ARTICLE INFO
Article ID: 12-06-03-0023

doi:10.4271/12-06-03-0023

Recognition Assistance Interface for Human-Automation Cooperation in Pedestrian Risk Prediction

Atsushi Kuribayashi,[1] Eijiro Takeuchi,[2] Alexander Carballo,[3] Yoshio Ishiguro,[4] and Kazuya Takeda[1]

[1]Nagoya University, Graduate School of Informatics, Japan
[2]TierIV Inc., Japan
[3]Faculty of Engineering, Gifu University, Japan
[4]The University of Tokyo, Interfaculty Initiative in Information Studies, Japan

Abstract

Autonomous driving systems (ADS) have been widely tested in real-world environments with operators who must monitor and intervene due to remaining technical challenges. However, intervention methods that require operators to take over control of the vehicle involve many drawbacks related to human performance. ADS consist of recognition, decision, and control modules. The latter two phases are dependent on the recognition phase, which still struggles with tasks involving the prediction of human behavior, such as pedestrian risk prediction. As an alternative to full automation of the recognition task, cooperative recognition approaches utilize the human operator to assist the automated system in performing challenging recognition tasks, using a recognition assistance interface to realize human-machine cooperation. In this study, we propose a recognition assistance interface for cooperative recognition in order to achieve safer and more efficient driving through improved human-automation cooperation. A simulator experiment with 18 participants is conducted to evaluate our recognition assistance interface in comparison with a conventional control intervention, in terms of driving safety, efficiency, and usability. Recognition of pedestrian crossing intention is selected for the cooperation task, and driving scenarios in which the automated system cannot reliably recognize the crossing intentions of pedestrians at non-signalized locations are selected as the driving scenario. Statistical analysis of our experimental results reveals that the proposed recognition assistance interface allowed more accurate operator intervention, was easier to use, and achieved more stable vehicle control than the control intervention. We also found that sharing the recognition information of the automated driving system with operators could divide their attention, impairing intervention performance. Our experimental results suggest that the unifying presentation of the system recognition information and the operator's manipulation target on the touchscreen of the user interface addresses this problem.

History

Received: 15 Dec 2022
Revised: 27 Feb 2023
Accepted: 07 Apr 2023
e-Available: 06 Jun 2023

Keywords

Automated driving, Human-computer interaction, Driver vehicle cooperation, Driving intervention, Cooperative perception

Citation

Kuribayashi, A., Takeuchi, E., Carballo, A., Ishiguro, Y. et al., "Recognition Assistance Interface for Human-Automation Cooperation in Pedestrian Risk Prediction," *SAE Int. J. of CAV* 6(3):345–363, 2023, doi:10.4271/12-06-03-0023.

ISSN: 2574-0741
e-ISSN: 2574-075X

I. Introduction

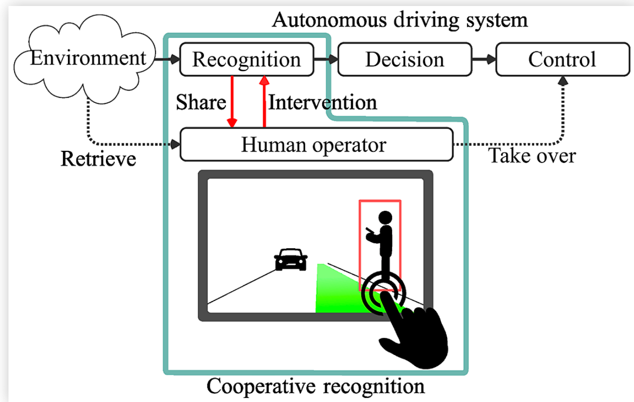
Level 3 autonomous vehicles [1] are now being tested and commercialized, bringing us closer to the realization of fully automated driving. Currently, when an autonomous driving system (ADS) is unable to appropriately control the vehicle, the operator must take over the driving task to ensure safety. However, these control phase interventions involve many challenges, which include requiring the operator to maintain a high level of situational awareness [2, 3], the need for excellent driver response and driving ability [4, 5], difficulty in smoothly handing over control [6], driver overconfidence in the automated system [7, 8], and misuse or disuse of the automated system [9].

To address these challenges, interfaces that allow quick and effortless takeover [6, 10, 11], or that facilitate driver understanding of the internal state of the autonomous system [12, 13, 14], have been proposed. However, as the performance of ADS improves and approaches Level 4 automation [1], intervention frequency decreases, making quick and accurate operator intervention more difficult. Therefore, recognition performance and driver intervention problems are inevitable, and we expect that human-automation cooperation during the ADS recognition phase (i.e., cooperative recognition) is a promising solution to them.

Automated driving systems operate vehicles in three stages: recognition, decision, and control. Both the decision and control phases are carried out based on the information obtained during the recognition phase. However, recognition of the risks is still difficult to accomplish with 100% accuracy for automation due to the recall-precision trade-off. For example, if a recall is prioritized, the ADS becomes more safety oriented, but driving performance in terms of efficiency is reduced due to overcautious misrecognitions [15, 16, 17]. If precision is prioritized, efficiency is improved, but safety is reduced due to omissions. It is therefore difficult to balance both safe and efficient driving. This problem exists in advanced tasks (e.g., behavior prediction of the other road users) and even simple tasks (e.g., object detection, traffic light classification).

To address this problem, in this study, we propose a method of human-automation cooperation during the recognition phase that utilizes human strengths such as intuition, prior experience, and understanding of appropriate interaction with other road users, abilities which allow more accurate processing of advanced recognition tasks. This cooperation is achieved through the use of a recognition assistance interface which shares recognition information generated by the ADS with the operator, allowing the operator to intervene in difficult recognition tasks. The interface is designed to present recognition results to the operator that the ADS recognizes as ambiguous or risky (i.e., possible recognition of false positives) and the operator is then tasked with eliminating recognition results that do not actually pose a risk. An overview of this process is shown in Figure 1. This design solves the recall-precision trade-off and eliminates the task of evaluating false negatives from the operator's task, which is recognized as being very difficult. Furthermore, tasks that are critical to driving safety are not allocated to the operator, a division of labor that takes into account the possibility of human error during interventions [18, 19, 20, 21].

FIGURE 1 Relationship between ADS and human operator during cooperative recognition. The recognition information of the automated system is shared with the operator. The operator interprets only the task-related context and then intervenes in the recognition process of the system (red arrows). The operator does not need to understand the entire driving environment or take control of the vehicle (dashed arrows).



Thus, the objective of our study is to realize safer and more efficient automated driving through better human-automation cooperation. "Efficient driving" in this study is driving that avoids unnecessary evasive action due to recognition errors and achieves smooth control of the vehicle. Therefore, the recognition assistance interface should allow quick and accurate intervention by the operator during the recognition phase as this allows the system to avoid drastic changes in vehicle control, such as when the ADS must perform emergency safety braking due to operator intervention errors. Recognition phase intervention also allows higher cruising speeds. In this study, we also clarify the design of cooperative recognition systems and how such approaches can be used to realize safe and efficient driving.

The proposed recognition assistance interface was implemented within an existing ADS, and a simulator experiment was conducted with 18 participants to evaluate differences in the performance of cooperative recognition and control intervention in terms of driving efficiency, safety, and usability (i.e., user workload, user preference, and time required for intervention). Pedestrian risk recognition [22, 23] at a non-signalized location was selected as the recognition task requiring cooperation between operators and the ADS. Although pedestrians are the most vulnerable road users, it is difficult for automated systems to accurately predict their behavior because they are not part of the flow of traffic. In contrast, humans can fairly accurately predict pedestrian behavior based on traffic conditions, prior experience, and subtle cues related to human behavior.

The recognition assistance interface proposed in this study consists of two user interfaces, a display for sharing recognition information from the ADS and an input device for operator intervention. Previous studies on ADS and the human factor have revealed that sharing the internal state [12, 13, 14, 24, 25] or intention [26] of an automated system with the user promotes an appropriate mental model and improves the situational awareness of the user. However, other studies [27, 28, 29] have revealed issues, such as information overload and the trade-off between display cluttering and divided attention, when system information is shared with users. The heads-up display (HUD) and head-mounted display have been considered as solutions to these issues, i.e., unifying information from the automated system and information from the outside world. This article also examines these issues in the context of cooperative recognition. We hypothesized that a touch display would be an appropriate solution because it is a familiar in-vehicle user interface which can be used to unify ADS recognition information, external world information, and input commands. To test this hypothesis, the effects of sharing recognition information with operators, as well as the effects of differences in the design of the user interface, were also evaluated.

The contributions of this article are as follows:

- We clarify the design of a recognition assistance interface for safer and more efficient automated driving, as well as for better human-automation cooperation.
- We implement a recognition assistance interface in an existing ADS, realizing cooperative recognition of pedestrian crossing intention.
- We reveal issues in the design of cooperative recognition systems from a cognitive aspect and propose a user interface that integrates the manipulation target and the information display. Such as a touch display could be the key to improving usability and to exploiting the potential of cooperative recognition.

II. Related Work

A. Human-Vehicle Interaction

Cooperative driving, in which the vehicle assists the human driver with steering and acceleration, has been widely studied. According to the classification system created by SAE International [1], these systems are classified as Level 2 automation. Although our research target is Level 3 systems and higher, we refer here to interface design issues explored in previous studies.

H-mode [30, 31, 32], named for the interaction between a rider and horse, is a cooperative driving method in which the driver is always in contact with the autonomous system through an interface and interactively controls the vehicle. A joystick, which is common in the aviation domain, and haptic feedback from the autonomous system have been used as the interfaces. Although different in modality, what H-mode and our recognition assistance interface have in common is that the autonomous system and the human operator are always connected and cooperate via a user interface.

Advanced driver assistance systems (ADAS), such as Adaptive Cruise Control and Automated Emergency Braking, are cooperative systems in practical use [33, 34, 35], and various interface methods and modalities have been proposed for these systems. Tactile vibration through the steering wheel is often used for control assistance [36, 37]. Head-down display [38], HUD [39], and augmented reality HUD (AR-HUD) [40, 41] are popular interfaces for cognitive assistance, such as displaying risks in the local environment. Although AR-HUD interfaces have become popular in the current literature, there are practical issues with this approach, such as binocular parallax, calibration, and limited field of view. As for the manipulation interface, in addition to the conventional steering wheel and pedals, joysticks [42, 43], voice and gaze [44, 45], hand gestures [46, 47], and touch displays [47, 48] have been studied.

As ADAS and automated driving systems have been developed, the out-of-the-loop state [34, 49] and driver trust in the autonomous system have become important issues. A higher level of trust often results in overconfidence, and low levels of trust lead to the disuse of these systems [9, 50]. Excessive trust, also known as “overconfidence” or “overreliance,” can also cause safety degradation due to misuse [7, 8]. Maintaining an appropriate level of user reliance on autonomous systems is crucial for safe operation. Sharing the internal information of the automated system, such as its confidence level [14, 24], capabilities [25], and perception information [13, 51], has been shown to be effective in maintaining an appropriate level of trust. Therefore, our proposed method shares the recognition information of the autonomous system with the operator, which is expected to facilitate a better understanding of the capabilities of the autonomous system and suppress overconfidence.

B. Human-Automation Cooperation in Autonomous Driving

During cooperative driving with Level 3 or 4 automated systems, the operators are generally assigned to monitor the driving of the autonomous system, taking over control of the vehicle only when necessary. The challenges in such control phase cooperation are maintaining the operator’s situational awareness [2, 3, 52], communicating takeover requests [53, 54, 55], and safety during control transitions [6, 11]. The design of takeover requests in emergency and nonemergency situations has been well studied, and auditory, visual, tactile, and multimodal interfaces have been investigated [56, 57, 58, 59]. However, surveillance and takeover are cognitively and physically demanding for operators [66]. Our approach to solving these challenges is to change the phase of human-automation cooperation from the control phase to the recognition phase. However, other cooperative driving approaches have focused on the decision and prediction phases of autonomous driving. The German Research Foundation (DFG) proposed the Conduct-by-Wire method, in which operators input discrete maneuver commands instead of control inputs [47, 60, 61, 62], which is an intervention in the decision-making phase of the ADS. PieDrive was proposed as a project outcome [63]. Walch et al. [48, 64] also proposed cooperation in the decision-making phase, abstracting the input commands of the conduct-by-wire method to allow operators to choose from among several commands. The advantage of these methods is that they do not require continuous vehicle control and thus require less physical load, but operators need to be able to quickly and accurately perceive all of the risks that are present in the surrounding environment when performing interventions. In contrast, when using a recognition assistance interface, operators are only required to assess specific elements within the surrounding environment.

Chao et al. [44] proposed and evaluated a framework for operator intervention during the prediction of the trajectories of surrounding vehicles in non-safety-related situations. Operators monitored autonomous driving and alerted the system using gaze and voice commands when they sensed danger. The vehicle trajectory prediction was then modified based on the operator's intervention, changing the driving behavior of the ego vehicle. The strengths of this method are that the operator does not need to intervene in maneuvering the vehicle, which can solve the drawbacks of conventional control intervention while also taking advantage of human decision-making abilities based on their experience and intuition. Major differences in our proposed method are that the operator is not required to monitor the autonomous system driving, and it is the system that requests intervention for safety-related driving situations, although the recognition information of the system is continuously shared with the operator. Furthermore, by designing the ADS to respond conservatively, we solve the problems of safety and time constraints during the intervention.

III. Recognition Assistance Interface

Here we explain the framework of our proposed cooperative recognition method, and how it realizes human-automation cooperation to achieve safer and more efficient driving through cooperative recognition.

A. Realization of Human-Automation Cooperation

In order to achieve human-automation cooperation, appropriate function allocation [18, 65] should be intentionally designed, in contrast, to control intervention approaches in which all of the unperformable tasks are simply allocated to the driver. In systems based on cooperative recognition, function allocation, i.e., selection of the operator's intervention targets, is based on the performance characteristics of humans and machines. The human intervention tasks involve high-level recognition such as long-distance recognition, prediction of the intentions of other road users, and occlusion-related risk assessment, because humans can perform these tasks using modalities that machines do not possess (intuition, driving experience, innate understanding of human psychology, and the resulting ability to predict the likely future behavior of other road users). On the other hand, the ADS is responsible for low-level recognition tasks such as short-range recognition for collision avoidance, lane keeping, and rule compliance, which do not require human intervention because machines are able to perform them more reliably. These simple tasks are also more critical to safety and harder for humans to perform due to the limited amount of time available for intervention.

Cooperative recognition also has advantages for resolving mental model-related issues in human-machine cooperation such as overconfidence [7, 8]. The current recognition information of ADS is shared with the operator via the user interface not only for intervention requests but also to allow the operator to comprehend the internal state of the ADS, which is intended to generate an appropriate level of operator trust in the system, preventing overconfidence [40].

B. Realization of Safety and Efficiency

The problem with integrating human modality into the process of an automated system is that humans are not good at continuous surveillance, and when compensating for system deficits, they are error prone, especially when reaction time is limited. Cooperative recognition systems need to be designed to compensate for these drawbacks. An ADS can be set to be more conservative (i.e., safety oriented) by instructing it to avoid false-negatives and accept false-positives. Conservative recognition requires the ADS to perform evasive action when encountering ambiguous recognition results by default, which provides more time for human intervention and ensures safety even when the operator fails to intervene. This division of labor is selected because it is difficult for humans to respond to omissions (false-negatives) and to perform continuous surveillance [66]. However, conservative recognition may result in unnecessary evasive action due to overestimation of risks, degrading driving efficiency. The operator's task, therefore, is to filter out erroneous threats within a limited amount of time (between the intervention request of the system and the last possible moment for initiating evasive behavior). Therefore, the user interface needs to be designed to allow quick operator intervention to allow higher cruising speeds, while also reducing unnecessary evasive behavior. Furthermore, recognition results with sufficiently high recognition likelihoods can be excluded as intervention targets, preventing degradation of the inherent performance of the automated system as the result of faulty human interventions.

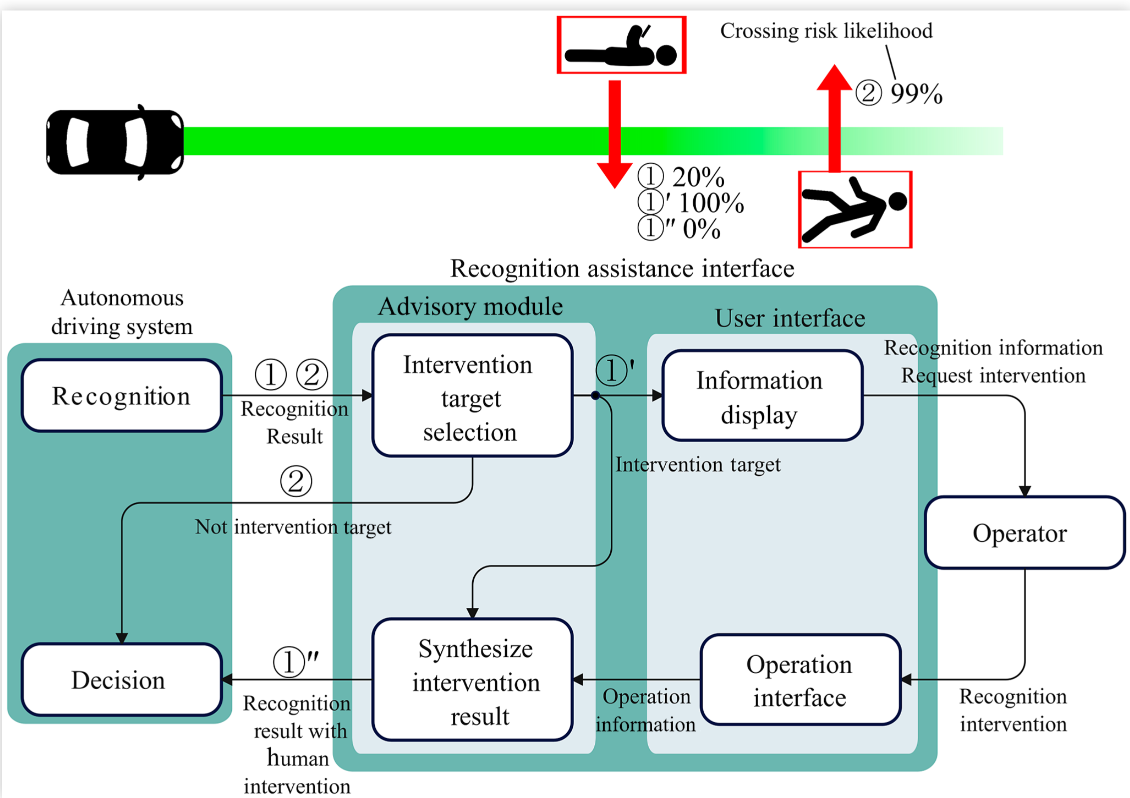
Cooperative recognition tasks are thus divided between the ADS and operator as described above. Although this division of labor is intended to improve driving performance, it can sometimes result in abrupt changes in driving behavior. For example, if the operator mistakenly determines from a distance that an intervention target is not a safety risk, the ADS is initially allowed to maintain the vehicle speed. But when it is a short range, a low-level recognition system detects the target as an imminent risk and the ADS must perform emergency braking. Therefore, even though the ADS is able to guarantee a minimum level of safety, the recognition assistance interface must allow accurate and timely operator intervention in order to avoid automated emergency responses. The operator performs interventions through the user interface, and the results of these interventions are then synthesized with the recognition results of the ADS. By weighing both inputs and taking into account the estimated reliability of the operator and the automated system, "adaptive automation" [65] can be realized. For example, if the driver monitoring system detects that the operator is not concentrating on the intervention, the recognition results of the ADS will be prioritized.

C. Design of the Recognition Assistance Interface

Here we explain how the framework described above (and shown in Figure 2) is realized. The system advisory module selects intervention targets from the recognition results. Results with insufficiently high or insufficiently low recognition likelihoods [Figure 2(1)] are considered to be ambiguous risks and become intervention targets. The remaining unambiguous recognition results (targets that are clearly risky or clearly not risky) are excluded as intervention targets [Figure 2(2)]. The ambiguous risks are conservatively recognized and flagged as risky [Figure 2(1)′].

These recognition results are shared with the operator via the user interface, and intervention is requested for one of the intervention targets. The operator's intervention input is then synthesized with the recognition results of the autonomous system [Figure 2(1)″]. The algorithms used for the selection of intervention targets, and for the synthesis of the human and recognition system evaluation results, are subjects for future study.

FIGURE 2 Block diagram of the proposed recognition assistance interface. The recognition results of the system are shared with the operator. Intervention targets are selected from among these recognition results by the system (1′), and intervention requests identifying the intervention target are delivered to the operator via the interface. The operator then performs recognition intervention using the interface. Finally, the system integrates the results of the operator's intervention with its original recognition results (1″).



IV. Evaluation of Recognition Assistance Interface

We conducted a subject experiment to compare the performance of cooperative recognition using the proposed recognition assistance interface and conventional control intervention, as regards their safety, driving efficiency, and usability. In addition, the recognition assistance interface was broken down into two elements, a display for sharing recognition information from the ADS and an input device for operator intervention. Both were experimentally evaluated to collect data for future development.

Our subject experiment was designed to comply with the requirements of the Nagoya University Ethics Committee (NUEC). The hypotheses of our experiment are as follows:

- H1: Cooperative recognition using the proposed recognition assistance interface allows safer and more efficient automated driving than conventional control phase intervention.
- H2: The proposed recognition assistance interface is easier to use than control intervention methods due to lower operator workload, shorter time required for intervention, and higher user satisfaction.
- H3: Sharing the recognition information of the automated system with the operator improves intervention performance.
- H4: Recognition intervention via a touch display is easier and more accurate than using a simple button interface.

Prediction of pedestrian crossing intention was selected as the recognition task requiring operator intervention, which is one of the most difficult tasks for automated systems to perform [22, 23]. On the other hand, humans can generally predict pedestrian behavior well based on their personal experience, understanding of human behavior, and intuition. Therefore, this task is well suited for our cooperative recognition framework.

To conduct the evaluation in a realistic driving scenario involving interaction between pedestrians and vehicles, a two-lane road in an urban area with sidewalks close to the roadways was constructed in the CARLA simulator environment [67]. The ego vehicle drives at a speed of 50 km/h, a common speed limit in Japan. The existing Autoware [68] was used to control the ego vehicle, and the recognition assistance interface was implemented as a component of the ADS.

A. Participants

We recruited 18 participants (10 males and 8 females) for our experiment, all of whom agreed to participate in the experiment and signed informed consent forms. All of the participants had a driver's license. Their average age was 32.8 (*SE* = 11.3, 22–55 years old), the average length of their driving experience was 11.1 years (*SE* = 9.7, 1–27 years), and their average driving frequency was 12.1 days in a month (*SE* = 14.2, 0–31 days). Seven participants had previously driven in a simulator, and five had ridden in a vehicle equipped with either an ADS or a driver assistance system.

B. Intervention Scenarios

The driving route was a 620-meter rectangle within the CARLA/Map03 urban road environment, as shown in Figure 3. While cruising along an urban road at 40 km/h to 50 km/h, the ADS encounters a pedestrian on the roadside along one of the red segments shown in Figure 3, whose intention is difficult to recognize. Due to its conservative risk tolerance setting, the ADS predicts that the target will cross the vehicle path and requests that the operator intervenes within 6 seconds of its activation of

FIGURE 3 Intervention methods. BASELINE and CONTROL are conventional control phase intervention methods in which operators use the accelerator to override the ADS, without and with the ability to view the system recognition information, respectively. BUTTON and TOUCH are recognition phase intervention methods in which operators use a push button or touchscreen, respectively, to toggle the recognition states.

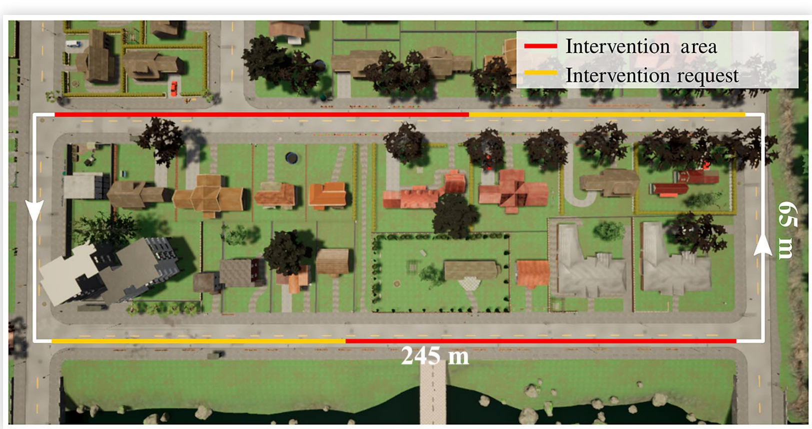
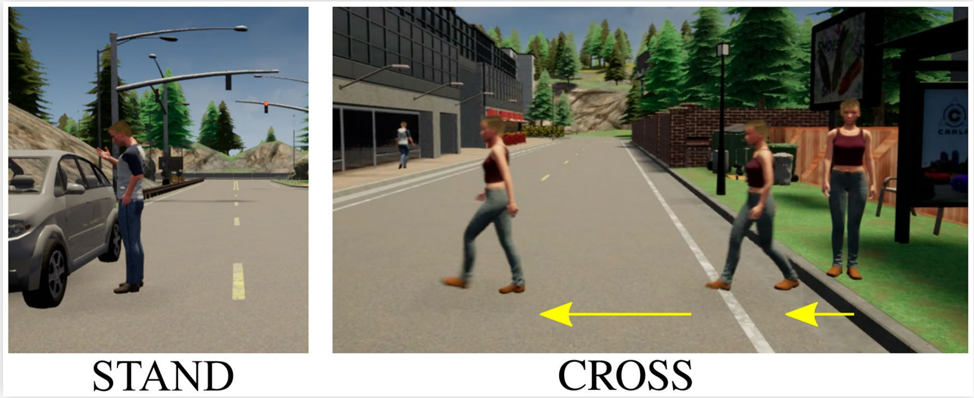


FIGURE 4 Simulated driving route used in the experiment. The intervention targets were placed along the red segments of the route, but the positions of the targets were changed for each trial. Requests for operator intervention were made within the yellow segments.



the intervention request alarm [53]. The vehicle begins decelerating at 0.25 G about 50 m from the pedestrian's location, in an attempt to stop 10 m before encountering the target.

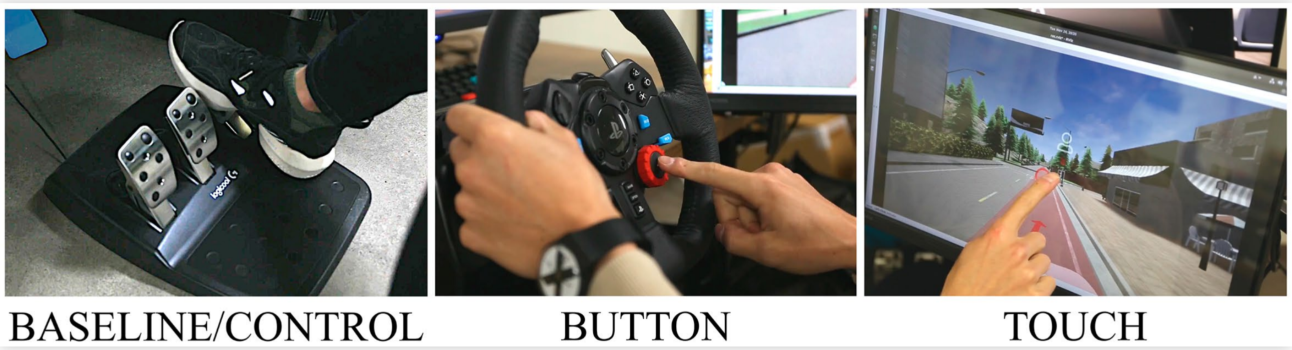
We created two pedestrian behaviors for our intervention targets, as shown in Figure 4. The STAND behavior features a pedestrian standing in the road next to a parked vehicle, without the intention to cross it. He is either using a cell phone or facing away from the road; thus, the ego vehicle can safely drive past without taking any evasive action. The operator needs to perform an intervention within the time limit to avoid unnecessary deceleration. The CROSS behavior features a pedestrian standing on the roadside with an apparent intention to cross the road. The ego vehicle must stop before hitting her and wait until she finishes crossing before continuing. The operator does not need to perform an intervention (see Section IV C) since the system default action is to decelerate, stop, and resume driving.

To simplify the experimental conditions, the automated evasive maneuver was limited to deceleration, and operator input during control intervention was also limited to speed control. Furthermore, only one intervention target appeared at a time; however additional "dummy" pedestrians and vehicles were placed along the route. All of these potential obstacles, including the intervention targets, were placed at different locations during each trial to prevent the participants from easily identifying the intervention targets by themselves.

C. Intervention Methods

The intervention methods compared in this experiment are shown in Figure 5. BASELINE and CONTROL are conventional, control interventions. The operator holds the steering wheel while autonomous driving. When the intervention was requested, the vehicle decelerates until the target pedestrian has been passed because the ADS always recognizes the intervention target as a risk. If the pedestrian had no risk and the deceleration is unnecessary, the operator pushes the accelerator pedal to keep the vehicle speed. The control is transferred back to ADS when the operator releases the pedal. Driver intervention under the CONTROL condition is the same as during the BASELINE condition, but the recognition information from the ADS is displayed to the operator in order to test H3.

FIGURE 5 Intervention scenarios. Pedestrians are placed along the vehicle route as intervention targets. STAND: The pedestrian on the left is using a smartphone or facing the roadside and has no intention of crossing the roadway. CROSS: The pedestrian on the right intends to cross the road and crosses after the vehicle stops.



BUTTON and TOUCH are recognition phase interventions. The participant toggles the predicted intention of the target pedestrian ("cross" or "not cross") by pressing a button on the steering wheel or touching the intervention target on the display screen, respectively, when the intervention is requested by the target pedestrian (the target is highlighted on the display). If the pedestrian has no crossing intention, the participant needs to perform the intervention since the default recognition state is always "cross." The intervention results are immediately reflected in the recognition state, and the driving behavior is calculated by the ADS. The participant toggles the state until the vehicle passes through the target, and they are encouraged to complete the intervention early as possible to avoid the unnecessary deceleration.

We designed BUTTON as the simplest interface manipulation method and TOUCH as a method that unifies the display of the recognition information from the ADS with the manipulation interface to test H4.

D. Implementation

The configuration of the simulator, ADS, and recognition assistance interface used in this experiment are shown in Figure 6. CARLA [67], the simulated driving environment used in this experiment, is an open-source, high-fidelity driving simulator that includes interfaces for the ADS, such as simulated sensors and vehicle controllers. The ADS and recognition assistance interface implemented in CARLA can also be easily installed in a real vehicle by replacing the simulator interface with an in-vehicle interface, and the simulated sensors with the real sensors of the vehicle.

FIGURE 6 Overview of the entire system used in the experiment. The proposed recognition assistance interface is shown incorporated into the ADS. The recognition results used in the experiment were obtained from the simulator environment (dashed lines) in order to reduce the real-time computational load.

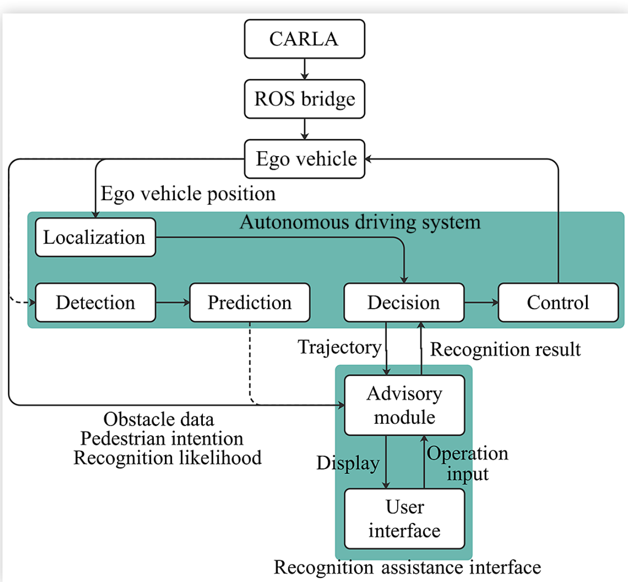
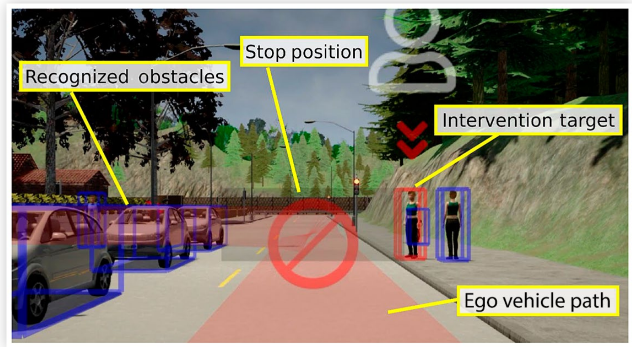


FIGURE 7 Example of the recognition information shared with the experiment participants. The bounding boxes indicate recognized obstacles. The red bounding box identifies the intervention target, which in our experiment is a pedestrian the ADS predicts has the intention of crossing the road. The red lane represents the planned trajectory of the ego vehicle. The red icon in the ego vehicle lane indicates the planned stopping position of the ego vehicle to avoid a possible collision with the intervention target.



Communication between CARLA and Autoware was realized using Python API provided by CARLA and ROS [69]. We obtained recognition information, including predicted pedestrian intentions, from CARLA instead of from the recognition functions of Autoware due to computational limitations. OpenPlanner [70] was used as the path-planning module. An example of the recognition information shared with the experiment participants is shown in Figure 7, where bounding boxes are used to indicate recognized obstacles. The red bounding box, which identifies the intervention target, signifies that the target pedestrian is predicted to have the intention of crossing the road. The red lane represents the planned trajectory of the ego vehicle. The red icon in the ego vehicle lane indicates the stopping position of the ego vehicle to avoid a collision with the intervention target.

Pure Pursuit and the Twist filter in Autoware were used as the vehicle control module. Behavior and recognition results for traffic agents other than the ego vehicle were controlled by Python API. The recognition likelihood of the intervention targets was set to less than 1.0, and likelihoods for the other obstacles were set to 1.0, in order to help the recognition assistance interface find the intervention targets.

The recognition assistance interface was implemented as described in Section III. The advisory module obtains front-view camera images and recognition information from the simulator, while the ego vehicle path is obtained from Autoware. The threshold for the recognition likelihood (0.0–1.0) was set to 1.0. Operator interventions were requested 6 seconds before automatic deceleration. The timing of the intervention requests was calculated using the future trajectory and current speed of the vehicle. Under recognition intervention conditions BUTTON and TOUCH, the recognition module updated the recognition results based on the operator's manipulation of the recognition assistance interface, and the results were then inputted into the ADS.

The target selection algorithm was not implemented because the driving scenarios were controlled to only show one intervention target at a time. The synthesis algorithm for merging user intervention and system recognition information was also not implemented, so the operator's intervention results were always prioritized.

The user interface, which included interactive markers and 3D visualizations from the Rviz tool, as shown in Figure 7, was implemented using ROS. Obstacles, the future trajectory of the vehicle, and the "stop" icon indicating the stopping position of the ego vehicle were all superimposed on the front-view camera image. The intervention target obstacle was surrounded by a red bounding box, while the bounding boxes of the other obstacles were blue. The touch display used to receive the operator's recognition input was also implemented using the Rviz tool. When the operator touched a target highlighted within a red bounding box, the box turned blue as verification feedback.

E. Equipment

The equipment used in this experiment is shown in Figure 8. The interfaces used for control intervention are a Logitech G29 steering wheel and pedal. A 21-inch, 1080 × 1980 pixel touch display was used for sharing the recognition information of the automated system, and as the manipulation interface for the TOUCH method. The size of the intervention target on the display, screen resolution, and available space inside a real vehicle were all taken into account when selecting the size of the touch display. The driver's view of the simulated driving environment was projected onto a 43-inch, 3840 × 1080 pixel display to obtain a sufficient field of view. The speed of the vehicle and current number of completed laps were displayed on the dashboard area of the simulation display. The speaker used to issue intervention requests was built into the forward display.

FIGURE 8 Driving simulator used in our experiment. The touch display shows the current recognition information of the system during the CONTROL, BUTTON, and TOUCH trials, and is the input device for TOUCH interventions. The large forward display shows the driver's view of the road ahead. The steering wheel is the home position for the hands of the participants. The button on the steering wheel is the interface for BUTTON interventions, while the accelerator is the interface for BASELINE and CONTROL interventions.

F. Non-driving-Related Tasks

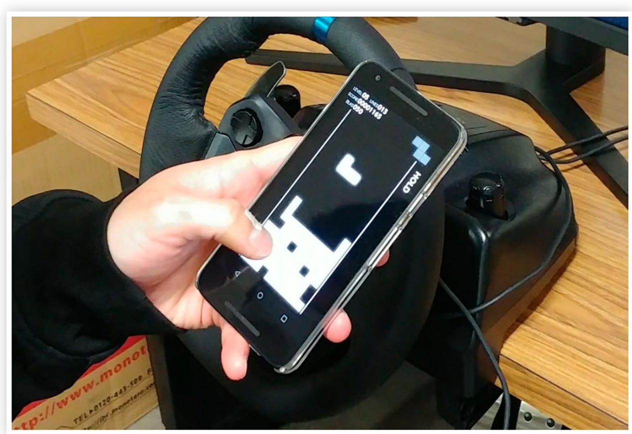
When evaluating human-machine interfaces, it is necessary to do so under conditions as close to actual usage conditions as possible. For example, it has been shown that operators prefer to engage in non-driving-related tasks (NDRT), such as reading a book, playing a game, or watching videos during automated driving [71, 72]. Lee et al. [73] investigated the effects of NDRTs on driving performance when drivers were asked to perform control interventions in response to a takeover request. They showed that cognitively demanding NDRTs negatively affected vehicle control by the drivers more than visually or physically demanding tasks.

In our experiment, we selected playing the game Tetris on a smartphone (Figure 9) as our NDRT [55]. The game requires the player to rotate puzzle pieces falling from the top of the screen in order to stack them with as few gaps as possible at the bottom. When an intervention request was alerted by the automated system, the participant paused the game and conducted the intervention. By deflecting the gaze away from the driving, the timing at which the participant starts to make decisions for intervention can be controlled.

G. Experimental Procedure

First, an outline of the experiment and its purpose was presented to the participants, and informed consent forms

FIGURE 9 The participants played Tetris during the experiment as an NDRT.



were signed. Snacks were then given to the participants as compensation for their participation. Next, they answered questions about their age, driving history, driving frequency, and experience with driver assistance systems, ADSs, and driving simulators. The participants then practiced using the four intervention methods (BASELINE, CONTROL, BUTTON, and TOUCH) while engaging in the NDRT during a simple driving scenario. They were allowed to practice repeatedly until they were able to meet the goal of "maintaining cruising speed as much as possible while driving safely." To encourage them to focus on the NDRT, we told them that additional snacks would be given to the three participants who achieved the highest Tetris scores.

An overview of the simulated driving task is shown in Figure 10. Three-minute driving sessions were consecutively repeated three times. Each driving session had NDRT and an intervention scenario using one of the intervention methods. The CROSS and STAND scenarios were randomly introduced. We allowed 90–120 seconds of the NDRT before participants received an intervention request to allow the participants to sufficiently concentrate on the NDRT [74, 75].

After the three simulated driving sessions using a particular intervention method were completed, workload surveys using NASA-TLX RAW [76] were conducted and breaks were allowed if necessary. Driving then resumed using a different intervention method. Each participant tested all four intervention methods in a randomized order. Intervention methods were counterbalanced to avoid the order effect.

After all of the simulated driving sessions were completed, participants ranked the intervention methods in order of their preferability and wrote down their impressions and opinions regarding the experiment and the intervention methods. The total time required for each participant to complete the entire process was about 90 minutes.

V. Results

The driving data, intervention operation data, and subjective evaluation results obtained from the participants during the experiment were statistically analyzed. Driving speed during the STAND intervention scenario was analyzed to evaluate differences in safety and driving efficiency when using the cooperative recognition intervention methods versus control interventions. Driving speed from 50 m in front of the target up to the target was also evaluated since the ADS began decelerating 50 m before reaching the targeted pedestrian. Their minimum and average speeds, as well as speed fluctuation (i.e., standard deviation of speed), were used as indicators of driving efficiency since our definition of higher efficiency is a lower level of unnecessary deceleration while maintaining stable control. Minimum time-to-collision (TTC) in the CROSS scenario was used as an indicator of driving safety. In our experiment, TTC was the time remaining until the ego vehicle collided with the crossing pedestrian if the vehicle maintained its current speed.

Our experimental results are shown in Table 1. After performing the Shapiro-Wilk test ($\alpha = 0.05$) and Levene test ($\alpha = 0.05$), the Friedman test was used as an analysis of variance (ANOVA) ($\alpha = 0.05$). Since significant differences were found only in speed fluctuation ($p = 0.028$, $F(3, 68) = 9.13$), a Conover test was used for post hoc analysis. The results showed significant differences between BASELINE-TOUCH ($p = 0.045$), CONTROL-TOUCH ($p = 0.018$), and BUTTON-TOUCH ($p = 0.010$). To compare the levels of driving efficiency and safety when using each intervention method, distributions of minimum speed and minimum TTC were investigated.

Figure 11 shows the percentage of driving at a speed higher than the threshold values (x-axis values) during the STAND scenario. We can see that when participants used the TOUCH intervention method, they were able to maintain a

FIGURE 10 Overview of the experiment. The intervention targets appeared in randomized and counterbalanced order.

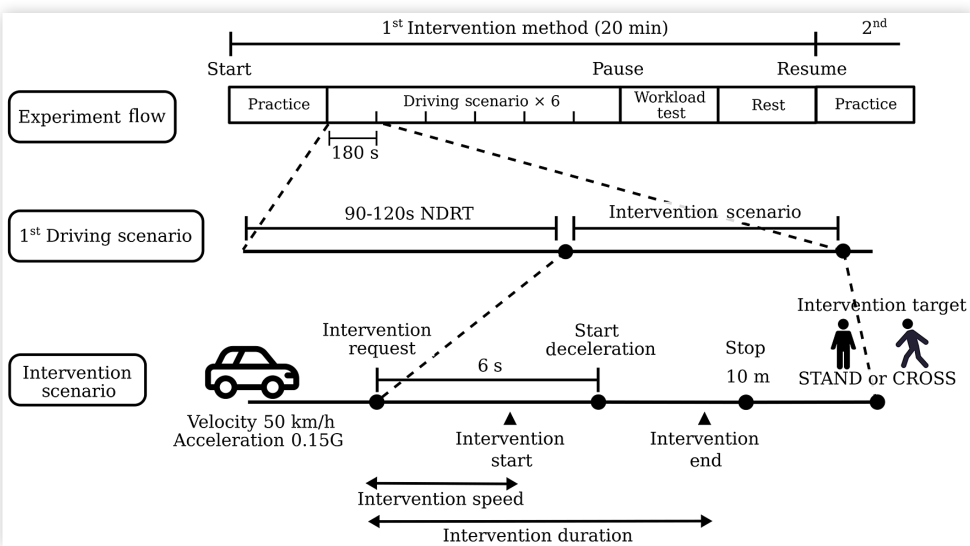


TABLE 1 Participant driving behavior during interventions.

	BASELINE		CONTROL		BUTTON		TOUCH		
	M	SD	M	SD	M	SD	M	SD	F(3, 68)
Minimum speed [km/h]	18.30	10.78	18.48	11.57	18.08	14.18	23.26	12.75	6.60
Average speed [km/h]	29.67	8.80	29.21	9.84	29.95	10.60	33.67	8.73	5.80
Speed fluctuation [km/h]	8.77	3.03	8.70	2.82	8.70	4.16	7.10	3.58	9.13*
Minimum TTC [s]	3.45	1.03	3.45	1.14	3.37	1.36	3.65	0.93	0.67

(*:$p < 0.05$) M = mean, *SD* = standard deviation, F = F-test distribution

higher speed than when using the other methods. BUTTON method, which employed the same cooperative recognition strategy as TOUCH, had a higher percentage of overthreshold driving at speeds of 25 km/h or higher than the control interventions, but also had a lower percentage at 0 km/h.

Figure 12 shows the percentage of driving which kept TTC to the pedestrian higher than the threshold values (x-axis values) in the CROSS scenario. Due to the parameters of the ADS, these percentages dropped sharply after 3.5 seconds. The TOUCH method achieved the highest TTC (i.e., was the safest). TTCs for the BUTTON method were higher than the control interventions at over 3.0 seconds, but lower at less than 3.0 seconds. No differences were found between the BASELINE and CONTROL methods.

Although we mentioned in Section III that safety should always be ensured by the ADS, even when intervention errors occur, such a capability was not implemented in this experiment. As a result, uncorrected interventions appear in the driving data. Therefore, intervention accuracy was evaluated using the participants' driving behavior in each scenario. Intervening to maintain cruising speed in the STAND scenario and not intervening (i.e., allowing the vehicle to yield to the pedestrian) in the CROSS scenario were the correct driving behaviors.

Table 2 shows the number of trials in which the vehicle yielded (i.e., $minimum\_speed \leq 1.0$ km/h) or maintained speed (i.e., $minimum\_speed > 1.0$ km/h) when encountering pedestrians in the two driving scenarios. ACCURACY is the percentage of correct driving behavior observed. Our experimental results show that the TOUCH intervention method was the most accurate, while the BUTTON method was the least accurate, with incorrect driving behavior observed twice as often in the CROSS scenario as when using the TOUCH method. Furthermore, the CONTROL method was less accurate than the BASELINE method.

To evaluate differences in the amount of time required for intervention, intervention speed (time elapsed from the intervention request to intervention initiation) and intervention duration (time elapsed from intervention initiation to termination) were compared. The results are shown in Figure 13. The Friedman test was used after the Shapiro-Wilk

FIGURE 11 Percentage of driving during which the operator maintained a speed higher than the threshold values. The percentages were only calculated during the STAND scenario, which required operator intervention to maintain vehicle cruising speed instead of automatically decelerating.

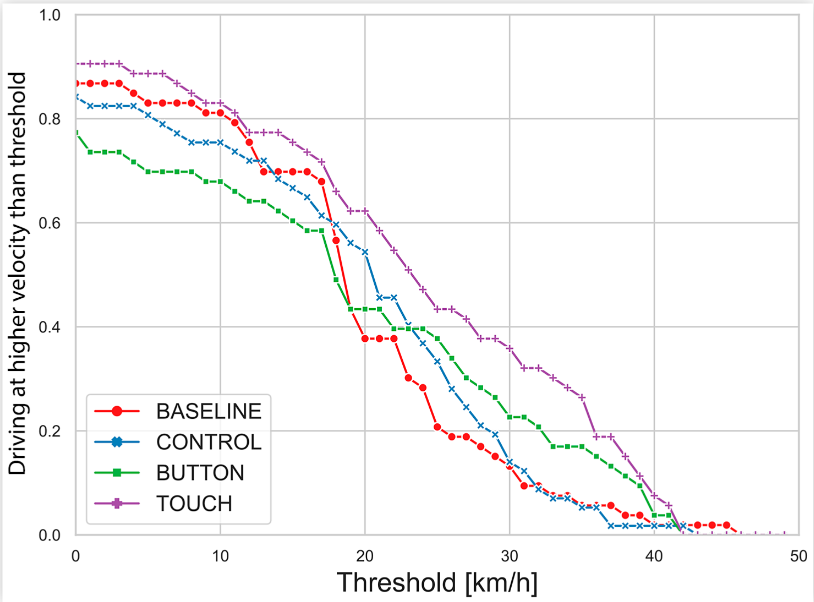


FIGURE 12 Percentage of driving that exceeded the threshold speed, by speed, for each intervention method during the STAND scenario, which required operators to intervene in order to maintain their speed. The y-axis indicates the percentage of driving during which minimum TTC was higher than the x-axis threshold value.

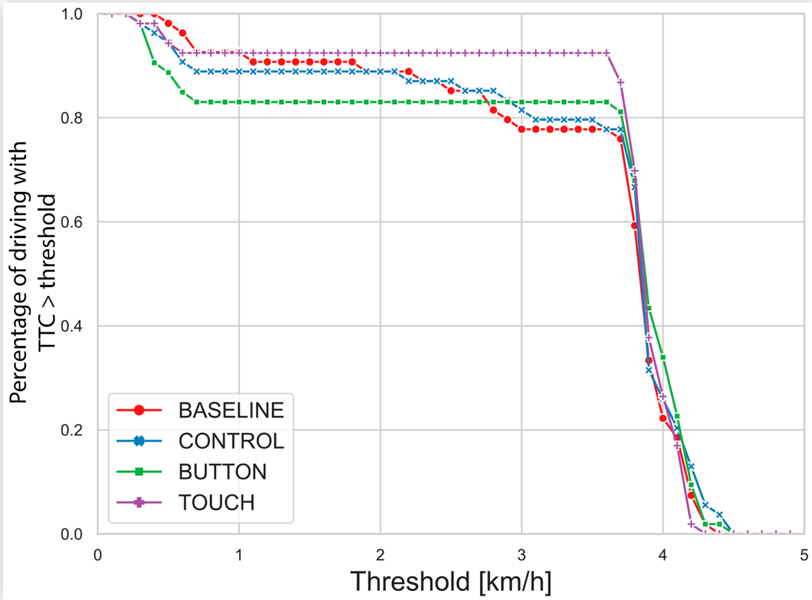
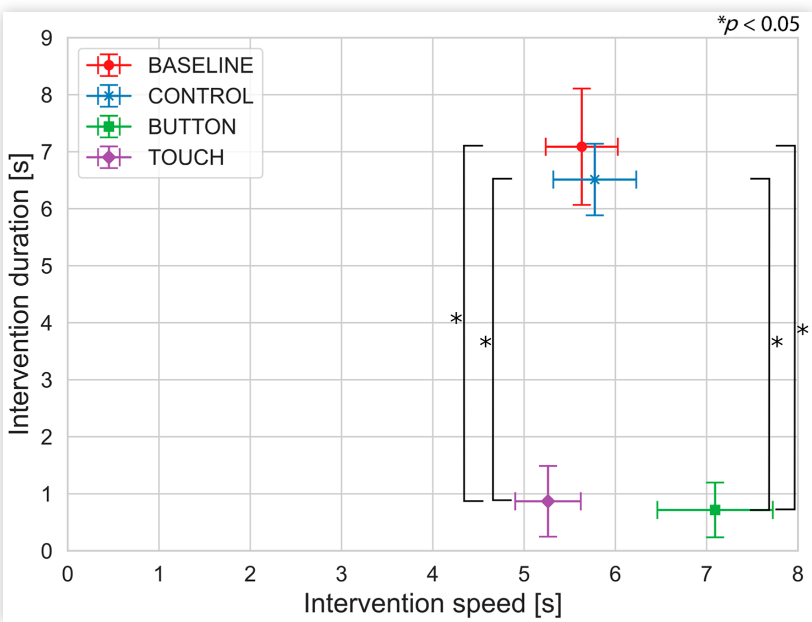


TABLE 2 Participant intervention behavior and accuracy rates for each intervention method.

		BASELINE	CONTROL	BUTTON	TOUCH
STAND	Yield	7	10	14	5
	Keep	46	47	39	48
CROSS	Yield	50	48	45	49
	Keep	4	6	8	4
ACC		0.89	0.87	0.79	0.91

Speed < 1.0 km/h is counted as yield. ACC is the rate of the correct behavior (stand-keep and cross-yield).

FIGURE 13 Average intervention speed (from ADS intervention request to operation initiation) and average intervention duration (from intervention initiation to termination). Shorter times are better. Error bars represent standard deviations.



test ($\alpha = 0.05$) and Levene test ($\alpha = 0.05$). No significant difference in intervention speed among the four intervention methods was observed ($p = 0.18$, $F(3, 68) = 4.86$). However, a significant difference was observed in intervention duration ($p = 0.001$, $F(3, 68) = 44.22$), so the Conover test was used as a post hoc test. The BUTTON and TOUCH methods had significantly shorter intervention durations than the BASELINE and CONTROL methods (BASELINE-BUTTON, $p = 0.001$; BASELINE-TOUCH, $p = 0.001$; CONTROL-BUTTON, $p = 0.001$; CONTROL-TOUCH, $p = 0.001$).

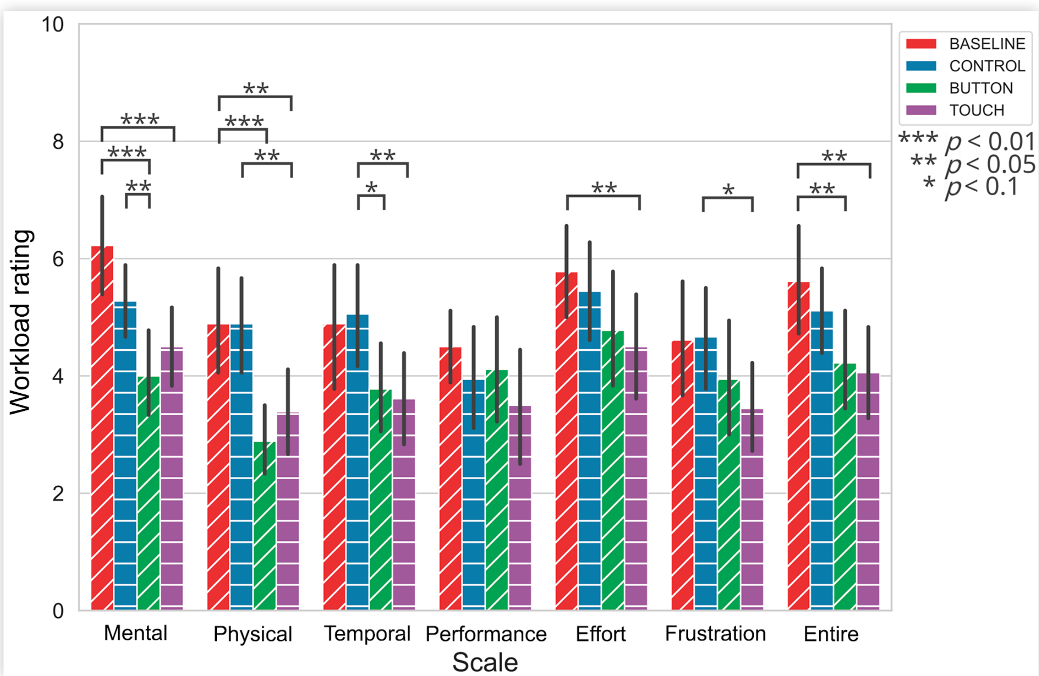
Subjective workloads while using each intervention method were collected from participants using the NASA-TLX RAW questionnaire on a scale from 1 to 10, where a larger number represents higher demand on the operator. We checked each scale for normality and equivariance using the Shapiro-Wilk test ($\alpha = 0.05$) and Levene test ($\alpha = 0.05$). When normality and equivariance were confirmed, we then conducted a one-way ANOVA ($\alpha = 0.05$) and multiple comparisons using Tukey's test ($\alpha = 0.05$). When normality and equivariance were not confirmed, Friedman's test ($\alpha = 0.05$) and Conover's test ($\alpha = 0.05$) were used. The results of the subjective workload survey are shown in Figure 14. Friedman's test showed significant differences in all of the workload categories: mental demand $p = 0.001$, $F(3, 68) = 17.96$; physical demand $p = 0.001$, $F(3, 68) = 22.35$; temporal demand $p = 0.0060$, $F(3, 68) = 12.44$; dissatisfaction with performance $p = 0.043$, $F(3, 68) = 8.15$; effort required for intervention $p = 0.001$, $F(3, 68) = 20.14$; frustration during intervention $p = 0.014$, $F(3, 68) = 10.57$; total demand $p = 0.001$, $F(3, 68) = 14.92$).

Then, based on the results of the Shapiro-Wilk and Levene tests, Tukey's test was conducted for physical demand, while Conover's test was used for the other scales. There were no differences between BASELINE-CONTROL or between BUTTON-TOUCH. For BASELINE-BUTTON, the following significant differences were observed: mental ($p = 0.001$), physical ($p = 0.001$), effort ($p = 0.001$), entire ($p = 0.012$). For BASELINE-TOUCH, the following significant differences were observed: mental ($p = 0.001$), physical ($p = 0.001$), temporal ($p = 0.019$), performance ($p = 0.001$), effort ($p = 0.001$), frustration ($p = 0.013$), entire ($p = 0.001$). For CONTROL-BUTTON, the following significant differences were observed: mental ($p = 0.001$), physical ($p = 0.001$), temporal ($p = 0.019$), effort ($p = 0.036$). For CONTROL-TOUCH, the following significant differences were observed: physical ($p = 0.001$), temporal ($p = 0.001$), effort ($p = 0.001$), frustration ($p = 0.0063$), entire ($p = 0.020$).

At the end of the experiment, users ranked their preferred intervention methods from first (most preferred) to fourth (least preferred). BUTTON was the most preferred method ($M = 1.61$, $SD = 0.60$), followed by TOUCH ($M = 1.83$, $SD = 0.70$), CONTROL ($M = 3.06$, $SD = 1.00$), and, finally, BASELINE ($M = 3.50$, $SD = 0.76$).

Feedback comments were also collected from the participants. Regarding the design of the experiment, some participants reported that it was difficult to judge the pedestrians' intentions. There was also a discrepancy between participants' perceptions of risk and the intervention requests of the automated system, leading to a decrease in trust in the system. As for the design of the user interface, some participants reported

FIGURE 14 Results of participant NASA-TLX RAW workload evaluations. Lower is better. Error bars indicate standard deviations.



that it took a long time for them to understand the cause of the automated deceleration of the vehicle when no recognition information was presented (i.e., when using the BASELINE method). Other comments noted that the bounding box hindered the visibility of the target pedestrian, that the TOUCH intervention method was easy to operate because of its familiarity, that TOUCH intervention errors led to frustration, and that magnification of the forward view image would be a useful function since it was difficult to recognize the intentions of pedestrians from a distance.

VI. Discussion

The safety of the participants' driving when using each intervention method was evaluated using TTC for pedestrians in the CROSS scenario, while driving efficiency was evaluated using the average, minimum, and fluctuation of speed in the STAND scenario. Statistical analysis revealed that when using the cooperative recognition methods, there was significantly less fluctuation in the ego vehicle speed than when using the control intervention methods. This suggests that cooperative recognition can reduce the level of disruption that occurs when drivers take over speed control during the control phase.

More detailed analyses of minimum speed and TTC were also conducted. No trend in minimum speed was observed between the cooperative recognition and control intervention methods. Regarding safety, TTC intervals of less than 1.5 seconds are considered to be critical, while TTCs of 2 seconds to 4 seconds are generally considered to be desirable [77]. At critical TTC, there was no observable difference between the cooperative recognition and control intervention methods; however, within the desirable TTC window, cooperative recognition was found to be superior.

Therefore, H1 was statistically proven only for driving efficiency. A trend was observed in driving safety, however, due to the huge difference between the BUTTON and TOUCH methods, indicating that the design of the interface influences performance during cooperative recognition, which will be discussed further below.

To evaluate the usability of the intervention methods, intervention time in the STAND scenario, operator workload reported in subjective evaluations, and order of preferred method were compared. These results show that the cooperative recognition intervention methods involved less operator workload, were more preferable, and took less time than the control intervention methods, confirming H2. We believe the reasons for the higher usability of the cooperative recognition methods are the absence of operator intervention in vehicle control, as well as their simplicity of operation.

As for intervention duration, it is obvious that the cooperative recognition methods involved shorter interventions than the control intervention methods since the latter require continuous intervention by the user until the vehicle passes the targeted pedestrian, as was quantitatively demonstrated in this experiment. The shorter time required for intervention during cooperative recognition not only improves usability but also leads to more efficient intervention, which would be especially useful in situations where one operator is intervening remotely in the control of multiple vehicles. However, in this study, intervention duration was measured based only on the duration of the intervention operation, while cognitive processing on the part of the operator was not taken into account. Therefore, more detailed verification using other modalities, such as eye gaze measurement, will be necessary.

Regarding H3, selected results for the CONTROL and BASELINE methods were compared, methods which only differed as to whether or not the recognition information of the automated system was shared with the operator, respectively. However, no differences were observed in intervention accuracy, intervention speed, or subjective evaluation. This is likely because the intervention target was always one pedestrian, and the participants could detect the intervention targets without any additional information from the ADS. However, intervention accuracy when using the CONTROL method was lower than when using the BASELINE method. This was likely due to the operators visually confirming the system recognition information before taking over, which we believe divided their attention [29]. This is consistent with the relationship observed between BUTTON and TOUCH, which will be described later.

Feedback from the participants indicated that the presentation of recognition information was helpful and that intervention was much easier as a result. On the other hand, some participants reported discrepancies between the targets of the intervention requests and their perceived level of risk, which reduced their reliance on the ADS. It should also be noted that previous studies have expressed concerns about increasing the amount of information drivers are required to process in complex driving environments; thus, the information displayed needs to be designed carefully to take into account the operator's cognitive load, mental model, clutter costs, and cognitive tunneling risks [27, 28].

To test H4, we compared the BUTTON and TOUCH intervention methods, two cooperative recognition intervention methods which use different interface manipulation methods. Although BUTTON intervention was expected to perform as well or better than the TOUCH intervention method, since it was the simplest manipulation method, BUTTON interventions had the worst accuracy and slowest intervention times of the four methods tested. TOUCH interventions, on the other hand, yielded the best performance in terms of both intervention speed (Figure 13) and accuracy (Table 2), maximizing driving safety. This may be due to the lower cognitive load when using the touchscreen interface, the intuitiveness of its operation, and the user's undivided focus since the cognitive and operational targets are unified [27, 29]. However, even when the TOUCH method was used for intervention, unsafe driving still occurred, e.g., failing to yield to the pedestrian who intended to cross in front of the vehicle. This was likely due to the operator's difficulty in predicting the intentions of the simulated pedestrians, which confirms the importance of the overall design of the ADS in order to ensure a minimum level of safety even when operator intervention errors occur, as discussed in Section III-A.

Regarding H3 and H4, our experimental results show that sharing the ADS recognition information during cooperative recognition results in dividing the attention of the operator, reducing intervention performance. However, integrating the information display and the manipulation target seems to effectively address this problem, as demonstrated by the superior performance of the TOUCH method.

VII. Conclusion

In this article, we have proposed a recognition assistance interface that realizes cooperative recognition during autonomous driving to resolve challenges in conventional control intervention methods. The proposed interface shares recognition information from the ADS with the operator and allows the operator to use this information to intervene during the recognition phase of the autonomous driving process. This approach is intended to solve control take-over issues in today's Level 3 ADS by conservatively designing the recognition system and having a human operator filter ambiguous recognition results. As demonstrated in this study, driving efficiency and safety increased when using the proposed cooperative recognition method, improvements which are difficult to achieve through modification of automated systems alone.

We first explained the cooperative recognition framework in detail, as well as the design of the proposed recognition assistance interface and the recognition system of the ADS. We then described our evaluation experiment, in which we implemented our proposed interface in a simulated driving environment using an existing ADS. Prediction of the road-crossing intentions of pedestrians at locations without a signal was selected as the operator task, and two driving scenarios were developed for our evaluation.

We then compared the intervention performance of 18 participants when using two cooperative recognition methods and two control intervention methods. The control and cooperative recognition intervention methods were compared with and without the sharing of recognition information from the ADS, and two different manipulation interfaces, a button and a touchscreen, were evaluated when using the cooperative recognition approach. Driving safety (minimum TTC), driving efficiency (the average, minimum, and fluctuation of vehicle speed during intervention), intervention performance (speed and accuracy), and usability (workload, user preference, and time required for intervention) when using each method were compared.

Our results showed that cooperative recognition resulted in more stable vehicle control during interventions and higher usability. However, sharing ADS recognition information with the operator resulted in divided attention, reducing intervention performance in terms of both accuracy and speed. But when the system recognition information and the manipulation target were unified through the use of a touch display, this problem was resolved. However, even when using this best-performing method (cooperative recognition with touch display intervention), unsafe driving behavior was still sometimes observed, confirming that cooperative recognition should be paired with a minimum level of safety provided through emergency control intervention by the ADS.

We also obtained important insights for improving the design of the proposed recognition assistance interface. The intervention targets could not be fully perceived when they appeared outside the displayed area or were displayed too small when distant, and they were sometimes located in occluded areas. It will be necessary to redesign the information display so that it enables the user to better understand the spatial information being presented (e.g., the relative positions of obstacles in relation to the ego vehicle) and allows the operator to perceive detailed information about the targets.

Further evaluation of our cooperative recognition approach, both of the recognition interface and of the ADS, will be necessary under more practical conditions. For example, the current limited recognition ranges may not allow a sufficient time window (6 seconds [53]) for operator intervention.

In addition, a variety of recognition tasks other than pedestrian intention need to be considered, such as predicting the future trajectories of other road users and occlusion risks in the surrounding environment, which are similar to pedestrian intention prediction tasks. And these tasks will need to be performed simultaneously. However, relatively simple tasks such as traffic light classification, and more complex tasks such as risk prediction in complex traffic situations, are likely to have different characteristics and may even require different interface designs. Therefore, it is also necessary to consider information display methods that take into account complex traffic situations, as well as the operator's cognitive load and processing capabilities. Since the positive effect of integrating recognition and manipulation targets is indicated by the results of our experiment, we will consider various interface modalities such as gaze input and VR/AR (virtual reality and augmented reality) displays.

The method of selecting targets for operator intervention requests, and synthesis of the operator's intervention inputs with the recognition results of the autonomous system, briefly discussed in Section III of this article, will also be the focus of future investigation.

Acknowledgments

The authors want to thank all of the study participants. The initial author would also like to take this opportunity to thank the Nagoya University Interdisciplinary Frontier Fellowship, which is supported by the Japan Science and Technology Agency and Nagoya University. This research was funded by the Program on Open Innovation Platform with Enterprises, Research Institutes, and Academia (OPERA) of the Japan Science and Technology Agency (JST), grant JPMJOP1612.

Contact Information

Atsushi Kuribayashi
kuribayashi.atsushi@g.sp.m.is.nagoya-u.ac.jp
Institution: Graduate School of Informatics, Nagoya University
Address: Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601, Japan

Eijiro Takeuchi
takeuchi@g.sp.m.is.nagoya-u.ac.jp
Institution: TierIV Inc.
Address: Mei-eki 1-chome, Nakamura-Ward, Nagoya, Aichi, 464-8601, Japan

Yoshio Ishiguro
ishiguro.yoshio@g.sp.m.is.nagoya-u.ac.jp
Institution: Interfaculty Initiative in Information Studies, The University of Tokyo

Alexander Carballo
alexander@g.sp.m.is.nagoya-u.ac.jp
Institution: Institute of Innovation for Future Society, Nagoya University
Address: Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601, Japan

Kazuya Takeda
kazuya.takeda@nagoya-u.jp
Institution: Graduate School of Informatics, Nagoya University
Address: Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601, Japan

References

1. SAE On-Road Automated Vehicle Standards Committee, "Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems (J3016 201806)," SAE International, accessed December 30, 2021, 2018, https://www.sae.org/standards/content/j3016201806/.
2. Lu, Z., Coster, X., and de Winter, J., "How Much Time Do Drivers Need to Obtain Situation Awareness? A Laboratory-Based Study of Automated Driving," *Applied Ergonomics* 60 (2017): 293-304.
3. Endsley, M.R., "Situation Awareness in Future Autonomous Vehicles: Beware of the Unexpected," in Bagnara, S., Tartaglia, R., Albolino, S., Alexander, T. et al. (Eds.), *Congress of the International Ergonomics Association* (Springer, Florence, Italy, 2018), 303-309.
4. Othersen, I., Petermann-Stock, I., Schoemig, N., and Fuest, T., "Method Development and Interaction Cognitive Driver Take-Over Ability after Piloted Driving," *ATZelektronik Worldwide* 13, no. 2 (2018): 28-33.
5. Sahaï, A., Barré, J., and Bueno, M., "Urgent and Non-urgent Takeovers during Conditional Automated Driving on Public Roads: The Impact of Different Training Programmes," *Transportation Research Part F: Traffic Psychology and Behaviour* 81 (2021): 130-143.
6. Johns, M., Strack, G., and Ju, W., "Driver Assistance after Handover of Control from Automation," in *21st International Conference on Intelligent Transportation Systems (ITSC)*, Maui, HI, IEEE, 2018, 2104-2110.
7. Itoh, M., Inahashi, H., and Tanaka, K., "Overtrust due to Lack of Informing Limit of Ability of Automation and/or Its Reason," *Transactions of the Human Interface Society* 5, no. 2 (2003): 283-290 [in Japanese].
8. Itoh, M., "Contributing Factors to Driver's Over-Trust in a Driving Support System for Workload Reduction," *Transactions of the Society of Instrument and Control Engineers* 45, no. 11 (2009): 555-561 [in Japanese].
9. Parasuraman, R. and Riley, V., "Humans and Automation: Use, Misuse, Disuse, Abuse," *Human Factors* 39, no. 2 (1997): 230-253.
10. Lindemann, P., Müller, N., and Rigolll, G., "Exploring the Use of Augmented Reality Interfaces for Driver Assistance in Short-Notice Takeovers," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, Paris, France, 2019, 804-809.
11. Lv, C., Li, Y., Xing, Y., Huang, C. et al., "Human–Machine Collaboration for Automated Driving Using an Intelligent Two-Phase Haptic Interface," *Advanced Intelligent Systems* 3, no. 4 (2021): 2000229.
12. Filip, G., Meng, X., Burnett, G., and Harvey, C., "Designing and Calibrating Trust through Situational Awareness of the Vehicle (SAV) Feedback," in *IET International Conference on Intelligent and Connected Vehicles (ICV 2016)*, Chongqing, China, 2016, 1-6.
13. Colley, M., Eder, B., Rixen, J.O., and Rukzio, E., "Effects of Semantic Segmentation Visualization on Trust, Situation Awareness, and Cognitive Load in Highly Automated Vehicles," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Yokohama, Japan, 2021, 1-11.
14. Beller, J., Heesen, M., and Vollrath, M., "Improving the Driver–Automation Interaction: An Approach Using Automation Uncertainty," *Human Factors* 55, no. 6 (2013): 1130-1141.
15. Agarwal, S., Terrail, J.O.D., and Jurie, F., "Recent Advances in Object Detection in the Age of Deep Convolutional Neural Networks," arXiv preprint arXiv:1809.03193, 2018.
16. Wang, C.-Y., Bochkovskiy, A., and Liao, H.-Y.M., "Scaled-YOLOv4: Scaling Cross Stage Partial Network," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, 2021, 13029-13038.
17. Redmon, J., Divvala, S., Girshick, R., and Farhadi, A., "You Only Look Once: Unified, Real-Time Object Detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, 2016, 779-788.
18. Fitts, P.M., Viteles, M., Barr, N., Brimhall, D. et al., *Human Engineering for an Effective Air-Navigation and Traffic-Control System* (National Research Council, Ohio, USA, 1951)
19. Joe, J.C., O'Hara, J., Hugo, J.V., and Oxstrand, J.H., "Function Allocation for Humans and Automation in the

Context of Team Dynamics," *Procedia Manufacturing* 3 (2015): 1225-1232.

20. Roth, E.M., Sushereba, C., Militello, L.G., Diiulio, J. et al., "Function Allocation Considerations in the Era of Human Autonomy Teaming," *Journal of Cognitive Engineering and Decision Making* 13, no. 4 (2019): 199-220.
21. Wright, P., Dearden, A., and Fields, B., "Function Allocation: A Perspective from Studies of Work Practice," *International Journal of Human-Computer Studies* 52, no. 2 (2000): 335-355.
22. Rasouli, A., Kotseruba, I., Kunic, T., and Tsotsos, J., "PIE: A Large-Scale Dataset and Models for Pedestrian Intention Estimation and Trajectory Prediction," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, South Korea, IEEE, 2019, 6261-6270.
23. Ghori, O., Mackowiak, R., Bautista, M., Beuter, N. et al., "Learning to Forecast Pedestrian Intention from Pose Dynamics," in *IEEE Intelligent Vehicles Symposium (IV)*, Changshu, China, IEEE, 2018, 1277-1284.
24. Helldin, T., Falkman, G., Riveiro, M., and Davidsson, S., "Presenting System Uncertainty in Automotive UIs for Supporting Trust Calibration in Autonomous Driving," in *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Eindhoven, the Netherlands, 2013, 210-217.
25. Khastgir, S., Birrell, S., Dhadyalla, G., and Jennings, P., "Calibrating Trust through Knowledge: Introducing the Concept of Informed Safety for Automation in Vehicles," *Transportation Research Part C: Emerging Technologies* 96 (2018): 290-303.
26. Detjen, H., Salini, M., Kronenberger, J., Geisler, S. et al., "Towards Transparent Behavior of Automated Vehicles: Design and Evaluation of HUD Concepts to Support System Predictability through Motion Intent Communication," in *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction*, New York, 2021, 1-12.
27. Yeh, M., Merlo, J.L., Wickens, C.D., and Brandenburg, D.L., "Head Up Versus Head Down: The Costs of Imprecision, Unreliability, and Visual Clutter on Cue Effectiveness for Display Signaling," *Human Factors* 45, no. 3 (2003): 390-407.
28. Ververs, P.M. and Wickens, C.D., "Head-Up Displays: Effect of Clutter, Display Intensity, and Display Location on Pilot Performance," *The International Journal of Aviation Psychology* 8, no. 4 (1998): 377-403.
29. Fadden, S., Ververs, P.M., and Wickens, C.D., "Costs and Benefits of Head-Up Display Use: A Meta-Analytic Approach," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 42, no. 1 (1998): 16-20.
30. Flemisch, F.O., Kelsch, J., Schieben, A., Schindler, J. et al., *Prospective Engineering of Vehicle Automation with Design Metaphors: Intermediate Report from the H-Mode Projects*. Vol. 7 (Zentrum Mensch-Maschine-Systeme (Hg.), Berlin, 2007)
31. Kienle, M., Damböck, D., Kelsch, J., Flemisch, F. et al., "Towards an H-Mode for Highly Automated Vehicles: Driving with Side Sticks," in *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Germany, 2009, 19-23.
32. Baltzer, M., Lopez, D., Kienle, M., Flemisch, F. et al., *Dynamic Distribution of Control via Grip Force Sensitive Devices in Cooperative Guidance and Control* (Berliner Werkstatt Mensch-Maschine, Berlin, 2015)
33. Winner, H., Witte, S., Uhler, W., and Lichtenberg, B., "Adaptive Cruise Control System Aspects and Development Trends," SAE Technical Paper 961010, 1996, doi:https://doi.org/10.4271/961010.
34. Okuda, R., Kajiwara, Y., and Terashima, K., "A Survey of Technical Trend of ADAS and Autonomous Driving," in *Technical Papers of 2014 International Symposium on VLSI Design, Automation and Test*, Hsinchu, Taiwan, IEEE, 2014, 1-4.
35. Adell, E., Várhelyi, A., and Dalla Fontana, M., "The Effects of a Driver Assistance System for Safe Speed and Safe Distance – A Real-Life Field Study," *Transportation Research Part C: Emerging Technologies* 19, no. 1 (2011): 145-155.
36. Inoue, S., Ozawa, T., Inoue, H., Raksincharoensak, P. et al., "Cooperative Lateral Control between Driver and ADAS by Haptic Shared Control Using Steering Torque Assistance Combined with Direct Yaw Moment Control," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Rio de Janeiro, Brazil, IEEE, 2016, 316-321.
37. Okamoto, K. and Tsiotras, P., "Data-Driven Human Driver Lateral Control Models for Developing Haptic-Shared Control Advanced Driver Assist Systems," *Robotics and Autonomous Systems* 114 (2019): 155-171.
38. Hegeman, G., Tapani, A., and Hoogendoorn, S., "Overtaking Assistant Assessment Using Traffic Simulation," *Transportation Research Part C: Emerging Technologies* 17, no. 6 (2009): 617-630.
39. Charissis, V. and Papanastasiou, S., "Human–Machine Collaboration through Vehicle Head Up Display Interface," *Cognition, Technology & Work* 12, no. 1 (2010): 41-50.
40. Wu, X., Merenda, C., Misu, T., Tanous, K. et al., "Drivers' Attitudes and Perceptions towards a Driving Automation System with Augmented Reality Human-Machine Interfaces," in *2020 IEEE Intelligent Vehicles Symposium (IV)*, Las Vegas, NV, 2020, 1978-1983.
41. Zimmermann, M., Schopf, D., Lütteken, N., Liu, Z. et al., "Carrot and Stick: A Game-Theoretic Approach to Motivate Cooperative Driving through Social Interaction," *Transportation Research Part C: Emerging Technologies* 88 (2018): 159-175.
42. Andonian, B., Rauch, W., and Bhise, V., "Driver Steering Performance Using Joystick vs. Steering Wheel Controls," SAE Technical Paper 2003-01-0118, 2003, doi:https://doi.org/10.4271/2003-01-0118.
43. Gil, J.J., Díaz, I., Ciáurriz, P., and Echeverría, M., "New Driving Control System with Haptic Feedback: Design and Preliminary Validation Tests," *Transportation Research Part C: Emerging Technologies* 33 (2013): 22-36.
44. Wang, C., Krüger, M., and Wiebel-Herboth, C.B., "'Watch Out!': Prediction-Level Intervention for Automated Driving," in *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, New York, 2020, 169-180.

45. Tall, M., Alapetite, A., San Agustin, J., Skovsgaard, H.H. et al., "Gaze-Controlled Driving," in *CHI '09 Extended Abstracts on Human Factors in Computing Systems, Ser. CHI EA '09*, New York, Association for Computing Machinery, 2009, 4387-4392, https://doi.org/10.1145/1520340.1520671.

46. Manawadu, U.E., Kamezaki, M., Ishikawa, M., Kawano, T. et al., "A Hand Gesture-Based Driver-Vehicle Interface to Control Lateral and Longitudinal Motions of an Autonomous Vehicle," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Budapest, Hungary, IEEE, 2016, 001 785-001 790.

47. Franz, B., Kauer, M., Blanke, A., Schreiber, M. et al., "Comparison of Two Human-Machine Interfaces for Cooperative Maneuver-Based Driving," *Work* 41, no. Supplement 1 (2012): 4192-4199.

48. Walch, M., Woide, M., Mühl, K., Baumann, M. et al., "Cooperative Overtaking: Overcoming Automated Vehicles' Obstructed Sensor Range via Driver Help," in *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, ser. AutomotiveUI '19*, Utrecht, the Netherlands, Association for Computing Machinery, 2019, 144-155.

49. Richardson, J., "Human Factors Research Priorities for ADAS Systems: A UK Perspective," in *2001 ADAS. International Conference on Advanced Driver Assistance Systems, (IEE Conference Publication No. 483)*, Birmingham, UK, 2001, 20-24. IET.

50. Saffarian, M., de Winter, J.C.F., and Happee, R., "Automated Driving: Human-Factors Issues and Design Solutions," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 56, no. 1 (2012): 2296-2300.

51. Schneider, T., Hois, J., Rosenstein, A., Ghellal, S. et al., "ExplAIn Yourself! Transparency for Positive UX in Autonomous Driving," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing* (New York: ACM, 2021), 1-12.

52. Olaverri-Monreal, C., Kumar, S., and Dìaz-Álvarez, A., "Automated Driving: Interactive Automation Control System to Enhance Situational Awareness in Conditional Automation," in *IEEE Intelligent Vehicles Symposium (IV)*, Changshu, China, IEEE, 2018, 1698-1703.

53. Gold, C., Damböck, D., Lorenz, L., and Bengler, K., "'Take Over!': How Long Does It Take to Get the Driver Back into the Loop?" *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 57, no. 1 (2013): 1938-1942.

54. Mok, B.K.-J., Johns, M., Lee, K.J., Ive, H.P. et al., "Timing of Unstructured Transitions of Control in Automated Driving," in *IEEE Intelligent Vehicles Symposium (IV)*, Seoul, South Korea, IEEE, 2015, 1167-1172.

55. Kuehn, M., Vogelpohl, T., and Vollrath, M., "Takeover Times in Highly Automated Driving (Level 3)," in *25th International Technical Conference on the Enhanced Safety of Vehicles (ESV)*, Detroit, MI, National Highway Traffic Safety Administration, 2017, 1-11.

56. Walch, M., Lange, K., Baumann, M., and Weber, M., "Autonomous Driving: Investigating the Feasibility of Car-Driver Handover Assistance," in *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Nottingham, UK, 2015, 11-18.

57. Scott, J. and Gray, R., "A Comparison of Tactile, Visual, and Auditory Warnings for Rear-End Collision Prevention in Simulated Driving," *Human Factors* 50 (2008): 264-275.

58. van Erp, J.B., Toet, A., and Janssen, J.B., "Uni-, Bi- and Tri-Modal Warning Signals: Effects of Temporal Parameters and Sensory Modality on Perceived Urgency," *Safety Science* 72 (2015): 1-8.

59. Bazilinskyy, P., Petermeijer, S.M., Petrovych, V., Dodou, D. et al., "Take-over Requests in Highly Automated Driving: A Crowdsourcing Survey on Auditory, Vibrotactile, and Visual Displays," *Transportation Research Part F: Traffic Psychology and Behaviour* 56 (2018): 82-98.

60. Winner, H. and Hakuli, S., "Conduct-by-Wire–Following a New Paradigm for Driving into the Future," in *Proceedings of FISITA World Automotive Congress*, vol. 22, Yokohama, Japan, 2006, 27-37.

61. Schreiber, M., Kauer, M., Schlesinger, D., Hakuli, S. et al., "Verification of a Maneuver Catalog for a Maneuver-Based Vehicle Guidance System," in *2010 IEEE International Conference on Systems, Man and Cybernetics*, Istanbul, Turkey, 2010, 3683-3689.

62. Kauer, M., Schreiber, M., and Bruder, R., "How to Conduct a Car? A Design Example for Maneuver-Based Driver-Vehicle Interaction," in *2010 IEEE Intelligent Vehicles Symposium*, La Jolla, CA, 2010, 1214-1221.

63. Franz, B., Kauer, M., Bruder, R., and Geyer, S., "pieDrive—A New Driver-Vehicle Interaction Concept for Maneuver-Based Driving," in *IEEE Intelligent Vehicles Symposium (IV)*, Alcala de Henares, Spain, IEEE, 2012.

64. Walch, M., Sieber, T., Hock, P., Baumann, M. et al., "Towards Cooperative Driving: Involving the Driver in an Autonomous Vehicle's Decision Making," in *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Ann Arbor, MI, 2016, 261-268.

65. Inagaki, T. et al., "Adaptive Automation: Sharing and Trading of Control," *Handbook of Cognitive Task Design* 8 (2003): 147-169.

66. Parasuraman, R. and Wickens, C.D., "Humans: Still Vital after All These Years of Automation," *Human Factors* 50, no. 3 (2008): 511-520.

67. Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A. et al., "CARLA: An Open Urban Driving Simulator," in *Proceedings of the 1st Annual Conference on Robot Learning*, Mountain View, CA, 2017, 1-16.

68. Kato, S., Tokunaga, S., Maruyama, Y., Maeda, S. et al., "Autoware on Board: Enabling Autonomous Vehicles with Embedded Systems," in *2018 ACM/IEEE 9th International*

Conference on Cyber-Physical Systems (ICCPS), Porto, Portugal, IEEE, 2018, 287-296.

69. Quigley, M., Conley, K., Gerkey, B., Faust, J. et al., "ROS: An Open-Source Robot Operating System," in *ICRA Workshop on Open Source Software*, vol. 3, no. 3.2, Kobe, Japan, 2009, 5.
70. Hatem, D., Eijiro, T., Kazuya, T., Yoshiki, N. et al., "Open Source Integrated Planner for Autonomous Navigation in Highly Dynamic Environments," *Journal of Robotics and Mechatronics* 29, no. 4 (2017): 668-684.
71. Cyganski, R., Fraedrich, E., and Lenz, B., "Travel-Time Valuation for Automated Driving: A Use-Case-Driven Study," in *Proceedings of the 94th Annual Meeting of the TRB*, Washington, DC, 2015, 19.
72. Fraedrich, E., Cyganski, R., Wolf, I., and Lenz, B., "User Perspectives on Autonomous Driving: A Use-Case-Driven Study in Germany," Technical Report 187, Geography Systems, Department, Humboldt University of Berlin, 2016.
73. Lee, S.C., Yoon, S., and Ji, Y.G., "Effects of Non-Driving-Related Task Attributes on Takeover Quality in Automated Vehicles," *International Journal of Human-Computer Interaction* 37, no. 3 (2020): 211-219.
74. Du, N., Kim, J., Zhou, F., Pulver, E. et al., "Evaluating Effects of Cognitive Load, Takeover Request Lead Time, and Traffic Density on Drivers' Takeover Performance in Conditionally Automated Driving," in *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2020, 66-73.
75. Yoon, S.H. and Ji, Y.G., "Non-driving-Related Tasks, Workload, and Takeover Performance in Highly Automated Driving Contexts," *Transportation Research Part F: Traffic Psychology and Behaviour* 60 (2019): 620-631.
76. Hart, S.G. and Staveland, L.E., "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Advances in Psychology* 52 (1988): 139-183.
77. Mahmud, S.S., Ferreira, L., Hoque, M.S., and Tavassoli, A., "Application of Proximal Surrogate Indicators for Safety Evaluation: A Review of Recent Developments and Research Needs," *IATSS Research* 41, no. 4 (2017): 153-163.