

Personalizing Driver Agent Using Large Language Models for Driving Safety and Smarter Human-Machine Interactions

Zixuan Xu, Tiantian Chen*, Zilin Huang, Yang Xing, Sikai Chen*

Abstract—Driver assistance systems have been shown to reduce crashes by providing real-time warnings or assistance, with effectiveness depending on the communication with driver. Due to their unique characteristics, human drivers possess varying hazard perception skills and interaction preferences, making personalized assistance crucial to improving user experience and system acceptance. However, how to leverage multimodal interfaces that dynamically adapt to warning contents and driver characteristics remains an open question. At the same time, Large language models (LLMs) have demonstrated advanced capabilities in knowledge acquisition, planning and human-machine collaboration, offering potential solutions for existing warning systems. Thus, we developed an LLM-based personalized driver agent (LLM-PDA), which provides personalized warnings through multimodal interactions (visual, voice, tactile). The agent’s architecture mimics human cognitive processes via four core modules: Memory, Perception, Control, and Action. Results from our experiments indicate that the LLM-PDA effectively customized warning contents for different drivers in various situations, providing enhanced safety and driver support. This study pioneers the integration of LLMs into automotive human-vehicle interaction and offers novel insights into personalized human-machine interaction in intelligent vehicles.

Index Terms—Large language models (LLMs); personalization; human-machine interaction; retrieval augmentation generation; multimodal warnings

I. INTRODUCTION

THE goal of driver assistance systems is to reduce traffic accidents caused by human error by either passively advising the driver of dangerous situations or, in some cases, actively intervening to control the vehicle [1]. With the development of sensors and computing techniques, driver assistance systems have evolved from simply improving safety to making drivers more comfortable and convenient [2]. Despite these advances, existing systems still exhibit limitations that may be addressed by integrating multimodal interaction capabilities and knowledge-driven approaches. Following is an overview of the algorithms and interaction tools used by current driver assistance systems.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (RS-2024-00351865). (Corresponding authors: Tiantian Chen and Sikai Chen)

Z. Xu and T. Chen are with the Cho Chun Shik Graduate School of Mobility, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, South Korea (e-mail: zixuan.xu@kaist.ac.kr; nicole.chen@kaist.ac.kr).

Z. Huang and S. Chen are with the Department of Civil and Environmental Engineering, University of Wisconsin-Madison, WI, USA (e-mail: zilin.huang@wisc.edu; sikai.chen@wisc.edu).

Y. Xing is with the School of Aerospace, Transport and Manufacturing, Cranfield University, MK43 0AL, UK (e-mail: yang.x@cranfield.ac.uk).

Manuscript

A. Knowledge-driven Driver Assistance Systems

The driver assistance systems receive sensor data, process the information, and then determine assistance contents [3]. A rule-driven or model-driven algorithm is most commonly used in these systems [4]. Rule-driven assistance systems operate according to predefined rules, thresholds, and conditions [5]. Rapid response systems, collision warning systems, cut-in warning systems and lane departure warning systems fall into this category. A model-driven driver assistance system, on the other hand, generates driver assistance strategies based on planning models [3]. Through the integration of statistical models and artificial intelligence (AI) techniques (e.g., computer vision [6]), model-driven systems can now provide more complex assistance features, such as driver monitoring systems and in-vehicle voice assistants [7]–[9].

Driver assistance systems that are based on rules and models are designed primarily to perform specific functions in specific scenarios. However, scenario-specific approaches are limited in terms of reasoning and personalization. In particular, real-world traffic environments are dynamic and contain corner cases posing potential hazards that algorithms cannot fully account for because they lack training data [10]. Additionally, human drivers have different hazard perception skills and are equipped with different driving experiences, so the warning contents they need in the same situation may differ greatly [11] [12]. For example, it is important to provide older drivers with more warnings that are relevant to their slower visual processing and impaired peripheral vision, such as sudden crossings by pedestrians. On the other hand, foreign drivers need comprehensive navigation aids, such as real-time translation of road signs and detailed turn-by-turn instructions, to navigate unfamiliar roads safely. Providing drivers with tailored warning content based on their hazard perception skills and experience level can greatly improve the effectiveness of driver assistance systems. The development of more sophisticated approaches to address the complexities of real-world driving and to customize assistance for individual drivers is therefore necessary.

A knowledge-driven approach that integrates generalization, cognitive abilities, and lifelong learning has shown promise in addressing these challenges [13]–[15]. In the field of driver assistance, such methods are still evolving, with earlier efforts focusing on simulating the human brain’s cognitive process, transferring and integrating knowledge (such as domain knowledge and common-sense knowledge). A novel brain-

inspired framework has been developed to simulate the human brain's cognitive process, incorporating perception, memory, attention, and decision-making modules for driver emotion detection [16]. Some studies in visual detection have attempted to transfer knowledge in a teacher-student manner, imparting the teacher's knowledge to the student by re-weighting distillation based on the teacher's loss [17]. For integrating knowledge into the model, several studies have used knowledge graphs to describe entities and their relationships in traffic scenarios, and employed knowledge-infused learning to predict situations or unobserved entities [18] [19]. Specifically, commonsense reasoning is a knowledge-driven approach to decision-making. As an emulator of the human mind, the system conducts reasoning based on common sense [20]. Nevertheless, there is still a lack of knowledge-driven systems that can interpret scenario reasoning based on individual driver preferences.

B. Multimodal Warnings for Driver Assistance

The effectiveness of warning systems relies on the design of the warning interface [21]. Poorly designed warnings may distract drivers and cause discomfort, while a user-centric approach enhances personalization and acceptability [22]. Drivers interact with warning systems via Human-Machine Interfaces (HMIs). These HMIs often utilize voice, visual, and tactile methods to convey warning information [23]. To enhance perceptibility and responsiveness, warnings are typically delivered through unimodal or multimodal signals [23] [24]. The use of visual warnings, such as icons and symbols, can increase driver awareness by providing detailed illustrations of information [25]. Auditory warnings can effectively attract drivers' attention and alert them to critical information, irrespective of their visual focus [26]. Although tactile warnings (such as vibrations from the seat, seatbelt, or steering wheel) are less common interaction methods, they are more effective when the driver's ability to process visual information is reduced or occupied [27] [28]. These unimodal warnings, however, may have limitations and may be perceived by drivers as not useful or not user-friendly under certain circumstances. For instance, it is not uncommon for drivers in busy traffic to find auditory warnings annoying and to show slower reaction times as a result [29] [30]. Furthermore, unimodal warnings are not equally favored by all drivers due to differences in their cognitive capabilities, physical conditions, and sensory abilities. Those with impaired vision may struggle with visual interfaces, while those with age-related sensory decline may not be able to react to auditory warnings [24] [31].

It has been proposed that multimodal warnings can overcome unimodal warnings' limitations by using more than one mode of interaction. A number of bimodal and trimodal warning types have been developed and applied to driver assistance systems, such as blind spot information, parking assistance, fatigue warnings, and collision warnings [32]–[34]. Studies also found that multimodal in-car warning systems lead to faster reaction times and enhanced traffic awareness than a unimodal system [22], [35], [36]. There is, however, a lack of approaches that could allow for tailoring interaction modalities

across different warnings in these multimodal systems. In addition, it is unclear which multimodal configuration is more effective for different driver groups. Though most advanced warning systems today offer personalization, they remain static based on the user's settings or require additional training, and cannot automatically update during ongoing communication with the driver. To the best of our knowledge, there has been no attempt to implement multimodal interfaces that take into account both warning content and driver characteristics when determining optimal interaction.

C. Problem Statement, Study Contribution, and Organization

In light of the limitations outlined above, there is a significant need for an adaptive and personalized warning system. This system should be able to decide which aspects to warn about and how to communicate with the driver based on the traffic environment (e.g., type of road hazard, level of emergency), while considering the driver's characteristics (e.g., age, driving experience). It has been demonstrated that Large Language Models (LLMs) have impressive capabilities for acquiring knowledge, reasoning, and interacting with humans [37] [38] [39]. Taking advantage of these advancements may pave the way for knowledge-driven driver assistance systems that are more personalized and capable of reasoning more effectively [40]. With the use of internal and external knowledge, LLMs can generate insights for input queries through human-like reasoning. Their ability to comprehend natural language, adapt to a variety of content, and generate humanlike responses further revolutionizes human-machine interaction [41]. Also, LLMs demonstrate expertise in generating personalized strategies, engaging in tailored interactions, and effectively addressing individual user needs and preferences across diverse applications [42] [43]. Vision-Language Models (VLMs) take the abilities of LLMs one step further by integrating visual information along with text. According to recent studies, the GPT 4.0 vision is capable of understanding different traffic scenarios, reasoning on corner cases, and making informed driving decisions [39] [44].

Our understanding of LLMs inspired us to develop an LLM-based personalized driver agent (LLM-PDA) for personalized assistance via multimodal interactions (visual, voice, tactile). The LLM-PDA is designed to determine the warning contents and select the most effective modalities (visual, voice, or tactile) for communication, based on the traffic environment and individual driver profiles. The proposed LLM-PDA aims to enhance driver safety and user-centric interaction by integrating internal and external knowledge, utilizing human-like reasoning, and leveraging multimodal tools. To our knowledge, this is the first study to apply LLMs to multimodal human-machine interaction for intelligent vehicles. The contributions of our work can be summed up as follows:

- We developed an innovative agent framework that allows LLMs to create personalized interaction strategies. Through this approach, not only does the system's operation become more understandable and expandable, but it also facilitates the implementation of a user-in-the-loop feedback loop.

- We proposed explainable knowledge-driven approaches to improve agent’s understanding, reasoning, and decision-making capabilities. Specially, a novel retrieval augmentation generation (RAG) method was developed for external domain knowledge learning. Through in-context learning, we invoke LLMs’ knowledge-driven reasoning capabilities to generate driver assistance content.
- Our experiment results show that LLM-PDA can generate explainable personalization profiles and assistance strategies. Across different scenarios, the LLM-PDA can accommodate the specific interaction requirements of different drivers.

This article is organized as follows: In Section I (Introduction), we provide an overview of knowledge-driven driver assistance systems, discuss the importance of multimodal warnings, and outline our contributions to the field. Section II (Related work) examines current applications of large language models (LLMs) and agent systems in intelligent vehicles, highlighting the potential of utilizing LLMs for personalized driver assistance. In Section III (Methodology), we present the proposed LLM-based Personalized Driver Agent (LLM-PDA). Section IV (Experiments: Case study with driver profiles) section details the case study setup and demonstrates the performance of the LLM-PDA. Section V (Discussion) discusses the limitations of the current work and explores future opportunities for LLMs in personalized human-machine interaction within intelligent vehicles. Finally, Section VI (Conclusion) summarizes our findings and suggests directions for future research.

II. RELATED WORK

A. Large Language Model-based AI Agent

AI agents are artificial entities that can sense their environment, make informed decisions, and take action in response [45]. Researchers have used LLMs as an AI agent’s brain to accomplish tasks through self-directed planning and actions since they demonstrate powerful knowledge acquisition capabilities, planning, and reasoning [46] [47]. LLM-based agents adopt different architectures depending on their functional scope and task execution requirements. Among them, the ‘Perception, Memory, Control, Action’ structure represents a type of LLM-based agent that simulates human interaction with multimodal environments, performing tasks while continuously learning [45] [48]. In this framework, the Perception module processes multimodal input from the environment, the Memory module stores data for retrieval and adaptation, the Control module acts as the ‘Brain’ for information processing and decision-making, and the Action module executes commands using integrated tools. With some advanced tools, such as LangChain, developers are able to chain together components from different modules and design interaction flows for LLM-based agents [4]. Recently, LLM-based autonomous agents have demonstrated expert-level performance in a wide range of fields, such as recommendation systems, embodied AI and intelligent monitoring [50]–[52]. For example, Pu et al. (2024) revealed the efficiency and

effectiveness of LLMs in intelligent monitoring by developing AutoRepo, a framework that combines multimodal LLMs with unmanned vehicles to generate automated inspection reports with expert-level accuracy while reducing operational time significantly [53]. Additionally, LLM-based agents will revolutionize how personalized systems generate content and interact with humans. These agents are dynamic systems that explore and reason about user profiles and requests, and then offer interactive and explainable answers. Enhancing personalization awareness of LLMs can be achieved through comprehensive prompts, memory augmentation, and user-oriented fine-tuning [54] [55]. Some research also integrates multiple LLMs into an agent to enable it to process and understand multimodal data [56]. In addition, conversational multi-agent systems, such as AutoGen, allow agents to plan, communicate, and cooperate to accomplish complex tasks effectively [57]. As a result of these advancements, LLM-based agents have evolved from single-function to sophisticated and cooperative systems, allowing them to be used in a wider range of applications.

B. Large Language Model-based Agent for Intelligent Vehicle

Enhanced commonsense knowledge and reasoning can significantly benefit a variety of human tasks, particularly driving-related activities such as driving assistance and autonomous driving. While machine learning algorithms, such as object detection and semantic segmentation, are excellent at observing and understanding the environment, explainable and interpretable reasoning is better at decision-making [58]. Additionally, while reinforcement learning (RL) and search-based methods are traditionally employed in autonomous driving for their efficiency in dynamic decision-making and path planning, LLMs uniquely incorporate commonsense reasoning to further enhance and optimize decision-making processes [59]. In recent studies, GPT 4.0-vision was demonstrated to be capable of understanding different traffic scenarios, reasoning on corner cases, and making informed decisions in driving situations [44]. Furthermore, prompt engineering [59] [60], fine-tuned pre-trained models [61]–[63], and driving world models [64] can reinforce reasoning abilities. For instance, DiLu proposed a knowledge-driven autonomous agent that can learn from experience and evolve over time [65]. Drivevllm explored the VLM for enhanced scene analysis, hierarchical planning, and integration into realistic autonomous driving systems [66].

Another application of LLMs is in hazard identification and providing recommendations across various driving scenarios for intelligent driving. This capability facilitates the integration of LLMs into Level 2 to Level 4 autonomous vehicles, potentially leading to broader and more immediate real-world deployment. One notable development is HazardVLM, a task-specific language model designed for describing hazards in real-time [67]. For driver assistance, a study proposed the idea of using ChatGPT as a vehicle co-pilot that comprehends human intentions and adapts vehicle operations accordingly [68]. With large-scale visual-textual pretraining, such as CLIP, vision-language models can be used to identify distracted driving behavior from naturalistic videos [69]. Another study

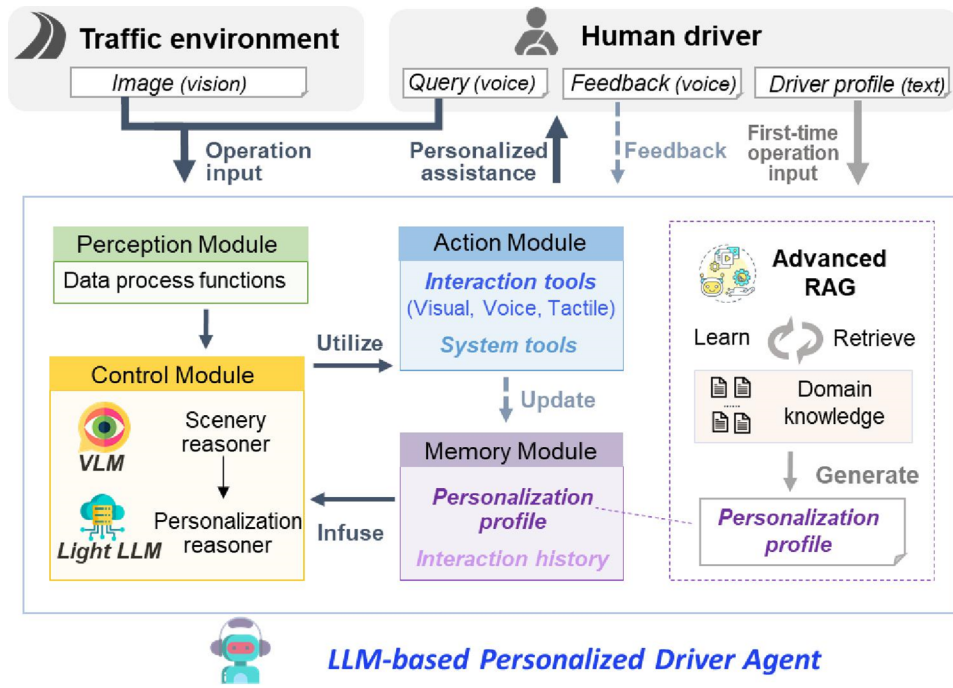


Fig. 1. Overview of the proposed LLM-PDA

introduced an end-to-end system named AccidentGPT, which offers safety warnings and driving recommendations to drivers through human-machine dialogue in a V2X environment [70]. Even though these systems can be applied in a variety of traffic scenarios, they cannot customize warning content for individual drivers. As previously noted, drivers' behaviors and characteristics can differ substantially, which necessitates different warnings and interactions tailored to their specific needs within the same traffic scenario. To fill this research gap, we employ LLMs to develop a personalized driver assistance agent (LLM-PDA), which can deliver personalized warnings and suggestions based on user profiles through a wide range of interaction tools.

III. METHODOLOGY

A. Overview

LLM-PDA interacts with the traffic environment and the human driver. The proposed LLM-PDA consists of four modules to provide personalized driver assistance: Memory, Perception, Control (the "Brain"), and Action modules. Without fine-tuning LLMs, complex reasoning requires considerable time from the driver agent. To ensure efficient response, drivers' personalization profiles are initially stored in the memory module upon first launching the system. These personalization profiles are initialized based on the driver's information and domain knowledge, encompassing feature analysis, detailed settings for three interaction methods, and other safety considerations. Using this personalization profile, the "Brain" will be guided to perform tasks. As the agent operates, the perception module receives images of traffic scenarios and text data from human drivers. To generate a scenario description, the control module reasoning will be carried out by a VLM. Based

on the personalization profile and the scenario description, a light LLM will generate personalized assistance contents and interaction strategies through step-by-step reasoning. Then, the control module will remind or warn the driver using the interaction tools of the action module. The next section will provide details on how these modules are implemented in LLM-PDA.

B. Memory Module

A Memory module contains two types of memory. Short-term memory stores temporary interaction histories, while long-term memory saves personalization profiles. In LLM-PDA, only long-term memory is actively used for personalized functionality, whereas the system does not employ a continuous learning mechanism utilizing short-term memory data. As mentioned above, the personalization profile for the driver is initially created when the system is operated for the first time. In this process, a knowledge-driven method, Advanced Retrieval-Augmented Generation (RAG), is applied to learn from external domain knowledge and mitigate hallucinations, thereby generating reliable and informative personalization profiles tailored to individual users.

The application of Large Language Models to domain-specific or knowledge-intensive tasks outside of their training data may result in erroneous outputs, commonly referred to as hallucinations [71]. The RAG technique is proposed as a way to mitigate this challenge by retrieving relevant information from a knowledge database and utilizing it as a reference to generate more accurate responses to LLM queries [72]. Several studies have demonstrated that RAG reduces LLM's hallucinations and enhances its adaptability to real-world applications, such as ChatPDF [73]. In general, RAG

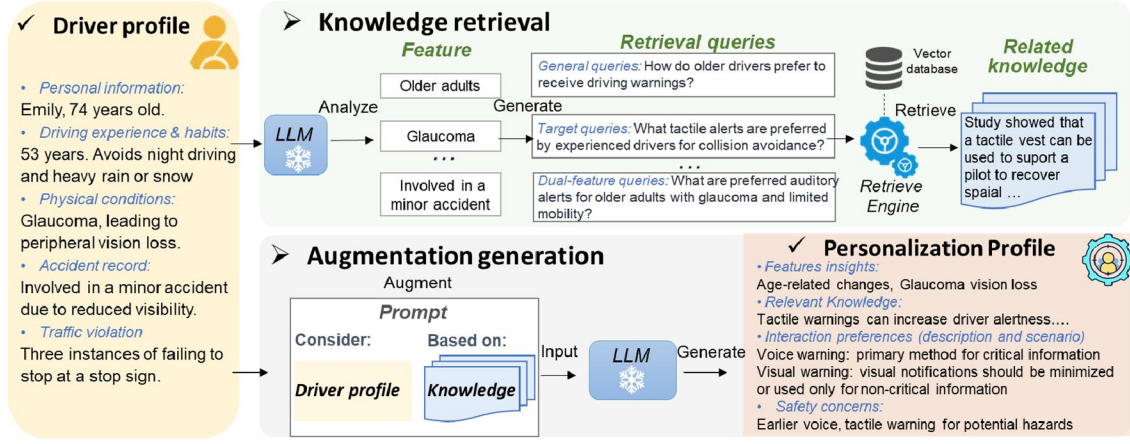


Fig. 2. An illustration of proposed MultiQuery-RAG

embeds input data in a vector database, retrieves its most relevant elements, and generates augmented outputs based on these elements. However, RAGs can sometimes be naive and retrieve misaligned or irrelevant knowledge. Therefore, an advanced RAG is proposed with the objective of improving the retrieval process and, as a result, the performance of the system. In Fig. 2, we illustrate our enhanced MultiQuery-RAG which makes LLM more effective in providing driver assistance and provides a better assessment of the needs and preferences of the driver. It is powered by our advanced query planning RAG and MultiHop-RAG [74].

The key idea of the feature-centered MultiQuery-RAG is to retrieve domain-specific knowledge related to the driver's feature (knowledge retrieval) and input this knowledge to LLM to generate a response (augmentation generation). The knowledge database consists of 84 peer-reviewed articles selected from academic literature based on relevance to driver interaction design, multimodal HMI, and safety considerations. The content was cleaned, embedded using *HuggingFace Embeddings*, and stored in a vector database for efficient retrieval. The input driver profile, denoted as $D_{profile}$, contains the basic information of the driver: personal information, driving experience and habits, physical conditions, and accident records. Based on the input driver profile, the first LLM (LLM_a), Llama 3.1 70B, analyses the features that influence the driving behavior and interaction preference (e.g., age, physical condition). To efficiently retrieve the knowledge related to driver's unique features, three types of queries are generated, represented as $Q = \{q_1, q_2, \dots, q_n\}$. General queries aim to understand broad interaction preferences for one specific feature. Target queries focus on one specific interaction method and one driver feature to extract detailed preferences. Additionally, dual-feature queries combine two features of the driver to identify more complex interaction preferences for driver assistance. For each query q , the retrieval engine searches the vector database to find the top- k most relevant knowledge contents. The de-duplicated knowledge is represented as $K = \{k_1, k_2, \dots, k_n\}$. Next, the knowledge, driver profile and prompts are combined and fed into another LLM (LLM_b), OpenAI-4o which reasons and generates the

final personalization profile ($P_{profile}$). The personalization profile contains analysis of the driver's feature, related external knowledge, interaction preference (description and application scenarios for visual, voice, and tactile interaction), and other safety concerns. The knowledge retrieval and augmentation generation process of multi-query RAG can be represented as in (1)(2):

$$LLM_a(D_{profile}, prompt_1) \rightarrow Q \rightarrow R(Q) \rightarrow K \quad (1)$$

$$LLM_b(D_{profile}, K, prompt_2) \rightarrow P_{profile} \quad (2)$$

C. Perception Module and Control Module

Perception module receives and processes input from a variety of sources, including audio queries from the driver and visual images of the surrounding environment. Specifically, it includes image processing and voice-to-text functions. As soon as the data has been processed, it will be input into the control module.

The control module is responsible for reasoning and decision-making for personalized assistance. Fig. 3 illustrates the reasoning and functionality in detail. We combined the processed scenario image and prompt and input them into VLM. The VLM analyzes traffic elements and predicts road users' behavior, thus generating a scenario description. This study uses LLaVA as the VLM due to its robust performance in complex image understanding, which is comparable to that of the GPT-4 vision. Additionally, LLaVA can run locally, ensuring enhanced privacy and security.

In the next step, a light LLM (Llama 3.1 70B, Groq) will reason and generate personalized assistance content and interaction strategies based on the scenario description and infused personalization profile. Chain-of-Thought (CoT) prompting is a structured approach that guides LLMs through a sequential reasoning process, enabling them to break down complex tasks into smaller, logical steps [76]. This method has demonstrated its effectiveness in intelligent transportation systems and vehicle-related applications, where structured reasoning helps LLMs interpret dynamic driving environments, assess situational factors, and generate contextual responses [65]

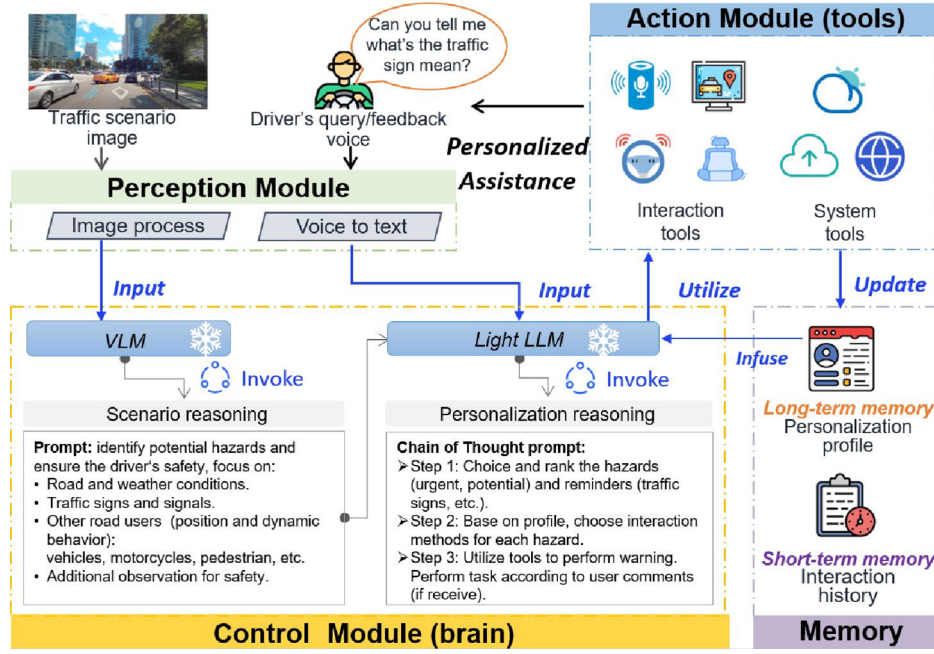


Fig. 3. Detailed functionality and reasoning of LLM-PDA



Fig. 4. Predefined tools in action module

[77] [78]. Inspired by zero-shot CoT [79], our CoT prompts are strategically designed to improve the LLM's multi-step reasoning for personalized driver assistance. The reasoning process begins with ranking potential hazards and reminders based on both the current driving scenario and the driver's profile. The model then analyzes the driver's interaction preferences to determine the most suitable communication modality for delivering these warnings. Afterward, the control module coordinates with the action module to provide personalized assistance to the driver. Additionally, driver-issued queries and feedback are also integrated into the LLM's decision-making process through a conditional execution branch ("if-receive"), allowing for dynamic response generation and tool activation. Through this mechanism, the LLM filters out irrelevant inputs, ensuring that only contextually meaningful interactions influence the system's behavior.

D. Action Module

The Action module contains a set of predefined tools that can be utilized by the control module to perform actions. Each tool contains a description, input parameters (Args), and execution functions. The brain, control model will decide which tools to use based on the tool description. When the tool is used, the tool name and input will be recorded in the log. As shown in Fig. 4, our LLM-PDA defines three simplified interaction tools for driver assistance action, including visual warning, voice warning, and tactile warning. In the case of the visual warning tool, the type of warning is the input parameter. Thus, a database of traffic signs and driving warning images was created. Using the tool input, the visual tool automatically finds and displays the corresponding image. In the voice warning tool, the input parameter is the message, and the voice warning is realized through the text-to-speech API of Elevenlabs. Tactile warnings require input parameters for warning type and intensity. Based on their

TABLE I
SAMPLE DRIVER PROFILES FOR CASE STUDY

	Driver #1	Driver #2	Driver #3	Driver #4
Personal information	Linda, 48 years old, estate agent.	George, 68 years old, doctor.	Thomas, 55 years old, worker.	Alex, 28 years old, graduate student visiting Jeju island.
Driving experience and habits	6 months. Anxious in heavy traffic. Nervous with honking; occasionally has difficulty starting the car.	25 years, drives 50 minutes daily for commuting. Adopts a cautious driving style.	6 years, drives 1 hour daily for commuting and family. Prefers aggressive driving style.	4 years in China, first time in South Korea. Overly cautious; struggles with traffic patterns; GPS-reliant but confused by information.
Physical conditions	Menopausal symptoms worsen after stressful driving, causing headaches and nausea.	Hearing impairment, with half-deafness in the left ear.	Poor vision (severe myopia), with eye injury from accident. Herniated disc, missing 3 toes on left foot. Occasionally has uncontrollable hand tremors.	Healthy, but anxious in unfamiliar driving situations.
Accident and violation record	No accidents; multiple records of diverse traffic violations.	One minor accident, rear-ended a car in stop-and-go; one instance of illegal parking.	No accidents; fined for running red lights.	No accidents; one near-miss at roundabout; fined for speeding on coastal road.

purpose, we defined 10 types of tactile warnings, such as collision warnings and valuable road user warnings. Aside from interaction tools, we set up system tools that provide features such as profile updates, web browsing, and weather searches. When operating the agent, the driver can provide oral feedback or commands, such as adjusting the intensity or frequency of interaction methods or warning contents. If this feedback is detected, the control module will utilize the profile update tool to update the personalization profile in the memory module. The tool operates through an LLM-based reasoning chain that identifies and updates the interaction settings of the personalization profile. Through this mechanism, the agent preliminarily realizes human-in-the-loop integration, enabling continuous and adaptive interaction with humans.

IV. EXPERIMENTS: CASE STUDY WITH DRIVER PROFILES

We utilized LangChain to integrate different modules and build the LLM-PDA. The pre-trained LLMs in LLM-PDA were accessed via APIs using LangChain’s ChatModel class, which provides a standardized interface to various models, such as those from OpenAI, Groq, and Ollama. In this research, we concentrated on the ability of pre-trained LLMs to generate personalized assistance content through in-context learning. Consequently, our experiments were designed to assess the system’s level of personalization and robustness across various drivers and scenarios, rather than evaluating the accuracy of the vision-language model. We collected 40 profiles of real drivers with diverse driving experiences, physical conditions, and accident histories through an online questionnaire. The questionnaire is provided in the appendix. In this section, we present case studies - using four of the samples, to demonstrate the results of LLM-PDA for personalized warnings (other examples are available on our project’s website). We first initialized their personalization profiles through the MultiQuery-RAG approach, and then tested LLM-PDA performance under different scenarios.

A. Personalization Generation

The input driver profile is provided in Table I, while the initialized personalization profile is presented in Fig. 5. Our results show that the MultiQuery-RAG approach can pinpoint drivers’ key characteristics, generate customized interaction strategies, and highlight safety concerns. By identifying and analyzing key characteristics of drivers, for instance, inexperience and anxiety in Driver #1, or severe myopia and hand tremors in Driver #3, the system analyzes how those factors affect driving behavior and the interactions between humans and vehicles. Based on this feature-centric approach, multimodal interaction strategies are tailored to meet each driver’s specific needs. These interaction strategies illustrate how the three interaction methods and their primary sequence are applied in warning systems. For example, in order to alleviate Driver #1’s anxiety and inexperience, clear, non-intrusive visual and voice warnings are given, as well as tactile feedback sparingly to avoid causing additional stress. Driver #2’s hearing impairment is also considered by prioritizing visual and tactile methods, ensuring George can receive warnings for critical scenarios despite his auditory limitations. For Driver #3, who experiences vision impairments and hand tremors, auditory interaction is the primary means of receiving essential information, with tactile feedback acting as a supplement. As Driver #4 is unfamiliar with local traffic and exhibits cautious behavior. The system’s visual and voice instructions will guide him through unfamiliar situations, providing clear, culturally relevant information.

Additionally, personalization profiles present some additional safety concerns. For instance, as Driver #3 suffers from uncontrollable hand tremors, he may benefit from specialized emergency features, as well as enhanced blind-spot monitoring to address vision issues. Driver #4, who has been fined for speeding and is unfamiliar with local roads as a visitor, can be helped with speed recommendations based on local traffic conditions. Based on the initialized personalization profiles, the proposed system will be able to tailor the driver-vehicle interaction to suit the unique needs and characteristics of each

Personalization profile				
	Driver #1	Driver #2	Driver #3	Driver #4
Features insights	<ul style="list-style-type: none"> Inexperienced Headaches and nausea Multiple traffic violations Anxious about driving 	<ul style="list-style-type: none"> Age-related changes Experienced driver Hearing impairment Accident record 	<ul style="list-style-type: none"> Driving affected by physical conditions: severe myopia, hand tremors, etc. 	<ul style="list-style-type: none"> Not familiar with local traffic Violation record Cautious style
Related knowledge	Auditory and haptic feedback can be more effective for drivers engaged in secondary tasks. Personalized, easy-to-understand warnings are crucial for drivers with limited experience...	Drivers with hearing loss may struggle with sound localization and detecting warnings. Older drivers may also have reduced sensitivity to touch and vibration, requiring stronger tactile stimuli...	Individuals with severe vision impairment may struggle with visual inputs. Tactile or haptic cues, along with auditory alerts, can better capture attention under varying conditions...	Content that aids in understanding traffic signs and maneuvers, such as merging and changing lanes, can increase his situational awareness and decision-making confidence...
Interaction strategies (Description- <i>D</i> ; Application scenario- <i>S</i> .)	<ul style="list-style-type: none"> Visual: <i>D</i>: most effective for non-urgent reminders and warnings. <i>S</i>: navigation, traffic sign and signal, road shape change. Voice: <i>D</i>: primary method for urgent warnings and driving instructions. <i>S</i>: urgent scenarios, such as sudden cut-in, adjusting speed, merging awareness. Tactile: <i>D</i>: should be used sparingly to avoid exacerbating discomfort. <i>S</i>: immediate attention but not cause stress, such as lane departure warning, detection of vulnerable road users, road shape change. 	<ul style="list-style-type: none"> Visual: <i>D</i>: supplementary methods, considering age and potential distraction. <i>S</i>: non-critical information: traffic sign reminders, road shape change. Voice: <i>D</i>: primary method for non-urgent situations, supplementary for urgent ones. Tailored to hearing capabilities. <i>S</i>: scenarios requiring attention but not emergency. Tactile: <i>D</i>: primary method for critical situations. Can be complemented with visual or auditory alerts. <i>S</i>: vehicle sudden cut-in warnings, merging awareness, lane departure warnings. 	<ul style="list-style-type: none"> Visual: <i>D</i>: limited utility of visual notification. <i>S</i>: may still be useful in non-conflict scenarios like navigation or vehicle status updates. Voice: <i>D</i>: primary method due to visual impairments; easily heard but not overwhelming. <i>S</i>: suitable for a wide range of scenarios like traffic sign reminders, sudden cut-ins, and merging. Tactile: <i>D</i>: vibrations in the seat and steer wheel can be the supplementary method. <i>S</i>: ideal for continuous feedback like lane keeping, detection of vulnerable road users. 	<ul style="list-style-type: none"> Visual: <i>D</i>: clear, and culturally familiar icons and symbols. <i>S</i>: navigation reminders: traffic signs/signals, road shape changes, contextual info (e.g., merging awareness). Voice: <i>D</i>: speech-based instructions; sound-only for general warnings. <i>S</i>: time-critical events: vehicle cut-in, detection of vulnerable road users, adjusting speed; driving instruction. Tactile: <i>D</i>: guides attention to specific hazards. <i>S</i>: specific warning: forward collision warnings, lane departure warning; and attention-directing for situations like vulnerable road user reminder.
Safety concerns	<ul style="list-style-type: none"> Providing clear, non-intrusive instructions and avoiding overwhelming alerts. Implement coaching content to improve driving habits and boost confidence. 	<ul style="list-style-type: none"> Earlier warnings for potential hazards. Customizable auditory alerts tailored to hearing capabilities. Enhance GPS with simplified visuals and right-side-focused audio. 	<ul style="list-style-type: none"> Specialized emergency features that can be easily activated in cases of hand tremors or other physical difficulties. Enhance blind-spot monitoring to compensate for vision issues. 	<ul style="list-style-type: none"> Enhance navigation: pair visual information with audio prompts for better awareness in unfamiliar areas. Provide speed recommendations based on local traffic and speed limits.

Fig. 5. Personalization profiles generated via MultiQuery-RAG (Note: only the key information is presented in this table to save space.)

driver while balancing safety and comfort.

B. Scenario Inference

LLM-PDA was tested using videos from YouTube showing real-life traffic scenarios in Seoul. The agent’s performance and robustness were evaluated using 53 diverse traffic scenarios, representing typical Korean driving conditions during daytime hours. The scenarios encompass a range of traffic situations on different types of roads, as illustrated in Fig. 6. The content analysis section demonstrates how the agent responds to different driving challenges through three specific scenarios. Based on the results of this study, we can conclude that the agent can handle real-world interactions effectively. The system’s ability to achieve personalized assistance is also illustrated in the quantitative analysis section, which presents quantitative results of the system’s performance across 53 different scenarios.

1) *Content analysis*: Fig. 7 presents the LLM-PDA’s response to three specific scenarios, along with an image of each scenario, a simplified description of each scenario generated

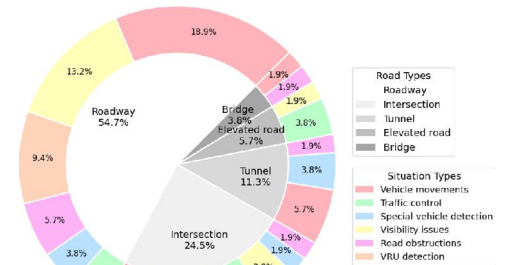


Fig. 6. Distribution of road types and driving situations across testing scenarios

by the Vision Language Model, and explanations of LLM-PDA’s driver assistance actions. As demonstrated through its responses to multiple scenarios, LLM-PDA not only identifies hazards, but also matches warnings with drivers’ individual needs. For example, in the first scenario, involving a bus cut-

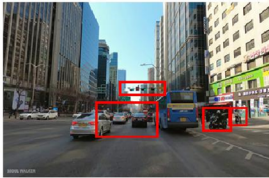
Traffic scenario & description	Driver #1	Driver #2	Driver #3	Driver #4
Scenario 1: Vehicle cut-in				
 <ul style="list-style-type: none"> Urban multi-lane road, traffic is moderate heavy. A red bus: left turn signal on - change lanes. A motorcycle: attempts to weave through traffic. Traffic is heavy, other vehicles brake light flashing, suggesting decelerating. No pedestrians or cyclists visible. 	<p>Hazard 1: bus with turn signal on - <i>Urgent</i> - Voice Warning - {"message": "Caution. Bus ahead cut into your lane."}</p> <p>Reason: The clear and concise voice alert on the bus helps drivers anticipate turns and merges, improving traffic flow and safety.</p> <p>Hazard 2: traffic condition - <i>Potential</i> - Visual Warning - {"type": "keep-safe-headway"}</p> <p>Reason: Visual notifications effectively communicate non-urgent warnings like vehicles decelerating ahead, providing information without being overwhelming.</p>	<p>Hazard 1: bus with turn signal on - <i>Urgent</i> - Tactile Warning - {"type": "sudden-cut-in"}</p> <p>Reason: Tactile feedback immediately alerts George to the bus cutting in, ensuring attention without visual or auditory reliance.</p> <p>Hazard 2: motorcycle - <i>Potential</i> - Voice Warning - {"message": "Cautious about the motorcycle."}</p> <p>Reason: Voice warnings provide clear, adjustable alerts for non-urgent situations, suiting George's hearing capabilities with distinct tones.</p>	<p>Hazard 1: bus with turn signal on - <i>Urgent</i> - Voice Warning - {"message": "Caution. Bus ahead cut into your lane."}</p> <p>Reason: For visually impaired drivers, voice alerts provide essential information about nearby hazards, allowing them to prepare for other vehicles' movement.</p> <p>Hazard 2: motorcycle - <i>Potential</i> - Tactile Warning - {"type": "vulnerable-road-user"}</p> <p>Reason: Tactile feedback is the ideal method for potential hazards that require attention but are not immediately critical.</p>	<p>Hazard 1: motorcycle - <i>Potential</i> - Tactile Warning - {"type": "vulnerable-road-user"}</p> <p>Reason: Tactile alerts effectively draw driver attention to nearby motorcycles without distracting them.</p> <p>Reminder: traffic condition - Visual Warning - {"type": "keep-safe-headway"}</p> <p>Reason: High-contrast visual alerts help the foreign driver quickly comprehend the need to reduce speed, overcoming his unfamiliarity with local traffic conditions.</p>
Scenario 2: Road near the tunnel				
 <ul style="list-style-type: none"> An urban street leading to a tunnel. Early morning or late afternoon. Traffic Signs: 30 km/h speed, left turn. Black sedan in the right lane: merge into the lane. Tunnels: bright lights at the entrance. 	<p>Hazard 1: black sedan merging - <i>Urgent</i> - Tactile Warning - {"type": "collision-prevention"}</p> <p>Reason: Tactile feedback with high intensity effectively draws the driver's attention to a hazard.</p> <p>Hazard 2: tunnel - <i>Potential</i> - Visual Warning - {"type": "tunnel-attention"}</p> <p>Reason: Visual notifications suit Linda's profile for non-urgent warnings, alerting her to approach tunnel without demanding immediate action.</p>	<p>Hazard 1: black sedan merging - <i>Urgent</i> - Tactile Warning - {"type": "vehicle-cut-in"}</p> <p>Reason: Tactile alerts provide direct, immediate feedback for urgent vehicle cut-ins, aligning with George's profile preference for critical situations.</p> <p>Hazard 2: tunnel - <i>Potential</i> - Voice Warning - {"message": "Stay cautious around the bend and the tunnel."}</p> <p>Reason: Voice alerts proactively warn George of approaching tunnels and visibility changes, suiting non-urgent situations with effective reminders.</p>	<p>Hazard 1: black sedan merging - <i>Urgent</i> - Voice Warning - {"message": "Caution, vehicle merging from the right."}</p> <p>Reason: Voice warnings promptly alert myopic Thomas to merging vehicles, aligning with his auditory alert preference for critical situations.</p> <p>Hazard 2: tunnel - <i>Potential</i> - Tactile Warning - {"message": "cautious-driving"}</p> <p>Reason: Tactile feedback at a moderate intensity provides a gentle, secondary alert to the vision-impaired driver in approaching the tunnel.</p>	<p>Hazard 1: bend and tunnel - <i>Potential</i> - Voice Warning - {"message": "Stay cautious around the bend and within the tunnel."}</p> <p>Reason: A verbal alert informs the unfamiliar driver to stay cautious due to reduced visibility around the bend and in the tunnel, providing necessary guidance.</p> <p>Reminder: speed limit sign - Visual Warning - {"type": "30-speed-limit"}</p> <p>Reason: Visual high-contrast icons help the foreign driver quickly comprehend the need to reduce speed, overcoming unfamiliarity with local traffic conditions.</p>
Scenario 3: Intersection crossing				
 <ul style="list-style-type: none"> Multi-lane road. Traffic signal green - vehicle can cross. Traffic sing: no -left turn. Vehicles: silver sedan, blue bus, black sedan- shows brake lights- slowing down. Motorcycle- road side. Pedestrian- road side. 	<p>Hazard 1: motorcycle - <i>Urgent</i> - Voice Warning - {"message": "Be cautious of motorcycle."}</p> <p>Reason: Auditory alerts quickly draw Linda's attention to risks from motorcyclists, aligning with her profile for urgent scenarios.</p> <p>Hazard 2: vehicles ahead decelerating - <i>Potential</i> - Tactile Warning - {"type": "keep-safe-headway"}</p> <p>Reason: Tactile alerts effectively inform Linda of non-urgent but immediate situations, ensuring awareness without causing stress.</p>	<p>Hazard 1: motorcycle - <i>Urgent</i> - Tactile Warning - {"type": "vulnerable-road-user"}</p> <p>Reason: Tactile feedback provides clear, directional vibrations to alert George to motorcyclist presence, enabling immediate awareness.</p> <p>Hazard 2: vehicles ahead decelerating - <i>Potential</i> - Voice Warning - {"message": "Vehicles ahead are slowing down. Maintain a safe distance." }</p> <p>Reason: Voice warning effectively draw George's attention to non-urgent traffic flow changes, ensuring clear and distinct hazard communication.</p>	<p>Hazard 1: motorcycle - <i>Urgent</i> - Voice Warning - {"message": "Be cautious of motorcycle."}</p> <p>Reason: Voice alerts are the most effective method to warn Thomas about the motorcycle, meeting his need for auditory warnings.</p> <p>Hazard 2: vehicles ahead decelerating - <i>Potential</i> - Tactile Warning - {"message": "keep-safe-headway" }</p> <p>Reason: Tactile feedback is the ideal alert method for notifying her of potential hazards that require attention but are not immediately critical, minimizing the distractions.</p>	<p>Hazard 1: motorcycle - <i>Urgent</i> - Voice Warning - {"message": "Be cautious of motorcycle."}</p> <p>Reason: Auditory alerts are the best option to immediately inform the foreign driver about potential hazards, unfamiliar motorcycle traffic patterns.</p>

Fig. 7. Inference on different scenarios from YouTube driving video

ting into the lane, LLM-PDA provides distinct voice warning for inexperienced driver like Driver #1, and tactile warning for hearing impaired drivers like Driver #2. According to the second scenario, the LLM-PDA warns drivers of approaching

tunnels and vehicle movements. For Driver #1, visual warnings provide non-urgent notifications that are suitable for her preference for clear, visual cues without causing her stress. On the other hand, Driver #3 receives tactile or voice warnings

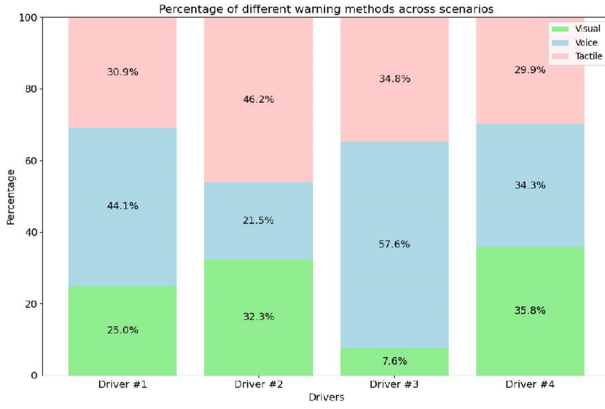


Fig. 8. Percentage of different warning methods across scenarios

based on his reliance on non-visual cues. The strategy shows how the system accommodates drivers' sensory preferences, ensuring that the warnings are supportive without becoming overwhelming. As shown in the Scenario 3, an intersection scenario, the system provides voice alerts for drivers who rely auditory cues to remind them of motors, as well as tactile feedback to drivers who need other types of support. Additionally, the system detects the brake lights of vehicles ahead and warns the driver.

In the examples above, LLM-PDA provides assistance and interacts with users in various ways. Personalization is achieved in the LLM-PDA system by tailoring assistance content and interaction strategies according to user characteristics and driving context. LLM-PDA's capability to improve driving experience and safety is demonstrated by the testing results. Furthermore, the reasons presented in Fig. 6 show that LLM-PDA's decision-making is both explainable and interpretable since it considers both the personalization profile and the current traffic situation. In this respect, our proposed LLM-PDA provides knowledge-driven and interpretable driver assistance and interaction.

2) *Quantitative analysis*: Fig. 8 shows the proportion of voice, visual, and tactile warnings or reminders that drivers receive across 53 traffic scenarios. For Driver #1, who is inexperienced and anxious, the system prioritizes voice warnings (44.1%) for critical situations, complemented by visual (30.9%) and tactile (25%) cues for other warnings. This distribution ensures that Driver #1 is not overwhelmed while still receiving crucial information. In the case of Driver #2, an older with hearing impairments, tactile feedback (46.2%) is the primary modality in critical situations, effectively capturing the driver's attention and compensating for voice limitations. Visual (32.3%) and voice (21.5%) feedback supplement the tactile cues, maintaining Driver #2's situational awareness despite sensory limitations. Driver #3 relies on voice cues and suffers from vision impairments. The results show that LLM-PDA significantly increases voice warning use (57.6%) over visual warnings (7.6%). As a secondary alert system, tactile feedback is provided at a relatively high frequency of 34.8%, which is particularly useful during complex driving situations such as navigating intersections or adverse weather

conditions. This enhances situational awareness without relying on degraded vision. For foreign Driver #4, unfamiliar with the local driving environment, the system provides all types of warnings, visual (29.9%), voice (34.3%), and tactile (35.8%) to support navigation and adapt to changing traffic conditions. Visual warnings help him recognize unfamiliar signs and signals, while voice warnings provide guidance in traffic situations. Tactile feedback adds an additional safety cue without overwhelming him. Overall, across 53 scenarios, the generated interaction strategies match the personalization profiles, demonstrating the robustness of the system. It is also evident from the distribution of different interaction modes that our LLM-PDA is customized to meet the needs and preferences of diverse users, enhancing driving safety and comfort.

V. DISCUSSION

A. Work Limitations

The proposed LLM-PDA can serve as an instructional co-pilot for novice and international drivers, providing real-time feedback and guidance on traffic regulations. With user-friendly interfaces and personalized settings, it facilitates a smooth transition to unfamiliar driving environments. Those with physical impairments and sensory decline can also benefit from the system's customized multimodal interaction [24]. This proposed agent, for example, can provide tactile alerts through the seat or steering wheel, which is especially beneficial to older drivers with hearing and vision challenges. Additionally, we will include fatigue and distraction detection features in the system, which will provide essential support to professional drivers who often perform demanding tasks behind the wheel [75]. Overall, different types of drivers can use this LLM-PDA, which caters to their unique challenges and needs. There is more explanation and intuitiveness in the assistance provided. This focus on individual needs is expected to improve driving safety and the overall driving experience.

Despite the LLM-PDA's ability to provide personalized interaction with drivers, it still has several limitations. Firstly, the LLM-PDA simplifies interaction types. The settings for driver-vehicle interactions, however, require more specific configurations and mechanisms. For example, in addition to providing warnings and guidance through textual information, voice assistance can be delivered in the form of pulsed tones or intermittent audio signals of various lengths and intensities. The intensity and duration of tactile alerts targeting specific areas (such as the steering wheel and seat) are also important characteristics. In the near future, AI-generated content (AIGC) technologies, such as Sora from OpenAI, will enable the creation of personalized vision content for human-machine interactions. Our future work will involve a more detailed analysis of HMIs and the integration of the above considerations.

Moreover, the experimental results presented in this paper need to be quantified further. In this study, we assessed the agent's personalization performance and robustness through case studies involving different drivers across a variety of traffic scenarios. However, as a user-centric system, the LLM-PDA should be evaluated from the user's perspective to

verify its effectiveness. Our plan is to gather user feedback through questionnaire surveys and driving simulator studies. Using survey data and driver behavior data, we can gain valuable insight into the effectiveness, usability, and real-world applicability of the system.

B. Future Research Opportunities

As part of our future work, we will develop a personalized driver assistance solution that better accommodates real-world driving scenarios and makes decisions based on multiple data sources, including traffic environment, driving behavior, and driver state. Through the LLM-based agent framework, we can easily enhance the capabilities of the system by adding tools, integrating memory mechanisms, and incorporating reflection functions. There are multiple technologies and methods that offer potential for improving the reasoning accuracy, robustness, and personalization level of LLM-PDA. In particular, we plan to explore approaches such as in-context learning, supervised fine-tuning, and knowledge-infusion in order to improve the system's accuracy and robustness. To enhance reasoning capabilities, models can be fine-tuned with domain-specific datasets like Reason2Drive [80] and DriveLM [81], facilitating structured reasoning patterns and domain-adapted refinements. The use of novel prompt engineering approaches, such as analogous prompting and Memory-of-Thought frameworks, can help LLMs leverage their pre-trained knowledge more effectively and adapt to different scenarios without extensive supervised training [82] [83]. Moreover, knowledge-infusion techniques, notably knowledge graphs (KGs) that represent entities and their interrelations to driving scenarios within a single graph, have been applied to autonomous driving and driving assistance [13]. These KG-based approaches are utilized to comprehend driving scenes, predict object behaviors, and plan motions based on semantic knowledge [84] [85]. With these techniques, future research could explore the integration of LLM's thought reasoning with knowledge-infusion mechanisms during training and inference phases. In order to improve LLM's reasoning, it is possible to develop traffic scene reasoning knowledge graphs (RKGs), which capture and streamline perception, reasoning, and decision-making processes.

The personalization of intelligent systems can be achieved through memory augmentation and data-driven fine-tuning, offering potential for more adaptive responses. Through memory-augmented approaches, agents are enhanced in their ability to meet the needs of their users and become more responsive to their preferences [55]. As needed, language models are able to retrieve information from short-term memory relevant to the ongoing dialogue. Consequently, advanced RAG approaches can facilitate context-aware personalization by obtaining information about recent interactions, driver states, and locations from memory. The long-term memory, on the other hand, is capable of storing user historical data over time. In order to improve the responsiveness of LLMs to different user needs and contexts, we can fine-tune them with specific individual or group data.

For the further deployment of such systems, a comprehensive fail-safe mechanism is imperative to address the inherent

randomness of user commands and potential hallucinations in LLM outputs [68]. Future work should explore adaptive and learning-based fail-safe protocols that continuously refine safety thresholds based on expert feedback and dynamic scenario analysis, thereby further reducing the likelihood of unsafe response in complex driving environments. Furthermore, privacy protection presents a critical challenge to LLM's deployment in automotive systems, including both data collection and operational aspects [86] [87]. The processing of sensitive user information—including demographic data, biometric indicators, and driving behavior patterns—presents substantial privacy risks [88]. Moreover, the real-time data exchange required for driver assistance systems introduces additional risks regarding unauthorized access. These challenges can be addressed by implementing comprehensive privacy-preserving frameworks and rigorous data governance protocols [89]. Such measures are essential to ensure the effective safeguarding of personal data throughout the system's lifecycle.

VI. CONCLUSION

The emerging capabilities of LLMs make them promising solutions for multimodal warning systems that have difficulty adapting to dynamic, real-world driving conditions and diverse drivers. This paper introduced a Large Language Model-based personalized driver agent (LLM-PDA) that provides warnings through multimodal interactions (visual, voice, tactile). Notably, the LLM-PDA operates on a reasoning-based architecture that emulates human cognitive processes through four core modules, capable of interacting with drivers based on the traffic environment and their characteristics. By implementing a MultiQuery-RAG method, we enabled the LLM-PDA to work as an expert in analyzing driver's characteristics. This approach leverages domain-specific knowledge to initialize a personalized driver profile that supports the agent's tailored functionalities. Through its agent-based workflow, the LLM-PDA can perceive the traffic environment, reason for personalized assistance, and issue customized warnings across various driving scenarios. We evaluated the LLM-PDA through case studies, demonstrating that it successfully adapts its warnings to the unique needs of different drivers, thereby enhancing both safety and the overall driving experience through personalized interaction strategies. Furthermore, the system has demonstrated the capacity to generate explainable personalization profiles and individualized assistance strategies based on domain-specific knowledge. This study explores a novel method of leveraging LLMs to facilitate human-machine interaction via multimodal communication, showing the potential for broader applications in enhancing the safe, comfortable, and personalized experience of intelligent vehicles.

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (RS-2024-00351865).

APPENDIX

DRIVER PROFILE AND EXPERIENCE SURVEY

A. Survey Introduction

This survey aims to gather comprehensive driver data for developing personalized driving assistance solutions. The collected information was used to improve user experience and promote safer driving, with all responses remaining strictly confidential and used exclusively for scientific research purposes.

B. Questionnaire Items

1) Basic Information

- Preferred name/nickname
- Age
- Gender
- Occupation (If retired, previous occupation)

2) Driving Experience

- Years of driving experience:
 - 6 months or less
 - 7 months - 1 year
 - 1 - 2 years
 - 2 - 5 years
 - 5 - 10 years
 - More than 10 years

3) Driving Frequency

- Average driving hours per week:
 - Less than 5 hours
 - 5-10 hours
 - 11-20 hours
 - More than 20 hours

4) Driving Purposes (Multiple selections allowed)

- Commuting
- Transporting family members
- Shopping
- Leisure
- Travel
- Work-related (not commuting)
- Other

5) Driving Habits and Preferences

- Open-ended response
- Example provided: prefer highways, avoid night driving

6) Health and Driving Records

- Do you have any health conditions or illnesses that affect your driving?
 - Open-ended response
 - Example provided: vision impairment, chronic pain
- Have you been involved in any traffic accidents in the past 5 years?
 - Open-ended response
 - Note: Please describe the incident(s), including details of the accident and any injuries or fatalities

7) Driving Style

- How would you describe your driving style?
 - Very cautious
 - Somewhat cautious
 - Balanced
 - Somewhat aggressive
 - Very aggressive

8) International Driving Experience

- Do you have experience driving in foreign countries? If yes, please describe your experience:
 - Open-ended response
 - Note: Please state which countries you have driven in and briefly describe your experience

REFERENCES

- [1] J. Piao and M. McDonald, "Advanced driver assistance systems from autonomous to cooperative approach," *Transport reviews*, vol. 28, no. 5, pp. 659-684, 2008.
- [2] M. Hasenjaeger, M. Heckmann, and H. Wersing, "A survey of personalization for advanced driver assistance systems," *IEEE Transactions on Intelligent Vehicles*, vol. 5, no. 2, pp. 335-344, 2019.
- [3] K. Bengler et al., "Three Decades of Driver Assistance Systems: Review and Future Perspectives," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 4, pp. 6-22, 2014.
- [4] Y. Wang et al., "Decision-Making Driven by Driver Intelligence and Environment Reasoning for High-Level Autonomous Vehicles: A Survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 10, pp. 10362-10381, 2023.
- [5] A. Khanafer, D. Balzer, and R. Isermann, "A rule-based collision avoidance system—scene interpretation, strategy selection, path planning and system intervention," *SAE International Journal of Passenger Cars-Mechanical Systems*, vol. 2, no. 2009-01-0156, pp. 389-397, 2009.
- [6] J. Horgan et al., "Vision-Based Driver Assistance Systems: Survey, Taxonomy and Advances," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015.
- [7] J. Park et al., "Driving Assistant Companion With Voice Interface Using Long Short-Term Memory Networks," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 582-590, 2019.
- [8] H. Wang et al., "FPT: Fine-Grained Detection of Driver Distraction Based on the Feature Pyramid Vision Transformer," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 2, pp. 1594-1608, 2023.
- [9] F. Damerow et al., "Intersection warning system for occlusion risks using relational local dynamic maps," *IEEE Intelligent Transportation Systems Magazine*, vol. 10, no. 4, pp. 47-59, 2018.
- [10] K. Yang et al., "Uncertainties in Onboard Algorithms for Autonomous Vehicles: Challenges, Mitigation, and Perspectives," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 9, pp. 8963-8987, 2023.
- [11] L. Li et al., "Cognitive Cars: A New Frontier for ADAS Research," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 395-407, 2012.
- [12] C. J. D. Patten et al., "Driver experience and cognitive workload in different traffic environments," *Accident Analysis & Prevention*, vol. 38, no. 5, pp. 887-894, 2006.
- [13] X. Li et al., "Towards Knowledge-driven Autonomous Driving," *arXiv preprint arXiv:2312.04316*, 2023.
- [14] G. Li et al., "Human-like decision making of artificial drivers in intelligent transportation systems: An end-to-end driving behavior prediction approach," *IEEE Intelligent Transportation Systems Magazine*, vol. 14, no. 6, pp. 188-205, 2021.
- [15] Z. Sheng et al., "Kinematics-aware multigraph attention network with residual learning for heterogeneous trajectory prediction," *Journal of Intelligent and Connected Vehicles*, vol. 7, no. 2, pp. 138-150, 2024.
- [16] W. Li et al., "Brain-Inspired Driver Emotion Detection for Intelligent Cockpits Based on Real Driving Data," *IEEE Intelligent Transportation Systems Magazine*, vol. 16, no. 4, pp. 62-80, 2024.
- [17] Q. Lan and Q. Tian, "Instance, scale, and teacher adaptive knowledge distillation for visual detection in autonomous driving," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 3, pp. 2358-2370, 2022.
- [18] R. Wickramarachchi, C. Henson, and A. Sheth, "Knowledge-based entity prediction for improved machine perception in autonomous systems," *IEEE Intelligent Systems*, vol. 37, no. 5, pp. 42-49, 2022.

- [0] L. Halilaj et al., "A knowledge graph-based approach for situation comprehension in driving scenarios," in *The Semantic Web: 18th International Conference, ESWC 2021*, Virtual Event, June 6–10, 2021, pp. 1–15.
- [1] S. Kothawade et al., "AUTO-DISCERN: autonomous driving using common sense reasoning," *arXiv preprint arXiv:2110.13606*, 2021.
- [2] F. N. Biondi et al., "The challenge of advanced driver assistance systems assessment: A scale for the assessment of the human-machine interface of advanced driver assistance technology," *Transportation research record*, vol. 2672, no. 37, pp. 113–122, 2018.
- [3] F. Biondi et al., "Advanced driver assistance systems: Using multimodal redundant warnings to enhance road safety," *Applied ergonomics*, vol. 58, pp. 238–244, 2017.
- [4] E. C. Haas and J. B. Van Erp, "Multimodal warnings to enhance risk communication and safety," *Safety science*, vol. 61, pp. 29–35, 2014.
- [5] J. Xu and A. R. Bowers, "Hazard warning modalities and timing thresholds for older drivers with impaired vision," *Accident Analysis & Prevention*, vol. 202, p. 107599, 2024.
- [6] Y. Cao et al., "On timing and modality choice with local danger warnings for drivers," in *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Essen, Germany, 2009, pp. 75–78.
- [7] D. C. Marshall, J. D. Lee, and P. A. Austria, "Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness," *Human factors*, vol. 49, no. 1, pp. 145–157, 2007.
- [8] A. Murata, T. Kuroda, and W. Karwowski, "Effects of auditory and tactile warning on response to visual hazards under a noisy environment," *Applied Ergonomics*, vol. 60, pp. 58–67, 2017.
- [9] J. Scott and R. Gray, "A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving," *Human factors*, vol. 50, no. 2, pp. 264–275, 2008.
- [10] E. Adell, A. Várhelyi, and M. Hja'Imdahl, "Auditory and haptic systems for in-car speed management – A comparative real life study," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 11, no. 6, pp. 445–458, 2008.
- [11] M. Bueno et al., "Behavioural adaptation and effectiveness of a Forward Collision Warning System depending on a secondary cognitive task," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 24, pp. 158–168, 2014.
- [12] C. L. Baldwin, B. A. Lewis, and P. M. Greenwood, *Designing transportation systems for older adults*. CRC Press, 2019.
- [13] C. Spence and C. Ho, "Multisensory warning signals for event perception and safe driving," *Theoretical Issues in Ergonomics Science*, vol. 9, no. 6, pp. 523–554, 2008.
- [14] T. Horberry et al., "Human-Centered Design for an In-Vehicle Truck Driver Fatigue and Distraction Warning System," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 5350–5359, 2022.
- [15] J. Lylykangas et al., "Responses to visual, tactile and visual-tactile forward collision warnings while gaze on and off the road," *Transportation research part F: traffic psychology and behaviour*, vol. 40, pp. 68–77, 2016.
- [16] A. Lundkvist and A. Nykänen, "Response times for visual, auditory and vibrotactile directional cues in driver assistance systems," *SAE International journal of transportation safety*, vol. 4, no. 1, pp. 8–14, 2016.
- [17] C. Ho, N. Reed, and C. Spence, "Multisensory in-car warning signals for collision avoidance," *Human factors*, vol. 49, no. 6, pp. 1107–1114, 2007.
- [18] X. Qu et al., "Envisioning the future of transportation: Inspiration of ChatGPT and large models," *Communications in Transportation Research*, vol. 3, p. 100103, 2023.
- [19] G. Li et al., "Camel: Communicative agents for "mind" exploration of large language model society," *Advances in Neural Information Processing Systems*, vol. 36, pp. 51991–52008, 2023.
- [20] S. Bubeck et al., "Sparks of artificial general intelligence: Early experiments with gpt-4," *arXiv preprint, arXiv:2303.12712*, 2023.
- [21] J. Chen et al., "When large language models meet personalization: Perspectives of challenges and opportunities," *World Wide Web*, vol. 27, no. 4, p. 42, 2024.
- [22] T. Wu, M. Terry, and C. J. Cai, "AI chains: Transparent and controllable human-ai interaction by chaining large language model prompts," in *Proceedings of the 2022 CHI conference on human factors in computing systems*, 2022.
- [23] B. Wang et al., "Task supportive and personalized human-large language model interaction: A user study," in *Proceedings of the 2024 Conference on Human Information Interaction and Retrieval*, 2024.
- [24] J. Zhang et al., "Prospect Personalized Recommendation on Large Language Model-based Agent Platform," *arXiv preprint arXiv:2402.18240*, 2024.
- [25] L. Wen et al., "On the Road with GPT-4V (ision): Early Explorations of Visual-Language Model on Autonomous Driving," *arXiv preprint arXiv:2311.05332*, 2023.
- [26] Z. Xi et al., "The rise and potential of large language model based agents: A survey," *arXiv preprint arXiv:2309.07864*, 2023.
- [27] S. Yao et al., "React: Synergizing reasoning and acting in language models," *arXiv preprint arXiv:2210.03629*, 2022.
- [28] W. X. Zhao et al., "A survey of large language models," *arXiv preprint arXiv:2303.18223*, 2023.
- [29] J. Zheng et al., "Lifelong Learning of Large Language Model based Agents: A Roadmap," *arXiv preprint arXiv:2501.07278*, 2025.
- [30] LangChain, "Introduction - langchain," 2023. [Online]. Available: <https://www.langchain.com/langchain>
- [31] A. Zhang et al., "On Generative Agents in Recommendation," in *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Washington DC, USA, 2024, pp. 1807–1817.
- [32] H. Zhang et al., "Building cooperative embodied agents modularly with large language models," *arXiv preprint arXiv:2307.02485*, 2023.
- [33] T. Driessen et al., "Putting ChatGPT vision (GPT-4V) to the test: risk perception in traffic images," *Royal Society Open Science*, vol. 11, no. 5, p. 231676, 2024.
- [34] H. Pu et al., "AutoRepo: A general framework for multimodal LLM-based automated construction reporting," *Expert Systems with Applications*, vol. 255, p. 124601, 2024.
- [35] A. Salemi et al., "LaMP: When Large Language Models Meet Personalization," *arXiv preprint arXiv:2304.11406*, 2023.
- [36] K. Zhang et al., "Memory-Augmented LLM Personalization with Short- and Long-Term Memory Coordination," *arXiv preprint arXiv:2309.11696*, 2023.
- [37] Y. Kuang, H. Lin, and M. Jiang, "OpenFMNav: Towards Open-Set Zero-Shot Object Navigation via Vision-Language Foundation Models," *arXiv preprint arXiv:2402.10670*, 2024.
- [38] Q. Wu et al., "Autogen: Enabling next-gen llm applications via multi-agent conversation framework," *arXiv preprint arXiv:2308.08155*, 2023.
- [39] C. Chen et al., "Edge intelligence empowered vehicle detection and image segmentation for autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 11, pp. 13023–13034, 2023.
- [40] D. Fu et al., "Drive like a human: Rethinking autonomous driving with large language models," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024.
- [41] C. Cui et al., "Drive as you speak: Enabling human-like interaction with large language models in autonomous vehicles," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024.
- [42] Z. Xu et al., "DriveGPT4: Interpretable End-to-End Autonomous Driving Via Large Language Model," *IEEE Robotics and Automation Letters*, vol. 9, no. 10, pp. 8186–8193, 2024.
- [43] L. Chen et al., "Driving with LLMs: Fusing Object-Level Vector Modality for Explainable Autonomous Driving," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024.
- [44] J. Mao et al., "Gpt-driver: Learning to drive with gpt," *arXiv preprint arXiv:2310.01415*, 2023.
- [45] Y. Wang et al., "Driving into the future: Multiview visual forecasting and planning with world model for autonomous driving," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024.
- [46] L. Wen et al., "Dilu: A knowledge-driven approach to autonomous driving with large language models," *arXiv preprint arXiv:2309.16292*, 2023.
- [47] X. Tian et al., "Drivevlm: The convergence of autonomous driving and large vision-language models," *arXiv preprint arXiv:2402.12289*, 2024.
- [48] D. Xiao et al., "HazardVLM: A Video Language Model for Real-Time Hazard Description in Automated Driving Systems," *IEEE Transactions on Intelligent Vehicles*, pp. 1–13, 2024.
- [49] S. Wang et al., "ChatGPT as your vehicle co-pilot: An initial attempt," *IEEE Transactions on Intelligent Vehicles*, 2023.
- [50] M. Z. Hasan et al., "Vision-Language Models Can Identify Distracted Driver Behavior From Naturalistic Videos," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 9, pp. 11602–11616, 2024.
- [51] L. Wang et al., "Accidentgpt: Accident analysis and prevention from v2x environmental perception with multi-modal large model," *arXiv preprint arXiv:2312.13156*, 2023.

- [0] Y. Zhang et al., "Siren's song in the AI ocean: a survey on hallucination in large language models," *arXiv preprint arXiv:2309.01219*, 2023.
- [1] Y. Gao et al., "Retrieval-Augmented Generation for Large Language Models: A Survey," *arXiv preprint arXiv:2312.10997*, 2023.
- [2] W. Fan et al., "A survey on rag meeting llms: Towards retrieval-augmented large language models," in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024.
- [3] Y. Tang and Y. Yang, "Multi-hop-rag: Benchmarking retrieval-augmented generation for multi-hop queries," *arXiv preprint arXiv:2401.15391*, 2024.
- [4] K. Zhang et al., "Integrating visual large language model and reasoning chain for driver behavior analysis and risk assessment," *Accident Analysis & Prevention*, vol. 198, p. 107497, 2024.
- [5] J. Wei et al., "Chain-of-thought prompting elicits reasoning in large language models," *Advances in neural information processing systems*, vol. 35, pp. 24824–24837, 2022.
- [6] S. Kuang et al., "Harnessing multimodal large language models for traffic knowledge graph generation and decision-making," *Communications in Transportation Research*, vol. 4, p. 100146, 2024.
- [7] X. Guo et al., "Towards explainable traffic flow prediction with large language models," *Communications in Transportation Research*, vol. 4, p. 100150, 2024.
- [8] T. Kojima et al., "Large Language Models are Zero-Shot Reasoners," *arXiv preprint arXiv:2205.11916*, 2023.
- [9] M. Nie et al., "Reason2Drive: Towards interpretable and chain-based reasoning for autonomous driving," in *European Conference on Computer Vision*, 2024, pp. 292–308.
- [10] C. Sima et al., "DriveLM: Driving with graph visual question answering," in *European Conference on Computer Vision*, 2024, pp. 256–274.
- [11] X. Li and X. Qiu, "MoT: Memory-of-Thought Enables ChatGPT to Self-Improve," in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 2023.
- [12] M. Yasunaga et al., "Large Language Models as Analogical Reasoners," *arXiv preprint arXiv:2310.01714*, 2023.
- [13] R. Wickramarachchi, C. Henson, and A. Sheth, "Knowledge-infused learning for entity prediction in driving scenes," *Frontiers in big Data*, vol. 4, p. 759110, 2021.
- [14] L. Halilaj et al., "Knowledge Graphs for Automated Driving," in *2022 IEEE Fifth International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, 2022.
- [15] Z. Yang et al., "LLM4Drive: A survey of large language models for autonomous driving," *arXiv preprint arXiv:2311.01043*, 2023.
- [16] C. Cui et al., "Large Language Models for Autonomous Driving (LLM4AD): Concept, Benchmark, Experiments, and Challenges," *arXiv preprint arXiv:2410.15281*, 2025.
- [17] M. Miranda et al., "Preserving privacy in large language models: A survey on current threats and solutions," *arXiv preprint arXiv:2408.05212*, 2024.
- [18] G. Feretzakis et al., "Privacy-Preserving Techniques in Generative AI and Large Language Models: A Narrative Review," *Information*, vol. 15, no. 11, p. 697, 2024.



Zixuan Xu is currently a Ph.D. student at Cho Chun Shik Graduate School of Mobility, Korea Advanced Institute of Science and Technology (KAIST). She received the B.S. degree in Transportation Engineering from Tongji University in 2022 and the M.S. degree in Data-Driven Modelling in the Department of Physics from the Hong Kong University of Science and Technology (HKUST) in 2023. Her research interests include Large Language Models (LLMs), human-machine interaction, and human factors in autonomous vehicles.



Tiantian Chen (nicole.chen@kaist.ac.kr) is an Assistant Professor at the Cho Chun Shik Graduate School of Mobility at KAIST. She received her PhD in Civil and Environmental Engineering from the Hong Kong Polytechnic University in 2021. She specializes in traffic safety, human factors, human-centered design, driving simulation, travel behavior, transport policy, and applications of statistical methods. She has over 20 journal publications in prestigious transportation & safety journals such as *Transportation Research Part A, C, E, F, Transport Policy, Analytic Methods in Accident Research, and Accident Analysis & Prevention*. She is now an Editor of the *Case Studies on Transport Policy*, and guest editor for the Special Issues in the *International Journal of Sustainable Transportation, Multimodal Transportation, and Research in Transport Business and Management*. Due to her expertise, she is also a member of the IEEE Emerging Transportation Technology Testing (ET3) Technical Committee.



Zilin Huang received the B.S. degree in mechanical engineering from the School of Electromechanical Engineering, Guangdong University of Technology in 2018. He received the M.S. degree in Communication and Transportation Engineering from South China University of Technology in 2021. He is currently pursuing the Ph.D. degree at the Department of Civil and Environmental Engineering, University of Wisconsin-Madison, USA. Before joining UW-Madison, he worked at the Center for Connected and Automated Transportation (CCAT), Purdue University, USA. His research interests include human-centered AI, autonomous driving, robotics, human-robot interaction, and intelligent transportation.



Yang Xing is currently a Senior Lecturer in Applied Artificial Intelligence at Cranfield University. He received his Ph. D. degree from Cranfield University, UK. He was a Research Associate at the University of Oxford and a Research Fellow at Nanyang Technological University. His research interests include human behaviour modelling, intelligent multi-agent collaboration, and intelligent vehicles. Dr Xing Serves as associate editor for *IEEE Transactions on Intelligent Vehicles* and *IEEE Transactions on Neural Networks and Learning Systems*.



Sikai Chen (sikai.chen@wisc.edu) is an Assistant Professor at the Department of Civil and Environmental Engineering and the Department of Mechanical Engineering (courtesy), University of Wisconsin-Madison. He received his Ph.D. degree in civil engineering with a focus on computational science & engineering from Purdue University in 2019. His research focuses on three major themes: human users, AI, and transportation. He aims to innovate and develop safe, efficient, sustainable, and human-centered transportation systems by cutting-edge methods and technologies. The focus is on incorporating human behaviors, interactive autonomy, digital infrastructure, and intelligent control frameworks. In addition, he is a member of two ASCE national committees: Connected & Autonomous Vehicle Impacts, and Economics & Finance; IEEE Emerging Transportation Technology Testing Technical Committee; and TRB Standing Committee on Statistical Methods (AED60).



DSPACE

<https://dspace.org/>

Personalizing driver agent using large language models for driving safety and smarter human–machine interactions

Xu, Zixuan; Chen, Tiantian; Huang, Zilin; Xing, Yang; Chen, Sikai

2025-07

Institute of Electrical and Electronics Engineers (IEEE)

<https://doi.org/10.1109/mits.2025.3551736>

Xu Z, Chen T, Huang Z, et al., (2025) Personalizing driver agent using large language models for driving safety and smarter human–machine interactions. IEEE Intelligent Transportation Systems Magazine, Volume 17, Issue 4, July-August 2025, pp. 96-111