Psychological Driving Style Estimation from GPS Sensor Data Alone

Hiroto Horimoto*, Ryusei Kimura*, Takahiro Tanaka† and Shogo Okada*
* School of Computer Science, Japan Advanced Institute of Science and Technology, Ishikawa, Japan
E-mail: s2310148@jaist.ac.jp, ryusei_kimura@jaist.ac.jp, okada-s@jaist.ac.jp
† Institutes of Innovation for Future Society, Nagoya University, Aichi, Japan
E-mail: takahiro.tanaka@mirai.nagoya-u.ac.jp

Abstract—**Automobiles are essential to society, but accidents involving older drivers are increasing. Driving Assistance Systems (DASs) have been developed to prevent such accidents. An essential goal is to build on this by implementing a system that provides adaptive assistance without burdening drivers with different characteristics. Accurately estimating driver characteristics is essential for this purpose. In contrast to studies that use in-vehicle sensor data through a Controller Area Network (CAN)-bus data, which requires additional equipment for data collection and is accessible to only a limited number of people, a more efficient and accessible approach is needed. Previous studies have shown that driving styles can be estimated using easily accessible Global Positioning System (GPS) data alone. We posit that the psychological driving style, which is the attitude and orientation for driving and affects the driver's behavior, can also be estimated from GPS data. In this study, we focus on developing a Driving Style Questionnaire (DSQ) as an indicator of the psychological driving style recognition model. The experimental results reveal differences in some features of all the GPS data between the high and low groups, which are divided by the median for each DSQ item on the two road types. Additionally, the findings indicate that the model with GPS data achieves the F1-macro greater than the random-assignment baseline on five out of eight of the DSQ items in the majority, and the average macro F1-score is 0.596, a difference of 0.036 from 0.632 for the model with CAN-bus data. The proposed GPS-based recognition model contributes to the development of personalized assistance systems without the burden of excessive sensor installation costs.**

I. Introduction

Automobiles are crucial to society, but they cause traffic accidents. This issue is more severe for older adults because as people age, their vulnerability to accidents or loss of life in a traffic accident increases. According to a Centers for Disease Control and Prevention report, about 7,500 older adults were killed in traffic crashes, and almost 200,000 were treated in emergency departments for crash injuries in 2020 [1].

Driving Assistance Systems (DASs), which provide vital information and warnings to drivers about circumstances that can improve driving safety and reduce traffic accidents, constitute a solution for preventing these issues.

However, these systems are generally designed for drivers in the current market [2]. This system leads to a potential mismatch between the driver and the system. One solution is to design support systems on the basis of understanding driver behavior [3], which is influenced by various factors, such as traffic, conditions, the environment, context, and personal characteristics [4]. Marafie et al. [5] reported that a personality-based driving agent improves the driver's experience. Kimura et al. [6] also reported that personalizing assistance and the basis of driver characteristics possibly improve drivers' behavior more effectively. These results indicate that an essential goal of DASs is to provide adaptive driving assistance suitable for drivers with different driving characteristics. To achieve this goal, using a Machine Learning (ML) model to accurately estimate drivers' characteristics, which are crucial in determining their behavior on the road, is the first step in implementing adaptive DASs without burden on a driver.

Some previous studies have focused on estimating driver characteristics using Controller Area Network (CAN)-bus data, a commonly used source for analyzing driving behavior [7], to provide a new direction for adaptive DASs. Wang et al. [8] proposed a prediction model for Big Five personality traits that with CAN-bus data. This study shows that driving data can automatically identify individual personality traits. Kimura et al. [9] estimated a driver's psychological characteristics, such as cognitive function, psychological driving style, and workload sensitivity from on-road driving data through a CAN-bus and a Global Positioning System (GPS) sensor. This study shows that the proposed model can accurately estimate a driver's cognitive function and characteristics.

While CAN-bus data are relatively accurate and frequent, obtaining driving sensor data through CAN-bus requires laborious processes like using data loggers for data collection [7]. Thus, it is difficult for all but a few companies or universities to collect scale data through a CAN-bus. A smartphone has a GPS sensor, and an accelerometer is a favorite platform for sensing a driver's behavior as another data source since it does not require additional equipment and is more straightforward than CAN-bus data. Furthermore, it is often used as hardware for a system to provide feedback according to analysis [5]. GPS sensor data are more reliable than accelerometer data since the orientation of a smartphone is not essential for obtaining GPS data [10]. Furthermore, they are easy to access and applicable to many vehicles/drivers [7].

In this context, some previous studies have focused on learning driving style representations, which generally aim to obtain latent representations of a driver's fine-grained driving habits [11] from GPS data and make use of such representations to identify drivers because the learned driving style

TABLE I
IN-VEHICLE SENSOR DATA THROUGH A CAN-BUS AND GPS SENSOR.

	Sensor Data	Unit
	CAN-bus	
1	Steering angle	deg
2	EPS torque	Nm
3	Forward acceleration	m/s^2
4	Lateral acceleration	m/s^2
5	Yaw rate	deg/s
6	Speed	km/h
7	Accelerator position	$\%$
8	Brake pressure	MPa
9	Fuel consumption	ml
	GPS sensor	
10	Longitude	deg
11	Latitude	deg

TABLE II
DRIVING STYLE QUESTIONNAIRE (DSQ).

	Item
1	Confidence in driving skill
2	Hesitation for driving
3	Impatience in driving
4	Methodical driving
5	Preparatory maneuvers at traffic signals
6	Importance of automobile for self-expression
7	Moodiness in driving
8	Anxiety about traffic accidents

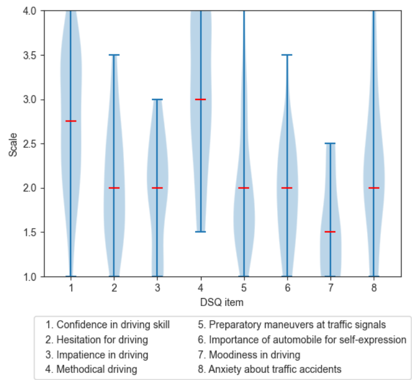

Fig. 1. Violin plots of the DSQ scores. The DSQ unit is on a scale from 1 to 4, and each violin plot represents each DSQ item. Each red line indicates the median of 23 drivers. The values are 2.5, 2.0, 2.0, 3.0, 2.0, 2.0, 1.5 and 2.0, respectively.

representations act as the “driver DNA” [12]. Chowdhury et al. [13] built an ML model for solving unique driver identification problems using smartphone GPS data as an indirect measurement of vehicle speed and acceleration. This study implies that a driver’s driving style can be analyzed using GPS data. Additionally, accelerator pedal and steering wheel signals, which can be measured directly by a CAN-bus, reflect the interaction between the driver and the vehicle. In contrast, speed and acceleration, which can be determined from both CAN-bus data and GPS data, are measures of driving styles that reflect specific driving preferences and habits [8].

On the basis of these results, we posit that psychological driving style, which is the attitude and orientation for driving and affects a driver’s behavior, can also be estimated from GPS data alone. In this study, we use the Driving Style Questionnaire (DSQ) [3], which is based on a self-report questionnaire, as an indicator of psychological driving style. In [6], the effectiveness of the driving support is suggested to depend on the drivers’ scores on the DSQ scale. We classified drivers with high and low scores of DSQ items.

Hence, as a novel challenge, we aim to develop a recognition model of psychological driving style from on-road driving data through a GPS sensor via ML models in this study.

II. DATASET

In this study, we use a dataset provided by the Institutes of Innovation for Future Society of Nagoya University [14], the same as that used in [9]. The dataset was collected from 24 older adults (12 males and 12 females) aged 50 to 79 years (the average age was 66 years). They drove a car equipped with several in-vehicle sensors through the CAN-bus and real-time kinematics GPS sensor (JAVAD Delta-3) twice each on the same public road in Japan, and 48 sets of driving session data were collected. We use 9 CAN-bus data and GPS sensor data for comparison experiments. Table I details 11 CAN-bus and GPS sensor data.

However, after removing incomplete or obviously inferior data, we retained 32 sets of driving session data from 23 drivers. Hence, for some drivers, one set of driving sessions is utilized. The Ethical Committee of Nagoya University approved all procedures in this study. Informed consent was obtained from all drivers before the experiments were conducted.

The driving tests were conducted on public roads. All participants first drove on an arterial road and then circulated in a residential area. The driving duration ranged from 2,324 s to 4,762 s, with an average of 2,907 s, and the mileage ranged from 10,079 m to 14,810 m, with an average of 12,109 m.

In addition to the driving tests, the drivers answered the DSQ, which is based on a self-report questionnaire and was introduced by [3] for characterizing driving style from a psychological aspect. Table II details the items of the DSQ. The DSQ includes eight items measured on a scale from 1 to 4. Each item is associated with two questions, and their mean value represents the score for the item. Fig. 1 shows violin plots of the scores for each DSQ item.

III. METHOD

This section presents a psychological driving style estimation method that uses GPS data alone. An overview of our model is shown in Fig. 2.

A. Data Segmentation

Estimating psychological characteristics is challenging because we do not know when or where differences in drivers’ characteristics are exhibited. Kimura et al. [9] hypothesized that it is possible to estimate drivers’ psychological characteristics by focusing on the segmentation of road types

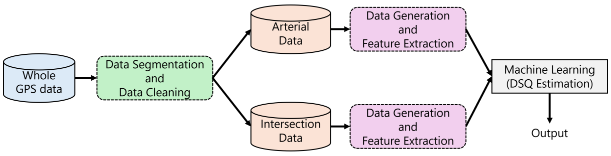

Fig. 2. Overview of our model. Arrows indicate GPS data flow. First, whole GPS data is segmented according to road type and cleaned. Then, primary and secondary GPS data are generated, and features are extracted. Finally, the DSQ is estimated via ML models with processed GPS data.

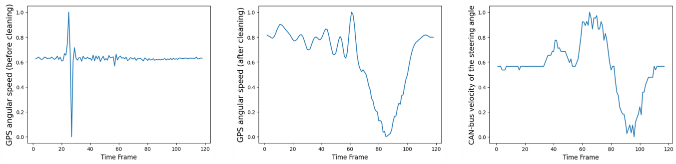

Fig. 3. Example of time-series GPS angular speed before (left) and after (middle) our data cleaning method and CAN-bus velocity of the steering angle (right) when a vehicle curves at an intersection

and partial driving data because the ability to drive safely varies across different road types. This study indicates that the estimation accuracy of the DSQ items by a method based on the hypothesis is higher than that of the method with whole driving data.

On the basis of this result, we largely followed the model introduced by [9]. First, we classified the whole GPS data into two road types, arterial roads and intersections, using the vehicle position obtained through the GPS sensor. The intersections are located in a residential area. Furthermore, we eliminated incomplete intersection data and different driving directions at the same intersection. We treated all these intersections as the same road types, but features were extracted separately because the visibility, ease of driving, and directions were different. Furthermore, when the data are arterial roads, we arranged the segment points so that the average number of seconds for each segment was in the set [ALL, 60, 30, 15, 10, 5, 3]. "All" means that no division is used. In [9], the intersection data were also segmented by brake sensor data. However, we did not do this because we only used GPS data in this study.

B. Data Cleaning

In the following section, we discuss the main challenges that arise from working with GPS data. The following speed, heading, longitudinal acceleration, and angular speed in this section are described in detail in section III-C

The first challenge is to process point cloud data. Point clouds mostly occur at very low speeds or if the vehicle stands still in one place for some time [15]. Those points do not represent any car movement and lead to high angular speed. To address this problem, when the speed was very low ($< 2km/h$), the points were removed, and then the points were filled via a linear interpolation method.

The second challenge is outliers. The set of driving session data we use contains an enormous number of GPS data points with inferior accuracy. GPS signals are affected by events such as urban tunneling and sensor error. Thus, we designed a filtering method to address this problem. First, outliers were removed, and then their points were filled via a linear interpolation method, the same as that used for the point cloud. We used an outlier detection method based on the Interquartile Range (IQR) considering longitudinal acceleration and angular speed data and set k = 2.25 to identify outliers. All the GPS data points were subsequently filtered with a Butterworth filter [16] because the smallest deviations in latitude and longitude can lead to high lateral and longitudinal acceleration [15]. The filter was a low pass filter with a degree of 2 and a cutoff frequency of 1 Hz. The filter was applied bidirectionally in the forward and reverse directions to cancel the phase delay in the data. Finally, for the arterial road data, we used the one with the smallest number of outliers detected from the filtered data for some drivers with two sets of driving sessions because large missing parts or GPS data with inferior accuracy were still found. On the other hand, we removed only the intersection data if outliers were still detected from the filtered data for the same reason as in the arterial road data.

We used one data at the arterial roads and four straight and two curve data at the intersections where all drivers' data were recorded after this cleaning in the experiments.

Fig. 3 shows one example of the time-series GPS angular speed before and after our data cleaning method and the CAN-bus velocity of the steering angle, which is the first derivative with respect to time for the steering angle when a vehicle curves at an intersection. These time-series data had point clouds ranging from 23 to 30 in each frame. Since the CAN-bus data and GPS data are not the same but similar, we scaled these data on a scale of 0 to 1 to focus only on the shape of these time series data. The point clouds considerably affect GPS angular speed before data cleaning. However, after data cleaning, the data behave like the CAN-bus velocity of the steering angle, which can be confirmed.

C. Data Generation and Feature Extraction

In this dataset, GPS data are collected at approximately 10 Hz. They consist of attributes such as the timestamp, longitude, and latitude.

We employed six primary and secondary basic data types to capture driving styles: speed, longitudinal acceleration, angular speed, lateral acceleration, longitudinal jerk, and lateral jerk. When two points' longitudes and latitudes are given, the distance (m) and the heading (deg) between the two points can be computed. The heading represents the current direction compared with North, and its range is $(0, 360]$. We computed the first derivative of the distance with respect to time to obtain the speed (m/s). Furthermore, we implemented a moving mean filter to make them smooth for speed and heading.

Few important secondary data are computed from these primary GPS measurements, namely, speed and heading. The longitudinal acceleration (m/s^2) can be computed from consecutive speed measurement samples via the equation expressed as:

$$acc_{long} = \frac{\Delta v}{\Delta t} \quad (1)$$

where Δv is the speed variation ($v_n - v_{n-1}$), and Δt is the temporal variation. The GPS data in our dataset have a sampling frequency of approximately 10 Hz, so Δt is approximately 100 msec. The angular speed (rad/s), which represents the rate of change in heading, can be computed from consecutive heading measurement samples via the expression specified by [13]:

$$ang_{speed} = \frac{\Delta\theta}{\Delta t} \cdot \frac{\pi}{180} \quad (2)$$

where $\Delta\theta$ is the heading variation. Given the vehicle's heading, before computing the angular speed we modified the heading variation via the following equation:

$$f(\Delta\theta) = \begin{cases} \Delta\theta - 360 & \text{if } \Delta\theta > 180, \\ 360 + \Delta\theta & \text{if } \Delta\theta < -180, \\ \Delta\theta & \text{otherwise.} \end{cases} \quad (3)$$

The lateral acceleration (m/s^2) can be evaluated via the expression specified in [13] and [17]:

$$acc_{lat} = \frac{v^2}{R} = v \cdot ang_{speed} \quad (4)$$

where v is the current time value and R is the turn radius, as introduced by [17]. The jerk, which represents the rate of change in acceleration, can be computed from consecutive accelerations. Therefore, we compute the first derivative with respect to time for the longitudinal acceleration and the lateral acceleration and use them as the longitudinal jerk (m/s^3) and the lateral jerk (m/s^3), respectively. Thus, the above equations can compute secondary data from primary GPS data.

The statistical features of the generated primary and secondary GPS data were extracted to estimate the DSQ via ML models. These statistical features are the mean, median, skewness, kurtosis, variance, and max. Finally, 36 (six data × six statistics) features were extracted. When a driver has multiple data at the same intersection, we computed the mean value of each feature.

D. Machine Learning Model

We used logistic regression with L2 regularization, a linear support vector machine, and random forest models to estimate the DSQ results. We split scores of the DSQ items on the basis of the median value to create binary classification labels as shown in Fig. 1 and then conduct binary classification. The model outputs the DSQ items' class categories (high or low).

IV. Experimental Settings

As an evaluation criterion for the classification model, we report the macro F1-score. We conduct binary classification for DSQ estimation. The parameter C values of the logistic regression and linear support vector machine are selected from the set [0.001, 0.01, 0.1, 1, 10, 100]. The maximum depth of the tree of the random forest is selected from the set [3, 5, 7, 9, 11]. We perform an exhaustive search of specified parameter values to select the best value in the training set. We compute the feature significance of a real-valued feature to a binary targets (high or low) as a p-value on the Kolmogorov-Smirnov (KS) test, which is a nonparametric hypothesis test that statistically examines differences in distribution. A $p < 0.05$ was considered statistically significant in this study. Therefore, we only include features that are statistically significant for each fold to avoid overfitting and reduce computational cost because there are 9036 features on arterial road data and 216 features on intersection data from the GPS sensor. We use leave-one-person-out cross-validation to evaluate the classification model.

In addition, we evaluate our model's effectiveness in estimating the DSQ from GPS features by comparing its feature set with those from the CAN-bus data detailed in Table I. The input signals from these CAN-bus data are resampled at 10 Hz to match the GPS data. In addition to the above CAN-bus data, we compute and use the first derivative with respect to time for the steering angle, forward acceleration, lateral acceleration, and accelerator position and use them as the velocity of steering angle, forward jerk, lateral jerk, and change of the acceleration position are used as secondary data. We also applied data segmentation and feature extraction to the CAN-bus data. After GPS data cleaning, all feature sets use the same driving data at arterial roads and intersections.

V. Results and Discussion

We applied the KS test to all drivers' statistical features to confirm the importance of GPS data for DSQ estimation (section V-A). Next, we investigate how the model with GPS data can estimate the DSQ results to validate our method and compare feature sets with GPS data and CAN-bus data (section V-B). This study evaluates the estimation macro F1-score by comparing the following two feature sets.

FS (GPS): a GPS feature set that is from all six primary and secondary GPS data, which is our method.

FS (CAN-bus): a CAN-bus feature set that is from all 13 primary and secondary CAN-bus data.

Fig. 4 shows bar plots of the ratio of GPS data with features highly related to the DSQ scores with $p < 0.05$ on the KS test. Table III shows the classification result of each DSQ item of a model with a GPS feature set and a CAN-bus feature set. The bold values represent the highest macro F1-score among each DSQ item in each road type with values greater than 0.5. We compute an average macro F1-score from the highest macro F1-score among each DSQ item from two road types within each feature set.

A. Importance of GPS Data

From the results of the KS test shown in Fig. 4, some features of all the GPS data were different between the high and low groups, with statistically significant differences in the arterial road data and intersection data when all the DSQ items were considered. In particular, the ratio of the speed with features was highest (0.275) on "importance of automobile for self-expression" on arterial road data, while

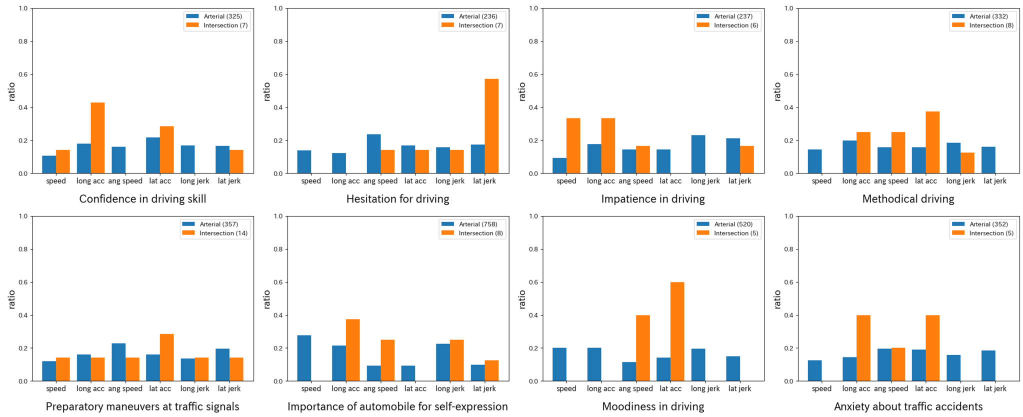

Fig. 4. Bar plots of the ratio of GPS data with features highly related to the DSQ scores with $p < 0.05$ on the Kolmogorov-Smirnov test. Blue (left) and orange (right) bars represent arterial roads and intersections. Legend numbers in parentheses show total statistically significant features. long acc = longitudinal acceleration, ang speed = angular speed, lat acc = lateral acceleration, long jerk = longitudinal jerk, lat jerk = lateral jerk.

the lateral acceleration with features was highest (0.600) on “moodiness in driving” on intersection data. This correlation between each DSQ item, which characterizes driving styles from a psychological aspect, and the primary and secondary GPS data on driving styles has significant implications for understanding and predicting driver behavior. This result also suggests that these important GPS features can be used to classify each DSQ item into groups via ML models while avoiding overfitting. Furthermore, 236 (2.61%) to 758 (8.39%) features from 9036 features on the arterial road data and 5 (2.31%) to 14 (6.48%) features from 216 features on the intersection data remained important GPS features for each DSQ item. This test contributes to reducing computational costs and improving cost-effectiveness.

B. Analysis of Classification Accuracy

From the classification result of each DSQ item of a model with FS (GPS) shown in Table III, three on each road type and five on two road types out of eight DSQ items were estimated, with the highest macro F1-score greater than 0.5, which is the random-assignment baseline. In particular, the macro F1-score of “importance of automobile for self-expression” and “moodiness in driving” were highest (0.625) on arterial road data, whereas “confidence in driving skill” was highest (0.775) on intersection data. Moreover, the model’s average macro F1-score was greater than 0.5. These results indicate that the majority of the DSQ items, which represent the psychological driving style, can be adequately estimated from primary and secondary GPS data on driving styles.

In contrast, a model with FS (CAN-bus) achieved the highest macro F1-score greater than 0.5: six on arterial road data, three on intersection data, and seven on two road types out of eight DSQ items. The model’s average macro F1-score was higher than the model with FS (GPS). These results indicate that the difference in the average macro F1-score between FS (GPS) and FS (CAN-bus) methods was close to 0.036. However, it also indicates that the model with FS (GPS) was considerably inferior to that with FS (CAN-bus) compared with the number of items greater than 0.5.

In addition, the results reveal a large difference between the macro F1-scores of the model with FS (GPS) and FS (CAN-bus) on “methodical driving” on the arterial road data. Therefore, we applied the KS test to all drivers’ FS (CAN-bus) and found that forward acceleration (0.151), lateral acceleration (0.123), and steering angle (0.114) presented high ratios. One possible reason why the macro F1-scores for FS (GPS) and FS (CAN-bus) differ significantly while the top three features in the CAN-bus are of a similar data type to the generated GPS data is that the generated GPS data, which represent the driving style, are not exactly identical in behavior to the CAN-bus data. Fig. III also shows that the model with FS (GPS) had higher macro F1-scores than did the model with FS (CAN) for some DSQ items, suggesting that the generated GPS data are important for these specific DSQ items.

VI. Conclusions

In this paper, we present the results of our study on the psychological driving style estimation approach using only GPS sensor data and compare the models’ results with CAN-bus data. The experimental results show differences in some features of all the GPS data between the high and low groups for each DSQ item on two road types via the KS test. Additionally, the findings indicate that the model with a GPS feature set can achieve a macro F1-score greater than 0.5 for five out of eight DSQ items in the majority and the average macro F1-score is 0.596 a difference of 0.036 from 0.632 for the model with a CAN-bus feature set. Our proposed GPS-based recognition model has the potential to significantly reduce the cost of developing personalized assistance systems, as it does not require expensive sensor installations. This study used GPS sensor data collected at approximately 10 Hz, whereas smartphone GPS data are typically collected at 1 Hz. To develop personalized DASs using smartphones alone in the future, it is recommended that we investigate whether a model with the data can estimate psychological driving style.

TABLE III
CLASSIFICATION MACRO F1-SCORE FOR DSQ ITEMS.

DSQ item	FS (GPS)						FS (CAN-bus)					
	Arterial			Intersection			Arterial			Intersection		
	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF	LR	SVM	RF
Confidence in driving skill	0.300	0.258	0.395	0.517	0.436	**0.775**	0.361	0.361	0.410	0.303	0.436	0.343
Hesitation for driving	0.329	0.425	0.361	0.487	0.384	0.343	0.617	**0.654**	0.395	0.281	0.436	0.487
Impatience in driving	0.417	0.327	0.425	0.343	0.361	0.462	0.517	**0.617**	0.425	0.436	0.436	0.395
Methodical driving	0.324	0.258	0.395	0.549	**0.662**	0.343	0.617	**0.744**	0.410	0.329	0.357	0.410
Preparatory maneuvers at traffic signals	0.394	**0.544**	0.324	0.654	**0.692**	0.521	**0.673**	0.617	0.462	**0.712**	0.487	0.673
Importance of automobile for self-expression	0.487	0.555	**0.625**	0.487	0.303	0.357	**0.635**	0.456	0.549	0.300	0.452	0.281
Moodiness in driving	**0.625**	0.487	0.462	0.258	0.324	0.462	**0.589**	**0.589**	0.378	0.673	**0.713**	0.625
Anxiety about traffic accidents	0.238	0.361	0.439	0.378	0.361	0.378	0.410	0.425	0.425	**0.546**	0.378	0.439
Average	0.596						0.632					

LR = Logistic Regression; SVM = Linear Support Vector Machine; RF = Random Forest.
The bold values represent the highest macro F1-score among each DSQ item on each road type with values greater than 0.5. The values in the last row represent the average macro F1-score from the highest macro F1-score among each DSQ item from two road types within each feature set.

REFERENCES

[1] Centers for Disease Control and Prevention, *Web-based Injury Statistics Query and Reporting System (WISQARS)*, https://wisqars.cdc.gov/. Accessed 21 Jul 2024, 2022.

[2] C. M. Martinez, M. Heucke, F.-Y. Wang, B. Gao, and D. Cao, "Driving style recognition for intelligent vehicle control and advanced driver assistance: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 3, pp. 666–676, 2017.

[3] M. Ishibashi, M. Okuwa, S. Doi, and M. Akamatsu, "Indices for characterizing driving style and their relevance to car following behavior," in *SICE Annual Conference 2007*, IEEE, 2007, pp. 1132–1137.

[4] M. Akamatsu, Y. Sakaguchi, M. Okuwa, T. Kurahashi, and K. Takiguchi, "Database of driving behavior measured with equipped vehicles in the real road environment," in *Proceedings of International Ergonomic Association*, 2003.

[5] Z. Marafie, K.-J. Lin, D. Wang, *et al.*, "Autocoach: An intelligent driver behavior feedback agent with personality-based driver models," *Electronics*, vol. 10, no. 11, 2021.

[6] R. Kimura, H. Horimoto, T. Tanaka, and S. Okada, "Investigating the effect of driving support robot depending on driver characteristics," in *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024, pp. 608–612.

[7] M. N. Azadani and A. Boukerche, "Driving behavior analysis guidelines for intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6027–6045, 2022.

[8] Y. Wang, N. Zhao, X. Liu, S. Karaburun, M. Chen, and T. Zhu, "Identifying big five personality traits through controller area network bus data," *Journal of Advanced Transportation*, vol. 2020, p. 8 866 876, 2020.

[9] R. Kimura, T. Tanaka, Y. Yoshihara, K. Fujikake, H. Kanamori, and S. Okada, "Estimating driver personality traits from on-road driving data," *IEEE Access*, vol. 11, pp. 93 679–93 690, 2023.

[10] A. Chowdhury, T. Chakravarty, and P. Balamuralidhar, "Estimating true speed of moving vehicle using smartphone-based gps measurement," in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2014, pp. 3348–3353.

[11] N. Lin, C. Zong, M. Tomizuka, P. Song, Z. Zhang, and G. Li, "An overview on study of identification of driver behavior characteristics for automotive control," *Mathematical Problems in Engineering*, vol. 2014, pp. 1–15, 2014.

[12] U. Fugiglando, P. Santi, S. Milardo, K. Abida, and C. Ratti, "Characterizing the "driver dna" through can bus data analysis," in *Proceedings of the 2nd ACM International Workshop on Smart, Autonomous, and Connected Vehicular Systems and Services*, 2017, pp. 37–41.

[13] A. Chowdhury, T. Chakravarty, A. Ghose, T. Banerjee, and P. Balamuralidhar, "Investigations on driver unique identification from smartphone's gps data alone," *Journal of Advanced Transportation*, vol. 2018, p. 9 702 730, 2018.

[14] Y. Yoshihara, E. Takeuchi, and Y. Ninomiya, "Accurate analysis of expert and elderly driving at blind corners for proactive advanced driving assistance systems," in *Transportation Research Board 95th Annual Meeting, no. 16-1992*, 2016.

[15] S. Würtz and U. Göhner, "Driving style analysis using recurrent neural networks with lstm cells," *Journal of Advances in Information Technology*, vol. 11, pp. 1–9, 2020.

[16] S. Butterworth, "On the Theory of Filter Amplifiers," *Experimental Wireless & the Wireless Engineer*, vol. 7, pp. 536–541, 1930.

[17] 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, pp. 88–96, 2014.