IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 10, 2018, accepted December 15, 2018, date of publication December 27, 2018, date of current version January 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2889751

A Comparative Study of Aggressive Driving Behavior Recognition Algorithms Based on Vehicle Motion Data

YONGFENG MA, ZIYU ZHANG, SHUYAN CHEN, YANAN YU, AND KUN TANG

[1]Jiangsu Key Laboratory of Urban ITS, Southeast University, Nanjing 210096, China
[2]Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University, Nanjing 210096, China
[3]School of Transportation, Southeast University, Nanjing 210096, China

Corresponding author: Kun Tang (kuntang_seu@seu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51478110, Grant 61374195, and Grant 51208100, in part by the Fundamental Research Funds for the Central Universities and the Jiangsu Province University Graduate Student Research and Innovation Program under Grant KYLX_0180, and in part by the Postgraduate Research and Practice Innovation Program of Jiangsu Province under Grant SJCX17_0038.

ABSTRACT Aggressive driving, amongst inappropriate driving behaviors, is largely responsible for leading to traffic accidents, which threatens both the safety and property of human beings. With the objective to reduce traffic accidents and improve road safety, effective and reliable aggressive driving recognition methods, which enables the development of driving behavior analysis and early warning systems, are urgently needed. Most recently, the research focus of aggressive recognition has shifted to the use of vehicle motion data, which has emerged as a new tool for traffic phenomenon explanation. As aggressive driving corresponds to sudden variations in data, they can be recognized based on the recorded vehicle motion data. In this paper, several kinds of anomaly recognition algorithms are studied and compared, using the motion data collected by the accelerometer and gyroscope of a smartphone mounted on the vehicle. Gaussian mixture model (GMM), partial least squares regression (PLSR), wavelet transformation, and support vector regression (SVR) are considered as the representative algorithms of statistical regression, time series analysis, and machine learning, respectively. These algorithms are evaluated by the three widely used validation metrics, including F_1-score, precision, and recall. The empirical results show that GMM, PLSR, and SVR are promising methods for aggressive driving recognition. GMM and SVR outperform PLSR when only single-source dataset is used. The improvement of F_1-score is almost 0.1. PLSR performs the best when multi-source datasets are used, and the F_1-score is 0.77. GMM and SVR are more robust to hyperparameter. In addition, incorporating multi-source datasets helps improve the accuracy of aggressive driving behavior recognition.

INDEX TERMS Aggressive driving recognition, Gaussian mixture model, partial least squares regression, wavelet transformation, support vector regression, vehicle motion data.

I. INTRODUCTION

Deaths caused by traffic accidents have been kept at a high level. According to the global status report on road safety 2015 [1] released by the World Health Organization, road traffic injuries claims more than 1.2 million lives each year and have a huge impact on health. Part of the reason for traffic injuries is due to the improper driving behavior of drivers. The aggressive driving behavior is an important factor that easily leads to dangerous driving. The media and the public have also paid enough attention to this dangerous behavior. It is worth noting that aggressive driving generally does not cause serious bodily harm to people directly, thus it is not obvious enough to arouse people's vigilance. However, it does not mean that aggressive driving behavior has no impact on traffic safety. On the contrary, the aggressive driving behavior is a hidden but important danger in modern road traffic safety.

At present, the research on aggressive driving behavior mainly focuses on two aspects. One is to conduct a questionnaire-based survey on drivers. The other is a small sample of real cars or simulator experiments. The main problem with the first method is that the questionnaire usually reflects the subjective views of the driver rather than the actual performance of the driver on the road. The second method is generally to control the driving environment manually to urge the driver to make more aggressive driving behavior. In contrast, the second analysis method can avoid the deviation caused by the drivers' own attitude. But there is an inevitable problem in either of the above methods at present, that is the definition and determination method of aggressive driving behavior is not clear until now.

Many vague definitions of aggressive driving behavior still exist. Among them, the definition of aggressive driving by NHTSA is 'a driving method that endangers or tends to endanger personal and property safety' [2]. And Tasca [3] propose a definition that is 'A driving behavior is aggressive if it is deliberate, likely to increase the risk of collision and is motivated by impatience, annoyance, hostility and/or an attempt to save time'. In particular, violent exchanges arising from traffic disputes where the intent is to harm another road user (i.e. "road rage.") is considered as a criminal act that should be dealt with by law instead of traffic regulations. In general, most people agree that the following behaviors are aggressive driving behaviors:1) Over-speed; 2) Frequent or sudden lane changes;3) Abnormal acceleration and deceleration;4) Inappropriate overtaking behavior, etc. [3]. In addition, some studies have recognized that the use of appropriate feedback can change the drivers' driving style [4], thus reducing the aggressive driving behavior. Based on the above point of view, to develop an aggressive driving behavior driving feedback system, the first problem is how to identify the aggressive driving state.

In the way of situational experiments, many studies have used simulators to simulate the behavior of drivers in aggressive driving to study the effectiveness of these methods [5]–[7]. Comparatively fewer studies using real vehicle data for analysis. Based on the data collected, many judgment methods have been put forward. However, most of these methods are aimed at continuous aggressive driving behavior, while there are few researches to compare these methods.

In addition, some researches [8], [9] based on the questionnaire draw lessons from the field of psychology. These studies generally focus on finding the influencing factors of aggressive driving behavior (e.g. driver age, driving years), rather than on identifying the moment when aggressive driving occurs. These results usually depend on the quality of the questionnaire. At the same time, some subjective deviation of drivers may affect the results of the analysis. In addition, the questionnaire-based method is mainly used for qualitative analysis of aggressive driving behavior, which has no advantage in quantitative analysis. Due to the complexity of the causes of aggressive driving behavior, the behavior evaluation method based on the subjective feelings of drivers is difficult to provide theoretical support for further research without systematic research and verification.

Besides the questionnaire-based method and the driving simulator method, the related studies can also be divided into two categories. One kind is the method of establishing a simplified physical road and vehicle model. Eboli *et al.* [10] constructs an aggressive driving behavior discrimination model which uses the speed and acceleration features of the vehicle. The model was verified by a set of real vehicle experiment data. Cerni and Bassani [11] considered the road design hypothesis that "drivers will follow the road alignment with trajectories centered in the lane". The possible driving errors or unexpected or undesired behaviors will be recognized by comparing the actual 'operating trajectory' with the 'designed trajectory'.

The other type is the data-driven methods. Lee and Jang [12] designed a driving behavior recognition framework for large-scale datasets. The driving event recognition was divided into three steps: abrupt detection, automatic encoding, and two-level clustering. The cluster with high aggressive driving potential is extracted from the vehicle sensor data of taxis. Feng *et al.* [13] utilized vehicle longitudinal jerk (change rate of acceleration) to identify aggressive drivers. The frequency of abnormal gas or brake pedal events are used to analyze aggressive driving behavior. Then the high correlation between these two groups of variables are obtained. González *et al.* [14] uses real vehicle data to verify that the GMM model can be used in the recognition of aggressive driving behavior. They built a classifier capable of detecting aggressive behavior from the driving signal. This classifier achieves a success rate up to 92%. Benavent *et al*. [15] established a logistic model of traffic accidents related to aggressive driving with road characteristics (road type, number of lanes, type of median, etc.). Eftekhari and Ghatee [16] built a system based on smartphone IMU (Inertial measurement unit). According to the results of the self-reported questionnaire of the Driver Anger Scale (DAS), the driver's behavior is classified into three categories: safety, semi-aggression and aggression. The mixture of Discrete Wavelet Transformation and Adaptive Neural Fuzzy Inference System is used to identify the driving behavior. The proposed system recognizes the whole driving behavior pattern with 92% accuracy without evaluating the maneuver one by one. At the same time, if there is no longitudinal acceleration data, the driver's behavior will not be recognized successfully, while the result will not be disturbed when the gyroscope is not available. These papers are characterized by the use of algorithms in different fields and have achieved some results on aggressive driving behavior. These papers tend to use accuracy as a measure and consider both normal and abnormal behavior. However, for the recognition of aggressive driving behavior, we actually pay more attention to abnormal behavior than normal behavior. At this time, the result of accuracy, which includes both normal behavior and abnormal behavior, may not be what we really want.

Aiming at the identification of driving behaviors, we can divide the driving behavior into two types. One is continuous behavior, such as over speeding. The observed indicators remained stable for a period. It must be determined by a certain threshold whether the state is abnormal or not. The other type is abrupt behavior, such as sudden lane change, abnormal acceleration, and deceleration, etc. The abrupt aggressive driving behavior means that the occurrence of this state will bring about a sudden change in the vehicle's motion state.

From the view of data, these behaviors are more likely to be recognized. Since the characteristics of the second type are more obvious, the results obtained through algorithm analysis are more reliable.

Considering that many previous studies on aggressive driving behavior do not focus on abnormal behavior during driving, but on the whole driving process, this paper attempts to pay more attention to unusual behaviors. Meanwhile, many classification algorithms are widely used in the field of abrupt aggressive driving behavior recognition, while other types of algorithms are less studied. This study will focus on the regression and clustering algorithms which have the potential to identify abnormal driving behavior. Therefore, starting from the abrupt aggressive driving behavior, the objective of this study is to 1) extracting abrupt aggressive driving behaviors from the whole driving behavior; 2) compare some anomaly detection algorithms in abrupt aggressive driving behaviors recognition; 3) find out the relationship between abrupt aggressive driving behavior and vehicle motion data. Specifically, the results of this study will be helpful to the development of a driving state recognition system, as well as to the further study of the mechanism, forecast, and early warning of aggressive driving behavior. The rest of the study is structured as follows. In the second section, a conceptual framework has been proposed. Also, some details of the methods used in this paper was given. The third part introduces the data sources and experimental results. Afterward, the performance of each algorithm is compared. Finally, the fourth part summarizes the main findings of this study and discusses the future work.

II. METHODOLOGY

A. CONCEPTUAL FRAMEWORK

For the purpose of selecting the best aggressive driving behavior recognition method from several representative algorithms and evaluate the prediction accuracy of several algorithms, we divide the work into four main parts.

1. Pre-processing the data. The data that located at the front and the back of the whole driving process will be eliminated.

2. Modeling the data by using several representative algorithms: GMM, PLSR, Wavelet, SVR. These algorithms include representative statistical regression, time series analysis, and machine learning algorithms. In addition, the residuals of the predicted values in a normal driving state should be mostly at a lower level, while only a small number of residual values deviate from normal values. Then the deviation of the predicted residual is calculated to determine whether the data is in an abnormal state. In that part, the Sum Squares Residual (SSR) was used as the metric, which is shown in Equation (1).

$$SSR = (y - \hat{y})^2 \tag{1}$$

where y is the real value, $\hat{y}$ is the predicted value.

3. When the modeling is completed, F_1-score, the harmonic average of precision and recall, followed by Equation (2), are used as criteria. F_1-score is specific in that it pays more attention to positive samples, while negative samples correctly classified will not affect it, which means F_1-score can be well applied to anomaly recognition, since abnormal samples are always far less than normal samples.

$$F_1 = \frac{2PR}{P + R} \tag{2}$$

where P is Precision, the fraction of relevant instances among the retrieved instances. R is Recall, the fraction of relevant instances that have been retrieved over the total amount of relevant instances.

4. The parameters of the algorithm will be adjusted until finding the optimal result. The conceptual framework is shown in FIGURE 1.

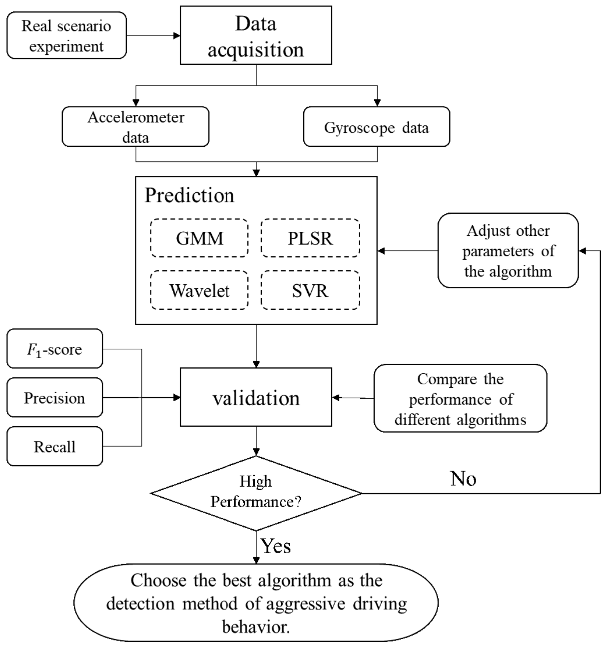


FIGURE 1. Conceptual framework.

B. ALGORITHMS INVOLVED IN COMPARISON

The algorithms selected in this paper should have at least one of the following characteristics: 1) the algorithm should be a regression or at least be able to calculate the predicted residual in order to facilitate the comparison of algorithms using a uniform standard, 2) the algorithm should have potential in driving behavior recognition or anomaly detection, and 3) the algorithm should be a representative algorithm in the corresponding field. Finally, four algorithms are applied to this comparative study.

1) GAUSSIAN MIXED MODEL

González *et al.* [14] use GMM model fitting data from a vehicle. They have proved that aggressive driving behavior can be detected by monitoring external driving signals, such as lateral and longitudinal acceleration and speed. According to their conclusion, this model has applicability to specific driving signals (speed, longitudinal and lateral acceleration), driver, and road type. The GMM model is defined as follows.

With the random variable X, the mixed Gaussian model can be expressed by Equation (3).

$$p(x) = \sum_{k=1}^{K} \pi_k N(x|\mu_k, \Sigma_k) \tag{3}$$

where $N(x|\mu_k, \Sigma_k)$ was called the k-th component in the hybrid model. The number K is the number of normal distributions used to fit the data.$\pi_k$is mixture coefficient and satisfies the condition shown in Equation (4).

$$\begin{cases} \sum_{k=1}^{K} \pi_k = 1 \\ 0 \le \pi_k \le 1 \end{cases} \tag{4}$$

where π_k corresponding to the weight of each component.

The process of using GMM to recognize abnormal condition is somewhat different from other algorithms. The most important thing is that GMM, as an unsupervised learning method, does not need to input dependent variables and independent variables at the same time. Therefore, the whole process of the algorithm is as follows.

1. Set the range of number components K
2. Use the number k in the range as the parameters of the GMM algorithm to fit the model
3. For the fitted data, we regard the distance from the current point to the classification center as the fitting residual and mark the data whose fitting residual exceeds a certain threshold as abnormal.
4. Calculate F_1-score, find the dividing threshold that maximizes the F_1-score value.

2) PARTIAL LEAST SQUARES REGRESSION

PLSR is a statistical regression method proposed by Wold *et al.* [17]. PLSR looks for a linear regression model by projecting prediction variables and observation variables into a new space. Similar to the PCR (Principal Component Regression), PLSR uses m principal components as new variables and carries out least square regression on this basis. Compared with PCR, PLSR essentially reduces the deviation at the expense of increasing a part of the variance. PLSR has already been used in anomaly detection. Wang *et al.* [18] applied PLSR to automatic traffic event detection algorithm to analyze and identify traffic events based on abrupt changes in traffic parameters. The same idea has also been applied to driver drowsiness recognition based on eye data [19]–[21]. Therefore, PLSR, as a statistical regression and a kind of anomaly detection algorithm, is one of the algorithms that we take part in comparison. The process of PLSR algorithm is shown in Table 1.

PLSR has two important characteristics:

1. PLSR is a multiple dependent features to multiple independent features regression modeling method. It can solve many problems that could not be solved by ordinary multiple regression in the past.

TABLE 1. PLSR Algorithm for aggressive driving behavior recognition.

PLSR Algorithm for aggressive driving behavior recognition
Input: Vehicle motion data matrix R with a timestamp; List of aggressive driving behaviors occur time T_o; The length of the aggressive driving behavior time window t.
Output: PLSR model $F_M(x)$; Residual error between the original value and the predicted value E_{residual}.
Step 1: Define l = the length of R, then define $X=$ the front $l-t$ term of R, $Y=$ the rear $l-t$ term of R.
Step 2: Standardize the input sample matrix X and output matrix Y, obtain matrices E_O, F
Step 3: let $n=0$
Step 4: Compute matrix expression $E_n^T F F^T E_n$, then compute the unit characteristic vector corresponding to the maximum characteristic root W_n of $E_n^T F F^T E_n$.
Step 5: let: $\begin{cases} t_n = E_n w_n \\ p_n = \dfrac{(E_n^T t_n)}{\lvert t_n \rvert^2} \\ r_n = \dfrac{(F^T t_n)}{\lvert t_n \rvert^2} \\ E_{(n+1)} = E_n - t_n p_n^T) \end{cases}$
Step 6: If the residual difference E_{n+1} achieves the satisfactory level, go to step 7; Otherwise, let $n=n+1$ and go to step 4.
Step 7: Get the regression equation $F = t_0 r_0^T + t_1 r_1^T + \cdots + t_n r_n^T + F_n$; Predictive equation $F_M(x) = E_O B$, where $B = \sum k_j r_j^T, k_h = [\prod_{j=0}^{h-1} I - w_j [_j^T] w_h$.
Step 8: After obtaining the PLSR model when the fitting is completed, Re-enter X as input, Get predicted value Y_{predict}, Calculate the residuals of Y and Y_{predict}.
Step 9: Calculate F_1-score, find the dividing point that maximizes the F_1-score value.

2. PLSR can realize regression modeling (multiple linear regression), data structure simplification (principal component analysis) and correlation analysis between two groups of variables (canonical correlation analysis) at the same time under one algorithm.

3) WAVELET TRANSFORMATION

For the wavelet transformation algorithm, the most widely used fields are signal processing and image processing. Especially the denoising and compression of signals and images. Since the selection of orthogonal basis in orthogonal wavelets is closer to the actual signal itself than traditional methods, noise can be separated more easily by wavelet transform. In a sense, an abnormal driving behavior is also a kind of noise. Some noise reduction algorithms also have application potential in the field of aggressive driving behavior recognition. Eftekhari and Ghatee [16] built a driving behavior recognize system by using Discrete Wavelet Transformation. The dbN, as one of the most classical wavelet, has been applied in the field of financial time series analysis [22], [23]. In this paper, db4 wavelet is used to decompose the time series signals of vehicles. Judging whether the data is noise or not by calculating the difference between the reconstructed data and the original data. The following is a brief flow of using wavelet algorithm to process driving data.

1. Standardize data.
2. Use db4 wavelet to decompose the time series
3. Set the highest frequency wavelet coefficients of each layer obtained through decomposition to zero. The noise in the data is removed at this step.
4. Reconstruct the denoised signal from the wavelet coefficients by inverse wavelet transform.
5. Calculating the residual between the original waveform and the reconstructed waveform.
6. Find the dividing point that maximizes the F_1-score value according to the residual.

In this way, we have obtained a method that can collect noise from the original data. This method also conforms to the time series characteristics of the driving data. Especially for the driving data of the non-stationary time series, compared with the traditional stationary time series analysis method, wavelet analysis can decompose the time series into different levels according to different scales. Then the difference between the sudden driving abnormality and the normal driving state are able to be found.

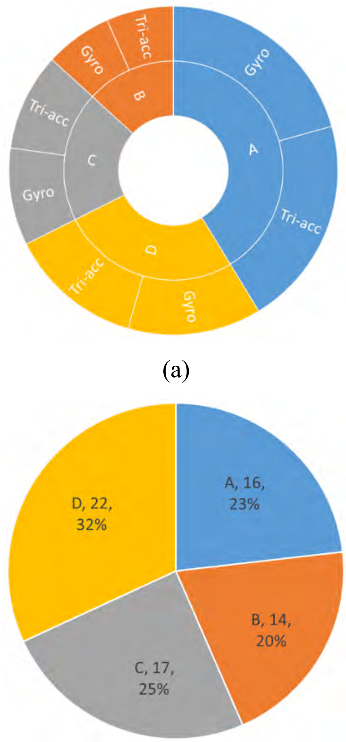
4) SUPPORT VECTOR REGRESSION

Support Vector Machine (SVM) is a common supervised machine learning algorithm. Modern SVM was proposed in the 1990s and is an extremely mature algorithm. In the field of driving behavior, many studies apply SVM to analyze drowsiness state [24]–[26]. In addition, the study of Chen *et al.* [27] shows that SVM can distinguish abnormal driving behaviors from normal driving behaviors. He points out that some driving behaviors have unique acceleration and directional patterns.

SVR algorithm is the application of SVM in regression problems [28]. It is also applied to the study of anomaly recognition and driving behavior. Wei and Liu [29] uses the track data of the vehicle to study the asymmetric driving behavior by using SVR. Liu *et al.* [30] used SVR as a benchmark system compared it with the fuzzy prediction and evaluation system for driving fatigue. Based on the above studies, we apply it to this comparative study.

III. DATA SOURCE

We use the dataset published by Ferreira *et al.* [31]. The dataset was collected by the smartphone IMU. The data from



(a)

(b)

FIGURE 2. Data composition (a) The Content and proportion of recorded data. "Tri-acc" for tri-axial accelerometer, "Gyro" for gyroscope; (b) The number of aggressive driving behaviors in each experiment and their proportion in the total dataset.

accelerometer, linear acceleration, magnetometer and gyroscope were recorded by a smartphone application. The details of the data are as follows:

1. It includes four driving segments. Each segment having an average length of 13 minutes.
2. The participants in the experiment included two drivers with more than 15 years of driving experience executed the driving events.
3. The sampling frequency of accelerometers and gyroscopes is 50Hz.
4. The weather was sunny. The experiment was carried out on a dry asphalt road.
5. The smartphone for collecting data is Motorola XT 1058, which is fixed on a 2011 Honda Civic.

The distribution and structure of the data are shown in FIGURE 2 below. In addition, each driving segment is evaluated manually and marked with aggressive driving events.

Data from accelerometers and gyroscope data are used as common inputs to the algorithm. These two datasets were collected by accelerometers and gyroscopes respectively. The acceleration includes three directions. The features named as xyz are the acceleration of the vehicle in three directions, in m^2/s, while heading, pitch, and roll are the rolling conditions of the vehicle in the corresponding axial direction, in rad/s. These six features can fully describe the motion

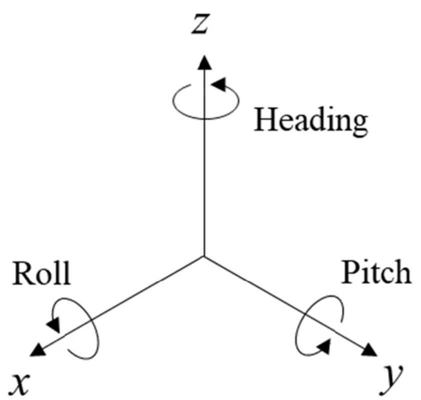


FIGURE 3. The Six features respectively describe the translation and rotation of the vehicle.

state of the vehicle. The relationship of these properties in space is shown in FIGURE 3.

IV. RESULTS OF DATA ANALYSIS

Since the distribution of abnormal and normal driving behavior in the data is extremely uneven and the dataset is with time series characteristics, it should be realized that simple up-sampling or down-sampling cannot be used when processing the data. The additional data filtering will not only lead to the loss of time information, but also does not adapt to the situation that the algorithm needs to be processed in the real-time practical application. In general, it is sufficient to standardize the data. In order to demonstrate the universality of the algorithm, we did not apply different parameters for different experiments, nor did we carry out additional processing on the data. The problem of imbalanced classes should be handled by the algorithm itself. Also, due to the imbalance between positive and negative samples, as common metrics, Precision, Recall, and F_1-score are introduced as evaluation metrics. Under the scenario of aggressive driving behavior recognition, the algorithm should be balance between precision and recall. Each algorithm was applied to accelerometer-only dataset, gyroscope-only dataset, and combined dataset respectively. Because the range of acceleration data and gyroscope data is different, we standardized the two groups of data separately and analyzed with the standardized data directly. Z-score was used as the standardized algorithm. The calculation method is expressed by Equation (5).

$$z = \frac{x - \mu}{\sigma} \tag{5}$$

where x and z are variables before and after standardization, μ is the mean of the sample, σ is the standard deviation of the sample.

Firstly, we used GMM algorithm to recognize the anomaly. GMM algorithm does not need to fit each feature sequentially. Instead, it can consider all features in a unified way. GMM algorithm has only one hyper-parameter, which is the number of categories. Theoretically, the prediction residual should decrease with the increase of this parameter, because the number of classifications will keep increasing until each point becomes the center of the classification. However, we found that when the number of classifications was small (less than 30 categories), the SSR of the results fluctuated within the range of $\pm 5\%$ of its average value. Due to the results are very similar and has little influence on the subsequent process, we selected the points that make the SSR smaller as the number of classifications. After obtaining the classification result, we used the distance between each point and its corresponding classification center as a residual. Then their SSR (Residual Sum of Squares) were calculated. Comparing the results with the actual abnormal events, we found that those locations with large SSR values are often the moments when aggressive driving behaviors occur. Therefore, we could use this indicator to screen the abnormal points in this way.

For PLSR algorithm, it can also consider all features in a unified way because it is the regression of multiple dependent variables to multiple independent variables. The number of components to keep was set as 2. After 500 rounds of iteration, the predicted values of all features through regression were obtained. The difference between the predicted values and the real values was used as the fitting residual. Then we calculated their SSR. It can be found that there are much fewer outliers in the results of PLSR than GMM.

Unlike GMM and PLSR, we applied db4 wavelet decomposition to each feature respectively. Next was setting the coefficient of the highest frequency to 0. Then the entire wave has been reconstructed. SSR was calculated from the reconstructed waveform and the original waveform. The result is very similar to that of PLSR. The points with high SSR are also relatively dense. Compare with the absolute value of residual calculated by GMM, that for Wavelet is smaller. The distinction between normal and abnormal States is not as obvious as PLSR algorithm. Even the data marked as normal also has some shorter peaks.

At the same time, another method is used to fit SVR. The Radial basis function was used as the kernel function. The Penalty Parameter C was 2. All the features of the dataset are taken as samples to fit each feature separately. After obtaining the fitting residuals, we found that SVR has achieved the lowest SSR among the four algorithms. Each segment of data is processed in the same way and the average of each piece of results is taken as the final result. Part of the results of these algorithms are shown in FIGURE 4. The vertical red line shows the moment when the aggressive behavior occurs. From this figure, we can find that the SSRs corresponding to most points are at a lower level, while the position with dense high SSR points are often the position where the aggressive event occurs. Also, we can find that the fitting residuals of most points are near zero, while only a few points are offset. The SSR obtained by SVR fitting has a relatively small number of points and a very concentrated distribution in the timestamp. From this point of view, SVR distinguishes the exception data most obviously.

After obtaining SSR, we still cannot directly judge whether abnormal events occur. Instead, we need to set a threshold as a dividing point. When a sufficient number of SSR points larger

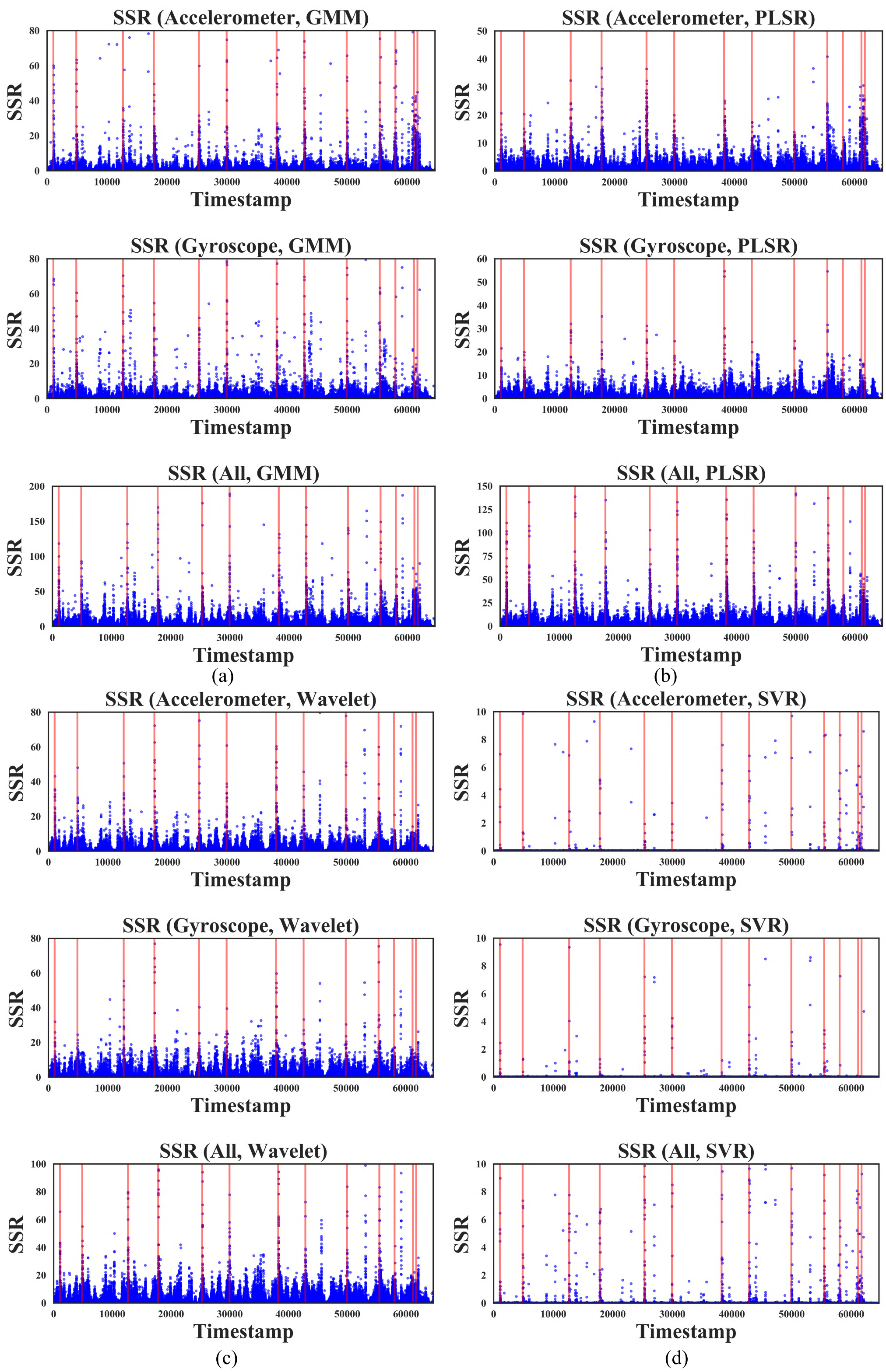


FIGURE 4. The predicted SSR values of different datasets and algorithm combinations. The timestamp is in tenths of a second. Only partial results are included. (a) GMM, (b) PLSR, (c) Wavelet, (d) SVR.

than the threshold occurs at a certain time, we believe that an abnormal event occurred at that time. However, the value of threshold will affect the result of the algorithm directly. So, we decided to try different thresholds and choose the value that maximizes a certain metrics as the best segmentation point. The scores are shown in FIGURE 5.

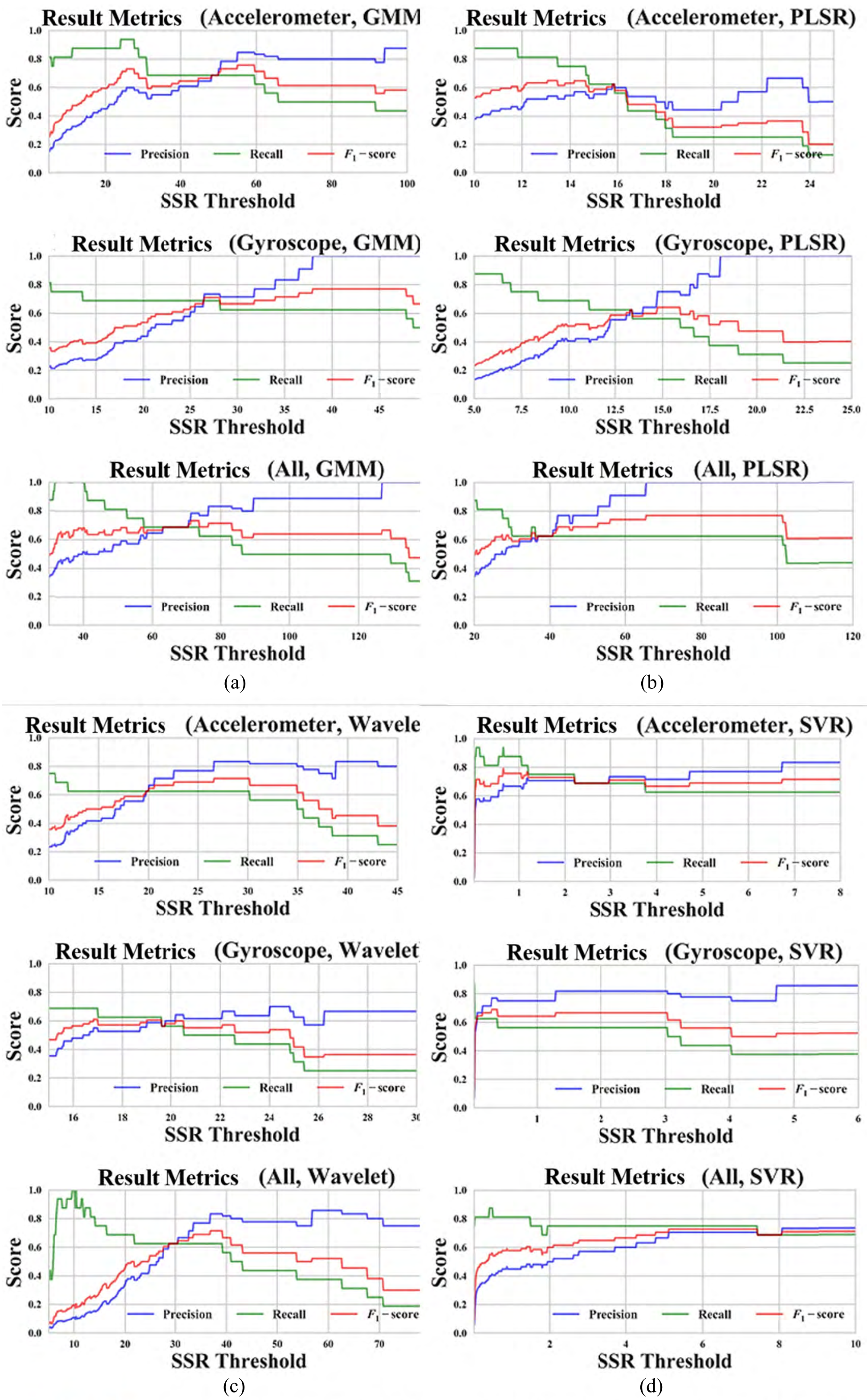


FIGURE 5. The results scores of the algorithms in different datasets. Only partial results are included. (a) GMM, (b) PLSR, (c) Wavelet, (d) SVR.

The horizontal axis in the figure is SSR thresholds, while the vertical axis is the result score corresponding to the thresholds. The figure includes the result curve of precision, recall, F_1-score. It is free to select the appropriate threshold according to their own requirements for precision and recall, or directly select the threshold that makes F_1-score the

TABLE 2. The Average optimal F1-score, Precision, and Recall value for each dataset and algorithm combination.

Combination of dataset and algorithm	F_1-score	Precision	Recall
Accelerometer			
GMM	0.73	0.79	0.69
PLSR	0.65	0.54	0.81
Wavelet	0.71	0.83	0.63
SVR	0.79	0.68	0.94
Gyroscope			
GMM	0.74	0.73	0.76
PLSR	0.64	0.75	0.56
Wavelet	0.61	0.55	0.69
SVR	0.69	0.77	0.63
All			
GMM	0.76	0.85	0.69
PLSR	0.77	1.00	0.63
Wavelet	0.71	0.83	0.63
SVR	0.73	0.71	0.75

highest as the best. In the figure, only those parts with scores exceeding 0.4 or with obvious changes in metrics are plotted.

From that figure, we can find that precision generally increases with the increase of the threshold and recall decreases with the increase of the threshold. F_1-score is positively correlated with the threshold when the threshold is small, while negatively correlated when the threshold is large. F_1-score is a more appropriate choice to evaluate the best results of the algorithm.

Therefore, we can draw a conclusion that the accuracy of the algorithm will obviously change with the change of the threshold value. For the four groups of algorithms, 80% or more of the results recognized by the algorithms are abnormal states at the best time, while the best recall is over 75%. But they cannot achieve the best result at the same time. As can be seen from FIGURE 5, a slight adjustment of threshold value will not affect the results of the three metrics since the curve is stepped, which shows it is robust to use these algorithms to recognize abrupt aggressive driving behavior. When the correct threshold is determined, small changes in the environment will not affect the result.

For comparing the results of the algorithms in these datasets, we computed the highest F_1-score values for each dataset and algorithm combination. The results are shown in Table 2. The table also includes the values of other indicators when F_1-score is optimal. For instance, the first row of data means that GMM algorithm obtains the highest F_1-score of 0.73 in accelerometer data. At that time, 79% of the abnormal events detected are real aggressive driving behaviors, while 69% of the real aggressive driving behaviors are detected.

Based on the F_1-score, GMM achieved the best results on average. However, SVR and PLSR can achieve a better result for different types of data sometimes. For example, the results of SVR are better than other algorithms in acceleration data. Compared with other algorithms, the evaluation metrics of SVR have several different characteristics: 1) The evaluation metrics all increase rapidly from 0 to more than 0.5 with the increase of threshold value in the position where SSR is close to zero, 2) After that, all three metrics remained at a high level, especially the change of recall was relatively small, which means that SVR may be more robust in identifying some specific aggressive driving behaviors and is not easily affected by parameters.

It is noteworthy that both GMM and PLSR perform better in the combined dataset than every single dataset. Especially PLSR does not perform satisfactorily in the acceleration data or gyroscope data but achieve the best results among the four algorithms in the combined dataset. At this time, all the identified abnormalities were real aggressive driving behaviors, while 77% of the aggressive driving behaviors were identified. In contrast, Wavelet performed worst in the other two groups of data except for the acceleration data. The performance of wavelet in the combined dataset is the same as that in the acceleration data, which shows that the change of acceleration masks the rotation of the vehicle. The performance of SVR algorithm in the combined dataset lies between the results of two separate datasets, which may be due to the poor effect of SVR on gyroscope data, leading to the final result being dragged down.

Finally, if considering selecting an appropriate algorithm in the actual scene, GMM and SVR will be the better choice if your dataset is limited because they perform better in insufficient data, which is very useful for practical applications. In addition, the two algorithms are not sensitive to SSR threshold selection, since you cannot get the algorithm's quality before the event occurs, a stable and effective algorithm will be a reliable option at this time. Meanwhile, when you expect the best results and have two datasets at the same time, PLSR is an option worth considering.

V. CONCLUSION

This research uses real vehicle experimental data to extract and identify aggressive driving behavior. According to the vehicle motion data collected by the smartphone sensor, the aggressive driving behavior is analyzed and recognized, thus providing a theoretical basis for effectively and dynamically recognizing abrupt aggressive driving behavior. The result shows 1) All four algorithms can recognize the aggressive driving behavior. According to the selection of SSR threshold, we can choose between high precision and high recall; 2) The results of several common anomaly detection algorithms are similar. The F_1-score of all algorithms can reach more than 0.7 in the combined dataset. Also, the performance of each algorithm to different datasets is different. In contrast, GMM and SVR are superior to the other two algorithms, while PLSR performed the best in the combined dataset; 3) The anomalies recognized from the acceleration data are more obvious than those recognized from the gyroscope data. For GMM and PLSR, the combined dataset is more helpful to the recognition of aggressive driving behavior, while Wavelet and SVR are not sensitive to this; 4) The algorithm is not sensitive to changes in a small range of threshold values. Therefore, when the threshold value is correctly selected, small changes in the environment will not affect the performance of the algorithm.

This article mainly uses data based on mobile phone sensors. In future work, we will use the datasets including dedicated high-precision GPS devices, vehicle distance radar, eye tracker data, and their combination to analyze the characteristics of aggressive driving behaviors and its recognition methods. At the same time, the weight of each feature in the dataset should also be considered. It is necessary to find the weight of features that make the evaluation criteria optimal.

REFERENCES

[1] World Health Organization. (2015). *Global Status Report on Road Safety*, p. 340. [Online]. Available: http://www.who.int

[2] J. H. Hedlund and W. A. Leaf, ''Aountermeasures that work: A highway safety countermeasure guide for state highway safety offices,'' Nat. Highway Traffic Saf. Admin., Washington, DC, USA, Tech. Rep. DOT HS 810 710, Jan. 2007.

[3] L. Tasca, ''A review of the literature on aggresive driving research,'' OECD Econ. Surv., Ireland, Tech. Rep., 2011, pp. 8–9. [Online]. Available: http://www.stopandgo.org/research/aggressive/tasca.pdf

[4] J.-H. Hong, B. Margines, and A. K. Dey, ''A smartphone-based sensing platform to model aggressive driving behaviors,'' in *Proc. 32nd Annu. ACM Conf. Hum. Factors Comput. Syst. (CHI)*, Apr. 2014, pp. 4047–4056. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2556288.2557321

[5] J. Wang, S. Zhu, and Y. Gong, ''Driving safety monitoring using semisupervised learning on time series data,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 3, pp. 728–737, Sep. 2010.

[6] H. Ning, W. Xu, Y. Zhou, Y. Gong, and T. S. Huang, ''A general framework to detect unsafe system states from multisensor data stream,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 1, pp. 4–15, Mar. 2010.

[7] J. Chai and G. Zhao, ''Effect of exposure to aggressive stimuli on aggressive driving behavior at pedestrian crossings at unmarked roadways,'' *Accident Anal. Prevention*, vol. 88, pp. 159–168, Mar. 2016.

[8] N. Zhao *et al.*, ''An investigation of the relationship between the driving behavior questionnaire and objective measures of highway driving behavior,'' *Transp. Res. F, Traffic Psychol. Behav.*, vol. 15, no. 6, pp. 676–685, 2012, doi: 10.1016/j.trf.2012.08.001.

[9] E. Constantinou, G. Panayiotou, N. Konstantinou, A. Loutsiou-Ladd, and A. Kapardis, ''Risky and aggressive driving in young adults: Personality matters,'' *Accident Anal. Prevention*, vol. 43, no. 4, pp. 1323–1331, 2011. [Online]. Available: https://www.ucy.ac.cy/psychophysiology/documents/8.pdf

[10] L. Eboli, G. Mazzulla, and G. Pungillo, ''Combining speed and acceleration to define car users' safe or unsafe driving behaviour,'' *Transp. Res. C, Emerg. Technol.*, vol. 68, pp. 113–125, Jul. 2016.

[11] G. Cerni and M. Bassani, ''Naturalistic driving data collection to investigate into the effects of road geometrics on track behaviour,'' *Transp. Res. C, Emerg. Technol.*, vol. 77, pp. 1–15, Apr. 2017.

[12] J. Lee and K. Jang, ''A framework for evaluating aggressive driving behaviors based on in-vehicle driving records,'' *Transp. Res. F, Traffic Psychol. Behav.*, to be published, doi: 10.1016/j.trf.2017.11.021.

[13] 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 Anal. Prevention*, vol. 104, pp. 125–136, Jul. 2017.

[14] A. B. R. González, M. R. Wilby, J. J. V. Díaz, and C. S. Ávila, ''Modeling and detecting aggressiveness from driving signals,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1419–1428, Aug. 2014.

[15] M. Benavente, M. A. Knodler, Jr., and H. Rothenberg, ''Analysis of the relationship between aggressive driving and roadway characteristics using linked data,'' presented at the ITE Annu. Meeting Exhibit, Pittsburgh, PA, USA, Aug. 2007.

[16] H. R. Eftekhari and M. Ghatee, ''Hybrid of discrete wavelet transform and adaptive neuro fuzzy inference system for overall driving behavior recognition,'' *Transp. Res. F, Traffic Psychol. Behav.*, vol. 58, pp. 782–796, Oct. 2018.

[17] S. Wold, H. Ruhe, H. Wold, and W. Dunn, ''The collinearity problem in linear regression. the partial least squares (PLS) approach to generalized inverses,'' *J. Sci. Stat. Comput.*, vol. 5, no. 3, pp. 735–743, Jan. 1984.

[18] W. Wang, S. Chen, and G. Qu, ''Incident detection algorithm based on partial least squares regression,'' *Transp. Res. C, Emerg. Technol.*, vol. 16, no. 1, pp. 54–70, 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X07000411

[19] D. Filev, J. Lu, K. Prakah-Asante, and F. Tseng, ''Real-time driving behavior identification based on driver-in-the-loop vehicle dynamics and control,'' in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2009, pp. 2020–2025.

[20] H. Su and G. Zheng, ''A partial least squares regression-based fusion model for predicting the trend in drowsiness,'' *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 38, no. 5, pp. 1085–1092, Sep. 2008.

[21] C.-S. Hsieh and C.-C. Tai, ''An improved and portable eye-blink duration detection system to warn of driver fatigue,'' *Instrum. Sci. Technol.*, vol. 41, no. 5, pp. 429–444, 2013.

[22] R. Gençay, F. Selçuk, and B. J. Whitcher, *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Amsterdam, The Netherlands: Elsevier, 2001.

[23] J. B. Ramsey, ''Wavelets in economics and finance: Past and future,'' *Stud. Nonlinear Dyn. Econometrics*, vol. 6, no. 3, 2002, doi: 10.2202/1558-3708.1090.

[24] R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, ''Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm,'' *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 121–131, Jan. 2011.

[25] T. Ersal, H. J. A. Fuller, O. Tsimhoni, J. L. Stein, and H. K. Fathy, ''Model-based analysis and classification of driver distraction under secondary tasks,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 3, pp. 692–701, Sep. 2010.

[26] Y. Liang, M. L. Reyes, and J. D. Lee, ''Real-time detection of driver cognitive distraction using support vector machines,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 340–350, Jun. 2007.

[27] Z. Chen, J. Yu, Y. Zhu, Y. Chen, and M. Li, ''D^3: Abnormal driving behaviors detection and identification using smartphone sensors,'' in *Proc. 12th Annu. IEEE Int. Conf. Sens., Commun., Netw. (SECON)*, Jun. 2015, pp. 524–532.

[28] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik, ''Support vector regression machines,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 1, 1997, pp. 155–161. [Online]. Available: http://papers.nips.cc/paper/1238-support-vector-regression-machines.pdf

[29] D. Wei and H. Liu, ''Analysis of asymmetric driving behavior using a self-learning approach,'' *Transp. Res. B, Methodol.*, vol. 47, pp. 1–14, Jan. 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S019126151200121X

[30] Y.-T. Liu, Y.-Y. Lin, S.-L. Wu, C.-H. Chuang, M. Prasad, and C.-T. Lin, ''EEG-based driving fatigue prediction system using functional-link-based fuzzy neural network,'' in *Proc. Int. Joint Conf. Neural Netw.*, vol. 1, Jul. 2014, pp. 4109–4113.

[31] J. Ferreira, Jr., *et al.*, ''Driver behavior profiling: An investigation with different smartphone sensors and machine learning,'' *PLoS ONE*, vol. 12, no. 4, p. e0174959, 2017.

YONGFENG MA received the Ph.D. degree in transportation engineering from Southeast University, Nanjing, China, in 2007. He is currently an Associate Professor with the School of Transportation, Southeast University.

He has been a Visiting Scholar with the Center for Road Safety, School of Civil Engineering, Purdue University, USA, from 2014 to 2015. He has authored or co-authored over 40 articles in international journals and conferences. His research interests include intelligent transportation systems, traffic safety, driving behavior analysis, and data mining.

ZIYU ZHANG received the B.S. degree in traffic and transportation engineering from Central South University, Changsha, China, in 2017.

He is currently a Research Assistant with the Jiangsu Key Laboratory of Urban ITS and the Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University. His research interests include intelligent transportation systems, driving behavior, traffic data mining, machine learning, and deep learning.

SHUYAN CHEN received the B.E. and M.E. degrees in computer science and engineering from Southeast University, Nanjing, China, in 1989 and 1992, respectively, and the Ph.D. degree in system engineering from Southeast University, in 2003.

She is currently a Professor with the School of Transportation, Southeast University. She has authored or co-authored more than 90 articles in international journals and conferences. Her research interests include traffic incident detection, traffic signal control, travel time prediction, route guidance, data mining, and machine learning and its applications in intelligent transportation systems. She has received two National Science and Technology Progress Awards of China.

YANAN YU received the B.S. degrees in traffic and transportation engineering from Dalian Maritime University, Dalian, China, in 2016.

She is currently a Research Assistant with the Jiangsu Key Laboratory of Urban ITS and the Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University. Her research interests include intelligent transportation systems, driving behavior, traffic data mining, machine learning, and deep learning.

KUN TANG received the B.S. degree in electronic information science and technology from Hebei Polytechnic University, Tangshan, China, in 2011, and the M.S. degree in electrical and communication engineering from the Jiangsu University of Science and Technology, Zhanjiang, China, in 2014. He is currently pursuing the Ph.D. degree in transportation engineering with the Jiangsu Key Laboratory of Urban ITS, School of Transportation, Southeast University, Nanjing, China.

He is currently a Research Assistant with the Jiangsu Key Laboratory of Urban ITS and the Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University. He has authored or co-authored more than 10 articles in international journals and conferences. His research interests include intelligent transportation systems, machine learning, traffic data mining, driving behavior, and traffic safety.