2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

Estimation of Drivers' Cognitive Load Through Foot Placement Analysis in a Car-Sharing Service

Takuya Sukegawa, Yasuhiro Hashimoto, Keisuke Hata
The University of Aizu
Aizu-Wakamatsu, Fukushima 965-8580, Japan
{m5261165, hashimo, hata}@u-aizu.ac.jp

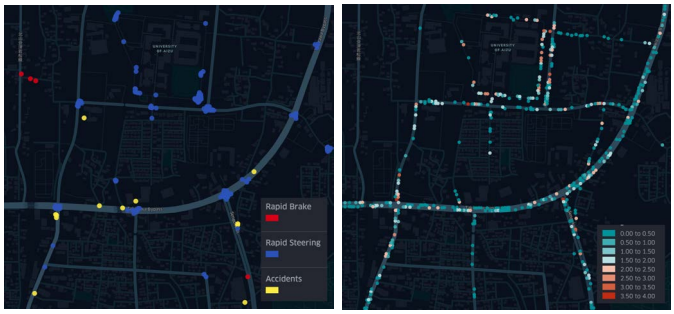
***Abstract*—Driver behavior plays a crucial role in mitigating traffic accidents. Unlike previous studies focusing on drivers' actual maneuvers, our study examines actions preceding a maneuver or those not manifest in conventional maneuvers. We gathered data from a car-sharing service frequented by university students, scrutinizing the frequency of pedal changes via foot camera images. These data were compared with the pedal depression measurements from Controller Area Network (CAN) bus data to discern instances of potentially reduced safety due to cognitive load. We identified locations with high pedal change frequency, even in the absence of any recorded pedal operation in the CAN bus data. This result suggests the existence of unrecorded driver behaviors that could precipitate traffic accidents. Our findings, therefore, bear substantial implications for enhancing traffic safety measures.**

***Index Terms*—Active safety, Foot tracking, Intelligent vehicle, Intelligent driving assistance, Driving maneuver**

I. Introduction

In recent years, the number of fatalities due to traffic accidents in Japan has been on a downward trend since its peak in 1992. This reduction has been attributed to factors such as passive safety technology, strengthening international automobile safety standards, and developing emergency medical care. However, the number of traffic accidents remains high, with approximately 340 thousand accidents recorded in 2020[1], underlining the need for further efforts to reduce traffic incidents.

The concept known as the "Heinrich Triangle" states that industrial accidents consist of *major accidents*, *minor accidents*, and a much higher occurrence of *near-misses*. Traditionally applied in industrial safety, this principle can be extrapolated to traffic accidents. In such a context, *major* and *minor accidents* would correspond to actual traffic incidents, while *near-misses* would denote situations requiring emergency braking or steering. In Japan, police authorities routinely analyze and disclose information about the locations and times of traffic accidents. Moreover, efforts are underway to identify high-risk locations using probe data for *near-miss* incidents [1]. Figure 1a provides a summary of traffic safety information for Aizu-Wakamatsu City in Japan, focusing on the vicinity of the University of Aizu. This includes locations of emergency braking (red), sudden steering (blue), and traffic accidents (yellow) that occurred between 2018 and 2020. However, determining the underlying cause or patterns of such incidents is a challenge based solely on the spatial distribution of recorded locations. While these incident maps can shed light on high-risk areas, their ability to accurately estimate potential risk in areas with lower incident rates or even no recorded incidents is limited.



(a) (b)

Fig. 1: Panel (a) depicts high-risk locations identified using conventional methods: emergency braking (red), sudden steering (blue) observed by the CAN bus system in the experimental vehicles, as well as traffic accident locations disclosed by police (yellow). Conversely, Panel (b) displays risk estimations based on our proposed method, which utilizes drivers' foot movements. Sites, where a high cognitive load is estimated are indicated by reddish dots.

Our study aims to estimate drivers' cognitive load that could lead to non-safety states—situations one level below *near-miss* states—by analyzing foot movements between accelerator and brake pedals. We capture these movements using a camera installed in the driver's footwell. This allows us to observe 'preliminary' actions, which are the driver's preparatory steps before executing a specific maneuver, such as depressing the accelerator or brake. The accepted model of driving behavior comprises three stages: *perception*, *judgment*, and *operation* [2]. While in-vehicle sensors capture the *operation* stage, a comprehensive understanding of the preceding *perception* and *judgment* processes requires an analysis of the driver's implicit behaviors, such as these foot movements. Our study juxtaposes actual maneuvers recorded by the Controller Area Network (CAN) bus system with non-explicit driver movements observed from foot video footage. We hypothesize that the frequency of pedal shifting and the nature of foot

[1]Road Safety Annual Report 2021: Japan.

978-0-7381-4408-5/23/$31.00 ©2023 IEEE

Manuscript 106 submitted to 2023 IEEE Symposium Series on Computational Intelligence. Received July 27, 2023.

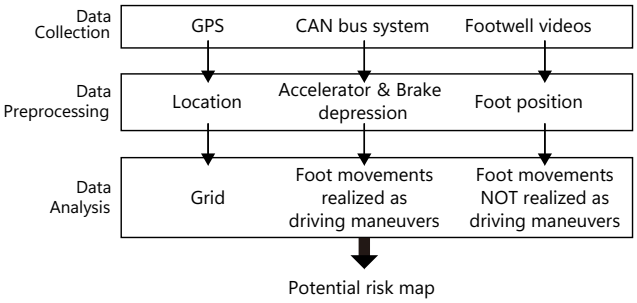


Fig. 2: Flow of our data analysis: Three data sources—the GPS, CAN bus system, and footwell camera—are integrated to create a risk map to illustrate areas of high cognitive load.

movements can reveal insights into the driver's cognitive load. Our approach is rooted in the belief that frequent pedal shifting may reflect stressful driving scenarios, increasing cognitive load. Similarly, foot movements between the accelerator and brake pedals, which do not result in actual depressions, could signal cognitive strain due to ambivalence or hesitance in decision-making. Though these movements do not directly influence vehicle control, they may capture subtle shifts in cognitive strain and serve as potential precursors to increased driving risk. In our work, we delve beyond observable driving maneuvers, scrutinizing these non-explicit driver behaviors as possible indicators of cognitive load and driving risk. Finally, we visualize discrepancies between foot movement data and CAN bus data on a map (Fig. 1b), facilitating further explorative analysis. The flow of this analysis is depicted in Fig. 2.

The paper is structured as follows: Section 2 reviews the existing literature in the field and distinguishes the unique contribution of this study. Section 3 describes the data collection and preprocessing methods utilized for our analysis. Section 4 outlines the experiment, detailing the methodologies deployed and key findings achieved. Following a discussion in Section 5, we conclude Section 6.

II. Related work

Prior research on cognitive load during driving has predominantly focused on the driving scene and the impact of in-vehicle tasks on the driver. For instance, Lian et al. employed the Gray correlation method to scrutinize the factors influencing workload and performance during highway driving [3]. They validated their findings by comparing them with actual traffic accident data. In another innovative approach, Nakayama et al. developed the so-called steering entropy method to gauge the smoothness of steering operation [4]. This method was utilized to assess the driver's workload when they were tasked with additional responsibilities such as conversations with fellow passengers. However, these studies focusing on drivers' cognitive load are primarily centered on behavioral analysis without considering the impact of regional, spatial, and traffic environmental factors within their scope of investigation.

Understanding driving behavior, including cognitive load, under specific situations traditionally involves experiments using driving simulators and test courses [5]. While these approaches guarantee high reproducibility of experimental conditions, they face challenges mimicking natural road environments and gathering many participants. Recent studies have addressed these issues by harnessing data from vehicles equipped with telematics devices driven by the general public [6]. Our study extends this approach, sourcing data from a car-sharing service to study everyday drivers in real-world conditions.

The 'human as a distributed sensor' concept capitalizes on collective intelligence and individual sensing capabilities. For instance, Aichinger et al. identified collision risk hotspots using smartphone GPS and motion sensor data [1]. Chen et al. estimated traffic accident risks through a deep learning approach, leveraging traffic accident data and GPS records from millions of users [7]. Similarly, Lee et al. identified high-risk locations by collating data from traffic accident statistics, road information, and user comments [8]. These studies share our natural state data collection approach, which minimizes experimental awareness. Unique to our study is incorporating behavioral data derived from novel foot placement analysis into the risk evaluation process.

Creating risk maps by aggregating traffic data has been a common approach in traffic safety research. For example, He et al. and Vandenbulcke et al. used various data sources to predict accident probabilities, highlighting the value of high-resolution mapping and resourceful use of limited data [9], [10]. In contrast, our study utilizes in-vehicle data from a shared car service to investigate how specific traffic situations may induce cognitive load in drivers, leading to states of decreased safety awareness or responsiveness. We use foot camera images to monitor pedal changes and CAN bus data to record pedal depressions. By associating areas where pedal change frequency and its discrepancy from actual maneuvers are high with heightened cognitive load, we offer a unique perspective on traffic safety, bridging the gap between traditional risk mapping techniques and granular behavioral data.

III. Data

A. Data acquisition

The data utilized in this study were sourced from a car-sharing service provided at the University of Aizu from December 2019 to February 2022. In this service, participants who registered were allowed to reserve one of the two experimental vehicles (Fig. 3a) at their convenience, and their driving time, purpose, destination, and route were left up to them. Out of 47 university students who registered for the service, we selected 39 subjects who had used the service for more than 10 hours. These 39 subjects were all males aged between 19 and 24, with less than five years of driving experience. All subjects were informed about the purpose of the study.[2]

[2]The approval of the experiment was obtained from the university's research ethics review committee.

Manuscript 106 submitted to 2023 IEEE Symposium Series on Computational Intelligence. Received July 27, 2023.

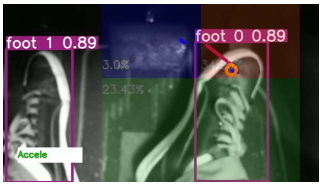
(a) Experimental vehicle

(b) Cockpit side camera

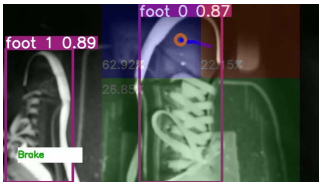
(c) Front camera

(d) Foot camera

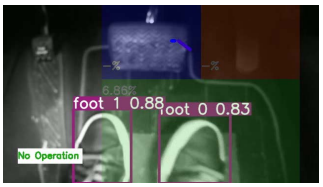
Fig. 3: (a) Appearance of an experimental vehicle: Two of these four-seat sedans were used in the experiment. (b) (c) (d) Still frames captured by the cameras mounted on the experimental vehicle. (c) and (d), in particular, are used in the present study.



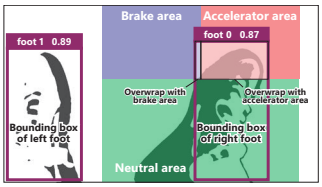
(a) Foot is on the accelerator.



(b) Foot is on the brake



(c) Foot is in the neutral position.



(d) Each area and bounding box.

Fig. 4: Detection of foot movements. The position of the foot—whether on the accelerator, brake, or in a neutral state—is determined by evaluating the overlap between the bounding boxes of the foot and each pedal area.

This study mainly deals with two types of data: CAN bus data, which captures actual driving operations, and footwell video images, which record implicit foot movements. To ensure that the driver's behavior was as natural as possible, all the installed sensors and cameras were discreetly placed to minimize any perception of the vehicle as an experimental unit. The CAN bus data comprise the amount of depression in accelerator and brake pedal operations, steering angle, and speed, as well as GPS data that recorded location information.[3] In this study, we use, in particular, the amount of depression in accelerator and brake pedals, speed, and GPS data from the CAN bus data. In addition, five cameras were installed in the vehicle, including a front camera, face camera, foot camera, front seat camera, and rear seat camera as in Fig. (3b), (3c), and (3d). The frame rate of the camera images was downsampled from 15 to 10 frames per sec (FPS) to synchronize with the frequency of the CAN bus data. The sampling interval for GPS data is 0.5 seconds, 2 FPS, which is used as is. One *trip* is defined as the duration of driving from the moment a subject enters the car until they exit. The dataset encompasses 437 trips undertaken in Aizu-Wakamatsu City between April 2020 and November 2021. This amounts to approximately 3.8 million records, corresponding to roughly 105 hours of video footage for each camera with 640×480 pixels.

B. Estimation of foot movements

We analyze foot movements using video images captured from the footwell. Our goal here is to assign each CAN bus data record a corresponding foot position label—either 'Accelerator,' 'Brake,' or 'Neutral'—for each time frame.

[3]We also recorded temperature, humidity, and CO2 levels as environmental factors inside the vehicle. Although this information could potentially inform an analysis of the driver's stress level, it is beyond the scope of the present study.

Traditional video analysis methods for estimating foot pedal operation, such as optical flow or frame-by-frame comparison, have limitations [11]. Specifically, they struggle with changes in video brightness, which can compromise accuracy under varying light conditions, for instance, when a vehicle emerges from a tunnel. To circumvent these issues, we adopted the YOLOv5 model [12], fine-tuning it for our experiment. This custom model effectively identifies feet in camera images, enabling the estimation of the foot's relative position to the accelerator and brake pedals.

We trained our model on a dataset of 500 images depicting various shoe types under different lighting conditions. These images were independent of the actual driving experiment. As illustrated in Fig. 4, the model identifies shoes by enclosing them within purple bounding boxes, each annotated with a confidence score in white.[4] We assigned operational zones to different areas within the foot camera image: an accelerator, a brake, and a neutral position corresponding to the red, blue, and green areas, respectively. Determining the foot's position involves identifying which area—either accelerator or brake—has the most overlap with the bounding box of the foot. If the foot's overlap with either area is less than 5% of the bounding box, we categorize the foot position as neutral. As a simple error correction, any transient change in the foot position label immediately followed by a return to the original label is treated as a detection error; in such cases, we preserve the initial label. Comparing the actual pedal depressions recorded by the CAN bus data with the foot placements estimated by our model, discrepancies or errors accounted for only 7% of the total driving time. This modest error rate, combined with the post-processing for error correction mentioned above, supports

[4]For reference, the model achieved a mean Average Precision (mAP@[.5:.95]) of 0.739, a comprehensive metric for evaluating object detection models.

Manuscript 106 submitted to 2023 IEEE Symposium Series on Computational Intelligence. Received July 27, 2023.

TABLE I

Time stamp	Latitude	Longitude	Velocity (km/h)	α	β	λ
2021-09-10 15:32:10.140	37.524317	139.939787	0	0	0	Brake
⋮	⋮	⋮	⋮	⋮	⋮	⋮
2021-09-10 15:34:31.940	37.51934	139.942558	57.91	41	0	Accelerator
2021-09-10 15:34:32.040	37.51934	139.942558	57.91	0	0	Brake
2021-09-10 15:34:32.140	37.51934	139.942558	57.69	0	3	Brake
⋮	⋮	⋮	⋮	⋮	⋮	⋮
2021-09-10 15:40:58.140	37.504927	139.938748	0	0	0	Brake

An example of preprocessed data for a typical trip. Variables α and β denote the degrees of depression for the accelerator and brake pedals from the CAN bus data. λ represents the label for the foot position from the foot camera. Each record, sampled at 10 FPS, with location sampled at 2 FPS, is consolidated into a single data frame per trip.

the practical applicability of our approach, enabling reliable analysis of foot positions.

We integrate the CAN bus data and the time series of foot position into a single time series for each trip, as shown in Table I. Each data record corresponds to one frame at a 10 FPS rate, meaning that the unit time for the series is 0.1 seconds. Finally, we exclude special pedaling activities related to parking, starting, and stopping to focus our analysis on standard driving operations. Therefore, we only consider operations performed at speeds ranging between 20 km/h and 60 km/h.

IV. Experiment

We propose to compare the frequency of pedal changes as captured by a foot camera to those measured via pedal depressions in CAN bus data. This approach aims to assess the cognitive load from driving situations or environments by quantifying alterations in pedal depression. If the count of pedal changes, based on actual foot movements, diverges from those recorded by the CAN bus system at specific locational points, this discrepancy could indicate unaccounted driving risks there that the conventional system may overlook, as mentioned in Section I

A. Map discretization

We begin by discretizing the area of interest into a grid, where each unit grid corresponds to a 10-meter square. The area of analysis, located within Aizu-Wakamatsu City, covers the region traversed by the shared cars. This region forms a rectangle defined by the coordinates of its southwest edge (37.3525, 139.8461) and its northeast edge (37.5813, 140.0439). Each 10-meter grid within this area is assigned a unique index, denoted by i. We then assign each record to the corresponding grid using the data records, as shown in Table I. The number of data records and unique users within each grid is represented by n_i and u_i, respectively. We found that around 33,000 grids, which cover an area of 3.3 km^2, contain at least one data record.

B. Frequency of pedal operations

We define a pedal change based on both foot camera data and CAN bus data, as illustrated in Fig. 5. From the

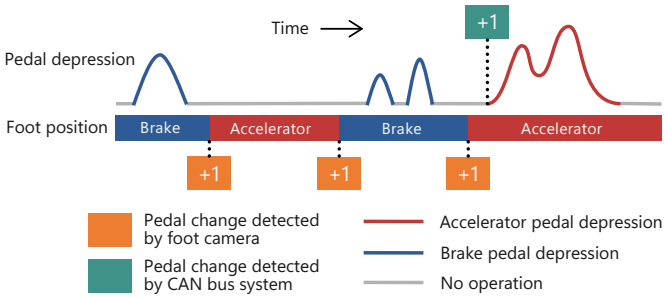


Fig. 5: Example of pedal change counting. The foot camera records three pedal changes in the above scenario, while the CAN bus data records only one.

foot camera's perspective, a pedal change is deemed to have occurred when the driver's foot moves from the 'Accelerator' to the 'Brake' or vice versa. However, since the camera cannot quantify the actual extent of pedal depression, such measurements are not taken into account in this context. Meanwhile, the CAN bus data records a pedal change when the depression level of one pedal (either the accelerator or brake) surpasses that of the other from one frame to the next. This definition implies that foot shifts without actual pedal depressions are not considered pedal changes.

In grid i, we calculate the number of pedal changes as recorded by both the CAN bus data and the foot camera, denoting these counts as m_i^{CAN} and m_i^{Foot}, respectively. We then use these counts to define the frequency of pedal changes within grid i: $\mu_i^{\mathrm{Foot}} = m_i^{\mathrm{Foot}}/n_i$ and $\mu_i^{\mathrm{CAN}} = m_i^{\mathrm{CAN}}/n_i$. To ensure the statistical reliability and robustness of our findings, we focus only on grids with a substantial amount of data, setting a threshold of $n_i \geq 10$. Using this criterion, we identify 3,670 grids that meet this requirement. Figure 6 depicts a histogram of μ_i constructed from 3,670 valid grids. There are 2,851 grids with $\mu_i^{\mathrm{CAN}} = 0$ and 1,076 grids with $\mu_i^{\mathrm{Foot}} = 0$. A comparative analysis of these histograms reveals that at nearly all points, the distribution of μ_i^{Foot} predominates over that of μ_i^{CAN}. This observation validates our understanding that the foot movements captured by the foot camera encompass a larger set of behaviors than the pedal depressions recorded by the CAN bus system. Furthermore, the histograms

Manuscript 106 submitted to 2023 IEEE Symposium Series on Computational Intelligence. Received July 27, 2023.

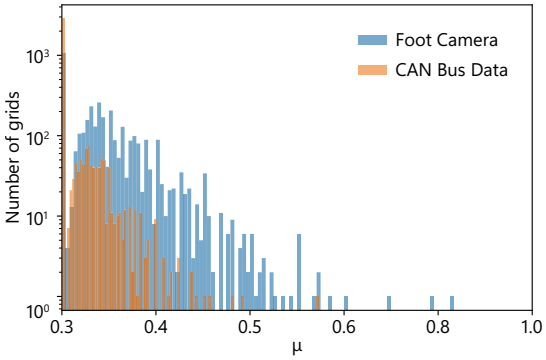


Fig. 6: Histogram of μ_i—the frequency of pedal changes at grid i. The orange distribution represents μ_i^{CAN}, and the blue one does μ_i^{Foot}.

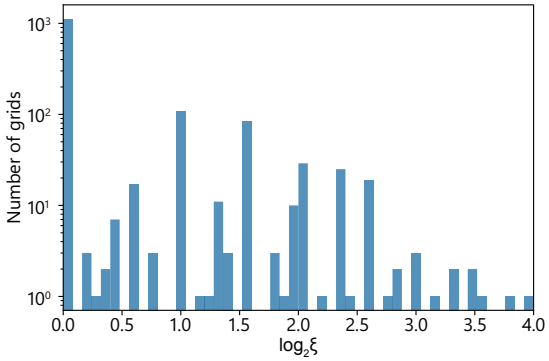


Fig. 7: Histogram of ξ_i—the discrepancy between μ_i^{CAN} and μ_i^{Foot}. On the x-axis, $\log_2 \xi_i$ of 4 signifies μ_i^{Foot} is 16 times larger than μ_i^{CAN}.

reveal a pattern in which a significant proportion of grids are associated with minimal foot activity, whereas only a few indicate a heightened frequency of foot movements. Notably, the histogram for μ_i^{Foot} highlights the presence of grids with relatively larger frequencies exceeding 0.5. This indicates the presence of specific grids or spots where foot movements between the accelerator and brake pedals occur in over half of the recorded instances.

C. Characteristic features revealed by the foot camera

We calculate the discrepancy between the frequency of pedal changes as captured by the CAN bus data and the foot camera. This discrepancy, denoted as ξ_i, is defined as the ratio $\mu_i^{\mathrm{Foot}}/\mu_i^{\mathrm{CAN}}$. A grid with $\xi_i = 1$ signifies that the frequency of pedal changes exactly matches the frequency of foot movements within that grid. In other words, each observed foot movement corresponds to an actual pedal alteration and vice versa. Although we expect $\xi_i \geq 1$ (since the CAN bus system only captures a subset of foot movements), instances of $\xi_i < 1$ can occur due to data loss or errors in foot position estimation. These grids, likely to contain less reliable data, are excluded from our analysis.

Figure 7 shows the histogram of ξ_i derived from 3,670 valid grids. There are 1,123 grids where $\xi_i = 1$, indicating a high degree of comparability between the pedal operations recorded by the CAN bus system and the foot camera across most grids. However, some grids show notably high ξ_i values. In these particular areas, a significantly higher frequency of foot movements does not translate to actual pedal operations. This finding suggests that the traffic environment within these areas may possess certain characteristics that trigger such implicit foot movements.

In Fig. 8 and Table II, we highlight with red and green circles several specific spots that exhibit high ξ_i values ($\xi_i \geq 8$). The red circles indicate locations of previous traffic accidents or near-miss incidents, as well as specific areas that, based on our local knowledge, could intuitively be anticipated as likely to induce cognitive load. For instance, Intersection (a) is the busiest traffic area in Aizu-Wakamatsu City. It reports the highest traffic accidents in Japan, earning a government designation as an accident-prone area. Locations (d) and (e) depict driving routes on the University of Aizu campus: Location (d) is near the campus entrance, a spot where drivers need to pay attention to pedestrians and other vehicles, while (e) is close to cones that drivers must circumvent. Location (h) represents a heavily trafficked T-junction. These spots could be readily identified as high-risk areas using conventional approaches.

On the other hand, our analysis revealed some unique spots represented by green circles—these were situated farther from busy intersections or without nearby traffic lights and had previously shown no signs of being 'special' regarding frequent foot movements. Location (b), for instance, is positioned just before Intersection (a), a spot we found prone to traffic queues during rush hour. The area near (c) features a straight, well-visible road with no traffic lights, aside from a roadside alone restaurant. Location (f) is a straight road adjacent to the university, where traffic typically moves at high speeds. Spot (g) marks an exit from a narrow side road onto the main road. Finally, location (i) represents an exit of a back road commonly used as a detour near the junction of this back road and the main road.

V. Discussion

Our study was predicated on the hypothesis that the frequency of pedal shifting could reflect cognitive load during driving. We substantiated our hypothesis by examining actual driving scenarios captured by front camera videos and leveraging our local knowledge of the area around the university campus. Indeed, we discovered a positive association between locations with high ξ_i values and real-world driving situations that induce cognitive stress. These situations mainly occur in areas subjected to heavy traffic, such as busy intersections and locations congested during rush hour. In future research, high-performance object recognition using front camera images could quantitatively and automatically assess this relationship. This approach could benefit cognitive load estimation not

Manuscript 106 submitted to 2023 IEEE Symposium Series on Computational Intelligence. Received July 27, 2023.

Fig. 8: Highlighted spots labeled from (a) to (i) represent areas where $\xi_i \geq 8$. Details about each spot are in Table II.

only through foot-movement detection but also by enhancing conventional methods like the steering entropy method.

Yet, there were locations where contributors to an increase in ξ_i remained elusive, such as those marked green in Fig. 8. While this might initially seem like a drawback, it paradoxically highlights a potential area of exploration. For instance, we observed spot (f) and found that it was occasionally illegally parked. It is possible that such illegal parking may have led to increased pedal activity on an otherwise fast, straight road. By focusing on these locations and re-examining them through front-camera imagery or on-site fieldwork, we can strive to understand better the unidentified factors contributing to cognitive load in drivers.

In addition to the ample number of data points for each spot highlighted in Sec. IV, the range of unique users at these locations—typically between 10 and 30, according to Table II—suggests that our results possess a degree of generalizability. The high average ξ_i values observed at these spots likely indicate a typical response to specific road conditions or driving scenarios among the subjects. However, our participant pool was limited to young male students with less driving experience. Expanding this pool to include a broader range of demographics could offer further insights into the impact of cognitive capacity on driving behavior, thereby paving the way for future research.

TABLE II

Spot	n_i	u_i	m_i^{Foot}	m_i^{CAN}	μ_i^{Foot}	μ_i^{CAN}	ξ_i
(a)	72	20	14	1	0.19	0.01	14.0
(b)	86	26	9	1	0.10	0.01	9.0
(c)	58	24	12	1	0.21	0.01	12.0
(d)	18	6	11	1	0.64	0.05	11.0
(e)	48	19	9	1	0.18	0.02	9.0
(f)	50	21	9	1	0.17	0.01	9.0
(g)	39	13	16	1	0.41	0.02	16.0
(h)	48	13	16	1	0.33	0.02	16.0
(i)	48	14	8	1	0.17	0.02	8.0

Details of extracted spots. μ_i values are rounded to two decimal places.

VI. Conclusion

In this study, we endeavored to elucidate the drivers' unsafe states induced by cognitive load. We utilized foot camera footage to track pedal changes and correlated this with the frequency of pedal depressions recorded in the CAN bus data. Our analysis successfully revealed areas of high pedal change frequency, indicating elevated cognitive load, even when these areas did not correspond to pedal operations captured by the CAN bus system. Significantly, we developed a practical map displaying the spatial distribution of cognitive load, as estimated through our foot placement analysis. The map was found to align well with our local knowledge and intuitively interpretable scenarios, thus assisting in identifying potential high-risk zones and contributing to efforts to enhance driving safety.

Acknowledgment

We gratefully acknowledge the substantial support provided by ALPS ALPINE CO., LTD. for the university share-car experiment. We also appreciate the developers of kepler.gl, used for explorative geographic data analysis in our research.

References

[1] C. Aichinger, P. Nitsche, R. Stütz, and M. Harnisch, "Using low-cost smartphone sensor data for locating crash risk spots in a road network," *Transportation research procedia*, vol. 14, pp. 2015–2024, 2016.

[2] K. INAGAKI, T. MARUNO, and K. YAMAMOTO, "Evaluation of eeg activation pattern on the experience of visual perception in the driving," *IEICE Transactions on Information and Systems*, vol. E103.D, no. 9, pp. 2032–2034, 2020.

[3] L. Xie, C. Wu, M. Duan, and N. Lyu, "Analysis of freeway safety influencing factors on driving workload and performance based on the gray correlation method," *Journal of advanced transportation*, vol. 2021, pp. 1–11, 2021.

[4] O. Nakayama, T. Futami, T. Nakamura, and E. R. Boer, "Development of a steering entropy method for evaluating driver workload," *SAE transactions*, pp. 1686–1695, 1999.

[5] Y. Chen, C. Hu, and J. Wang, "Human-centered trajectory tracking control for autonomous vehicles with driver cut-in behavior prediction," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8461–8471, 2019.

[6] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on intelligent transportation systems*, vol. 6, no. 2, pp. 156–166, 2005.

[7] Q. Chen, X. Song, H. Yamada, and R. Shibasaki, "Learning deep representation from big and heterogeneous data for traffic accident inference," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, no. 1, 2016.

[8] T. Lee, S. Matsushima, and K. Yamanishi, "Traffic risk mining using partially ordered non-negative matrix factorization," in *2016 IEEE international conference on data science and advanced analytics (DSAA)*. IEEE, 2016, pp. 622–631.

[9] S. He, M. A. Sadeghi, S. Chawla, M. Alizadeh, H. Balakrishnan, and S. Madden, "Inferring high-resolution traffic accident risk maps based on satellite imagery and gps trajectories," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 11 977–11 985.

[10] G. Vandenbulcke, I. Thomas, and L. I. Panis, "Predicting cycling accident risk in brussels: a spatial case–control approach," *Accident Analysis & Prevention*, vol. 62, pp. 341–357, 2014.

[11] C. Tran, A. Doshi, and M. M. Trivedi, "Pedal error prediction by driver foot gesture analysis: A vision-based inquiry," in *2011 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2011, pp. 577–582.

[12] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

Manuscript 106 submitted to 2023 IEEE Symposium Series on Computational Intelligence. Received July 27, 2023.