WILEY | Hindawi

Research Article

Heterogeneity in the Driver Behavior: An Exploratory Study Using Real-Time Driving Data

Jahnavi Yarlagadda [ID] and **Digvijay S. Pawar** [ID]

Transportation Systems Engineering, Department of Civil Engineering, Indian Institute of Technology Hyderabad, Kandi, Medak 502285, India

Correspondence should be addressed to Digvijay S. Pawar; dspawar@ce.iith.ac.in

Driver behavior heterogeneity is a significant aspect to understand the individual behavioral variations and develop driver assistance systems. This study characterizes the heterogeneity in driving behavior using real-time driving performance features. In this context, the study investigates the extent of variations in the individual's driving styles during routine driving. The driving styles are conceptualized using the vehicle kinematic data, that is, speed and accelerations performed during longitudinal control. The data is collected for 42 professional drivers using instrumented vehicle over a defined study stretch. An algorithm is developed for data extraction and total 7548 acceleration and 6156 braking maneuvers and corresponding driving performance features are extracted. The driving maneuver data are analyzed using the unsupervised techniques (PCA and K-means clustering) and three patterns of acceleration and braking are identified, which are further associated with two patterns of speed behavior. The results showed that each driver is found to exhibit different driving patterns in different driving regimes and no driver shows constantly safe or aggressive behavior. The aggression scores are found to be different among drivers, indicating the behavioral heterogeneity. This study results demonstrate that, driver's level of aggression in different driving regimes is not constant and characterizing the driver by means of abstract driving features is not indicative of the diversified driving behavior. The proposed method identifies the individualized driving behaviors, reflecting the driver's choice of driving maneuvers. Thus, the insights from the study are highly useful to design driver-specific safety models for driver assistance and driver identification.

1. Introduction

Road traffic accidents are one of the leading causes of death resulting in approximately 1.35 million deaths every year [1]. The factors associated with road crashes are studied over decades and driver behavior is concluded to be the major contributory factor [2–7]. Therefore, understanding the driver behavior is important for many applications like driver assistance or personalized feedback provision for enhancing the driving safety, economy, and comfort. In addition, the implications of driver behavior research are significant inputs for the design of autonomous vehicles.

Driver behavior indicates the manner of executing various driving tasks, which can be perceived as controlling the vehicle in the longitudinal and lateral directions. The habitual way of performing driving maneuvers is considered as driving style, which characterizes the individual driver or a group of drivers [8]. Many researchers have attempted to classify the drivers and the driving styles based on the outcomes of driving tasks in the perspective of driving safety [9–20]. Most of the studies considered the same predefined criteria (thresholds for abstract features of driving data or for safety-critical events) for all the drivers neglecting the variation in driver attributes and driving skills. The differences in the individual driving characteristics among drivers might result in different driving responses toward a set of driving conditions [21]. Thus, classifying the driver behavior based on the predefined thresholds common to all the drivers may not yield to the proper evaluation of driver behavior. Moreover, the driving styles are assumed to be

constant or consistent in the entire trip, which is unrealistic as the external driving conditions tend to change time to time [22, 23]. In other words, the driver classified as aggressive may not continuously exhibit aggressive behavior in all the driving maneuvers performed in a trip. Therefore, this study proposes a framework to explore the behavioral heterogeneity in the longitudinal control by segmenting the driving profiles into driving maneuvers and identifying different driving patterns. The analysis is conducted using unsupervised machine learning techniques to group and interpret the driving patterns, without any fixed thresholds.

In this study, a high-frequency (10 Hz) GPS instrumentation is used to collect the real-time driving data for passenger car drivers. The longitudinal control exhibited by drivers in terms of speeding, accelerating, and braking were considered to explore the driving pattern variations in short-term driving decisions. The next section of the paper presents a summary of the literature review on driving style identification. The third section details the data collection techniques and the fourth section presents the adopted methodology. The fifth section details the results and discussion. The last section of the paper presents the conclusions, and future work.

2. Literature Review

Over the last three decades, many researchers have conducted the safety analysis of driving behavior that prominently involved identifying and classifying driving styles. The driving style detection and classification were mainly regarded as a means of differentiating drivers in the perspective of driving safety and to evaluate individuals. The existing literature on driving style classification is synthesized as per the nature of the data used for the analysis and presented in the following subsections.

2.1. Predefined Thresholds of Safety Critical Events. With the advancement in the data collection techniques, the driving styles were conceptualized using different driving performance features depending on the research motive and method of data collection. Majority of the studies classified the driving styles based on the predefined thresholds of safety-related events, such as sudden acceleration, sudden braking, and sharp turning [9, 11–13, 16, 17, 20, 24–29]. Johnson and Trivedi [11] developed smartphone-based application, which categorizes the driving styles as aggressive and nonaggressive based on the nature of detected driving maneuvers. The sudden accelerations and braking, swift swerves, and hard right and left turns were the typical maneuvers considered, for which the reference thresholds were recorded by a single driver and a vehicle. Aljaafreh et al. [12] classified driving styles into four categories using the fuzzy logic inference system. The fuzzy rules for classification were predefined based on the driving performance of three expert drivers in terms of speed and longitudinal/lateral accelerations. Similarly, Feng et al. [24] developed a fuzzy logic driver model to simulate different driving styles. The fuzzy rules were framed based on experts' knowledge on

drivers' decisions using the parameters such as vehicle speed, headway distance, pedal position, and gear selection. Van Ly et al. [13] explored the possibility of CAN bus data to build driving profiles for characterizing individuals. Authors represented the driving maneuvers using acceleration, braking, and turning events. The k-means clustering and support vector machine were used in the training algorithm and showed an accuracy of 60%. Vaiana et al. [16] classified driving behaviors into safe and aggressive using a g-g diagram developed by referencing a single driver. Eboli et al. [17] constructed a g-g diagram and defined the safety domain for driver behavior classification. The frequency of data points out of the safety domain was considered to classify drivers as safe or unsafe. Mantouka et al. [20] categorized the trips into six distinct groups with the increasing level of safety. The driving performance over the trip was defined using the frequencies of harsh accelerations and harsh brakes per kilometer traveled and the percent of speeding and mobile usage during the trip. Chen et al. [25] developed a supervised hierarchical Bayesian model to understand the latent driving styles using the labeled data. The vehicle motion data, including acceleration, speed, and turning signals were categorized into multiple levels ranging from very low to very high and encoded to input the model. In these studies, the predefined thresholds were needed to detect the driving events, but the thresholds used for defining the driving events were not consistent. Paefgen et al. [26] used ± 0.1 g (1 m/s^2) for identifying critical acceleration and braking events, whereas Fazeen et al. [27] used the thresholds of ± 0.3 g (3 m/s^2). Bagdadi [28] used ± 0.48 g (4.8 m/s^2), and Bergasa et al. [29] considered ± 0.4 g (4 m/s^2) as thresholds for critical events.

2.2. Abstract Driving Features. Some studies classified the driving styles based on the abstract features of driving performance data observed during the study period [10, 15, 30]. Constantinescu et al. [10] classified driving styles as per the risk-proneness of drivers in the context of road safety. The hierarchical cluster analysis was performed on the abstract features of driving performance data, such as the mean and standard deviation of speed, acceleration, deceleration, percentage of time speeding over speed limit, and the mechanical work over the entire trip. The identified driving styles were compared against the test driver's performance and categorized into five groups ranging from nonaggressive to aggressive. Similarly, Kalsoom and Halim [30] classified driving styles into slow, normal, and fast categories based on the abstract features such as maximum and average speed, number of brakes, and number of horns during the trip. Given that, the performance aggregated over an entire trip was used for classification, the results of these studies cannot directly indicate the unsafe practices or the individual's driving faults to diagnose through assistance.

2.3. Driving Profiles. A few studies explored the driving styles using continuous driving profile data, without any prior ground truth [19, 22, 23, 31, 32]. Li et al. [31] considered the transition probabilities between driving

maneuvers as indicators of driving styles and analyzed the transitions among 12 types of maneuvers. All the maneuver episodes were manually identified from the naturalistic data corresponding to 28 drivers. The authors developed random forest classifier based on the optimized set of five transition features and classified drivers into three risk groups, that is, high, moderate, and low-risk representative of the entire trip. Further, the obtained classification was compared against the subjective evaluation scores given by expert team. The authors concluded that transition probabilities were resulting in better estimation accuracy (overall recognition rate of 93%) compared to other traditional methods. Chen et al. [22] developed driving behavior graph for each driving performance feature by considering the driving pattern exhibited over every three seconds. However, the univariate approach in behavior recognition does not account the interdependencies of driving decisions. Higgs and Abbas [23] considered the combination of driving performance features to segment and cluster the car-following periods and explore the driving patterns of individual drivers. Authors analyzed the driving data of 10-truck and 10-car drivers collected as a part of the 100-car naturalistic driving study. The state-action variables corresponding to the segmented car-following periods were used for clustering. A total of 30 clusters were identified representing different driving patterns and the proportion of patterns varied among drivers. Similarly, Chen and Chen [19] analyzed the base-line events (recorded as a part of SHRP-2 study) concerning the driving performance over the event. Three clusters of events were identified representing three different driving styles. Li et al. [32] proposed an unsupervised framework to segment the driving sequences into fragments and cluster the fragments into descriptive driving patterns. The authors utilized two Bayesian algorithms for segmentation, and two extended latent Dirichlet allocation (LDA) models for clustering the driving patterns. Total four driving patterns were identified common to all the driving maneuvers, where each pattern was a combination of both longitudinal and lateral behavioral characteristics. Although, these studies evaluated the driving maneuvers and driving patterns using the driving profile data, the variation in the individual driving characteristics is still unclear and not conclusive. The objectives of studies, instruments used for data collection, the sampling frequency, the driving features, and the kind of data used for analysis across the previous studies is summarized in Table 1.

2.4. Research Gaps. The literature review shows that the number of classes or groups to which the drivers or driving styles are classified was not consistent among the studies. Also, the definition of the class or group was inconsistent across the studies and varied concerning the study methodologies. Based on the in-depth review of the literature on driving style classification, the research gaps observed are highlighted as follows:

(1) Most of the studies classified driving styles based on either the abstract performance features or the predefined thresholds. The aggregated features over

entire trip does not indicate the nature of short-term driving decisions and the respective pattern variations in the individuals. Whereas in case of predefined thresholds, the thresholds used across the studies were not uniform.

(2) In majority of the studies, the drivers and driving styles were characterized by a single classification of safe or aggressive for the entire trip. Very limited research is available which speaks of variations in the individual's driving styles.

(3) None of the previous studies presented the intradriver and interdriver behavioral heterogeneity in the instantaneous driving decisions in different driving regimes.

2.5. Research Objectives. Given the research gaps, the study is aimed to pursue the following objectives to explore the extent of variability in individual's driving patterns in short-term driving decisions.

(a) Segmenting the high-frequency (10 Hz) driving profile data into driving maneuvers (acceleration and braking), and extracting the respective driving performance data.

(b) Identifying different driving patterns using unsupervised machine learning techniques, without using predefined ground truth.

(c) Assigning the driving style classification to individuals and quantifying the driver performance using relative aggression score.

(d) Exploring the individual's behavioral heterogeneity in instantaneous driving decisions.

3. Methodology

The proposed framework for the current study is presented in Figure 1. The framework is divided into five steps: (1) Data acquisition, (2) Maneuver detection and feature extraction, (3) Dimensionality reduction, (4) Maneuver clustering and driving style classification, and (5) Driver behavior heterogeneity. First, the real-time driving profiles of passenger car drivers were collected using high-frequency (10 Hz) GPS instrumentation. Second, the driving profiles were segmented into acceleration and braking maneuvers, and the respective driving performance features were extracted. As the extracted dataset was unlabeled concerning the nature of maneuver, the unsupervised learning techniques were used to explore the underlying patterns. Therefore, in the third step, the principal component analysis was conducted on acceleration and braking datasets, to reduce the feature dimensionality and to improve the clustering efficiency. Fourth, the k-means clustering was used to group the similar patterns of maneuvers. Then, the identified groups were assigned a driving style classification based on the respective characteristics of performance features. Finally, the proportion of each driving pattern was computed for individuals, and assigned a relative aggression score based on the frequency of the aggressive maneuvers exhibited per

TABLE 1: Summary of the literature on driving style classification.

Studies	Objective	Instruments	Sampling frequency	Driving features	Nature of the data
Murphey et al. [9]	Driving style classification—calm, normal, aggressive, no speed	—	—	Acceleration and jerk features from 11 standard driving cycles	Pre-defined thresholds of safety critical events
Johnson and Trivedi [11]	Driving style classification—aggressive, nonaggressive	Smartphone sensors	1 Hz	Speed, acceleration, rotation	
Eboli et al. [17]	Driver behavior classification—safe, unsafe	GPS	1 Hz	Speed, acceleration	
Feng et al. [18]	Driver classification—aggressive, nonaggressive	CAN bus	10 Hz	Speed, acceleration, jerk	
Mantouka et al. [20]	Trip classification—six categories from aggressive to nonaggressive	Smartphone sensors	1 Hz	Speed, acceleration, percentage of mobile usage while driving	
Feng et al. [24]	Simulating driving styles—aggressive, normal, defensive	On board data logger (OBD-II), radar	1 Hz	Speed, headway, pedal movement, gear position	
Chen et al. [25]	Driving style modeling—aggressive, moderate, careful	GPS	10 Hz	Speed, acceleration	
Bagdadi [28]	Detection of safety critical braking events	GPS	10 Hz	Acceleration, jerk	
Constantinescu et al. [10]	Driver classification—six categories from very aggressive to nonaggressive	GPS	1 Hz	Speed, acceleration	Abstract driving features
Hong et al. [15]	Driving style classification—Aggressive, nonaggressive	Smartphone sensors	2 Hz	Self-reports, speed, acceleration, engine RPM	
Kalsoom and Halim [30]	Driver classification—slow, normal, fast	Driving simulator	0.1 Hz	Speed, acceleration, frequency of brakes, horns, gear, indicators	
Chen and Chen [19]	Driving style classification—three groups	GPS	10 Hz	Speed, acceleration	Driving profiles
Chen et al. [21]	Driving style modelling—Individual driving habit graphs	Driving simulator	—	Speed, acceleration, steering wheel angle, throttle actuation	
Chen et al. [22]	Identifying individual driving patterns	GPS	1 Hz	Speed, acceleration	
Higgs and Abbas [23]	Driving pattern identification—30 patterns	GPS, cameras, radar	10 Hz	Speed, acceleration, yaw rate, range, lane offset	
Li et al. [31]	Driving-style classification—low-risk, moderate-risk, high-risk	CAN bus, cameras	10 Hz	Speed, acceleration, headways, video data	
Li et al. [32]	Identifying descriptive driving patterns—four patterns	CAN bus, cameras	10 Hz	Speed, acceleration, headways, video data	

kilometer traveled. Then, the behavioral heterogeneity was presented based on the driving style variations observed with-in, and among the individuals.

3.1. Data Acquisition

3.1.1. Study Stretch. Driving behavior depends on several external factors, among which the road geometry and driving environment play a significant role. To keep the driving environment (with respect to road geometry) uniform for all the participants, a predefined study stretch was used for collecting the driving data. The identical test route ensures each participating driver faces similar geometrical elements and other road infrastructure features. The study stretch is of total 23 km length on the four-lane national highway (NH-64) near Hyderabad city. The selected route consists of 12 intersections, 13 mid-block openings, and four gentle curves.

3.1.2. Driver Participation. The recruitment of drivers for the present study was primarily conducted among the fleet management companies surrounding the Hyderabad city. The participation was voluntary, and the interested candidates were asked to fill a short questionnaire consisting of demographic questions of suitability. The knowledge of the route and how frequent the driver travels on a particular route influences the driving outcomes [33].

Considering the route familiarity and minimum driving experience of one year, total 42 drivers were chosen who were frequent travelers on the study route. The age of the

Step 1: Data acquisition

Selection of study stretch

Participant recruitment
- Route familiarity
- Demographics

Instruments acquisition & installation

Driving data collection

Kinematic data
- Speed
- Acceleration

Positional data
- Longitude & Latitude
- Distance and heading

Video data
- Road view
- Driver view

Step 2: Maneuver detection and feature extraction

- Acceleration maneuvers
- Braking maneuvers

Segmenting the driving profile

Removing errors and inappropriate maneuvers

Extracting driving performance features

Step 3: Dimensionality reduction

PCA

Feature reduction

Step 4: Maneuver clustering & Driving style classification

Multi-variate K-means clustering

- Acceleration clusters
- Braking clusters
- Speed clusters

Classifying driving styles

Step 5: Driver behavior heterogeneity

Driving pattern proportions of individuals

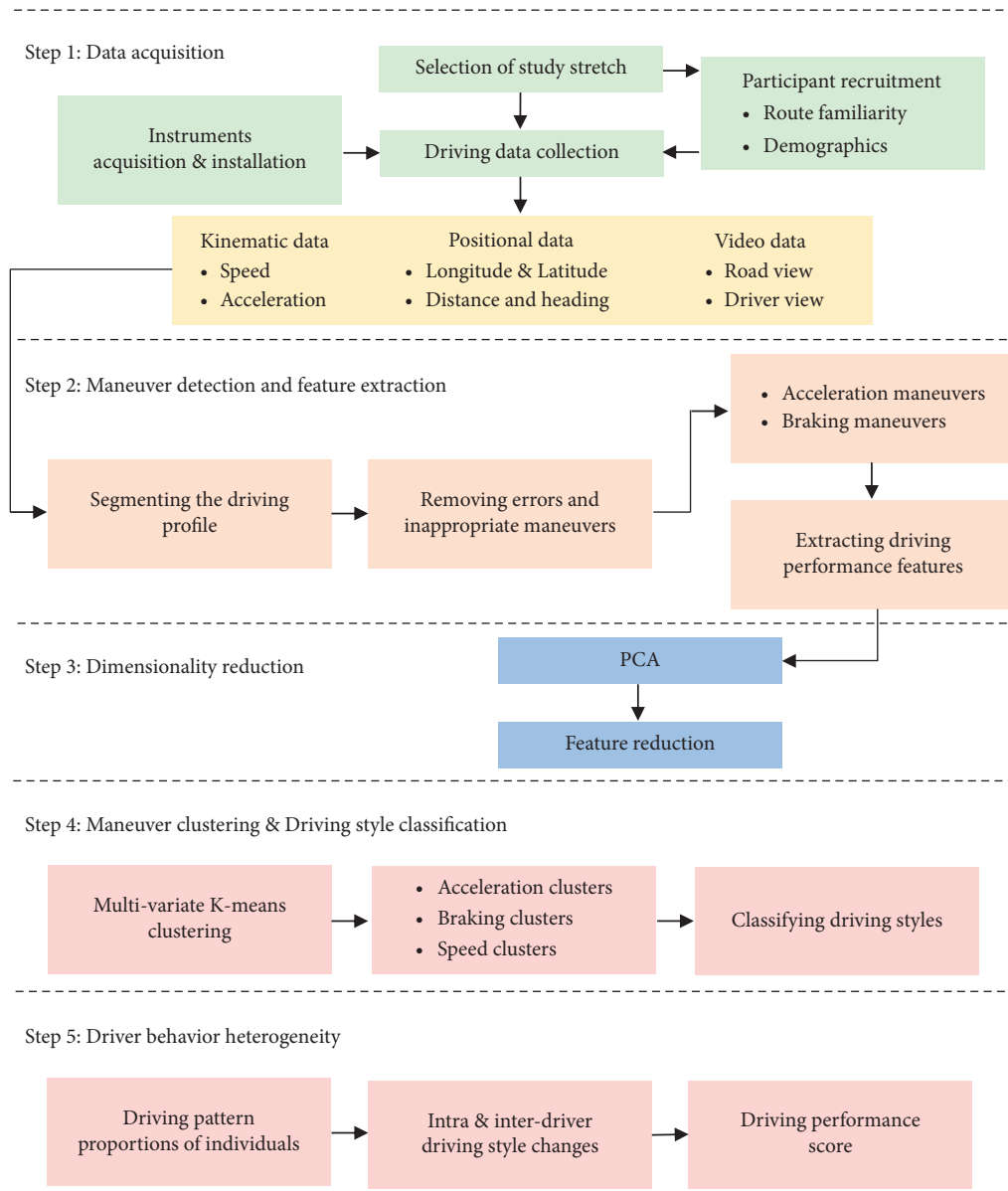Intra & inter-driver driving style changes

Driving performance score

FIGURE 1: Proposed framework of the study methodology.

selected drivers ranged from 19 to 45 years with a mean of 29.9 (sd = 6.8) years and an average driving experience of 7.4 (sd = 5.4) years. The participants were introduced to the goals of the study and informed to drive in their natural way as they would generally drive in day-to-day routine. The study was ethically approved by the Institutional Ethics Committee, Indian Institute of Technology Hyderabad, and a declaration of consent was obtained from each participant before the commence of trip. The data were collected for a minimum of two trips for each participant and were compensated by a gift after the study period.

3.1.3. Driving Data Acquisition System. A passenger car was used to collect the driver behavior data and was equipped with a high-frequency GPS instrumentation and four video-cameras (see Figure 2). The vehicle was instrumented before

starting the scheduled trip and removed after completion of the trip. The instrument captures the vehicle speed, longitudinal/lateral acceleration, heading, and positional coordinates at a frequency of 10 Hz along with the synchronized video data. The measured absolute position and speed are accurate to ±3 m and 0.1 km/h, respectively. To minimize the effect of weather and traffic conditions, the data was collected in the dry weather conditions, for the trips (ride requests) that were scheduled during off-peak hours. A total of 98 trips data were collected, comprising 65 hours of driving data over 2254 kilometers.

3.2. Maneuvre Detection. The speed and acceleration are the most commonly used kinematic parameters to distinguish the driver behaviors with respect to the level of aggressiveness [11, 12, 17, 26, 34–37]. Considering the knowledge

FIGURE 2: (a) Instrumentation details; (b) Study vehicle; (c) Recorded speed and acceleration profiles; and snapshots of collected video data.

from previous studies, present study explores the changes in driving styles based on the driving patterns exhibited in acceleration and braking maneuvers. An algorithm was designed to segment the driving profiles into acceleration and braking maneuvers and extract the respective driving performance features. The algorithm works at three levels to extract the significant acceleration and braking maneuvers. The conceptual representation of the segmentation process is depicted in Figure 3. (i) In the first level, the acceleration and braking maneuvers are identified, and the respective driving performance features are extracted. In this level, the instantaneous rate of change in the speed profile was computed and the segments with the positive and negative rate of change were categorized under acceleration and braking segments, respectively. Keeping the sign of rate of change as a reference, the consecutive positive change points/negative change points were grouped (speed bins) to represent a segment. However, this resulted in many number of segments corresponding to minor fluctuations in the high-frequency GPS data.

To eliminate the fluctuations and identify the actual maneuvers, a threshold was applied on the minimum size of the speed bin ranging from 1 to 6 time stamps (0.1 sec to 0.6 sec). To finalize the appropriate minimum bin size, the maneuvers were manually identified for five data files and checked for the identification rate. The minimum bin size of 3 resulted in higher efficiency in the maneuver identification. Whereas, the higher bin size resulted in losing more number of maneuvers and smaller size lead to identification of more inappropriate maneuvers. Thus, all the bins of size less than 3 were eliminated as GPS fluctuations and the resulting sequential positive bins/negative bins were conglomerated representing the acceleration/braking maneuvers. (ii) In the second level, the insignificant maneuvers resulting due to small speed fluctuations (change in speed <5 kmph) and lower speed values (speed <15 kmph) were eliminated. (iii) In the third level, the free decelerations resulted from the release of acceleration pedal were separated from the braking maneuvers. The part of the deceleration segments with a deceleration value less than 0.05 g were deducted. The above
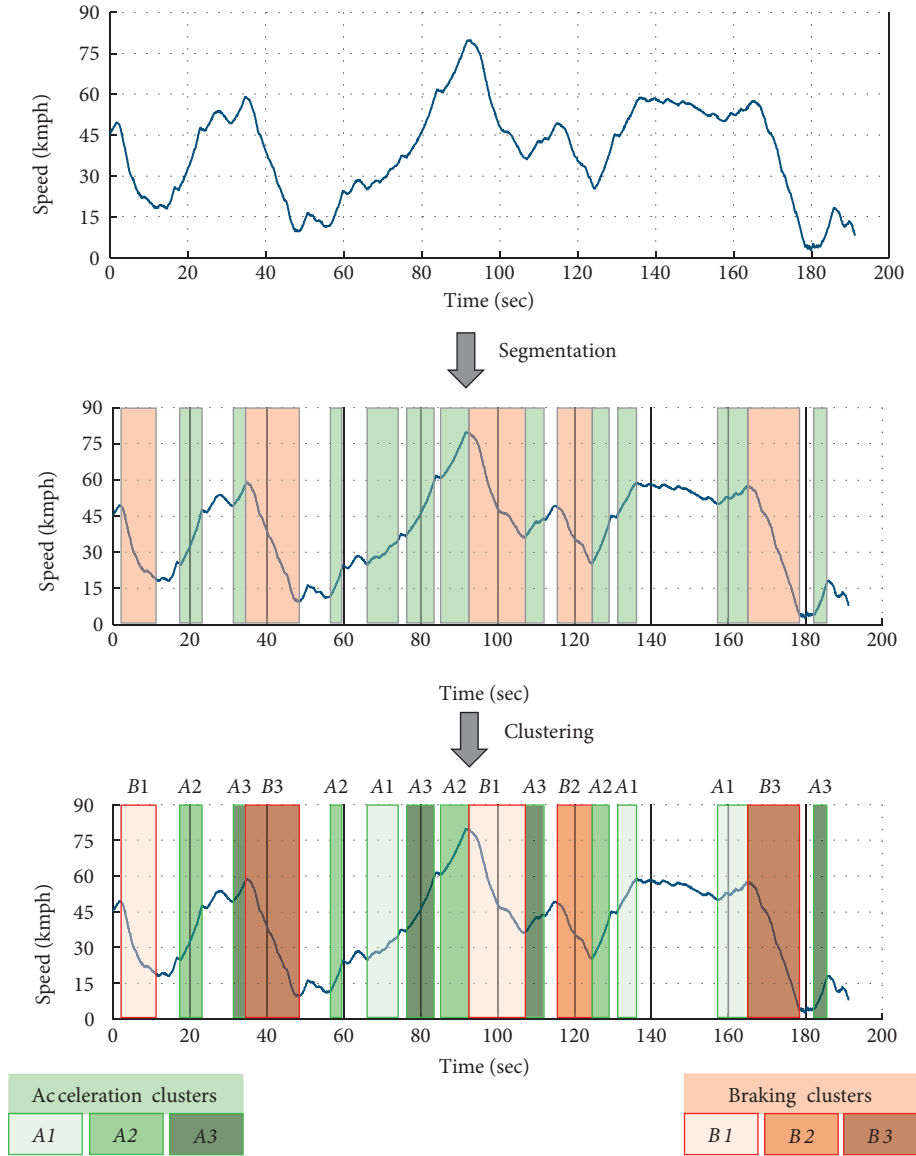
FIGURE 3: Illustration of segmentation and clustering process.

thresholds used for data extraction were framed after the manual video observation for 300 randomly chosen acceleration and braking maneuvers.

3.3. Feature Extraction.

The driving performance features of the final acceleration/braking segments were extracted from the respective driving profile data. The driving performance features include the motion data and the respective derived statistical features (see Table 2). For each identified maneuver, the speed and acceleration performance were extracted along with the respective maximum yaw rate. The yaw rate is the rate of change of heading angle, which indicates the vehicular lateral movement while maneuvering. The acceleration/braking maneuvers associated with high yaw rates tend to reflect the aggressive driving styles. Thus, the maximum yaw rate corresponding to each maneuver was

TABLE 2: Driving performance features of acceleration and braking maneuvers.

Feature Description	Variable notation
Minimum speed (kmph)	V_{min}
Maximum speed (kmph)	V_{max}
Change in speed (kmph)	$\Delta\mathbf{V}$
Duration of maneuver (sec)	$\Delta\mathbf{T}$
Mean speed (kmph)	V_{mean}
Standard deviation of speed	V_{sd}
Maximum longitudinal acceleration/deceleration (g)	LA_{max}/LA_{min}
Mean longitudinal acceleration/ deceleration (g)	LA_{mean}
Standard deviation of longitudinal acceleration/ deceleration	LA_{sd}
Maximum yaw rate (°/sec)	Yr_{max}

computed from the raw data of heading values. The algorithms for segmentation and feature extraction were coded in Python 3.7.

3.4. Dimensionality Reduction Using PCA.

The identified acceleration and braking maneuvers are characterized by a set of driving performance features. Each feature is representative of different driving behaviors and interpreting the nature of maneuver by combining all the features is a difficult or more of a subjective task. Thus, a dimensionality reduction technique was used to reduce the number of original variables to a more interpretable combination. The principal component analysis (PCA) is an unsupervised dimensionality reduction technique, which linearly transforms the correlated variables to a new set of uncorrelated principal components (PC). The number of successive PCs explaining 80–90% of total variance are considered to be a good representative of the dataset [10]. To understand the meaning of PCs, the loadings are computed, which shows the correlation between the PCs and original features. Higher loadings represent stronger correlations between features and the respective PCs. Based on the results of the PCA, the feature dimensionality was reduced prior clustering.

3.5. Clustering and Driving Style Classification.

To explore the underlying patterns and group based on the similarity, we performed the k-means clustering on the feature data. K-means is an unsupervised technique which groups the data based on the intrinsic similarity in the dataset. K-means clustering is the simplest and most widely used machine learning algorithm, which finds similar groups by minimizing the Euclidean distance between centroids. As the algorithm works on unlabeled datasets, the target number of centroids/clusters (k) should be defined prior to clustering. In this study, we used elbow and silhouette methods to decide the optimal k-value and validate the clusters.

Further, the identified clusters of driving maneuvers were categorized based on the respective cluster characteristics. Then, the driving style classification was assigned to each maneuver as per the cluster number under which it is grouped.

3.6. Driver Performance Score.

After assigning the driving styles to each driving maneuver, the proportion of acceleration and braking maneuvers corresponding to different clusters were computed for individuals. The drivers exhibiting higher proportions of aggressive driving maneuvers indicate unsafe driving behavior and thus need to be identified. In this context, the driving performance of each driver was quantified based on the number of aggressive maneuvers exhibited over the observed period. However, the thresholds for safe or unsafe driving behaviors have not been established in the literature. Therefore, in this study, the performance score was computed relative to the maximum number of aggressive maneuvers per kilometer traveled among all the drivers. The drivers exhibiting higher aggression while accelerating, braking, and speeding were considered as a benchmark to assign the relative aggression score to other drivers. The number of aggressive patterns per kilometer traveled is normalized using the reference maximum value, such that each driver takes a score between zero to 100. The lower scores represent lower levels of aggression on a relative scale.

4. Results and Discussion

4.1. Dataset Details.

The acceleration and braking maneuvers were identified using the designed algorithm and the final dataset consists of 7548 acceleration maneuvers and 6156 braking maneuvers corresponding to 42 drivers. Each maneuver is characterized by ten driving performance features, which includes minimum speed, maximum speed, mean speed, standard deviation of speed, maximum acceleration/deceleration, average and standard deviation of acceleration/deceleration, change in speed during the maneuver, duration of the maneuver, and maximum yaw rate. Since the units of features are not uniform, the data were prepared for analysis by scaling through the Z-score standardization technique.

4.2. Dimensionality Reduction.

The PCA was performed on the scaled data (Z-score standardization) of both acceleration and braking datasets. The correlation circles were developed to understand the correlations between the PCs and the original variables. The PC1 and PC2 loadings against each feature for both datasets are shown in Figures 4(a) and 4(b). The direction of feature lines represents the positive or negative correlations and lengths indicate the strength of the correlation. In other words, the features closer to the circumference of the circle are more important to interpret the respective components [38].

For both acceleration and braking datasets, most of the features are cross-loaded on each PC and are showing higher correlations with PCs. To aid the interpretation, the loadings were rotated using Varimax rotation; the respective rotated component (RC) loadings are shown in Table 3. Four components were chosen as the cumulative explained variance is greater than 85% for both the datasets. The loadings with high magnitude are highlighted in bold (see Table 3), as the higher absolute loadings indicate stronger correlations between features and components.

In case of acceleration dataset (Table 3), RC1 is showing significant and negative correlation with all the longitudinal acceleration features. RC2 shows stronger positive correlations with the mean and standard deviation of speed during the maneuver. RC3 is associated with change in speed and duration of the maneuver and RC4 is highly correlated with the speed features and yaw rate. The similar correlations can be observed for braking dataset, in which RC1 is associated with longitudinal deceleration features and RC2 shows high correlation with speed features and yaw rate. RC3 is showing strong association with change in speed and duration of maneuver and RC4 is significantly correlated with mean and standard deviation of speed in the maneuver.
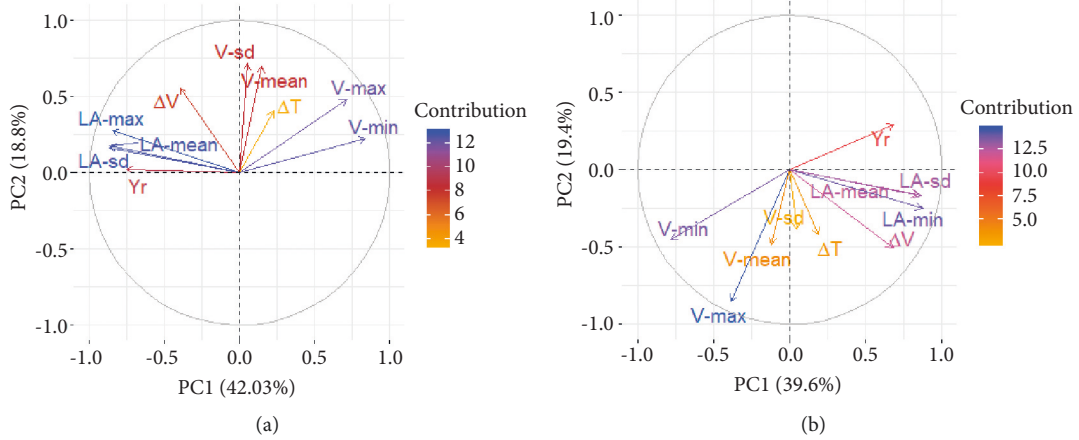
Figure 4: Correlation circles for (a) acceleration maneuvers (b) braking maneuvers.

Table 3: Varimax rotated loadings of acceleration and braking datasets. Note: RC stands for rotated components.

| | RC loadings of acceleration maneuvers | | | | | RC loadings of braking maneuvers | | | |
Features	RC1	RC2	RC3	RC4	Features	RC1	RC2	RC3	RC4
V_{min}	0.321	0.116	0.078	0.908	V_{max}	−0.007	−0.888	−0.352	−0.179
V_{max}	0.171	0.142	−0.290	0.911	V_{min}	−0.320	−0.879	0.198	−0.131
ΔV	−0.383	0.043	−0.849	−0.110	ΔV	−0.495	−0.112	0.815	0.051
ΔT	0.260	−0.037	−0.909	0.106	ΔT	−0.111	0.025	−0.955	0.044
V_{mean}	0.010	0.904	0.015	0.121	V_{mean}	−0.018	−0.184	−0.028	−0.753
V_{sd}	−0.052	0.913	−0.018	0.045	V_{sd}	0.082	0.018	0.032	−0.829
LA_{max}	−0.920	0.008	−0.124	−0.215	LA_{min}	−0.922	−0.164	0.179	0.020
LA_{mean}	−0.869	0.056	0.045	−0.319	LA_{mean}	−0.934	−0.184	−0.002	0.055
LA_{sd}	−0.915	0.001	0.047	−0.261	LA_{sd}	0.957	0.173	−0.011	−0.020
Yr_{max}	−0.386	0.004	−0.214	−0.692	Yr_{max}	0.256	0.764	−0.260	−0.026

The PCA resulted in four components, each indicating different driver behavior aspects. However, clustering each component individually (respective features) would result in four sets of clusters for both acceleration and braking datasets, making the interpretation of results a tedious and non-conclusive task. Hence, the rotated components are further grouped based on the interpretation of correlations between RCs and original features. In case of both the datasets, RC2 and RC4 significantly represent the speed choices and speed variability exhibited during the maneuver, whereas RC1 and RC3 are explaining the acceleration/braking behavior and the respective change in speed over the full duration of the maneuver. It is possible to exhibit the aggressive acceleration/braking behavior at lower speeds, and similarly, the high-speed behavior may not be always associated with aggressive accelerations/braking behavior. In other words, the individuals might perform different frequencies of aggressive low-speed maneuvers, aggressive-high speed maneuvers, nonaggressive low-speed maneuvers, and nonaggressive high-speed maneuvers. Each combination speaks a different aspect of driver such as nonaggressive low-speed indicating base-line behaviors, and aggressive-low speed maneuvers indicating rear-end collision tendencies, whereas the nonaggressive high-speed signifies the driving efficiency, and aggressive-high speed indicates the high-risk/faulty behaviors. Thus, to aid the detailed analysis of longitudinal control, the acceleration/braking behavior, and the speed behavior are analyzed independently for each maneuver.

To this end, the original variables are divided into two groups based on the observed correlations between rotated components and the driving performance features. The first group of variables ((LA_{max}/LA_{min}), LA_{mean}, LA_{sd}, ΔV, ΔT) represents the acceleration/braking behavior, and the second group of variables (V_{max}, V_{min}, V_{mean}, V_{sd}, Yr_{max}) indicates the speed behavior exhibited while accelerating or braking. Further analysis to identify the underlying patterns of acceleration/braking behavior, and speed behavior is explained in the following section.

4.3. Clustering and Driving Style Classification

4.3.1. Level-Wise Clustering. Considering the results of principal component analysis, the k-means clustering analysis was performed at two levels on the acceleration and braking datasets. In the first level of clustering, the driving patterns in terms of the acceleration/braking performance are detected. In the second level of clustering, the speed patterns exhibited during acceleration/braking are identified by clustering the speed variable data. The features used for each level of clustering are given in Table 4.

4.3.2. Level-One Clustering. To characterize the longitudinal control concerning the acceleration and braking behavior, the k-means clustering was performed on level-1 features

TABLE 4: Driving performance features used for level-wise clustering.

Clustering level	Features
Level-1	$(LA_{\max}/LA_{\min}), LA_{\text{mean}}, LA_{\text{sd}}, \Delta V, \Delta T$
Level-2	$V_{\max}, V_{\min}, V_{\text{mean}}, V_{\text{sd}}, Yr_{\max}$

(see Table 4) for acceleration and braking maneuvers. The optimal $k = 3$ was chosen by considering the results of both elbow and silhouette methods. To visualize the compactness and separation of clusters, the individual silhouettes were computed for each cluster and presented in Figure 5. Although few data points under second and third clusters (Figures 5(a) and 5(b)) are showing the negative silhouettes, the average silhouettes of all the clusters of acceleration and braking maneuvers are positive indicating well-separation from neighboring clusters. Hence, the k-means clustering was performed to group the acceleration and braking maneuvers into three patterns each.

The identified clusters are further interpreted using the bi-plots of acceleration and braking clusters against the principal components (see Figure 6). The correlations between the performance features and the principal components are shown in Table 5. PC1 is indicating the nature of acceleration/braking behavior, and PC2 is representing the speed surge/speed reduction and the duration of maneuver. In Figure 6(a), the acceleration cluster-1 is taking lower values in both dimensions, showing the smooth acceleration behavior associated with low-speed surge and longer duration of maneuvers. Acceleration cluster-2 is ranging over low-to-moderate accelerations with high-speed surge and longer maneuver duration. The cluster-3 is dispersed over high acceleration values and low change in speed over smaller durations. Considering the cluster separations associated with feature values, the three acceleration clusters are characterized as smooth, moderate, and aggressive acceleration patterns, respectively. The centers of the clusters for each feature and the respective classification is shown in Table 6. Similarly, the braking clusters are interpreted which are showing smooth, moderate and aggressive patterns of braking behavior (see Figure 6(b)). The characterization of braking clusters and the respective centers are shown in Table 6. The centers of all the features are notably different among the cluster categories, affirming the classification.

4.3.3. Level Two Clustering. The second level of clustering is performed to identify the speed patterns exhibited during acceleration and braking. The level-2 features, as mentioned in Table 4 were used for clustering. The k-means was performed on both the datasets with an optimum $k = 2$, as obtained through elbow and silhouette methods. The cluster separation is shown in Figure 7, indicating well-separated speed clusters with an average silhouette index equal to 0.29. Thus, k-means clustering was performed to group the speed data of acceleration and braking maneuvers into two clusters each.

The correlations between the speed performance features and the principal components are shown in Table 7. For acceleration maneuvers, PC1 is showing strong positive correlation with the minimum and maximum speeds and negative correlation with the yaw rate. As the speeds are higher, the steering action is low and at lower speeds drivers are exhibiting high steering activity. PC2 is indicating significant and positive correlation with the mean and standard deviation of speeds during the maneuver. The PC loadings of braking dataset are showing similar correlations with the respective performance features, with the negative correlation. To summarize PC1 is indicating the speed choices and steering action and PC2 speaks of speed variability during the maneuver. Given the correlations, the results of level-2 clustering are presented as bi-plots against PCs in Figure 8, to aid the cluster interpretation.

The speed variability appears to be nearly equal in both the clusters with cluster-2 taking relatively higher values (see Figure 8). However, the speed choices are significantly higher in cluster-2 and cluster-1 appears to take low values of speed variables and higher value of yaw rate. Given the interpretation, the cluster-1 is characterized to represent the nominal-speed behavior and cluster-2 indicating the high-speed behavior for both acceleration and braking maneuvers. The cluster centers of all the performance features and the cluster categories are presented in Table 8. The speed clusters of acceleration dataset, and braking dataset are representing the similar speed behavior, which indicates the uniformity in speed choices irrespective of the type of maneuver.

The proportion of maneuvers shared by each acceleration/braking pattern and the respective share of nominal and high-speed clusters are presented in Table 9. The major proportion of maneuvers (56.9% accelerations and 61.6% braking maneuvers) are observed to be classified under smooth driving styles, in which 19.8% accelerations and 18.1% braking maneuvers are also associated with nominal speeding patterns, indicating the safe driving behavior. In case of moderate and aggressive driving maneuvers, 18.2% accelerations and 15% braking maneuvers are associated with high-speed patterns indicating unsafe behaviors. Further, the aggressive maneuvers characterized by aggressive acceleration/braking related with high-speed behavior are observed in 6.2% of accelerations and 8.3% of braking maneuvers, reflecting the actual faulty/critical driving behaviors. Although, the proportions of critical maneuvers are relatively lower, the role of such maneuvers is vital in assessing the driving performance of individuals in day-to-day driving. Since, both the speed and acceleration behavior defines the level of driving safety, further, we analyzed the variations in accelerations, braking, and speeding at individual driver level, which shows the in-detail driving nature of individuals.

4.4. Validation of Driving Patterns. Based on the insights from literature, the ground-truth about driving style of a driver is determined by means of traffic accident data, or subjective evaluation data (self-report/expert scoring), or safety-critical event data [11, 15, 20, 22, 31]. In this study, neither the crash-related information nor the self-reported
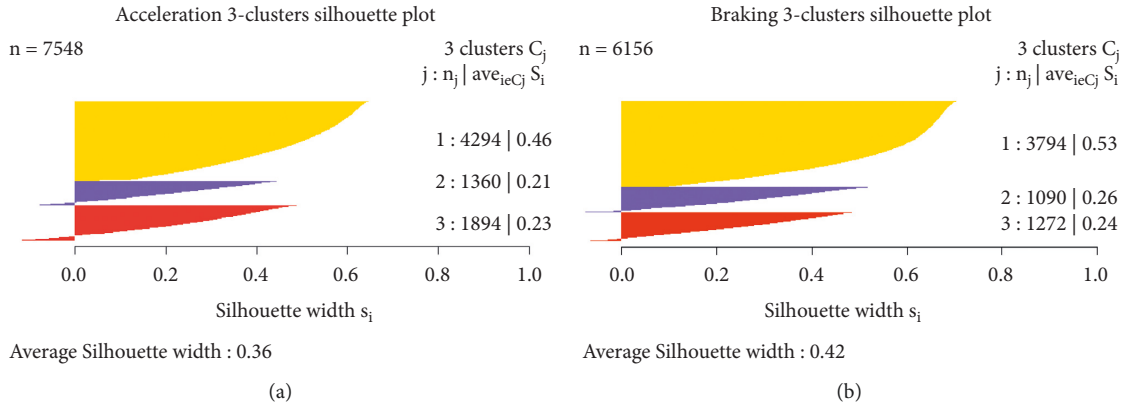
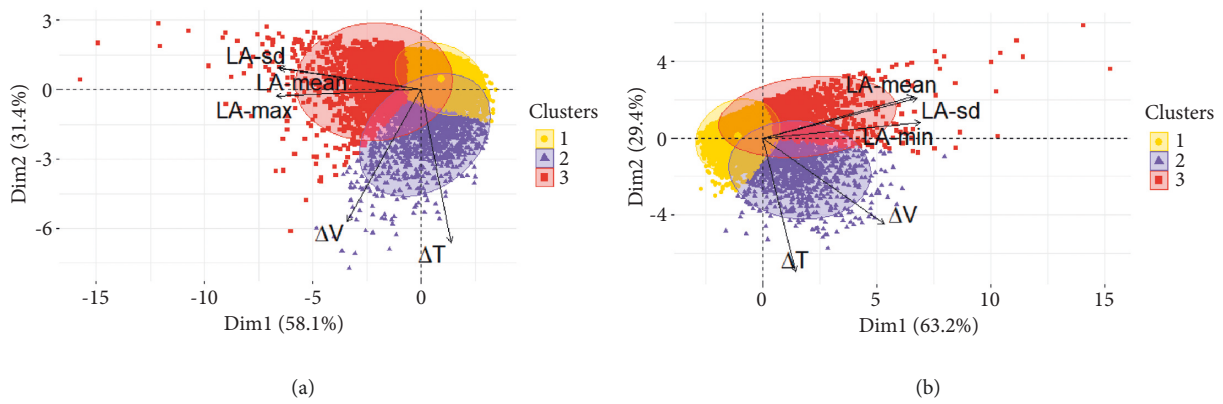FIGURE 5: Silhouettes of (a) acceleration clusters and (b) braking clusters.

FIGURE 6: Bi-plots of (a) acceleration clusters and (b) braking clusters.

TABLE 5: The correlation between the features and components (PC loadings).

Features	Acceleration maneuvers		Braking maneuvers	
	PC1	PC2	PC1	PC2
ΔV	−0.483	**−0.811**	**0.731**	**−0.616**
ΔT	0.195	**−0.937**	0.196	**−0.956**
(LA_{max}/LA_{min})	**−0.948**	−0.039	**0.949**	0.112
LA_{mean}	**−0.924**	0.129	**0.907**	0.295
LA_{sd}	**−0.939**	0.135	**0.931**	0.283

TABLE 6: Level-1 Cluster centers of acceleration and braking maneuvers.

Acceleration clusters	ΔV	ΔT	LA_{max}	LA_{mean}	LA_{sd}	Classification
1	9.00	5.51	0.15	0.06	0.036	Smooth
2	21.08	12.27	0.18	0.07	0.038	Moderate
3	15.65	4.79	0.26	0.11	0.059	Aggressive
Braking clusters	ΔV	ΔT	LA_{min}	LA_{mean}	LA_{sd}	Classification
1	10.35	6.44	0.17	0.06	0.041	Smooth
2	35.35	13.41	0.29	0.10	0.065	Moderate
3	18.17	4.93	0.33	0.13	0.085	Aggressive

driving violation data could be obtained for the participated professional drivers. Hence, the frequency of safety-critical event data is used as a means to describe the driving style of individuals. The safety-critical patterns are obtained through unsupervised techniques and needs to be verified prior proceeding to driving style classification. To this end, the cluster characteristics of critical patterns are compared against the feature thresholds used across studies.
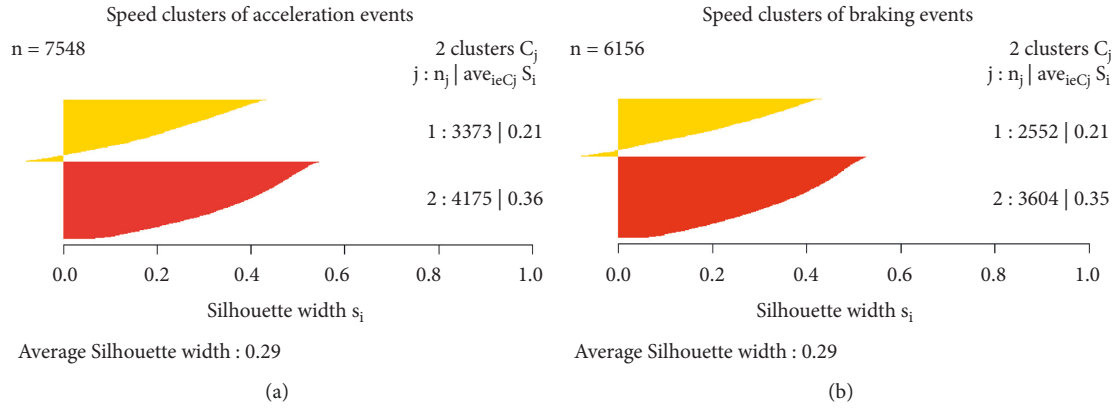
FIGURE 7: Silhouettes of (a) Speed_acceleration clusters (b) Speed_braking clusters.

TABLE 7: The PC loadings of speed variables for acceleration and braking datasets.

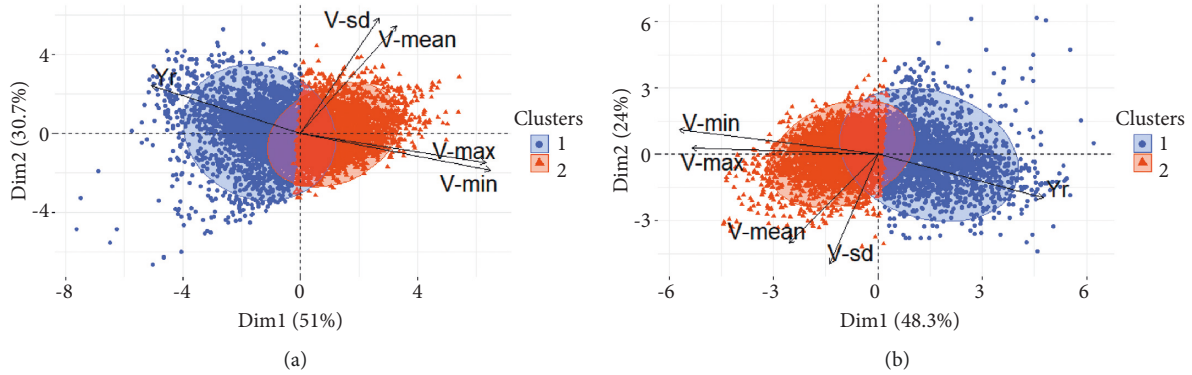| Features	Acceleration maneuvers		Braking maneuvers	
	PC1	PC2	PC1	PC2
V_{min}	0.923	−0.262	−0.923	0.179
V_{max}	0.900	−0.208	−0.866	0.045
V_{mean}	0.466	0.780	−0.414	**−0.648**
V_{sd}	0.382	0.833	−0.224	**−0.803**
Yr_{max}	−0.726	0.348	0.769	−0.316

FIGURE 8: Bi-plots of speed clusters identified in (a) acceleration maneuvers (b) braking maneuvers.

TABLE 8: Speed cluster centers for acceleration and braking maneuvers.

Speed clusters$_{acceleration}$	V_{min}	V_{max}	V_{mean}	V_{sd}	Yr_{max}	Classification
1	22.13	36.05	31.38	15.54	13.22	Nominal
2	49.15	61.13	44.84	16.64	5.45	High speed
Speed clusters$_{braking}$	V_{min}	V_{max}	V_{mean}	V_{sd}	Yr_{max}	Classification
1	17.83	37.76	31.93	16.37	14.53	Nominal
2	48.08	61.97	45.59	17.13	5.29	High speed

TABLE 9: Proportion of maneuvers shared by each pattern.

Acceleration	Smooth	Moderate	Aggressive	Total
Nominal speed	1491 (19.8%)	453 (6%)	1429 (18.9%)	3373 (44.7%)
High speed	2803 (37.1%)	907 (12%)	465 (6.2%)	4175 (55.3%)
Total	4294 (56.9%)	1360 (18%)	1894 (25.1%)	7548 (100%)
Braking	Smooth	Moderate	Aggressive	Total
Nominal speed	1116 (18.1%)	677 (11%)	759 (12.4%)	2555 (41.5%)
High speed	2678 (43.5%)	413 (6.7%)	513 (8.3%)	3601 (58.5%)
Total	3794 (61.6%)	1090 (17.7%)	1272 (20.7%)	6156 (0%)

TABLE 10: Cluster characteristics of data-driven critical driving patterns.

Driving pattern	Performance feature	Mean	SD	Upper limit (Mean + SD)	Lower limit (Mean - SD)
Aggressive acceleration	LA^{*}_{max}	0.26	0.06	**0.32**	**0.20**
Aggressive braking	LA_{min}	0.33	0.08	**0.41**	**0.25**
High-speed	V^{**}_{max}	61.97	11.68	**73.66**	**50.29**

Note: *The units of accelerations are in "g" (9.8 m/s^2); ** the units of speed are in "kmph"

The clustering results presented in Table 9 indicates that the aggressive acceleration and aggressive braking maneuvers associated with high-speed behavior are representing the critical/unsafe patterns exhibited by individuals. These patterns are quantitatively represented by means of the cluster centroid plus/minus one standard deviation (SD), and shown in Table 10. The identified aggressive acceleration pattern is in the range of 0.2 to 0.32 g acceleration, and the aggressive braking pattern is varying over the maximum deceleration of 0.25 to 0.41 g, corresponding to the vehicle speed of 50.20 to 73.66 kmph. These aggressive maneuvers are observed to be falling in the range of thresholds used in the previous studies by Johnson and Trivedi [11], Chen et al. [22], Paefgen et al. [26], Fazeen et al. [27], and Bergasa et al. [29]. Thus, the obtained clusters can be effectively used to assess the driving styles of individuals.

4.5. Driver Behavior Heterogeneity. The identified driving patterns are representing the acceleration and braking behavior ranging from smooth to aggressive and the respective speed patterns are identified to be nominal and high-speed behavior. The proportion of maneuvers under each identified pattern for all the 42 drivers are computed and shown in Figure 9. Drivers are not executing driving maneuvers in a single pattern, rather each individual is showing variations in acceleration, braking, and speeding behavior during the study period. In other words, drivers are exhibiting each driving pattern at a certain proportion, which is different from one driver to another driver, and also from one driving regime to another. For instance, the driver D1 is exhibiting aggressive behavior in 16% of the accelerations, whereas 35% of the braking maneuvers are associated with aggressive behavior. Similarly, each individual driver is showing heterogeneity in the patterns of longitudinal control. No driver was found to constantly exhibit safe or aggressive behavior in the longitudinal control during the entire trip. The differences in driving conditions may result in different driving responses by the same driver, which leads to changing driving styles within the driving period. Thus, the individual's driving styles are found to be inconsistent and heterogeneous in both acceleration and braking regimes. The study findings are consistent with the existing literature on individual driving characteristics [19, 22, 23]. In the study of Higgs and Abbas [23], the car drivers have shown different proportions of behavior patterns (ranging from 0 to 30 patterns) in the car-following behavior, that differs among individuals. Chen and Chen [19] and Chen et al. [22] also showed that each driver was exhibiting a set of driving patterns that are different from one driver to another.

As shown in Figures 9(a) and 9(b), the smooth acceleration and braking behavior is the dominating driving pattern in most of the drivers, which indicates the base-line or "normal" driving behavior of each individual. The proportion of moderate and aggressive behaviors are changing from one driver to another, which indicates different levels of aggression among individuals. In addition, the speed behavior in accelerations and braking maneuvers (see Figures 9(c) and 9(d)) are notably different among the drivers.

4.5.1. Driving Performance Score. The level of aggression of each acceleration and braking maneuver depends on the combination of acceleration/braking pattern and the respective speed pattern. The aggressive behavior is represented by the aggressive acceleration/braking maneuvers, which are further associated with high-speed behavior. In order to separate the aggressive behavior of each driver, the number of maneuvers of aggressive acceleration, aggressive braking, and the high-speed patterns are computed per kilometer traveled. The speed clusters of acceleration and braking maneuvers are observed to be representing similar characteristics in terms of the driving performance features (see Table 8). Thus, the speed clusters of acceleration and braking maneuvers are aggregated and the total number of high-speed maneuvers per kilometer traveled during acceleration and braking are summed up to propose a single speed score as shown in Figure 10.

The computed relative driver aggression in acceleration, braking, and speed behavior are shown in Figure 10. The driver D24 has exhibited the highest number of aggressive maneuvers per kilometer traveled, in acceleration and braking behavior and the driver D25 has shown the dominance in high-speed behavior. The level of aggression of individuals was found to be different in acceleration and braking regimes, and changing among drivers. For example, the drivers (D6, D18, D27, D29, D42) are showing higher aggression score (>60) in accelerations, whereas the respective individuals' braking aggression scores are observed to be lower than 40. Similarly, the drivers D1, D11, and D22 are exhibiting higher levels of aggression (>50) in braking maneuvers and moderate-to-lower aggression in acceleration behaviors. The aggression scores are indicative of drivers' propensity to exhibit harsh driving maneuvers and are particularly relevant in driving assistance.

Each individual is showing a unique combination of driver aggression in acceleration, braking, and speed behavior. The variation in the interdriver aggression might be due to the individual driving characteristics or the external influencing factors such as traffic and road geometry. In the
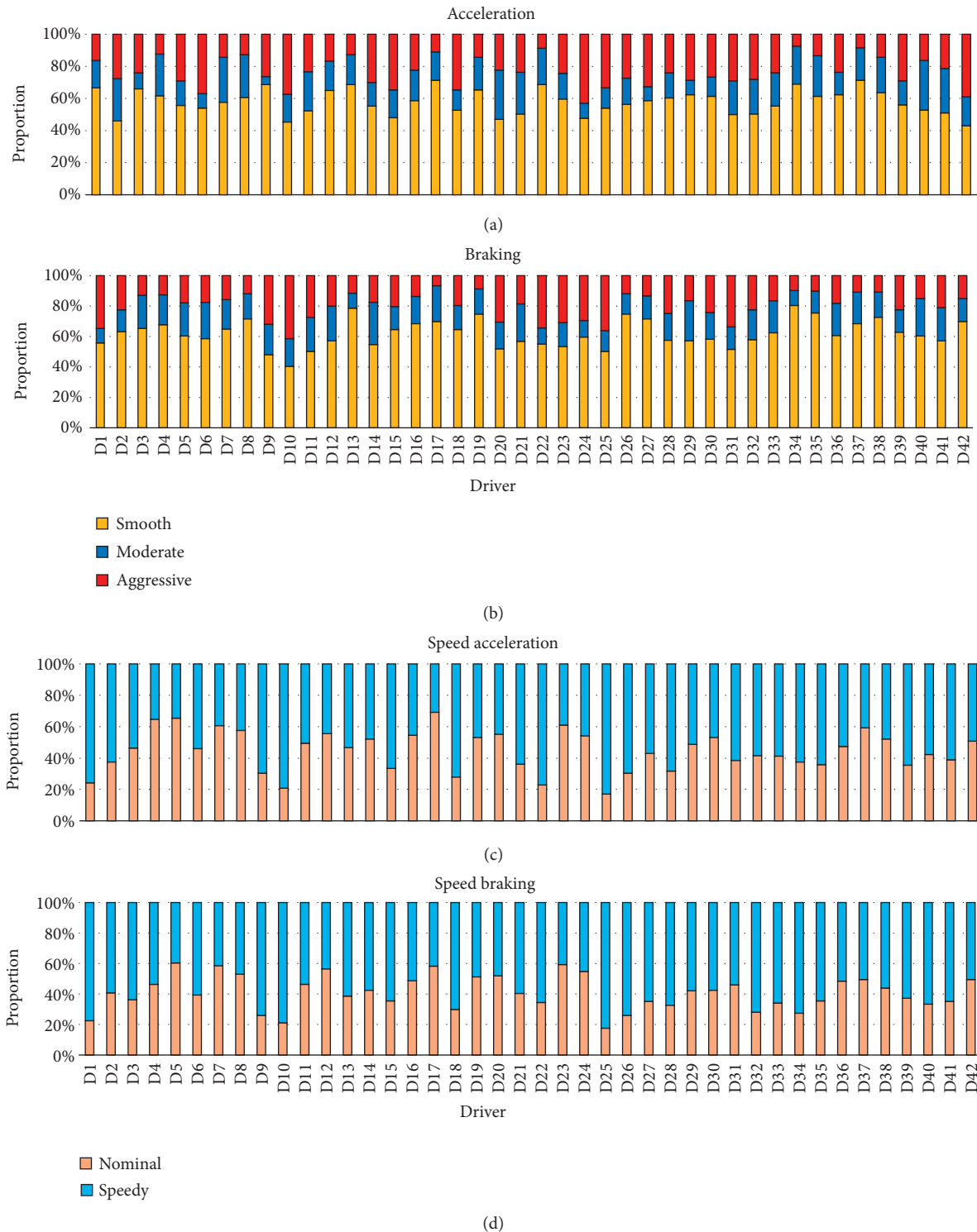
FIGURE 9: The proportion of (a) acceleration patterns; (b) braking patterns; (c) and (d) speed patterns for individuals.

present study, all the drivers made trips on the same road stretch resulting in similar exposure to the geometric elements. Although, the participants drove vehicle during same time of the day, the observed variations may be due to varying traffic conditions. Thus, the identified patterns are considered to be representing more of driver specific variation or the habitual driving styles of each individual corresponding to the exposed traffic.
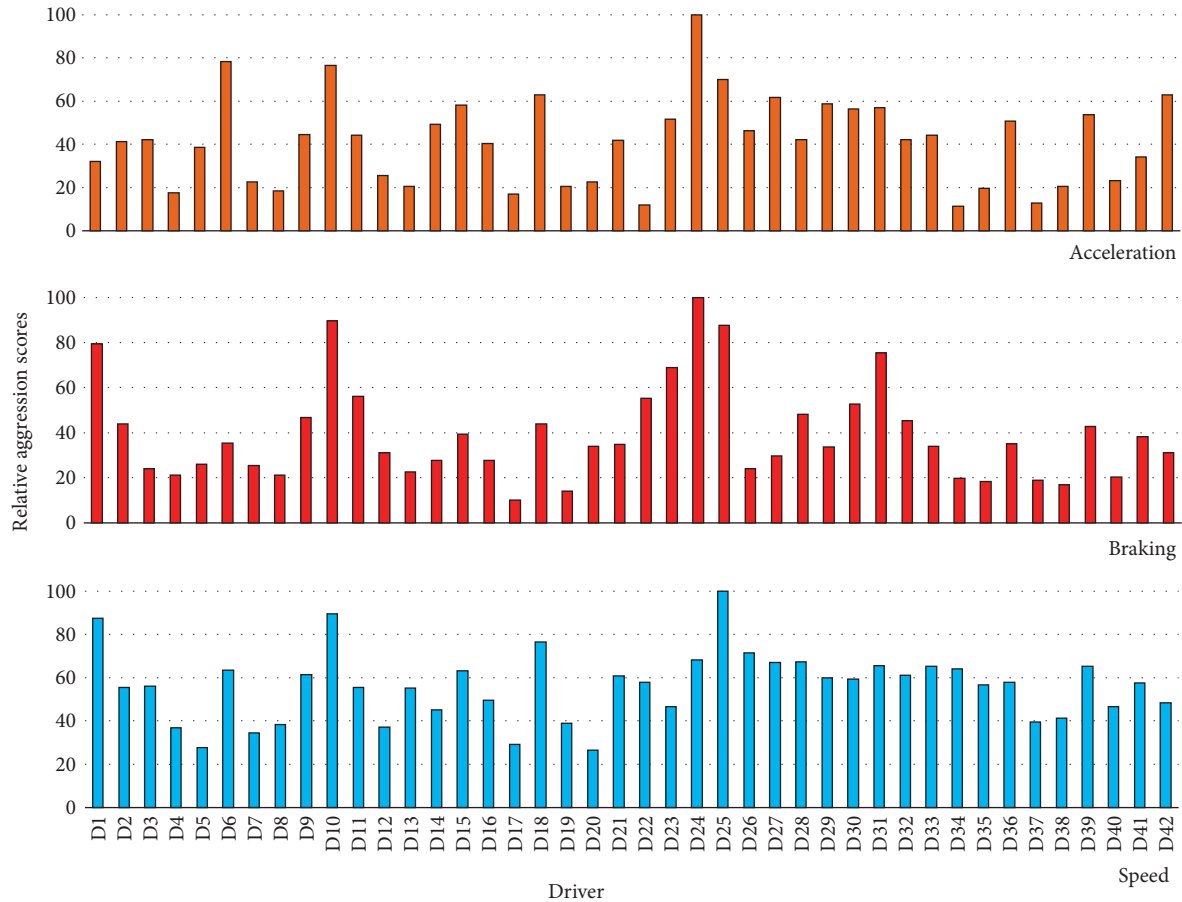
FIGURE 10: The relative driver aggression in acceleration, braking, and speeding.

5. Conclusion

In the context of assisting the driver to optimize the driving behavior, it is essential to understand the individual's instantaneous driving decisions. Most of the studies in this context were undertaken in the developed countries and no studies analyzed the behavioral heterogeneity of Indian drivers. Moreover, majority of the existing studies conceptualized the driving styles using predefined thresholds of abstract driving features, and the drivers were characterized as safe or aggressive. This study takes the advantage of continuous driving profiles collected using high-frequency GPS instrumentation to explore the extent of driving heterogeneity in different driving regimes. The framework is designed to conceptualize the driving styles exhibited in the longitudinal control by extracting the short-term driving decisions and group the similar behaviors by means of unsupervised techniques.

The methodology is implemented on the driving profiles of 42 professional car drivers, captured in the naturalistic driving conditions over a defined study stretch. An algorithm is developed to extract the acceleration and braking maneuvers and the respective driving performance features as representation of driving decisions in longitudinal control. The similar patterns of decisions are identified using k-means clustering technique and interpreted using the

principal component analysis. Total three patterns of acceleration and braking behavior each, characterized by smooth, moderate, and aggressive behaviors, which are associated with nominal and high-speed behavior are identified. The proportion of each pattern observed during the study period of individuals revealed the interesting theories about the extent of driver behavioral variations. The drivers showed varying proportions of patterns in each driving regime, indicating the heterogeneity in driving behaviors. In addition, the driving performance scores were different among individuals in both acceleration and braking maneuvers. No driver was found constantly exhibiting either safe or aggressive driving decisions over the observed driving period.

6. Research Contribution

(a) In the existing literature, most of the studies derived driving styles using predefined thresholds of safety-critical events or the abstract driving performance aggregated over the study period. The thresholds were not consistent across studies and abstract performance did not represent the faulty behaviors in short-term driving decisions. In this context, this study analyzed the driving profile data using unsupervised techniques, without any prior ground

truth. Also, the study proposed a methodology to segment the continuous driving profiles as maneuvers of instantaneous driving decisions in the longitudinal control.

(b) Further, the previous studies assigned a single characterizing driving pattern to an individual. The heterogeneity in driving behaviors pertained to the changes in driver attributes and driving environment was not considered. In the light of these findings, this study explored the individual's behavioral variations with-in driving period. Rather assigning a single representative classification, the actual at-risk behaviors exhibited by each driver are identified. Moreover, the unsupervised approach presented in this study would theoretically account for the unobserved influencing factors effecting the heterogeneity in individual's driving patterns.

The findings of the study emphasize the need for continuous monitoring of driver behavior for driver assistance and personalized feedback provision. Thus, the study methodology and study findings are useful for safety risk profiling of drivers, and also risk scoring of roadway segments for hot-spot analysis.

7. Study Limitations and Future Scope

This study accounts few limitations which provides scope for future research.

Experimental design: As the study is short-term instrumented vehicle study, the driving data is collected using instrumented vehicle over a short-period of time. The presence of instrumentation might influence the actual driver behavior even though the drivers were informed that the collected data would be utilized for only research purposes and not for any legal enforcement. The instrumentation used in this study recorded the continuous vehicle kinematic data and the video-data at a frequency of 10 Hz. The advanced sensors such as LiDAR and eye trackers can be used to capture the other prominent driving parameters like headway choices and distractions, that would help to understand the comprehensive nature of various driving patterns.

Influence of amount of the data: Driving data for longer duration and for more number of drivers from different age groups and gender would give detailed insights about the variations in the driving patterns. With the multiple trip data of individuals, the researchers can evaluate the stability in driving behaviors and consistency in different driving patterns.

The methodologies presented in the current study are replicable by future researchers to assess the driving patterns of individuals. However, the study results are subjected to change concerning the local driving habits and the traffic laws corresponding to the study location. The future scope of the study can be extended to understand the driving pattern variations in other driving regimes like car-following or lane-changing scenarios for different vehicle types. The future research efforts can also be focused on certain critical driving patterns to better understand the influencing factors behind such typical behaviors.

Data Availability

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] World Health Organization, *Global Status Report on Road Safety 2018*, Summary (No. WHO/NMH/NVI/18.20), 2018.

[2] J. R. Treat, N. S. Tumbas, S. T. McDonald, D. Shinar, and R. D. Hume, "Tri-level study of the causes of traffic accidents," *Executive Summary*, DOTHS034353579TAC (5) Final Rpt), University of Indiana, Bloomington, IN, USA, 1979.

[3] J. N. Ivan, C. Wang, and N. R. Bernardo, "Explaining two-lane highway crash rates using land use and hourly exposure," *Accident Analysis & Prevention*, vol. 32, no. 6, pp. 787–795, 2000.

[4] J. Weng and Q. Meng, "Effects of environment, vehicle and driver characteristics on risky driving behavior at work zones," *Safety Science*, vol. 50, no. 4, pp. 1034–1042, 2012.

[5] J. Lee, M. Abdel-Aty, and K. Choi, "Analysis of residence characteristics of at-fault drivers in traffic crashes," *Safety Science*, vol. 68, pp. 6–13, 2014.

[6] H. M. Hassan, M. Shawky, M. Kishta, A. M. Garib, and H. A. Al-Harthei, "Investigation of drivers' behavior towards speeds using crash data and self-reported questionnaire," *Accident Analysis & Prevention*, vol. 98, pp. 348–358, 2017.

[7] Basic Road Statistics of India, "Ministry of road transport and highways, transport research wing," *'Road Accidents in India'*, 2019.

[8] F. Sagberg, G. F. Selpi Bianchi Piccinini, G. F. Bianchi Piccinini, and J. Engström, "A review of research on driving styles and road safety," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 57, no. 7, pp. 1248–1275, 2015.

[9] Y. L. Murphey, R. Milton, and L. Kiliaris, "Driver's style classification using jerk analysis," in *Proceedings of the 2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems*, pp. 23–28, IEEE, Nashville, TN, USA, March 2009.

[10] Z. Constantinescu, C. Marinoiu, and M. Vladoiu, "Driving style analysis using data mining techniques," *International Journal of Computers, Communications & Control*, vol. 5, no. 5, p. 654, 2010.

[11] D. A. Johnson and M. M. Trivedi, "Driving style recognition using a smartphone as a sensor platform," in *Proceedings of the 2011 14th International IEEE Conference on Intelligent*

Transportation Systems (ITSC), pp. 1609–1615, IEEE, Washington, DC, USA, October 2011.

[12] A. Aljaafreh, N. Alshabatat, and M. S. N. Al-Din, "Driving style recognition using fuzzy logic," in *Proceedings of the 2012 IEEE International Conference on Vehicular Electronics and Safety (ICVES 2012)*, pp. 460–463, IEEE, Istanbul, Turkey, July 2012.

[13] M. Van Ly, S. Martin, and M. M. Trivedi, "Driver classification and driving style recognition using inertial sensors," in *Proceedings of the 2013 IEEE Intelligent Vehicles Symposium*, no. IV, pp. 1040–1045, IEEE, Gold Coast, QLD, Australia, June 2013.

[14] S. W. Chen, C. Y. Fang, C. T. Tien, and B. Y. Wu, "Dangerous Driving Event Prediction on Roadways," *Stud. Syst. Sci*, vol. 2, 2014.

[15] J. H. Hong, B. Margines, and A. K. Dey, "A smartphone-based sensing platform to model aggressive driving behaviors," in *Proceedings of the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 4047–4056, ACM, April 2014.

[16] R. Vaiana, T. Iuele, V. Astarita et al., "Driving behavior and traffic safety: an acceleration-based safety evaluation procedure for smartphones," *Modern Applied Science*, vol. 8, no. 1, p. 88, 2014.

[17] L. Eboli, G. Mazzulla, and G. Pungillo, "Combining speed and acceleration to define car users' safe or unsafe driving behaviour," *Transportation Research Part C: Emerging Technologies*, vol. 68, pp. 113–125, 2016.

[18] F. Feng, S. Bao, J. R. Sayer, C. Flannagan, M. Manser, and R. Wunderlich, "Can vehicle longitudinal jerk be used to identify aggressive drivers? An examination using naturalistic driving data," *Accident Analysis & Prevention*, vol. 104, pp. 125–136, 2017.

[19] K.-T. Chen and H.-Y. W. Chen, "Driving style clustering using naturalistic driving data," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 6, pp. 176–188, 2019.

[20] E. G. Mantouka, E. N. Barmpounakis, and E. I. Vlahogianni, "Identifying driving safety profiles from smartphone data using unsupervised learning," *Safety Science*, vol. 119, pp. 84–90, 2019.

[21] S.-W. Chen, C.-Y. Fang, and C.-T. Tien, "Driving behaviour modelling system based on graph construction," *Transportation Research Part C: Emerging Technologies*, vol. 26, pp. 314–330, 2013.

[22] C. Chen, X. Zhao, Y. Zhang, J. Rong, and X. Liu, "A graphical modeling method for individual driving behavior and its application in driving safety analysis using GPS data," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 63, pp. 118–134, 2019.

[23] B. Higgs and M. Abbas, "Segmentation and clustering of car-following behavior: recognition of driving patterns," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 1, pp. 81–90, 2015.

[24] Y. Feng, P. Iravani, and C. Brace, "A fuzzy logic-based approach for humanized driver modelling," *Journal of Advanced Transportation*, vol. 2021, pp. 1–13, 2021.

[25] D. Chen, Z. Chen, Y. Zhang, X. Qu, M. Zhang, and C. Wu, "Driving style recognition under connected circumstance using a supervised hierarchical bayesian model," *Journal of Advanced Transportation*, vol. 2021, pp. 1–12, 2021.

[26] J. Paefgen, F. Kehr, Y. Zhai, and F. Michahelles, "Driving behavior analysis with smartphones: insights from a controlled field study," in *Proceedings of the Proceedings of the*

11th International Conference on mobile and Ubiquitous Multimedia, pp. 1–8, December 2012.

[27] M. Fazeen, B. Gozick, R. Dantu, M. Bhukhiya, and M. C. González, "Safe driving using mobile phones," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 3, pp. 1462–1468, 2012.

[28] O. Bagdadi, "Assessing safety critical braking events in naturalistic driving studies," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 16, pp. 117–126, 2013.

[29] L. M. Bergasa, D. Almería, J. Almazán, J. J. Yebes, and R. Arroyo, "Drivesafe: an app for alerting inattentive drivers and scoring driving behaviors," in *Proceedings of the 2014 IEEE Intelligent Vehicles Symposium Proceedings*, pp. 240–245, IEEE, Dearborn, MI, USA, June 2014.

[30] R. Kalsoom and Z. Halim, "Clustering the driving features based on data streams," in *Proceedings of the INMIC*, pp. 89–94, IEEE, Lahore, Pakistan, December 2013.

[31] G. Li, S. E. Li, B. Cheng, and P. Green, "Estimation of driving style in naturalistic highway traffic using maneuver transition probabilities," *Transportation Research Part C: Emerging Technologies*, vol. 74, pp. 113–125, 2017.

[32] G. Li, Y. Chen, D. Cao, X. Qu, B. Cheng, and K. Li, "Extraction of descriptive driving patterns from driving data using unsupervised algorithms," *Mechanical Systems and Signal Processing*, vol. 156, Article ID 107589, 2021.

[33] T. Toledo, H. N. Koutsopoulos, and M. Ben-Akiva, "Estimation of an integrated driving behavior model," *Transportation Research Part C: Emerging Technologies*, vol. 17, no. 4, pp. 365–380, 2009.

[34] A. S. Zeeman and M. J. Booysen, "Combining speed and acceleration to detect reckless driving in the informal public transport industry," in *Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pp. 756–761, IEEE, The Hague, Netherlands, October 2013.

[35] C. Wang, R. Fu, J. S. Peng, and J. Mao, "Driving style classification method for lane change warning system," *Jiaotong Yunshu Xitong Gongcheng Yu Xinxi*, vol. 14, no. 3, pp. 187–193+, 2014.

[36] Y. Feng, S. Pickering, E. Chappell, P. Iravani, and C. Brace, "Driving style analysis by classifying real-world data with support vector clustering," in *Proceedings of the 2018 3rd IEEE International Conference on Intelligent Transportation Engineering (ICITE)*, pp. 264–268, IEEE, Singapore, September 2018.

[37] G. Qi, J. Wu, Y. Zhou et al., "Recognizing driving styles based on topic models," *Transportation Research Part D: Transport and Environment*, vol. 66, pp. 13–22, 2019.

[38] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: Computational Statistics*, vol. 2, no. 4, pp. 433–459, 2010.