movement of face, eye, and head. Installation of body sensors to capture ECG, EEG, and EOG under different driving conditions.
iii. More recently, data is being captured from onboard diagnostics(OBD) devices that many cars come re-fitted with, and others can be retrofitted for the same. OBD is plugged into OBD-II port of the car to capture data from built-in GPS, GPRS, accelerometer sensors, fuel sensors etc. Data collection via OBD is more reliable and allows data to be captured seamlessly.

Table 8 elaborates on different percentages of data attributes being consumed in eco-driving and safe-driving approaches. Driving behavior is the most used data attribute being used in both approaches. An environmental contextual attribute is the second most used feature especially for safe driving approaches, followed by vehicle attributes and trip attributes. Some of driving approaches also use driver's attributes and physiological attributes. Though many of these features have been utilized in existing approaches, these approaches fail to study the combined effects of these attributes.

Table 9 elaborates on driving safety approaches based on ML techniques. Driver risk has been identified as events related to driver's distraction, drowsiness or risky driving pattern. SVM, DT, and RF have been popular ML techniques in this domain. Table 10 describes DL techniques pertaining to driving safety approaches. Driver risk has been identified as events related to a driver's drowsiness or risky driving pattern. NN, ANN, and CNN have been popular DL techniques in this domain. Even though driving safety approaches elaborated in Table 9 and Table 10 discuss the identification of driving risk, however, approaches relating driving risk to crash prediction are lacking. Furthermore, driver profiling approaches are coarse-grained and do not lead to any real-time alerts.

Table 8. Comparative study of data attributes considered in the models

	Reference		Considered attributes for safe-driving	Considered attributes for eco-driving
	Safe driving	Eco-driving		
Vehicle attributes	[1],[6],[22]	[2],[11-12], [14],[21],[24], [30-32],[34-41]	Fuel consumed-5% Engine speed-16%	Fuel consumed-100% Engine speed-61%
Driver attributes	[6],[10],[17]	[32]	Age/gender-16%	Age/gender-5%
Trip attributes	[1],[9-10],[17]	[11],[32],[34-35]	Trip length-16% Number of trips-16 %	Trip length-16% Number of trips-11%
Environment context	[4],[7-10],[17], [19-20],[26],[43-44]	[11],[21],[31-32],[34-35]	Time of travel-37% Road type-42% Traffic flow-21%	Time of travel-11% Road type-22% Traffic flow-28%
Physiological attributes	[4],[6-7],[19-20], [26],[33],[43]	[32]	Eye blink/facial expression-16% ECG/EEG/EOG-42%	ECG/EEG/EOG-5.5%

Table 9. Comparative study of ML approaches - safe driving

References	Purpose of study	Computational model for analytics	Evaluation metrics
Gwakl et al. 2018 [7], Barua et al., 2018 [26]	Driver's drowsiness detection	KNN, SVM, RF, LR	Accuracy, precision rate, recall rate
Darji [6], 2018, Bian[9], 2018	Driver's risky driving detection	SVM, LR, DT	Accuracy, MAE, RMSE
Osman et al., 2019 [22]	Driver's distraction identification	DT, RF	Accuracy

Table 10. Comparative study of DL approaches - safe driving

References	Purpose of study	Computational model for analytics	Evaluation metrics
Abdennour et al., 2021 [1], Ferreira et al., 2017 [8], Yu, 2017 [28]	Driving pattern identification and driver profiling	NN, RF, SVM	Accuracy, precision, recall rate, FPR, AUC
de Naurois et al., 2018 [23]	Driver's drowsiness detection	ANN	RMSE
Liu et al., 2020 [30]	Risky driving behavior detection	CNN	Average precision

Table 11 elaborates ML-based driving fuel economic approaches to make fuel predictions or study the effect of training on fuel consumption optimization. SVM, K-Means clustering, and RF have been popular ML techniques in this domain. Table 12 describes DL techniques pertaining to driving fuel economic approaches. Fuel-economical driving approaches either optimize fuel economy by adjusting speed or predict instantaneous fuel economy. GRU and LSTM represent popular DL techniques for fuel-efficient driving. Many fuel consumption prediction approaches elaborated in Table 10 are coarse-grained and consider average values of velocity or acceleration parameters. Approaches specified in Table 11 do not use multi-variate data for fuel consumption prediction.

Table 11. Comparative study of ML approaches - fuel economical driving

Reference	Purpose of study	Computational model for analytics	Evaluation metrics
Boggio-Marzet et al., 2021 [11]	To study effect of training on fuel consumption	K-Means Clustering	t-statistics
Abukhalil et al., 2020 [12],Yao et al., 2020 [21], Hsu et al., 2017 [40]	To predict fuel consumption	SVM, regression, RF, model tree	RMSE, MAPE, R-squared

Table 12. Comparative study of DL approaches - fuel economical driving

References	Purpose of study	Computational model for analytics	Evaluation metrics
Ping et al., 2019 [2]	To predict fuel consumption based on timeseries data	LSTM	Accuracy
Ozkan et al., 2021 [25]	To optimize fuel economy based on speed of preceding vehicle	GRU	% efficiency gain, RMSE

Table 13 describes approaches to engage drivers by applying game theory based techniques such as Nash equilibrium or Stackelberg algorithms. GLM or fuzzy PSO models have also been used in selective approaches. Many of these approaches are not robust to address incomplete information about fellow drivers and don't provide effective feedback. Furthermore, many of these approaches consider only partial aspects either for safety or fuel economy. There is scope for developing robust approaches that can cater to safety, fuel economy, comfort etc. and deal with incomplete information about fellow drivers.

Table 13. Comparative sudy of game theory based approaches - fuel economical driving

Reference	Purpose of study	Computational model for analytics	Evaluation metrics
Michieli et al., 2018 [19] Li et al., 2020 [20]	To model traffic interaction	Nash and Stackelberg	Accuracy, reduced # of risk occurrences
Xuewu et al., 2018 [45]	To develop cooperative steering systems	Nash and Stackelberg	Box and whisker plot
Yin et al., 2018 [15] Payyanadan et al., 2018 [17]	Study the effect of driver's attributes on driving behavior risk analysis and profiling	Generalized linear mixed-effects regression model & Fuzzy PSO model	Standard error, t-value, MAE, MSE

Limitations of existing approaches can be exploited to enhance research in the connected car domain.

i. Lack of appropriate data abstraction; the driving behavior of a driver should be assessed in alignment with contextual features. It has been observed that many of the approaches either considered partial contextual features or lacked considering them. Models for driver profiling considered average or maximum coarse-grained values, whereas models considering every second's data were computationally intensive. A balanced approach is an approach where driving and contextual features are considered at an appropriate granularity level that not only leads to accurate estimates but is also not computationally intensive.
ii. Lack of real-time driving risk assessment; existing approaches for driver's state detection or driving abnormal pattern detection can be associated with driving risk and crash prediction.
iii. Lack of comprehensive approach; a comprehensive approach that considers all aspects of driving utility, such as comfort, safety, and fuel economy is lacking. There is a need to develop reliable soft computational methods for maintaining eco-safe driving.
iv. Need for developing instantaneous detection of anomalous conditions; any non-optimal driving behavior should be notified to the driver in real-time; however, it must be ensured that feedback must be given only for appropriate anomalous conditions. There is a need to develop computational models that can accurately detect anomalies.
v. Effective feedback; methods for reactive feedback help to only rectify driver's behavior post any event. Secondly, repetitive and excessive feedback may be non-obtrusive to the driver. Driver's actions can be controlled by providing proactive and effective feedback.
vi. Need to deal with incomplete information about a fellow driver; there is a lack of computational models that can deal with incomplete information about fellow drivers. Driving safety depends not only on self-behavior but also depends on how driver responds to fellow driver's actions in a driving event.

8.0 CONCLUSIONS

Driving behavior is an important factor that impacts fuel economy and driving safety. Eco-safe driving is a driving practice to reduce fuel consumption while maintaining driving safety. As part of our analysis, it is recommended to use random forest, decision tree, SVM, or neural networks based supervised classification method. Deep learning models like LSTM, DeepRCN, and 3Dresnet are suitable for classifying time series based data. Ensemble learning approaches such as bagging, and ADABoost are also widely used. Accuracy rate, precision rate, and recall rate are the standard metrics used for classification.

It has been observed from the studied literature that many authors focused only on the driving pattern, style, or driving profile detection without correlating the same with driving crash risk. Secondly, driving behavior of a driver is not assessed at appropriate granularity according to contextual conditions. Regression-based methods are not sufficient to model non-linear data and are not sufficient to address issues in driving risk analysis and fuel economy prediction alone. Existing deep learning approaches do not learn long term temporal dependencies appropriately for multi-variate data processing in real-time. Existing approaches are either giving excessive feedback, which disengages drivers or not giving timely feedback. There is a need to separate anomalous driving behavior from acceptable deviation so that the driver remains engaged in economical and safe driving without being distracted or disengaged. A comprehensive approach that provides relevant driving feedback and considers all aspects of driving utility, such as comfort, safety, and fuel economy, is lacking. Furthermore, driving safety and fuel economy gets affected not only due to drivers' own actions but also gets affected due to the actions of neighbouring drivers. The information about neighbouring drivers is often missing and incomplete; existing approaches do not provide generic mechanisms to deal with missing information. Identified research areas and opportunities can be further exploited to optimize power distribution for electric and hybrid vehicles for a cleaner and more sustainable transportation solution, contributing to air quality improvement and combating climate change while providing potential long-term cost savings and advancements in technology.

Furthermore, a large amount of driving data along with driver's characteristics gives data monetization opportunities in terms of providing roadside assistance, micro-insurance, or alliances with third parties for various offers based on the current location of customers, such as providing value-added services such as food, travel, service coupons. V2V communication via the exchange of information between vehicles can help in enhancing overall safety and convenience. Cooperative strategies can be employed for joint driving decisions, thereby reducing operational costs and increasing traffic efficiency. There can be different cooperation strategies, centralized or decentralized. Cooperative perception helps exchange information about road conditions such as potholes, speedbumps, and accident spots, enabling drivers to opt for optimal routes. It also helps in smooth traffic management and optimal routing by exchanging congestion information. Existing work on driving behavior would be a fundamental base on which further research on the above areas can be conducted.

9.0 ACKNOWLEDGEMENT

We would like to acknowledge the support provided by the Department of CSE&IT, JIIT Noida for their support in carrying out the research reported in this work. We sincerely thank the editors and reviewers for their time, expertise, and invaluable input. Their feedback has been instrumental in improving the quality and rigor of the article.

10.0 REFERENCES

[1] N. Abdennour, T. Ouni, and N. Ben Amor, "Driver identification using only the CAN-Bus vehicle data through an RCN deep learning approach," *Robotics and Autonomous Systems,* vol. 136, p. 103707, 2021.

[2] P. Ping, W. Qin, Y. Xu, C. Miyajima, and K. Takeda, "Impact of driver behavior on fuel consumption: Classification, evaluation, and prediction using machine learning," *IEEE Access*, vol. 7, pp. 78515-78532, 2019.

[3] S.L. Jamson, L.D. Hibberd, and A.H. Jamson, "Drivers' ability to learn eco-driving skills; effects on fuel-efficient and safe driving behavior," *Transportation Research Part C: Emerging Technologies,* vol. 58, pp. 657-668, 2015.

[4] J. Pekkanen, O. Lappi, P. Rinkkala, S. Tuhkanen, R. Frantsi, and H. Summala, "A computational model for driver's cognitive state, visual perception and intermittent attention in a distracted car following task," *Royal Society Open Science,* vol. 5, no. 9, p. 180194, 2018.

[5] X. Ji, K. Yang, X. Na, C. Lv, and Y. Liu, "Shared steering torque control for lane change assistance: a stochastic game-theoretic approach," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3093-3105, 2018.

[6] A. Darzi, S.M. Gaweesh, M.M. Ahmed, and D. Novak, "Identifying the causes of drivers' Hazardous States Using driver characteristics, vehicle kinematics, and physiological measurements." *Frontiers in neuroscience,* vol. 12 pp. 568, 2018.

[7] J. Gwak, M. Shino, and A. Hirao, "Early detection of driver drowsiness utilizing machine learning based on physiological signals, behavioral measures, and driving performance," presented at 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, Hawaii, USA, 2018.

[8] J. Ferreira, E. Carvalho, B.V. Ferreira, C. de Souza, Y. Suhara, et al., "Driver behavior profiling: An investigation with different smartphone sensors and machine learning," *PLoS One*, vol. 12, no. 4, p. e0174959, 2017.

[9] Y. Bian, C. Yang, JL. Zhao, and L. Liang, "Good drivers pay less: A study of usage-based vehicle insurance models," *Transportation Research Part A: Policy and Practice*, vol. 107, pp. 20-34, 2018.

[10] F. Knoefel, B. Wallace, R. Goubran, and S. Marshall, "Naturalistic driving: a framework and advances in using big data," *Geriatrics*, vol. 3, no. 2, pp.16-26, 2018.

[11] A. Boggio-Marzet, A. Monzon, AM. Rodriguez-Alloza, and Y. Wang, "Combined influence of traffic conditions, driving behavior, and type of road on fuel consumption," *International Journal of Sustainable Transportation*, vol.16.4, pp. 1-13, 2021.

[12] T. Abukhalil, H. AlMahafzah, M. Alksasbeh, and BA. Alqaralleh, "Fuel consumption using OBD-II and support vector machine model," *Journal of Robotics*, vol.2020, pp. 1-9, 2020.

[13] R. Coppola, and M. Morisio, "Connected car: technologies, issues, future trends," *ACM Computing Surveys (CSUR)*, vol. 49 no. 3, pp. 1-36, 2016.

[14] A. Bätz, H. Gimpel, S. Heger, and M. Wöhl, "Driving sustainably–The influence of IoT-based eco- Feedback on driving behavior," In Proceedings of the 53rd Hawaii International Conference on System Sciences, 2020, pp. 912-921.

[15] J.L. Yin, B.H. Chen, K.H. Lai, and Y. Li, "Automatic dangerous driving intensity analysis for advanced driver assistance systems from multimodal driving signals," *IEEE Sensors Journal*, vol. 18, no. 12, pp. 4785-4794, 2017.

[16] F. Bellotti, R. Berta, P. Paranthaman, G. Dange, and A. De Gloria, "REAL: Reality-enhanced applied games." *IEEE Transactions on Games,* vol. 12, no. 3, pp. 281-290, 2019.

[17] P. Payyanadan, A. Rashmi, S. Fabrizzio, and J. D. Lee, "Influence of familiarity on the driving behavior, route risk, and route choice of older drivers," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 1, pp. 10-19, 2018.

[18] X. Chang, H. Li, L. Qin, J. Rong, Y. Lu, and X. Chen, "Evaluation of cooperative systems on driver behavior in heavy fog condition based on a driving simulator," *Accident Analysis & Prevention,* vol. 128, pp. 197-205, 2019.

[19] U. Michieli and L. Badia, "Game theoretic analysis of road user safety scenarios involving autonomous vehicles," In 2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Bologna, Italy, 2018, pp. 1377-1381.

[20] N. Li, DW. Oyler, M. Zhang, Y. Yildiz, I. Kolmanovsky, and AR. Girard, "Game-theoretic modelling of driver and vehicle interactions for verification and validation of autonomous vehicle control systems," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 5, pp. 1782-1797, 2017.

[21] Y. Yao, X. Zhao, C. Liu, J. Rong, Y. Zhang, et al., "Vehicle fuel consumption prediction method based on driving behavior data collected from smartphones," *Journal of Advanced Transportation*, vol.2020, pp.1-11, 2020.

[22] O.A. Osman, M. Hajij, S. Karbalaieali, and S. Ishak, "A hierarchical machine learning classification approach for secondary task identification from observed driving behavior data," *Accident Analysis & Prevention,* vol. 123, pp. 274-281, 2019.

[23] C.J. de Naurois, C. Bourdin, C. Bougard, and JL. Vercher, "Adapting artificial neural networks to a specific driver enhances detection and prediction of drowsiness," *Accident Analysis & Prevention*, vol. 121, pp. 118-128, 2018.

[24] N. Guo, X. Zhang, Y. Zou, L. Guo, and G. Du, "Real-time predictive energy management of plug-in hybrid electric vehicles for coordination of fuel economy and battery degradation," *energy*, vol. 214, 119070, 2021.

[25] M.F. Ozkan, and Y. Ma, "Eco-driving of connected and automated vehicle with preceding driver behavior prediction," *Journal of Dynamic Systems, Measurement, and Control*, vol. 143, no. 1, p. 011002, 2021.

[26] S. Barua, M.U. Ahmed, C. Ahlström, and S. Begum, "Automatic driver sleepiness detection using EEG, EOG and contextual information," *Expert Systems with Applications*, vol. 115, pp. 121- 135, 2019.

[27] Y. Zhang, and T. Kumada, "Automatic detection of mind wandering in a simulated driving task with behavioral measures," *PloS one*, vol. 13, no. 11, p. e0207092, 2018.

[28] J. Yu, Z. Chen, Y. Zhu, Y. Chen, L. Kong, and M. Li, "Fine-grained abnormal driving behaviors detection and identification with smartphones," *IEEE Transactions on Mobile Computing*, vol. 16, no. 8, pp. 2198-2212, 2016.

[29] K. Lin, C. Li, G. Fortino, and J. Rodrigues, "Vehicle route selection based on game evolution in social internet of vehicles," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2423-2430, 2018.

[30] Y.C. Liu, Y.A. Hsieh, M.H. Chen, C.H. Yang, J. Tegner, and Y.C. Tsai, "Interpretable self-attention temporal reasoning for driving behavior understanding," In ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020, pp. 2338-2342.

[31] F.K. Adamidis, E.G. Mantouka, and E.I. Vlahogianni, "Effects of controlling aggressive driving behavior on network-wide traffic flow and emissions," *International Journal of Transportation Science and Technology*, vol. 9, no. 3, pp. 263-276, 2020.

[32] A. Barbado, and Ó. Corcho, "Vehicle fuel optimization under real-world driving conditions: An explainable artificial intelligence approach," *arXiv preprint arXiv*, vol. 2107, p. 06031, 2021.

[33] J. Hu, X. Li, H. Xin, and M. Wuqiang, "Abnormal driving detection based on normalized driving behavior," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 6645-6652, 2017.

[34] B. Sarkan, S. Semanova, V. Harantova, O. Stopka, M. Chovancova, and M. Szala, "Vehicle fuel consumption prediction based on the data record obtained from an engine control unit," *MATEC Web of Conferences*, vol. 252, p.06009, 2019.

[35] L. Mubarak, and A. Al-Samari, "Investigation of driving behavior on performance and fuel consumption of light-duty vehicle," *Diyala Journal of Engineering Sciences*, vol. 13, no. 4, pp. 102-113, 2020.

[36] J. Rios-Torres, J. Liu, and A. Khattak. "Fuel consumption for various driving styles in conventional and hybrid electric vehicles: Integrating driving cycle predictions with fuel consumption optimization," *International Journal of Sustainable Transportation*, vol. 13, no. 2, pp. 123-137, 2019.

[37] R.S.K. Gadde, R. Sriganesh, C. Karthick, S. Lalasure, and K. Rajakumar, "Determination of principal variables for prediction of fuel economy using principal component analysis," *SAE Technical Paper:* No. 2019-26-0359, 2019.

[38] A.J. Hoffman, and M. van der Westhuizen, "An empirical model for the assessment of truck driver fuel economy," In IEEE Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 2019, pp. 637-642.

[39] J. Lasocki, and K. Boguszewski, "Environmental effects of driving style: impact on fuel consumption," In 11th Conference on Interdisciplinary Problems in Environmental Protection and Engineering EKO-DOK 2019, vol. 100, p. 00043, 2019.

[40] C.Y. Hsu, S.S. Lim, and C.S. Yang, "Data mining for enhanced driving effectiveness: An eco-driving behavior analysis model for better driving decisions," *International Journal of Production Research*, vol. 55, no. 23, pp. 7096-7109, 2017.

[41] V.C. Magaña, and M.M. Organero, "GAFU: using a gamification tool to save fuel," *IEEE Intelligent Transportation Systems Magazine*, vol. 7, no. 2, pp. 58-70, 2015.

[42] S.L. Jamison, L.H. Daryl, and A.H. Jameson, "Drivers' ability to learn eco-driving skills; effects on fuel-efficient and safe driving behavior," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 657-668, 2015.

[43] G. Castignani, R. Frank, and T. Engel, "Driver behavior profiling using smartphones: A low-cost platform for driver monitoring," *IEEE Intelligent Transportation Systems Magazine*, vol. 7, no. 1, pp. 91-102, 2015.

[44] T. Sehyun, S. Kim, and H. Yeo, "Development of a deceleration-based surrogate safety measure for rear-end collision risk," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2435-2445, 2015.

[45] X. Ji, K. Yang, X. Na, and C. Lv, "Shared steering torque control for lane change assistance: a stochastic game- theoretic approach," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3093-3105, 2018.

[46] P. Malhotra, A. Ramakrishnan, G. Anand, L. Vig, P. Agarwal, and G. Shroff, "Lstm-based encoder-decoder for multi-sensor anomaly detection," *arXiv preprint arXiv*, vol. 1607, p. 00148, 2016.

[47] A. Das, A. Ghasemzadeh, and M.M. Ahmed, "Analyzing the effect of fog weather conditions on driver lane-keeping performance using the SHRP2 naturalistic driving study data," *Journal of Safety Research,* vol. 68, pp. 71-80, 2019.

[48] Wikimedia Commons contributors, File:Simulator vožnje.jpg, Wikimedia Commons, 2 September 2020, [Online]. Available: https://commons.wikimedia.org/w/index.php?title=File:Simulator_vo%C5%BEnje.jpg&oldid=446827200 [Accessed July, 2023]