Exploring Eye Movement Patterns and Driver Cognitive State in Partially Autonomous Driving Simulation

Yifan Wang[a], Shixiao Wang[b], Zhiwu Dong[a,b], Chenlei Liao[b], Xiqun (Michael) Chen[b,*], Der-Horng Lee[a,*]

[a] Zhejiang University/University of Illinois Urbana-Champaign (ZJU-UIUC) Institute, Zhejiang University, International Campus, Haining 314400, China

yifan3.23@intl.zju.edu.cn, zhiwu@zju.edu.cn, dhlee@intl.zju.edu.cn

[b] Institute of Intelligent Transportation Systems, College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, China

wangshixiao@zju.edu.cn, liaochenlei@zju.edu.cn, chenxiqun@zju.edu.cn

* Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)

September 02-03, 2024, Crete, Greece

April 29, 2024

Keywords: Eye-tracking, LightGBM, Markov chain, Autonomous driving, Cognitive state.

1 INTRODUCTION

In the rapidly evolving field of autonomous vehicles, the imperative to discern the congnitive state of drivers by monitoring their physiological data has become increasingly apparent (Reimer and Mehler, 2011). Among a range of physiological indicators, eye movement data, particularly pupil diameter, directly reflects a driver's focus and is utilized to evaluate the psychological load during the driving process (Radhakrishnan et al., 2023). Bitkina et al. (2021) demonstrated that pupil diameter and X-axis glance fixation were the most relevant metrics for assessing driver congnitive state. Zhou et al. (2022) built a machine learning model for predicting situational awareness via light gradient boosting machine (LightGBM) using only eye-tracking data. Although research on eye movements and driver states has been well established, in-depth eye movement studies of drivers during autonomous driving are scarce, and the scenario design and data quality of most of these studies leave much to be desired.

Our research collected eye movement data from 12 drivers under various driving conditions, including non-driving-related tasks (NDRT), light intensity, and driving mode. LightGBM was employed to classify load states, and Shapley Additive exPlanations (SHAP) were utilized to interpret results. The analysis underscored the significant role of fixation point data in determining the load states of the drivers. Consequently, following the clustering of fixation point data, a Markov chain based on fixation regions was established for analysis. This approach facilitated the identification of eye movement state transition patterns under different cognitive conditions.

2 METHOD

2.1 Experimental Design and Data Collection

We are recruiting 12 experienced drivers as participants, who will wear Tobii Pro Glasses 3 to collect eye-related data while driving in a simulator. The simulated environment is an urban road, constructed using CARLA, an open-source autonomous driving simulator (Dosovitskiy et al., 2017), includes variables such as the presence of non-driving-related tasks (NDRT), light intensity, and driving mode. Based on these three variables, six experimental scenarios were designed as shown in Table 1. We require the driver to focus on driving, therefore, no NDRT were set up under manual driving conditions. Notably, the autonomous mode corresponds to Level 2 autonomy, where drivers maintain a driving posture and continuous road awareness. The Twnty Questions Task (TQT) serves as the NDRT for assessing cognitive status.

Table 1 – *Experimental Scenario Design*

Scenario	Driving mode	NDRT	Light intensity
1	Manual	×	Daytime
2	Manual	×	Nighttime
3	Autonomous	×	Daytime
4	Autonomous	×	Nighttime
5	Autonomous	√	Daytime
6	Autonomous	√	Nighttime

Ultimately, we arranged six experimental scenarios in a Latin square design to randomize the order of conditions. The collected raw data undergo preprocessing—including assisted mapping and selection of times of interest—to facilitate effective data analysis and modeling. The nine features reserved and their explanations are presented in Table 2.

Table 2 – *Features Reserved and Their Explanations*

Features	Explanations
ADWF	Average duration of whole fixations
NWF	Number of whole fixations
APD	Average whole-fixation pupil diameter
NA	Number of saccades
APVS	Average peak velocity of saccades
MPVA	Maximum peak velocity of saccades
SDPVA	Standard deviation of peak velocity of saccades
AAS	Average amplitude of saccades
MAS	Maximum amplitude of saccades

2.2 Significant Feature Analysis

In our study, LightGBM is employed to classify distinct categories of experimental variables, such as differentiating between driving conditions with or without NDRT. LightGBM is an advanced ensemble machine learning framework that leverages gradient-boosted decision trees for enhanced predictive accuracy (Zhou et al., 2022). Binary cross-entropy is used as the loss function during the training process. A gradient-based optimization algorithm is employed to minimize the loss function.

SHAP is a method for interpreting machine learning models (Lundberg and Lee, 2017). The SHAP values are calculated for each feature to determine its contribution to the model's classification result.

2.3 Fixation Area Division and Markov Chain Model

The mini-batch K-means algorithm is applied in the research. It is a variant of the K-means clustering algorithm that uses small and random subsets of the dataset for each iteration, improving speed and scalability for large datasets. The optimal number of clusters is determined by comparing three indicators, i.e., Silhouette coefficient, Calinski-Harabasz index, and Davies-Bouldin index.

The Markov chain model is employed to examine the regularities in the transitions of eye movement states. Markov chain is a stochastic model that describes a sequence of possible events in which the probability of each event only depends on the state attained in the previous event.

Define the areas formed by clustering fixation points as the states in a Markov chain, compute the transition probabilities to generate a one-step transition matrix. After a finite number of sequential transitions, the state transition matrix will eventually reach a stable probability distribution. Assuming that $X = \{X_n, n = 0,1,2 \ldots\}$ is a homogeneous Markov chain with state space S and transition probability P_{ij}, there exists a steady-state probability distribution $\{\pi_j, j \in S\}$, if it satisfies:

$$\begin{cases} \pi_j = \sum_{j=1} \pi_j P_{ii} \\ \sum_{j=1} \pi_j = 1, \pi_j \geq 0 \end{cases} \tag{1}$$

The probability distribution $\{\pi_i, j \in S\}$ is referred to as the stationary distribution of the Markov chain.

3 RESULT

3.1 SHAP Result

Figure 1 presents the SHAP results for the LightGBM models, utilizing three variables as independent outputs and nine eye movement features as inputs. Specifically, label 'a' indicates the presence of NDRT, label 'b' pertains to different light intensities, and label 'c' represents various driving modes. In Figure 1-a, ADWF shows significantly higher SHAP values compared to other variables, underscoring the importance of fixation-related information. In Figure 1-b, APD achieves the highest SHAP values mainly due to physiological reactions to brightness variations between daytime and nighttime, followed closely by SDPVA and NWF, which indicate the significance of the standard deviation of peak saccadic velocity and the total number of fixations, respectively. In Figure 1-c, the critical variables are AAS and NWF.

Overall, whether ADWF or NWF, fixation-related information proves to be an extremely important feature for researching driver load states, consistent with the findings of El Khatib et al. (2020).

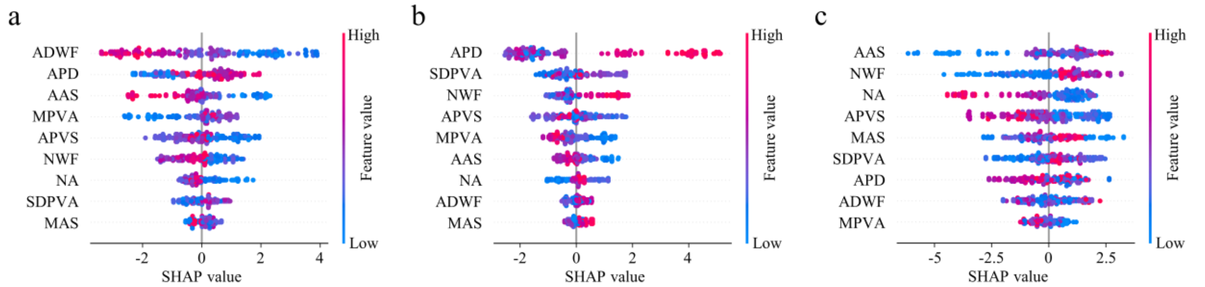

Figure 1 – *SHAP value distributions*

3.2 Clustering Results

Having established the significance of fixation data, mini-batch K-means clustering is used to visualize eye fixation information. Figure 2 illustrates the clustering results of fixation point coordinates, displaying a diverse set of colored dots where each color represents a distinct cluster. Seven clusters, denoted as Areas 1 through 7, are identified; Areas 1, 2, 3, and 4 are centered on the windshield, Area 5 focuses on the vehicle's dashboard, and Areas 6 and 7 correspond to the right and left rearview mirror areas, respectively. The analysis reveals that users' fixation is most frequently centered on the front road and diminishes gradually towards the sides. Fixations on the rearview mirrors are primarily concentrated within the lane markings, indicating strategic monitoring of surrounding traffic.

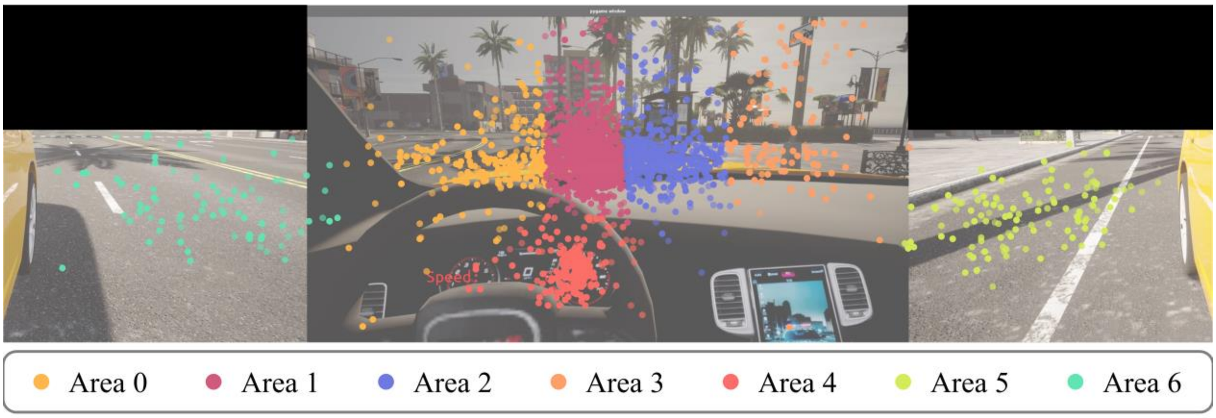

Figure 2 – *Fixation point clustering result*

3.3 State Transition Matrix

Table 2 shows the steady-state probability distribution of fixation points in each area under different driving conditions based on the Markov chain. The following conclusions can be drawn from Table 2:

1. An increase in cognitive load levels (such as with NDRT, nighttime, and manual driving) results in drivers looking at the dashboard (Area 5) more frequently; conversely, drivers tend to focus more on the road ahead (Area 1, 2, 3) under opposite conditions;

2. When cognitive load is high, drivers are more likely to focus on the area directly ahead and the dashboard, whereas under lower cognitive loads, drivers are more likely to look at other areas;

3. The effects of NDRT and light intensity on drivers are similar, but the impact of driving mode is less significant.

Table 2 – *Steady-state Probability Distribution Matrix*

Variables	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7
Without NDRT	0.07	**0.30**	0.17	0.08	0.28	0.07	0.04
With NDRT	0.10	0.24	0.12	0.04	**0.44**	0.03	0.03
Daytime	0.11	**0.32**	0.15	0.07	0.28	0.04	0.03
Nighttime	0.07	0.22	0.13	0.05	**0.43**	0.05	0.04
Autonomous	0.07	**0.30**	0.17	0.08	0.28	0.07	0.04
Manual	0.10	0.32	0.13	0.04	**0.32**	0.05	0.03

Figure 3 shows the state transition matrices under 6 driving scenarios by Markov chain. The number in each cell represents the probability of transitioning from the region on the y-axis to the region on the x-axis. The assignment of matrices across different conditions is as follows: Matrix 'a' corresponds to the non-NDRT condition, matrix 'b' to daytime driving, and matrix 'c' to autonomous driving. Conversely, matrix 'd' is associated with NDRT, matrix 'e' with nighttime conditions, and matrix 'f' with manual driving. The following inference can be drawn:

1. As the cognitive load increased (a to d, b to e, c to f), the driver's fixation preference for the dashboard was confirmed by an increase in the probability of several all areas to the dashboard (Area 5).

2. When the cognitive load was low, drivers were more distracted, mainly in the form of a greater sweep of the front right-side area. This is reflected in a higher probability of moving from Area 3 to Area 4 and from Area 4 to Area 3.

3. The above findings are further confirmed by the fact that drivers pay more attention to the dashboard when cognitive load is high. The presence of NDRT and light intensity had a more significant effect on the driver's fixation, while the presence of autopilot had a lesser effect.

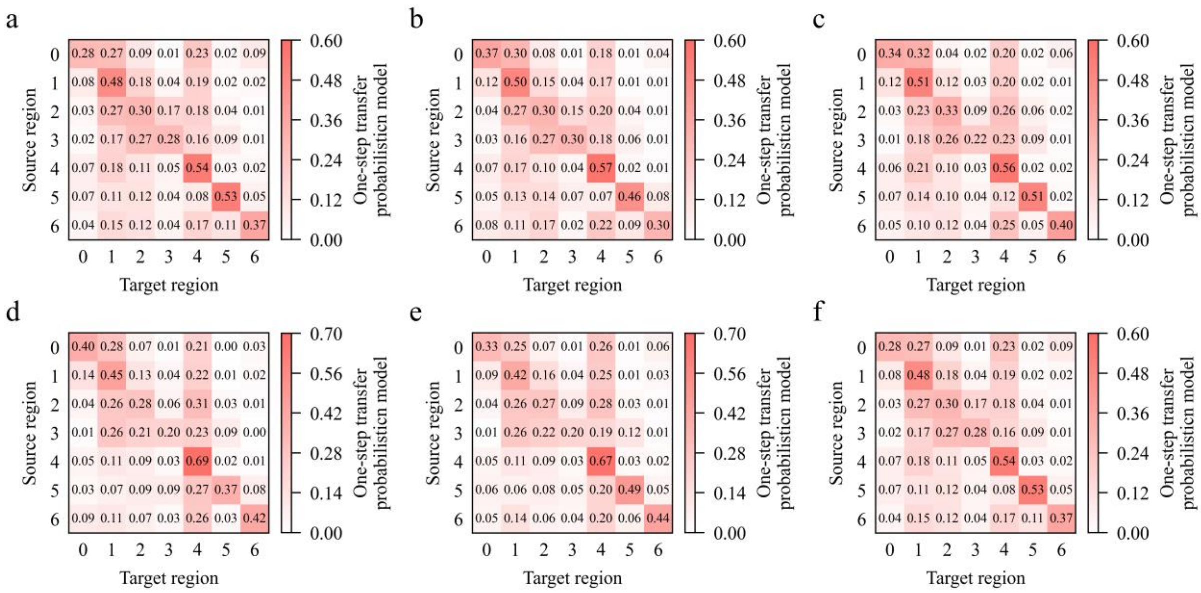

Figure 3 – *State transition matrices under 6 driving scenarios*

4 CONCLUSION

Eye movement data were collected from 12 drivers under various driving conditions and analyzed using LightGBM classification, enhanced by SHAP interpretations. We specifically focused on fixation data, applying mini-batch K-means clustering and Markov chain models to gain a detailed understanding of fixation behaviors under different cognitive loads. By integrating the results from the steady-state probability distribution and the state transition matrices of the fixation points, we can draw several potential conclusions.

Cognitive load levels during driving affect drivers' eye movement characteristics. As cognitive load increases, drivers' attention shifts from the road ahead to the dashboard, providing a rapid method to gauge vehicle safety and operational status. This shift offers drivers a more intuitive overview than observing the road ahead. Conversely, a low cognitive load may lead to distractions, causing their gaze to wander around the forward area. Our study gains novel insights into the dynamics of driver's gaze and attention, thereby contributing to improving safety features and the design of cognitive load management for autonomous vehicles.

This extended abstract may be considered as a foundation for more intricate modeling approaches, we acknowledge the limitations associated with conclusions derived from modeling analyses that utilized data from 12 drivers. Future work will conduct experiments on a larger dataset, considering the inclusion of more physiological data, such as electroencephalogram (EEG) and electrocardiogram (ECG), to form a multimodal dataset for studying driver cognitive load. Additionally, models with stronger feature extraction capabilities, such as deep neural networks, will be applied for modeling.

References

Bitkina, O.Vl., Park, J., Kim, H.K., 2021. The ability of eye-tracking metrics to classify and predict the perceived driving workload. *International Journal of Industrial Ergonomics*, **86**, 103193.

Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., Koltun, V., 2017. CARLA: An open urban driving simulator. *Conference on Robot Learning*, **78**, 1-16.

El Khatib, A., Ou, C., Karray, F., 2020. Driver inattention detection in the context of next-generation autonomous vehicles design: A survey. *IEEE Transactions on Intelligent Transportation Systems*, **21**, 4483-4496.

Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. *In Proceedings of the 31st International Conference on Neural Information Processing Systems*, **30,** 4768-4777.

Radhakrishnan, V., Louw, T., Cirino Gonçalves, R., Torrao, G., Lenné, M.G., Merat, N., 2023. Using pupillometry and gaze-based metrics for understanding drivers' mental workload during automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, **94**, 254-267.

Reimer, B., Mehler, B., 2011. The impact of cognitive workload on physiological arousal in young adult drivers: A field study and simulation validation. *Ergonomics*, **54**, 932-942.

Zhou, F., Yang, X.J., de Winter, J.C.F., 2022. Using eye-tracking data to predict situation awareness in real time during takeover transitions in conditionally automated driving. *IEEE Transactions on Intelligent Transportation Systems*, **23**, 2284-2295.