A comprehensive review of various data collection approaches, features, and algorithms used for the classification of driving style

G.Priyadharshini1, J.S.Femilda Josephin2
1Department of Computer Science and Engineering, 2Department of Software Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai-603203, India
priyadhg1@srmist.edu.in, femildaj@srmist.edu.in

Abstract: In the transport sector problems, road safety is a prime concern in emerging nations. Applications about driving assistance are being actively studied to address road safety matters, including humanistic performance defined as one of the principal causes for problems on roadsafety, which confirms why driving style is currently experiencing extensive research attention. Future driving style prediction will form the basis for eco-driving and energy management strategies. From this aspect, analyzing drivers' behavior is necessary to improve road safety. The comprehensive survey provides a summary, outline, and structure a large collection of work from various sources and proposes a comparison of the devices used, parameters studied, and classification algorithms used for analyzing driving style. The researches done on the driving style analysis uses a wide variety of devices and several parameters. This analysis shows that a diverse set of parameters that can be used to analyze the style of driving and seeks to understand the various machine learning classification methods and metrics for the classification of driving style.

1. Introduction
Transportation has a great impact on daily lives with a vital role in economic growth and globalization. Transportation and enterprise sector account for a dominant portion of the entire final energy requirement, promoted by urbanization [1]. Driving style is described as a way the driver chooses to drive a vehicle which can be measured by accumulating driving experience. Vehicle drivers possess notable boundaries to gain safety and decrease emissions [2] [3] [4]. More than 25% of the fuel can be saved by good driving styles [5] [6]. Driving style analysis plays an important part in a wide variety of fields such as human-centred vehicle control systems, road safety, intelligent transportation systems, and power administration for vehicles. For example, in fleet management, the administrator is keen to review the driving style of the drivers which helps them to identify how efficiently a driver uses the resources. Currently the insurance firms rely on driving data to calculate insurance premiums. Furthermore, identifying individual driving styles is necessary for the use of smart transportation assistance such as the employment of artificial intelligence in the connected autonomous vehicles and the transportation sector, for more reliable, cleaner, more intelligent and more effective transport modes.

The proportion of population, road traffic deaths and vehicle registered is shown in figure 1, extracted from the status report released by World Health Organisation (WHO) [7] on road safety in the year 2018. According to report, WHO points out that

- The road traffic death rate in low and middle income countries is out of proportion compared to population size and the circulation of vehicles.
- Road traffic injuries are the eighth leading cause for death among all the age groups.
- By comparing the proportion of deaths among the drivers /passengers of four, three or two wheelers and pedestrians, cyclists the former account for 29% and the later 26 % throughout the world.
As per the global status report, 1.35 million people die every year due to road fatalities. This rate remains to be saturated instead of getting reduced. Also, nowadays purchasing a fuel-efficient vehicle is a trend among most of the users but only a few know the fact that a good driving style can contribute to save cost and benefit the environment. Therefore higher priority is given to improve traffic safety by government and automobile industries and hereby the innovation in driver assistance comes into the picture. Therefore, over the past year, interest has been raised to keep an eye on the vehicle and driving data which aims at identifying the driving patterns and series of movements that are used for driving style classification. The main aim of this study is to examine research in the area of driver style classifications to address problems related to road safety. This is accomplished by understanding the relationship between the driving style and road safety issues and classifying the drivers using their inclination to risk [8]. The link between the vehicle control and driver, and the responsiveness of the driver to difficult driving circumstances (e.g., weather, traffic) are important relevant factors towards traffic collisions [9].

The main aim of this study is to examine research in the area of driver style classifications to address problems related to road safety. This is accomplished by understanding the relationship between the driving style and road safety issues and classifying the drivers using their inclination to risk [8]. The link between the vehicle control and driver, and the responsiveness of the driver to difficult driving circumstances (e.g., weather, traffic) are important relevant factors towards traffic collisions [9].

Research in the fields of driver behavior analysis, driving style classification, usage-based insurance and anti-theft detection is uniformly growing as the demand for innovative, quick and precise solutions are needed in the development of driver assistance and self-driving cars. The basis of data about driver style research is divided according to differences in data collection platforms and the features used across a variety of disciplines and industries. The machine learning (ML) and deep learning algorithms applied in the classification of driving style have been discussed with a correlation between the metrics that were used to analyze the classification results. This was done in order to gather the range of feasible algorithms to review the proposed problem of driving style classification. A small variation in driving can have a notable impact on the driving style. To classify the driving style in an efficient way there arise a necessity for the research of characteristic features in driving cycles of various drivers. To pay attention, this paper seeks to identify the features
from various data collection methods and machine learning based classification methods for
driving style classification for driver assistance applications and to address road safety
problems.

This article is arranged as follows. Section 2 gives the search keywords that were used to
identify the relevant works. Section 3 followed by 4 outlines the comparative analysis of
various devices used for data collection and presents the parameters observed from those
devices. Section 5 explains the various methods used to study the various driving styles.
Section 6 concludes the survey.

2. Search methodology

This study critically examines the parameters that were been used by various
researches in identifying the driving style. To identify the relevant study the keywords were
limited to a combination of words like eco-driving, fuel consumption, fuel efficiency, in-
vehicle systems, fuel emission, and fuel economy in the context of driving. Various database
like Sciencedirect, Institute of Electrical and Electronics Engineers(IEEE), Google scholar,
Association for Computing Machinery(ACM) and many other were used to identify the
relevant papers. The keywords that are used are defined briefly in Table 1.

Since 1978, several researches have been carried out to analyse the different driving style,
so there are a number of works related to this topic. For this paper around 60 research
articles that are closely related have been considered to be the most suitable one for eco-
driving classification.

Table 1. Search methodology

| Keywords                          | Description                                      |
|-----------------------------------|--------------------------------------------------|
| Eco-Driving                       | Energy efficient driving techniques               |
| Fuel Efficiency/ Fuel Economy     | Distance travelled per unit of fuel              |
| Fuel Consumption                  | Rate at which the engine uses fuel               |
| In-Vehicle Systems                | Combination of hardware and software sensor data with a information delivery to drivers |
| Fuel Emission                     | Emissions produced by vehicles                   |
| Advanced Driver Assistance System (ADAS) | Electronic devices to control drivers[10] |

3. Driving Data Collection devices

Advancements in wireless and mobile sensing technologies have a great impact on
societies. Furthermore, various compact computations, transmission devices, and peculiar
computerized devices have become filled with a mixture of sensing circuits to develop user
communication and facilitate context-sensitive operation. Certain wireless and mobile
sensing devices integrate humans with their surroundings stronger than ever imagined [11].
Due to the advancement in those devices, it is likely to accumulate a distinct set of data
regarding car and driving style (e.g. fuel consumption shift gear state, acceleration, braking,
and manoeuvre, engine Revolutions per Minute). The collective study of the devices and the
parameters that can be observed from these devices gives an abstract overview of applying to
many valuable applications like vehicle financing, relating vehicle and driver insurance,
government and law enforcement and various other applications in industries.

The data collected from these devices can be categorized based on whether the data
belong to numeric or multimedia type. These devices in turn have their own advantages and
disadvantages which differ from the facts that range from costly to critical data accuracy.
Table 2. Comparison of different data collection approaches

| Collection method                  | Data type         | Advantages                                      | Disadvantages                                      |
|------------------------------------|-------------------|-------------------------------------------------|---------------------------------------------------|
| OBD                               | Numeric           | Inexpensive option, Easiness and flexibility     | Susceptible to damage                              |
| Driving simulator                 | Numeric           | Controlled Environment, Ease of Access, Handle unpredictable tasks | May involve unrealistic driving style leading to invalid research outcomes. |
| Sensors, Hardware, camera-based method | Both numeric and multimedia | Real-world data, Higher precision, Access to both driver and vehicle data | Require more computing resources to process images, Expensive, Interference problem |
| Smartphone method                 | Both numeric and multimedia | Low expense, Easier access                      | Data accuracy becomes critical due to position, clarity |
| Traffic surveillance              | Multimedia       | Real time data, Collect diminished value        | Costly affair, Algorithms used, Easily abused     |

From Table 2 it is inferred that researchers have used instrumented vehicles to collect the natural driving data which also has in-vehicle mounted cameras to obtain the videos of drivers. The videos contain data like speed, car following distance, acceleration/deceleration. While others used some special hardware such as On-Board Diagnostic (OBD) and sensors [14-19] to collect data's like yaw rate, throttle opening acceleration, jerk, Magnetic Anomaly Detection (MAD) signal and many more. However, the use of expensive devices and sensors to collect the driving data can be a major obstacle to collect the naturalistic driving data. Some researchers have used driving simulators [33-36] e.g., IPG’s Truck Maker vehicle dynamics simulation software to accumulate driving style data. The simulator data may not match to real-world driving. In addition to these, traffic videos can be obtained from surveillance cameras fitted on the roadside. This also would be challenging from the fact that it depends on the quality and the algorithms used. The data's from OBD and smartphone are considered to be the inexpensive options, but in terms of accuracy the OBD becomes the reliable option.

4. Features interpreted for driving style classification

A thorough study of features is essential in order to classify a driving style. Too much features increase the computation cost and causes degradation of the models developed. On the other hand, too little features cannot correctly classify a driving style. There are several ways to gather driver and vehicle data as discussed in the previous section. A collective summary of the parameters that were used to analyse the driving style is summarized under this topic.

4.1. On-Board Diagnostic (OBD)

Researchers [12] stated that OBD holds a collection of measures for executing a computer-based system to analyse engines major performance components. OBD [13] is an in-vehicle system that detects and notifies the state of vehicle well-being. The repair technicians and the after-sales service providers can get access to data like vehicle acceleration, engine data, and emission data.

Lee et al [14] analysed the driving style and road type from the OBD data. The road type classification is done by comparing the measures obtained from real driving states with...
an open-sourced map. Jerk and acceleration based methods are judged and compared with the identical set of measures for the driving style classification.

OBD2 device is plugged into the car and a smart phone app name torque [15] is connected through Bluetooth to receive the driving data. These data’s are used to develop a model of the driving style.

Feng et al. [16] used on-board diagnostics data and forward-looking radar and a monocular dash cam to collect the vehicle position and traffic data of three drivers. The parameters collected from these devices were used to classify the driving style and extract the most cost-effective and eco-friendly driving styles. In [14] and [15] parameters from more than 2 devices are used to analyse the driving style.

Ferreira et al. [17] analysed the influence of driving styles on fuel consumption, by considering the parameters like engine rotation speed, acceleration, etc collected from the vehicle Control Area Network(CAN) bus and Global Positioning System devices. Karsten et al [18] developed eco-driving based driver advice. The data for the analysis was collected from the CAN bus and GPS data. The GPS data include latitude, longitude position, and direction.

Ferreira et al [17] proposed a driving style classification from the comfort perspective in which the data from the CAN bus, pressure sensor and IMU unit are used. It is been identified that comfort comes from a different perspective but frequency and amplitude of acceleration are the key factors.

4.2. Accelerometers

Accelerometers are used to sense static and dynamic forces of acceleration of vehicles [20]. In [21] the authors used 3-axis accelerometer data to classify the driver style. The author also mentioned that it is low-cost, ubiquitous and captures driving style, skill and driving style information. In this paper only the accelerometer signals are used, there is a need to extract useful data to avoid redundancy. After the extraction of useful data the feature calculation is done to recognize the driving style.

4.3. GPS receivers, Magnetometers, Gyroscopes

GPS receivers receive information from satellites and calculate the devices geographical location. The driver's speed and position were collected with Ublox Neo 6 M GPS module in [22] which was selected due to low cost and high accuracy. Fan et al [23] analysed the taxi driving style and driving risk based on trajectory data. For this analysis the author used the GPS data by extracting the useful data from it like longitude, latitude, speed. In addition to this data the authors also have integrated the weather information.

Magnetometers which work by passive sensing technologies are used to detect large objects. Tracking and classification of object plays a major role in intelligent transportation system. A magnetic anomaly detection (MAD) system with a single three-axis magnetic sensor is illustrated in [24]. MAD has been recently used in intelligent traffic control [25], on-destructive testing [26], and intelligent parking [27]. Gyroscopes is used to provide stability or maintain directions.

4.4. Smartphone, cameras

Driver monitoring with mobile devices is an emerging trend that accommodates many market needs. The smart phones have applications installed so that it can be used as sensing parameters to identify driver characteristics. Instead of using separate sensors like accelerometer sensors, gyroscope sensors etc. a single app can be used to detect all the data. The latest mobile phones are provided with a magnetometer, camera, microphone, GPS, 3-axis gyroscope, 3-axis accelerometer, microphone, ambient light, proximity.

Castignani et al [28] proposed SenseFleet, a driver profile platform to identify unsafe driving style. The fuzzy system is adopted to measure the score using real-time context data like route topology or climatic conditions for different drivers. Johnson et al. [29] and
Castignani et al. [30] categorized driving style into typical (non-aggressive) and aggressive. Among the various inputs from a mobile, accelerometer, only a rear-facing camera, gyroscope, and GPS are considered. By using the sensor fusion and dynamic time wrapping algorithm [29], a complete driver monitoring system which determines both driving style and the kind of driving manoeuvre is done. In [30] Multivariate Normal distribution (MVN) is used to build a mathematical model of the user’s driving styles.

You et al [31] developed the first dual-camera application for android users. This application fuses the input from both front and back camera images and other embedded sensors to recognize and warn drivers on critical driving conditions. Castignani et al. [32] proposed an evaluation study of driver profiling based on the smart phone sensor data of accelerometer, GPS, magnetometer, gravity sensor.

4.5. Driving simulator

Driving simulators form an excellent educational and training tool to impart safety driving technique to the drivers. It allows researchers to reliably control, standardize and replicate certain driving conditions like weather, traffic flow etc.

Mumcuoglu et al. [33] developed a driving style recognition algorithm based on Long Short term Memory (LSTM). For this purpose an artificial training road is designed with a simulator to extract different driving profiles.

Yang et al. [34] used EEG features and 14 driving style variables extracted from the simulator for the identification of driving state prediction performance. Wang et Al. [35] proposed a pattern recognition approach to characterize the drivers curve negotiating style. The input to the method includes vehicle speed and throttle opening which are extracted from driving simulator. Wang et Al [36] used a driving simulator to collect the speed and throttle opening data for statistical pattern recognition of driving style.

| Authors          | Collection method                                                                 | Feature parameters                                                                 | Purpose                                      |
|------------------|-----------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|----------------------------------------------|
| [14]             | OBD                                                                               | Intake manifold pressure, Engine revolutions, Air-fuel ratio                         | Road type and driving style                  |
| [15]             | OBD2 device and smartphone                                                       | Speed data, GPS coordinates                                                        | Driving Style                                |
| [16]             | OBD and Instrumented vehicle                                                      | Vehicle speed, Throttle pedal position, Engine speed, Headway distance               | Driving style influencing fuel efficiency    |
| [17]             | CAN-bus and GPS                                                                   | Acceleration, Braking, Clutch use, Engine RPM, Ignition on/off.                      | Impact of driving style on fuel consumption  |
| [18]             | OBD                                                                               | Latitude, Longitude position, Direction, Position, fuel consumption per second with Revolution per minute | Eco-Driving Advice                          |
| [19]             | CAN-bus, Pressure sensors, IMU unit                                               | Steering wheel speed, Percent gas pedal, Brake pedal pressure, Vehicle speed, Gas pedal pressure, X.Y.Z axis accelerometer, Steering wheel angle, Engine RPM | Driving correction for safety and comfort improvement |
| [21]             | 3-axis-accelerometer                                                            | lateral, longitudinal and vertical acceleration                                     | Aggressive and normal driving style          |
4.6. Overview of Parameters used to recognize driving style

Feature collected through other devices like sensors and hardware, simulators are explained in table 3. The study of features for driving style started in 1978 with an introduction of 10 features [37] based on acceleration-deceleration changes. In [38] more than 62 features are used for identifying driving style which is classified into 16 groups out of which 9 are meant to affect fuel efficiency and economy. But, numerous features make higher hardware costs and higher computational time, which could not be achieved in on-board vehicles. Recent studies have shown to lessen the number of features that have used for driving style analysis. However, the initial step for driving style classification is to identify the variable for robust classification. Choosing the correct variable is an important task where further results are dependent on it. The research performed so far reveals that there is no guaranteed set of features for this purpose. This diversity leads to the plurality of applications in eco-driving, driving correction, fuel consumption, and insurance premium calculation. This also paves the way of research to identify the best method to filter out the feature for driving style classification.

5. Driving style classification

Driver style is a common term used to describe different notions linked to a driver’s driving styles and driving characteristics which include countless variables including acceleration, deceleration, engine Revolution per Minute, and many more. Hence, it provides an inference that there is no accepted explanation for the driving style. From fig 2 it can be seen that the driving style classification can be done either objectively or subjectively
Subjective and objective measures are used for assessment in all the fields likely educational, research, news reporting etc. Subjective information is based on individual views, commentaries, emotions and analysis. Extended-response questions, questionnaires, surveys and essays come under subjective measures. Amongst the subjective methods, expert evaluation and questionnaires/surveys are used most widely [40-43]. Objective information can be scored or expressed. Objective information does not differ, whereas subjective information can change considerably from person to person or day to day. Subjectivity can really be wrong, or far from the fact, whereas objectivity signifies staying as close to the fact as possible. Usually objectivity is applied in a decision-making process, whereas subjectivity should be admitted, but less strongly so.

O. Taubman-Ben-Ari et al. [40] conducted a series of studies applying multivariate statistical models to study the association between objective measures and self-reporting two samples of young drivers. S. Amado [41] investigated the driver style based on a questionnaire. During the driving session an expert evaluated the driving style and violations of traffic. At the end of the session the driver was asked to answer questionnaires on self-evaluation and Driver Style. A comparison of the report from the experts and drivers were evaluated for the study. O. Taubman-Ben-Ari, D. Yehiel [42] developed questionnaires based on the association between big-five personality traits and costs and driving style. The questionnaires were completed by 320 drivers (170 women and 150 men). The conclusion revealed that the driving style depends on sociodemographic, personality, and motivational factors. This technique was used to evaluate the style of drivers in Israel and elsewhere [43-45]. To the evaluation criterion, L. Eboli et al. [46] added the record of driving violations.

Eventhough subjective evaluation is useful for classifying driving style, this method is extremely labour-intensive and needs authority to be in the vehicle always. This is usually not possible, particularly if the expected sample size is huge or the focus is on higher-level automation, where an authority cannot be present in the vehicle always. Hence, to estimate the driving style there is a need to find another way which means through the objective method. Most of the stated objective works on driving style classification were based on vehicle movement signals and driver operation. The objective evaluation can be broadly classified into three types based on the algorithms used namely heuristic rule-based, fuzzy logic-based, and machine learning-based.

5.1. Heuristic rule based approaches

The simplest strategy for objective based Driving Style recognition is named as the heuristic rule-based approach. The rules can be the predefined thresholds to cluster driving style into groups. The rule based driving style classification was first introduced by Stoichkov [47], combines the data from an accelerometer, gyroscope, and geomagnetic field sensor and shows it as less noisy. Upon setting thresholds, driving events were identified and classified referring to those events the driving style of the driver can be assessed and classified.

Castignani et al. [28] fused the motion sensors and GPS data to detect the acceleration, braking, steering and over-speeding events. Fuzzy set limits were used to detect the driving
events from various devices and vehicles for the event detection algorithm. Then a scoring algorithm was proposed which not only depends on the events of driving but on weather and time of day.

Saiprasert et al [48] proposed three algorithms namely rule based algorithm, pattern matching algorithm and self triggered pattern matching algorithm with the first algorithm depends on the data from GPS receiver whereas the other two algorithms depends on the accelerometer data. Then it uses pattern matching models with dynamic time wrapping for driver style analysis.

Joubert et al [49] used the velocity and accelerometer data from in-vehicle data recorder (IVDR) from 124 drivers. The evaluation of the driver style risk levels are based on critical events that can occur. Murphey et al [50] proposed an innovative way to classify the driving style by examining the jerk profile of the driver. Applying Rule-Based algorithms unifies easy analysis, simplicity, and implementation, but restricts the number of parameters that can be handled. The preceding examples are generally based on driving event calculation and the fixed thresholds. With the evolution of machine learning-based approaches, there is increasing attention due to more limited expert knowledge involved and high performance compared to rule-based approaches.

5.2. Machine learning based approaches

Machine Learning provides the capability to automatically learn and improve on its own through experience. These need not be explicitly programmed rather these are programs that can access data and learn by themselves. The machine learning techniques are categorized into supervised and unsupervised. Supervised learning indicates the ability of the algorithm to infer knowledge from available data which are labelled so that the algorithm can be used to predict new events. When the data is not classified or categorized, the unsupervised learning is done. In unsupervised learning the data is being clustered using some algorithms or automated methods to learn the underlying similarities or characteristics from the data. Machine learning based approaches are gaining interest because it requires less expert knowledge. The table 4 explains various algorithms under the machine learning category and the metrics used for driving style classification.

5.2.1. Supervised learning

Supervised learning algorithms require labelled datasets for the prediction of new labels. The algorithms include the simplest algorithms like K-nearest neighbour to Support vector machines which is most widely used in the classification of driving style.

Vaitkus et al [21] used data from accelerometer and used k-nearest neighbour algorithm for driving style classification. The samples are classified based on the majority votes of the k-nearest neighbour. Among 117 features only 7 features were considered for classification and accuracy of 100% was achieved. Woo et al [51] used Support vector machine algorithm for driving style classification which depends on the probability of each class. The accuracy was 71% with the F1 score. Wang et al [52] suggested the use of semi supervised support vector machine to overcome the time consuming problem in labelling the data. In this work few data were chosen and manually labelled using k-means clustering then quasi-Newton algorithms were used to assign an optimal label to all the training data. Then the semi supervised support vector machine was used to classify the driver into aggressive and normal style.

Random forest (RF) proves to be the best classifier out of all the classifiers as stated in [53]. This result also suits to the case of driving style classification. I et al [54] used Random forest for the classification of the driving style from low-risk to moderate-risk and the accuracy was 93%. In [55] the authors have used most of the classifier to develop a personalized driving state recognition. Among all the classifiers used Random forest showed the highest accuracy.
Markov chains means memorylessness which means that the current state purely depends on the previous state[59]. The hidden markov models [HMM] also have been applied to analyse the driver style. Gadepally et al [60] used HMM to analyse the driver style at intersections. Similarly Guardiola et al. [61] employed HMM to model the driver style and analyse its performance on energy management. Augustynowicz [62] ranked the drivers in terms of most active to extremely mild which would further lead to dynamic classification of driving style in the form of neural network.

Table 4: Driving style classification algorithms and the metrics

| Algorithm                  | References | Type Of Algorithm | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|----------------------------|------------|-------------------|---|---|---|---|---|---|---|---|
| Fuzzy logic and random forest | [3]        | S                 |   |   |   |   |   |   |   | * |
| Fuzzy logic                | [2] [15]   | S                 |   | * | * |   | * |   |   |   |
| K nearest neighbour        | [21] [34]  | S                 |   |   |   | * |   |   |   |   |
| Support Vector Clustering  | [16]       | US                |   |   |   |   |   |   |   |   |
| LSTM                       | [33]       | S                 | * |   |   |   |   |   |   |   |
| Bayesian probability       | [36]       | S                 | * |   |   |   |   |   |   |   |
| Support Vector Machine     | [51] [52]  | S                 | * |   |   |   |   |   |   |   |
| Random forest              | [54]       | S                 | * |   |   |   |   |   |   |   |
| Neural network             | [62]       | S                 | * |   |   |   |   |   |   |   |
| HMM                        | [60] [61]  | S                 |   |   |   |   |   |   |   |   |
| Bayesian regression        | [56] [57]  | US                | * |   |   |   |   |   |   |   |
| Gaussian mixture model     | [58]       | US                |   |   |   |   |   |   |   |   |

S: Supervised US: Unsupervised

1- Mean-Squared-Error 2- Recall 3- Accuracy 4- Score/Rank 5- Percentage 6- Time 7- ROC Curve 8- Squared Correlation Coefficient (R2)

5.2.2. Unsupervised learning

An unsupervised algorithm does not require supervision for the model construction. Instead, the model learns on its own to analyse the information. The unsupervised algorithms include Bayesian network, clustering algorithm, principal component analysis and so on. Support vector clustering was used by Feng et al [16] to study the driving style of three drivers. The data consists of 12 trips from 3 drivers. The classifier evaluated the driver style in terms of percentage. Furthermore, a positive correlation was confirmed between fuel consumption and driving aggressively. Moreover, it was found that weather conditions, the driver’s excitement and time of the day can cause notable changes in driving style.

Hierarchical Bayesian regression analysis was used by Mudgalet al [56] to study the driving style at roundabouts. McCalland Trivedi [57] used Bayesian framework to develop a novel method for braking assistance which fuse the vehicle and surround information to develop a driver style prediction. Drivers vary in the style they hit the brake and gas pedals, the way they turn the steering wheel, and the car-following distance to maintain safety and
well-being. Miyajima et al [58] used Gaussian mixture modelling to model the pedal operation and the car-following pattern.

5.2.3. Combination of Supervised and Unsupervised learning

Supervised and Unsupervised machine learning methods have been used together because of its advantages and to improve the overall performance of the system. Del Campo et al [19] proposed a hybrid machine learning approach which combines unsupervised hierarchical clustering and supervised Extreme learning machine (ELM). The hierarchical clustering is used to identify the driving style from the ride comfort perspective whereas the ELM is used to model the driving style classifier.

An approach to recognize the drivers curve negotiating style was developed in [35]. The authors used a combination of k-means clustering and support vector machines were used to reduce the recognition time. Van Ly [63] used data from inertial sensors of CAN bus to model the driver style. The authors used k-means clustering and support vector machines to reduce the complexity. Yi [64] have used almost all the classification algorithm to model a personalized driving state of a driver in which the random forest showed higher accuracy than the other algorithms.

Bender et al [66] converted the naturalistic driving data to high level driving data using unsupervised method. This work includes two parts. The first part divides the data into short segments. The second part matches the segments to high level driving data.

5.2.4. Discussion

A large variety of supervised, unsupervised and combination of these algorithms are used in the classification of driving style. Table 5 shows the description of metrics that were used to analyse the performance of the algorithms as stated in [66-68].

| Metric                  | Description                                                   |
|-------------------------|---------------------------------------------------------------|
| Mean-Squared-Error      | Difference between the actual and predicted values            |
| Recall                  | Number of positives the model returns                         |
| Accuracy                | Number of accurate predictions                                |
| Score/Rank              | Generating values or ranks based on predictions               |
| Percentage              | Proportions under each class                                  |
| Time                    | Time taken for processing                                     |
| ROC Curve               | Capability of the model                                       |
| Squared Correlation Coefficient (R2) | Goodness of fit                                        |

The mostly used metric to evaluate the algorithms are accuracy but they differ from the fact they use distinct features. The time denotes the time-taken for prediction is taken as performance measure to analyse the algorithm. On the other hand the score / rank classify the driver based on some prefixed thresholds. Clustering based approach uses percentage as the measure of performance to categorize the data under each class. Regression based methods use squared correlation coefficient to indicate the achievement.

6. Conclusion

In this paper, we have investigated the various devices and a bunch of features customized by industries and research papers that are used to identify the driving style. Related work in terms of objectives of the research, data collection approaches with advantages and disadvantages, parameters selected, purpose, and inferences are discussed. The algorithms used and the metrics used to analyze the performance of driving style classification was also discussed in detail. Since there is a progressive development in ADAS
systems and autonomous vehicles there is a need for more extensive driver style analysis methods. Most of the researches discussed in this paper consider vehicle-based measurements. The researches done on the driving style analysis uses a wide variety of devices and several parameters. This analysis shows that a diverse set of features that can be used to analyze the style of driving and seeks to identify the best set of features through feature selection algorithms. As a future work, this paper provides a foundation to identify the driving style with the machine learning techniques on larger datasets and its application. In addition, the driver’s psychological and physiological factors which are least discussed in this paper will be considered while developing driver style classification models. This survey places the driving style classification as a critical concept in intelligent vehicle design and its integration with the market.

References

[1] https://niti.gov.in/sites/default/files/2019-07/India%E2%80%99s-Energy-and-Emissions-Outlook.pdf accessed 4 December 2019.

[2] Massoud, R., Poslad, S., Bellotti, F., Berta, R., Mehran, K., & De Gloria, A. (2018). A fuzzy logic module to estimate a driver’s fuel consumption for reality-enhanced serious games. International Journal of Serious Games, 5(4), 45-62.

[3] Massoud, R., Bellotti, F., Berta, R., De Gloria, A., & Poslad, S. (2019, April). Exploring fuzzy logic and random forest for car drivers’ fuel consumption estimation in IoT-enabled serious games. In 2019 IEEE 14th International Symposium on Autonomous Decentralized System (ISADS) (pp. 1-7). IEEE.

[4] Barkenbus, J. N. (2010). Eco-driving: An overlooked climate change initiative. Energy policy, 38(2), 762-769.

[5] Barbé, J., & Boy, G. (2006, September). On-board system design to optimize energy management. In Proceedings of the European Annual Conference on Human Decision-Making and Manual Control (EAM 2006), Valenciennes, France (pp. 27-29).

[6] Koskinen, O. H. (2008). Improving vehicle fuel economy and reducing emissions by driving technique. In 15th World Congress on Intelligent Transport Systems and ITS America’s 2008 Annual Meeting ITS AmericasERTICOITS JapanTransCore.

[7] World Health Organization. Global status report on road safety. 2018-https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/

[8] Hong, J. H., Margines, B., & Dey, A. K. (2014, April). A smartphone-based sensing platform to model aggressive driving behaviors. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 4047-4056).

[9] Mitrovic, D. (2005). Reliable method for driving events recognition. IEEE transactions on intelligent transportation systems, 6(2), 198-205.

[10] https://en.wikipedia.org/wiki/Advanced_driver_assistance_systems accessed 16 December 2019.

[11] Derhab, A., Guerroumi, M., & Younis, M. (2017). Wireless and mobile sensing technologies for the future internet.

[12] Le Nguyen, D., Lee, M. E., & Lensky, A. (2012, December). The design and implementation of new Vehicle Black Box using the OBD information. In 2012 7th International Conference on Computing and Convergence Technology (ICCCCT) (pp. 1281-1284). IEEE.

[13] https://www.embitel.com/automotive-insights/on-board-diagnostics-obd-ii-stack accessed 18 December 2019.

[14] Lee, C. F., & Öberg, P. (2015). Classification of road type and driving style using obd data (No. 2015-01-0979). SAE Technical Paper.

[15] Peraud, I. S., & Narango, J. E. (2019, January). Are you a Good Driver? A Data-driven Approach to Estimate Driving Style. In Proceedings of the 14th International Conference on Computer Modeling and Simulation (pp. 3-7).

[16] Feng, Y., Pickering, S., Chappell, E., Iravani, P., & Brace, C. (2019). A support vector clustering based approach for driving style classification. International Journal of Machine Learning and Computing, 9(3), 344-359.

[17] Ferreira, J. C., de Almeida, J., & da Silva, A. R. (2015). The impact of driving styles on fuel consumption: A data-warehouse-and-data-mining-based discovery process. IEEE Transactions on Intelligent Transportation Systems, 16(3), 2653-2662.

[18] Jakobsen, K., Mouritsen, S. C., & Torp, K. (2013, November). Evaluating eco-driving advice using GPS/CANBus data. In Proceedings of the 21st ACM SIGSPATIAL international conference on advances in geographic information systems (pp. 44-53).

[19] del Campo, I., Asua, E., Martínez, V., Mata-Carballeira, Ó., & Echanobe, J. (2018, November). Driving style recognition based on ride comfort using a hybrid machine learning algorithm. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 3251-3258). IEEE.
[20] https://learn.sparkfun.com/tutorials/accelerometerbasics/all accessed 18 December 2019

[21] Vaitkus, V., Lengvenis, P., & Žylis, G. (2014, September). Driving style classification using long-term accelerometer information. In 2014 19th International Conference on Methods and Models in Automation and Robotics (MMAR) (pp. 641-644). IEEE.

[22] Abu-Bakar MH, Razuan R, Kamal S. (2019) Teenage Driving Style Modelling Using Deep Learning for Driver Style Classification. In Proceedings of the Third International Conference on Computing, Mathematics and Statistics (ICCMS2017) (pp. 449-455). Springer, Singapore.

[23] Fan, J., Li, Y., Liu, Y., Zhang, Y., & Ma, C. (2019, June). Analysis of taxi driving behavior and driving risk based on trajectory data. In 2019 IEEE Intelligent Vehicles Symposium (IV) (pp. 220-225). IEEE.

[24] Zhang, Q., Li, X., Pan, H. L., Wang, J. T., & Zhao, Z. J. (2018). Detection of vehicle tracks by a three-axis magnetometer. Sensors and Actuators A: Physical, 276, 83-90.

[25] Saus, R., & Varaiya, P. (2008). Wireless magnetic sensors for traffic surveillance. Transportation Research Part C: Emerging Technologies, 16(3), 294-306.

[26] Vacher, F., Alves, F., & Gilles-Pascaud, C. (2007). Eddy current nondestructive testing with giant magneto-impedance sensor. NDT & E International, 40(6), 439-442.

[27] Zhang, Z., Li, X., Yuan, H., & Yu, F. (2013). A street parking system using wireless sensor networks. International Journal of Distributed Sensor Networks, 9(6), 107975.

[28] Castignani, G., Derrmann, T., Frank, R., & Engel, T. (2015). Driver behavior profiling using smartphones: A low-cost platform for driver monitoring. IEEE Intelligent Transportation Systems Magazine, 7(1), 91-102.

[29] Johnson, D. A., & Trivedi, M. M. (2011, October). Driving style recognition using a smartphone as a sensor platform. In 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC) (pp. 1609-1615). IEEE.

[30] Castignani, G., Derrmann, T., Frank, R., & Engel, T. (2017). Smartphone-based adaptive driving maneuver detection: A large-scale evaluation study. IEEE Transactions on Intelligent Transportation Systems, 18(9), 2330-2339.

[31] You, C. W., Montes-de-Oca, M., Bao, T. J., Lane, N. D., Lu, H., Cardone, G., & Campbell, A. T. (2012, September). CarSafe: a driver safety app that detects dangerous driving behavior using dual-cameras on smartphones. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (pp. 671-672).

[32] Castignani, G., Frank, R., & Engel, T. (2013, October). An evaluation study of driver profiling fuzzy algorithms using smartphones. In 2013 21st IEEE International Conference on Network Protocols (ICNP) (pp. 1-6). IEEE.

[33] Mumcuoglu, M. E., Alcan, G., Unel, M., Cicek, O., Mutluergil, M., Yilmaz, M., & Koprubasi, K. (2019, July). Driving Behavior Classification Using Long Short Term Memory Networks. In 2019 AET International Conference of Electrical and Electronic Technologies for Automotive (AET AUTOMOTIVE) (pp. 1-6). IEEE.

[34] Yang, L., Guan, W., Ma, R., & Li, X. (2019). Comparison among driving state prediction models for car-following condition based on EEG and driving features. Accident Analysis & Prevention, 133, 105296.

[35] Wang, W., & Xi, J. (2016, July). A rapid pattern-recognition method for driving styles using clustering-based support vector machines. In 2016 American Control Conference (ACC) (pp. 5270-5275). IEEE.

[36] Wang, W., Xi, J., & Li, X. (2016). Statistical pattern recognition for driving styles based on Bayesian probability and kernel density estimation. arXiv preprint arXiv:1606.01284.

[37] Kuhler, M., & Karstens, D. (1978). Improved driving cycle for testing automotive exhaust emissions. No. 780650. SAE Technical Paper.

[38] Ericsson, E. (2001). Independent driving pattern factors and their influence on fuel-use and exhaust emission factors. Transportation Research Part D: Transport and Environment, 6(5), 325-345.

[39] Li, G., Zhu, F., Qu, X., Cheng, B., Li, S., & Green, P. (2019). Driving style classification based on driving operational pictures. IEEE Access, 7, 90180-90189.

[40] Taubman–Ben-Ari, O., Eberenfreund–Hager, A., & Prato, C. G. (2016). The value of self-report measures as indicators of driving behaviors among young drivers. Transportation research part F: traffic psychology and behaviour, 39, 33-42.

[41] Amado, S., Arkan, E., Kaça, G., Koyuncu, M., & Turkan, B. N. (2014). How accurately do drivers evaluate their own driving behavior? An on-road observational study. Accident Analysis & Prevention, 63, 65-73.

[42] Taubman–Ben-Ari, O., & Yehiel, D. (2012). Driving styles and their associations with personality and motivation. Accident Analysis & Prevention, 45, 416-422.

[43] Tao, D., Zhang, R., & Qu, X. (2017). The role of personality traits and driving experience in self-reported risky driving behaviors and accident risk among Chinese drivers. Accident Analysis & Prevention, 99, 228-235.

[44] Brown, T. G., Ouimet, M. C., Eldred, M., Tremblay, J., Vingilis, E., Nadeau, L., & Bechara, A. (2017). The effect of age on the personality and cognitive characteristics of three distinct risky driving offender groups. Personality and individual differences, 113, 48-56.

[45] Ryder, B., Gahr, B., Egolf, P., Dahlinger, A., & Wortmann, F. (2017). Preventing traffic accidents with in-vehicle decision support systems: The impact of accident hotspot warnings on driver behaviour. Decision support systems, 99, 64-74.
