2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)
June 28-30, 2023. Seattle, Washington, USA

A Novel Human-Machine Collaboration Approach for Autonomous Driving with Hand Gesture-based Guidance

Yiran Zhang, Zhongxu Hu, Chen Lv

Abstract— **In highly automated driving vehicles, a human-vehicle interface might still be required for individualization and emergency intervention. We propose a tactical human-vehicle collaboration framework by leveraging the hand-landmark extraction algorithm and the augmented reality visual feedback. The proposed vision-based interface projects the gesture, as the driver's intention, onto the ground and feeds the projection back to the driver through the AR-HUD interface. The projected intention functions as a strategic decision or planning suggestion to the vehicle while collision avoidance, traffic rules compliance, and precise control are realized by the automation algorithm. The feasibility of the framework is validated through an integrated self-driving algorithm combining the risk field, learning-based trajectory prediction, and model predictive control. Comparisons with the conventional manual driving scheme demonstrate that high-level collaboration vastly reduces human physical burdens without compromising driving performance and driver mental workloads.**

I. INTRODUCTION

For future mobility, the necessity of the conventional human-vehicle interface is yet to be decided [1]. As highly automated self-driving does not require continuous supervision and intervention, the driver-vehicle interface seems redundant to normal driving tasks [2], [3]. As for emergent takeovers, the conventional steering-wheel-and-pedal interface requires the driver to directly intervene at the control level [4]. When an emergent intervention requirement arises, the driver might not be well prepared to jump directly into the control loop; it is irresponsible to abruptly throw the entire control task to the driver[5], [6], [7], [8]. Additionally, humans in autonomous vehicles might not be "qualified" to drive, as a human should be a passenger instead of a driver in a highly autonomous vehicle, and enforcing the passenger to acquire a driving license is paradoxical. On the other hand, humans have the inherent instincts to make appropriate plans and decisions in most scenarios [9], [4] regardless of the driving license, which inspires us to design a novel collaboration framework. Regarding the malfunctions of the highly automated intelligence, such as miss detection, incorrect prediction, and over-risky/conservative decision [10], technical issues still exist, while moreover, passengers may prefer to drive in their own styles [11]. Therefore, using a more inclusive, expressive, and transparent interface [12] to coalesce the human's high-level intelligence and the vehicle's lower-level accuracy and swiftness might shed light on the future mobility.

Scarcely any extant autonomous driving algorithm claims to cover all the uncertainties that could possibly take place, in reality [13]. When an unknown emergent situation occurs, most of the automation algorithms are still designed to hand over the final judgment to the human driver. Obviously, the human driver is not always the panacea. For example, in a highly automated self-driving vehicle, the driver is allowed to conduct secondary tasks, which consequently slows down the human's reaction to unanticipated events. Therefore, instead of a brute-force takeover, the authority is steadily released to the driver based on the monitored driver status [14], [15], [16], [17] and driving performance [18], [19] in the shared control scheme. In this scheme, haptic [5], visual [20], and acoustic feedback is suggested to enhance the overall driving performance and assist the driver back to a sufficiently situation-aware status [7], [4], [21]. Nevertheless, the control authority shifting process is essentially a trade-off between the response time and system robustness. Besides, collaboration at the control level requires expertise, which consequently limits the massive deployment of this scheme. The direct control command on the conventional interface exchanges the least amount of interactive information [22]; thus, the shared scheme requires enduring interactions thereupon aggravating human burdens.

Consequently, more expressive and inclusive human vehicle interfaces have been explored [23]. The gesture-based interface [24] interprets the driver's gestures into high-level commands [25], [26], such as changing lanes, acceleration, and deceleration. However, these types of commands can be easily replaced by physical buttons and the commands do not include the driver's comprehension of the scene. On the other hand, brain-computer interface using Electroencephalogram (EEG) [27], [28] is a sophisticated but invasive approach for human-vehicle interaction [29], thus, is only pragmatic for lab-exclusive environments. Both the EEG and gesture interfaces do not comprise tactical information, such as the driver's initiatives, or the exact geographical information. Hence, the visual feedback, especially AR-HUD, is supplemented for more transparent interactions [20]. These two techniques inspired us to form a closed-loop human-vehicle interaction framework.

Some conventional collaboration frameworks are summarized in Figure 1. At the bottom, human and machine inputs are fused via the steering wheel, while above that, monotonous higher-level commands are conveyed to the

This work was supported in part by the SUG-NAP Grant, Nanyang Technological University, the A*STAR AME Young Individual Research Grant (No. A2084c0156), the ANR-NRF joint grant (No.NRF2021-NRF-ANR003 HM Science), and A*STAR MTC Individual Research Grants (No.M22K2c0079) .

The authors are with School of Mechanical and Aerospace Engineering, Nanyang Technological University, 639798 Singapore. The corresponding author: Chen Lv (lyuchen@ntu.edu.sg).

vehicle through a gesture recognition module. Compared to these approaches, our contributions are: 1) a high-level interaction framework using gesture recognition and visual feedback is proposed. More gesture information is extracted representing more tactical and expressive commands imitating a human pointing directions to another human. 2) The model predictive control (MPC) and risk field (RF) coalesce to achieve the proposed collaboration scheme. 3) Human-in-the-loop experiments are conducted to evaluate human efforts using the suggested algorithm.

II. High-level Collaboration Framework

Imitating the approaches through which a human manifests direction to another human via gestures, a natural human-vehicle interaction framework is proposed. The gesture recognition algorithm first extracts the landmarks on the human hand and projects these landmarks onto the ground. Next, the projections are fed back to the driver using the AR-HUD techniques. As is same to the human-human interaction, we demonstrate three types of gesture shown in Figure 1:1) Avoiding obstacle (shared perception/prediction). The human driver uses a fist to highlight the possible obstacles alarming the vehicle to get away from it. 2) Pointing destination (shared decision). The human driver uses an index finger to point to the preferred location. 3) Demonstrating trajectory (shared planning). The driver uses his forearm and hand to demonstrate a possible trajectory that the vehicle could adopt. The gesture is identified using another camera on the right-hand side of the driver. The human driver interacts with the vehicle using gestures and obtains feedback from the windshield interface. The gesture is filtered and projected via the proposed algorithm.

A. Gesture Projection

The camera pixel is projected to the geographical location denoted by (x_t, y_t) as is shown in Figure 2. The projection is given by equation (1). To simplify the problem, we use MediaPipe, which is an open-source library, for human body landmark extraction. The filtered land-mark output in UV coordinate is represented by (x_g, y_g). $c1$, $c2$, $q1$, and $q2$ are the scaling and offset parameters.

$$\begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix} \begin{bmatrix} y_g \\ x_g \end{bmatrix} + \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} x_t \\ y_t \end{bmatrix} \tag{1}$$

where $(x_g, y_g) \in \{(x_{obs}, y_{obs}), (x_{des}, y_{des}), (x_{traj}, y_{traj})\}$ can be the target obstacle position (x_{obs}, y_{obs}), the target destination (x_{des}, y_{des}), and the demonstrated trajectory (x_{traj}, y_{traj}). Then, the car-based coordinate is converted to the global coordinate using

$$\begin{bmatrix} \cos\phi_{ego} & -\sin\phi_{ego} \\ \sin\phi_{ego} & \cos\phi_{ego} \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} + \begin{bmatrix} x_{ego} \\ y_{ego} \end{bmatrix} = \begin{bmatrix} x_r \\ y_r \end{bmatrix} \tag{2}$$

where ϕ_{ego} is the yaw angle of the vehicle; x_{ego} and y_{ego} are the ego vehicle's current position; $(x_r, y_r) \in \{(x_r^{obs}, y_r^{obs}), (x_r^{des}, y_r^{des}), (x_r^{traj}, y_r^{traj})\}$ are the projected gesture position. A sample image of the gesture projection is depicted in Figure 3.

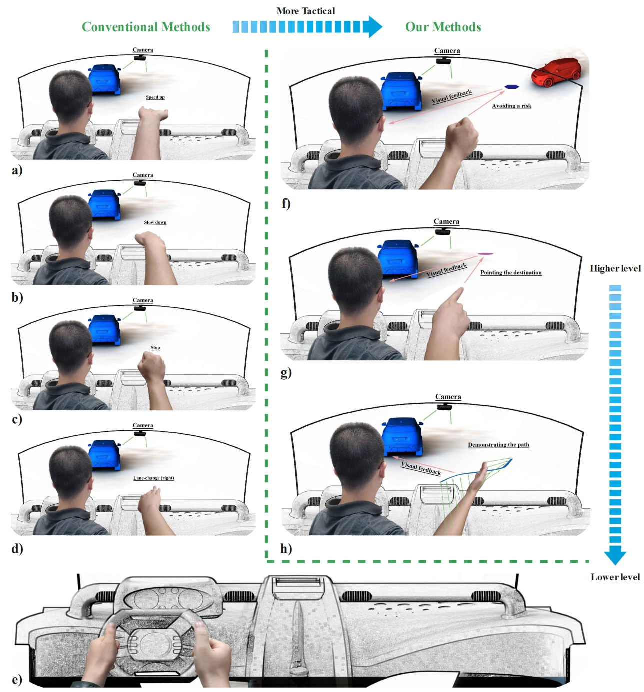

Fig. 1. Comparison of high-level collaboration interface with conventional human-vehicle interaction interface: a) use the palm to issue a speed-up command; b) use the palm to issue a slow-down command; c) use the fist to issue a stop command; d) use the hand to suggest a lane-change; e) collaborate on the steering wheel issuing a maneuver command; f) use the fist to point at a possible obstacle (shared prediction/perception); g) use the index finger to point at a destination (shared decision); h) use the entire arm to demonstrate a possible trajectory (shared planning).

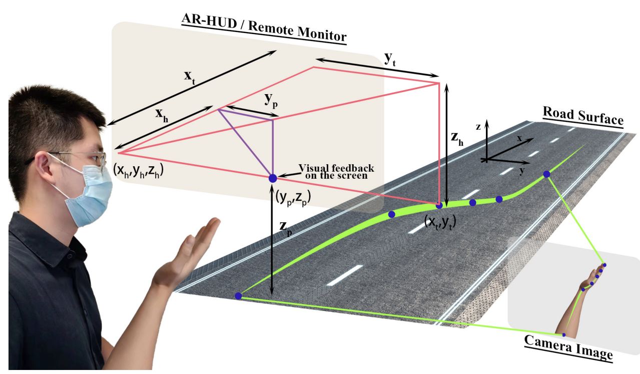

Fig. 2. Human vehicle interface: human driver's gesture is projected on the ground, which is then shown on the AR-HUD interface (for simulation).

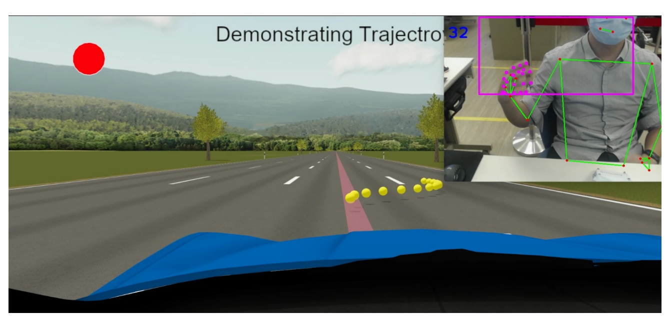

Fig. 3. Gesture projection under the shared planning mode.

Shared perception/prediction: Pointing obstacles can indicate positions where the driver feels unsafe. A case in point is when the driver realizes that the vehicle is not slowing down and heading directly into an obstacle that the driver observes or predicts. In this case, the user can use the fist to point at the obstacles (x_{obs}, y_{obs}), which is miss detected or incorrectly predicted. The obstacle position at the road surface can be projected using equations (1) and (2).

Shared decision: The driver uses the index finger to point at the desired location (x_{des}, y_{des}) for a preferred location. This gesture is opposite to the aforementioned occasion but the UV coordinate to geographical coordinates projection algorithm is the same.

Shared planning: Planning level interaction is more tricky, as it comprises a set of coordinates and the dynamic constraints should be considered. During human-human interaction, one human would use the entire arm to present the complex trajectory. Therefore, we extract the elbow joint and the mid-finger landmarks (a total of 6 points) to represent this trajectory. Considering the moving space of the arm [24] and to make the driver feel comfortable, we use the elbow joint as the origin of the other landmarks. Additionally, these points are not smooth enough to track, thus, are mainly used to control the trajectory generated by B-spline. Therefore, the driver can use gesture to control the shape of the desired trajectory $(x_{r,i}^{traj}, y_{r,i}^{traj})$, which is smooth and physically feasible for the vehicle. The positions indicated by the driver's gestures are only references in this framework. Here, we must consider three circumstances: 1) the positions might not be extremely accurate because of the error generated by the head-pose estimation algorithm; 2) we cannot force the driver to point at a precise location in a moving vehicle; 3) the driver might prefer the adjacent area as well. Therefore, the Gaussian distribution function is used to represent the driver's indicated position but these representations can be replaced by other functions based on the actual precision requirements. These functions can be easily combined with the artificial potential field.

III. Lower Level Controller

A GRU-FRENET-based prediction module forecasts surrounding vehicles' future movements, which are fed to the risk field to describe the future scene. The risk field combined with the interpreted driver intention is then given to the model-based controller, which fuses the above information and casts the final control on the vehicle. The details of the projections are presented in the rest of this section.

A. Trajectory Prediction

To obtain the future positions of the obstacle vehicles, we train a GRU neural network for lightweight maneuver-based classification tasks. The input and output of the neural network are

$$X_{in} = [x^{(t-t_h)}, ..., x^{(t-1)}, x^{(t)}] \tag{3}$$

$$X_{out} = [p_1, p_2, ..., p_{45}] \tag{4}$$

respectively. And $x^{(t)} = [\sum_{m=1}^{M} \frac{1}{\rho_{x,t,m}}, \sum_{m=1}^{M} \frac{1}{\rho_{y,t,m}}, \delta x^{(t)}, \delta y^{(t)}]$. $\sum_{m=1}^{M} \frac{1}{\rho_{x,t,m}}$ is the sum of the inverse of x distance while $\sum_{m=1}^{M} \frac{1}{\rho_{y,t,m}}$ is the sum of the y distance between the predicted vehicle and its surrounding vehicles. Considering the sensory range of the ego vehicle for surrounding vehicle predictions, the positions of the surrounding vehicles are first rotated and represented by the ego-vehicle-based coordinates. δx is obtained through $\delta x^{(t-t')} = x_{obs,m}^{(t-t')} - x_{ego,t}^{(t)}, t' \in [0, t_h], m \in [0, M]$. δy is obtained through $\delta y^{(t-t')} = y_{obs,m}^{(t-t')} - y_{ego,t}^{(t)}$. The output of the neural network is the probability distribution of 45 possible future trajectories, which are generated using [30] with the maneuver based [31] terminal statuses that are determined by 15 longitudinal behaviors and three lane-changing behaviors. The start and terminal status are represented by FRENET states as

$$\begin{gathered}\left[d(t_S), \dot{d}(t_S), \ddot{d}(t_S), s(t_S), \dot{s}(t_S), \ddot{s}(t_S)\right] \\ \left[d(t_T), \dot{d}(t_T), \ddot{d}(t_T), s(t_T), \dot{s}(t_T), \ddot{s}(t_T)\right]\end{gathered} \tag{5}$$

The trajectories are given by solving the parameters of

$$\begin{cases} d(t) = \alpha_{d_0} + \alpha_{d_1} t + \alpha_{d_2} t^2 + \alpha_{d_3} t^3 + \alpha_{d_4} t^4 + \alpha_{d_5} t^5 \\ s(t) = \alpha_{s_0} + \alpha_{s_1} t + \alpha_{s_2} t^2 + \alpha_{s_3} t^3 + \alpha_{s_4} t^4 + \alpha_{s_5} t^5 \end{cases} \tag{6}$$

The training data are the first 50000 samples from Argoverse [32]. The trajectory prediction algorithm will be fed to the risk field combined with the gestures so that the trajectory prediction errors could be compensated by human guidance.

B. Risk Field

The risk field [33] is simplified by considering only three components in Equation (7), i.e., stationary obstacles, dynamic objects, and road boundaries.

$$R = R_c + R_b + R_{nc} \tag{7}$$

where R is the superposition of all the two factors at time t. R_c and R_r are generated by dynamic objects and road boundaries, respectively. R_{nc} is the driver's intention given by the gestures.

C. Model Predictive Control

A simplified vehicle model [34] is used and the integrated solutions are obtained through repetitively solving the following constrained optimization problem

$$\begin{aligned} \arg\min_{u} \sum_{k=1}^{H_p} \{ &||u||_{Q_u}^2 + ||y - y_{ref}^*||_{Q_{\alpha_1}^*}^2 + Q_r R + \\ &Q_{\alpha_2}^*(U_{Neg} + U_{Traj}) + ||v - v_{goal}||_{Q_v}^2 + \\ &||\phi - \phi_{road}||_{Q_\phi}^2 \} \\ s.t.\, &u_{min1} \leq u_1 \leq u_{max1} \\ &u_{min2} \leq u_2 \leq u_{max2} \\ &vmin \leq v_{ego} \leq v_{max} \end{aligned} \tag{8}$$

where H_p is the optimization horizon; v_{goal} is the desired velocity constrained by traffic rules; Q_ϕ is obtained from the road preview information. Q_u, Q_r, Q_ϕ, $Q^*_{\alpha_1}$, $Q^*_{\alpha_2}$, and Q_v are the weights.

IV. Experiment

The human-in-the-loop simulation was conducted to validate the feasibility and effectiveness of the proposed method. We compared the higher-level cooperation mode with manual driving. The study protocol and consent form were approved by the Nanyang Technological University Institutional Review Board (protocol number IRB-2018-11-025). The physical workload was evaluated using EMG criteria and the mental workload was evaluated using EEG as is shown in Figure 4. The experiment results are shown in Table I.

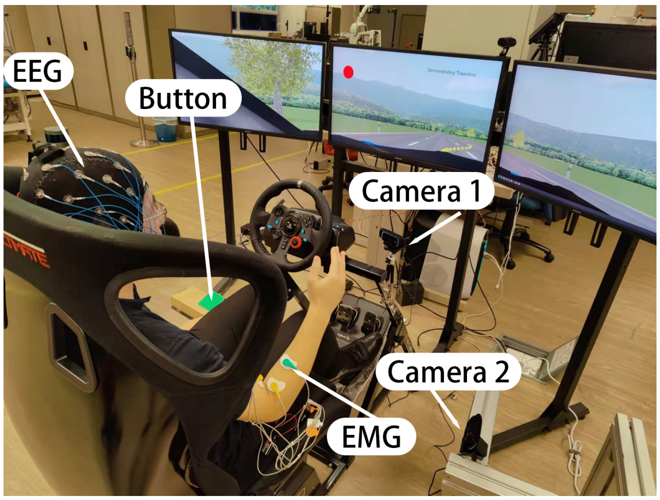

Fig. 4. Experiment Environment

TABLE I
Experiment Results

	EMG 1	EMG 2	EEG F3	EEG F4
Conventional Method	39.5607	149.5317	4170.6	4167.0
Proposed Method	44.0287	116.9875	4174.9	4165.7

Physical workload: The average muscle intensities of the upper arm (EMG 1) are 39.5607 and 44.0287 for the Manual mode and HL mode, respectively. As for the forearm (EMG 2), the average muscle intensity of the Manual mode is 149.5317, which is higher than the HL mode of 116.9875. The results indicate that the forearm physical workload can be remarkably alleviated $(p < 0.05)$ while the muscle intensities of the upper arm from both modes are not significantly different $(p > 0.05)$.

Mental workload: EEG comparison indicates that the proposed interface does not escalate the driver's mental workload. The RMS of F3 and F4 and the RMS spectral power of AF3/AF4 gamma frequency are used as indicators for mental workload estimation. The average F3 RMSs are 4174.9 and 4170.6 for the HL mode and Manual mode, respectively, while the average F4 RMSs are 4165.7 and 4167.0, respectively. The mean AF3/AF4 gamma frequency spectral power of the Manual mode are 0.2309 and 0.2677, which are both lower than the RMS mean of the HL mode of 0.2710 and 0.3014, respectively. The difference between the two modes is not significant $(p > 0.05)$ because, during the experiment, the driver had to continuously monitor the automation algorithm and the traffic scene in both manual driving and high-level collaboration mode.

Vehicle performance: The safety criteria were evaluated after the experiment using the risk field by giving the exact position of the surrounding vehicles and the obstacles (the obstacle positions were not viable to the vehicle during the experiment; the obstacles were deemed as non-crossable stationary objects). The jerk and the lateral acceleration of the high-level collaboration model are significantly better than the manual mode $(p < 0.05)$, as the jerk of the HL mode is only 15.8949 and the lateral acceleration is 0.1441, compared to 19.9156 and 0.2441 of the manual mode. On the other hand, the safety is not deteriorated using the high-level collaboration $(p > 0.05)$, though the risk of the manual mode is 5.4034, which is lower than the mean risk of the HL mode (5.5152).

In brief, the algorithmic proposition is competent in mitigating physical exertion imposed on human drivers, without compromising their mental workload and ensuring consistent driving performance. The algorithm accomplishes this by offloading the responsibility of driving-related tasks, such as speed regulation, adherence to traffic rules, environmental information processing, and collision prevention, from the driver to the automated system. Consequently, the driver's workload is significantly reduced. Our proposed scheme further eliminates the requirement for drivers to possess a high degree of expertise or experience, instead enabling them to issue commands using gesture-based communication, emulating natural human interaction patterns.

V. Conclusion

A novel high-level human-vehicle collaboration framework is proposed for fully or highly autonomous vehicles. Using the proposed framework, a human driver can deliver hints to the autonomous driving algorithm using gestures and obtain feedback through the VR-HUD interface. A lower-level control framework using a neural network for trajectory prediction, risk-field for collision avoidance, and the MPC to fuse the inputs from the human driver and the autonomous driving algorithm are designed. Finally, a human-in-the-loop experiment verifies that the proposed framework can reduce human physical burden without compromising mental workload and vehicle performance.

References

[1] S. Mahmud, X. Lin, and J.-H. Kim, "Interface for human machine interaction for assistant devices: A review," in *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0768–0773, 2020.

[2] Q. Zhao, H. Zheng, C. Kaku, F. Cheng, and C. Zong, "Safety spacing control of truck platoon based on emergency braking under different road conditions," *SAE International Journal of Vehicle Dynamics, Stability, and NVH*, vol. 7, no. 1, pp. 69–81, 2022.

[3] N. M. Negash and J. Yang, "Anticipation-based autonomous platoon control strategy with minimum parameter learning adaptive radial basis function neural network sliding mode control," *SAE International Journal of Vehicle Dynamics, Stability, and NVH*, vol. 6, no. 3, pp. 247–265, 2022.

[4] Y. Xing, C. Lv, D. Cao, and P. Hang, "Toward human-vehicle collaboration: Review and perspectives on human-centered collaborative automated driving," *Transportation Research Part C: Emerging Technologies*, vol. 128, p. 103199, 2021.

[5] D. A. Abbink, M. Mulder, and E. R. Boer, "Haptic shared control: smoothly shifting control authority?," *Cognition, Technology & Work*, vol. 14, no. 1, pp. 19–28, 2012.

[6] J. Nilsson, P. Falcone, and J. Vinter, "Safe transitions from automated to manual driving using driver controllability estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 1806–1816, 2014.

[7] C. Lv, Y. Li, Y. Xing, C. Huang, D. Cao, Y. Zhao, and Y. Liu, "Human–machine collaboration for automated driving using an intelligent two-phase haptic interface," *Advanced Intelligent Systems*, vol. 3, no. 4, p. 2000229, 2021.

[8] M. Heydrich, V. Ivanov, A. Bertagna, A. Rossi, M. Mazzoni, and F. Bücner, "Hardware-in-the-loop testing of a hybrid brake-by-wire system for electric vehicles," *SAE International Journal of Vehicle Dynamics, Stability, and NVH*, vol. 6, no. 4, pp. 477–487, 2022.

[9] C. Huang, C. Lv, P. Hang, Z. Hu, and Y. Xing, "Human machine adaptive shared control for safe driving under automation degradation," *IEEE Intelligent Transportation Systems Magazine*, pp. 2–15, 2021.

[10] M. Schilling, S. Kopp, S. Wachsmuth, B. Wrede, H. Ritter, T. Brox, B. Nebel, and W. Burgard, "Towards a multidimensional perspective on shared autonomy," in *2016 AAAI Fall Symposium Series*, 2016.

[11] C. Huang, H. Huang, P. Hang, H. Gao, J. Wu, Z. Huang, and C. Lv, "Personalized trajectory planning and control of lane-change maneuvers for autonomous driving," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 5511–5523, 2021.

[12] D. Omeiza, H. Webb, M. Jirotka, and L. Kunze, "Explanations in autonomous driving: A survey," *CoRR*, vol. abs/2103.05154, 2021.

[13] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. A. Sallab, S. Yogamani, and P. Pérez, "Deep reinforcement learning for autonomous driving: A survey," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–18, 2021.

[14] S. Martin, S. Vora, K. Yuen, and M. M. Trivedi, "Dynamics of driver's gaze: Explorations in behavior modeling and maneuver prediction," *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 2, pp. 141–150, 2018.

[15] S. Gaglio, G. L. Re, and M. Morana, "Human activity recognition process using 3-d posture data," *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 5, pp. 586–597, 2014.

[16] Y. Xing, C. Lv, Z. Zhang, H. Wang, X. Na, D. Cao, E. Velenis, and F.-Y. Wang, "Identification and analysis of driver postures for in-vehicle driving activities and secondary tasks recognition," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 1, pp. 95–108, 2017.

[17] L. Fridman, D. E. Brown, M. Glazer, W. Angell, S. Dodd, B. Jenik, J. Terwilliger, J. Kindelsberger, L. Ding, S. Seaman, *et al.*, "Mit autonomous vehicle technology study: Large-scale deep learning based analysis of driver behavior and interaction with automation," *arXiv preprint arXiv:1711.06976*, vol. 1, 2017.

[18] Z. Wang, R. Zheng, T. Kaizuka, and K. Nakano, "Relationship between gaze behavior and steering performance for driver–automation shared control: a driving simulator study," *IEEE Transactions on Intelligent Vehicles*, vol. 4, no. 1, pp. 154–166, 2018.

[19] X. Chang, H. Li, L. Qin, J. Rong, Y. Lu, and X. Chen, "Evaluation of cooperative systems on driver behavior in heavy fog condition based on a driving simulator," *Accident Analysis & Prevention*, vol. 128, pp. 197–205, 2019.

[20] X. Wang, X. Zheng, W. Chen, and F.-Y. Wang, "Visual human–computer interactions for intelligent vehicles and intelligent transportation systems: The state of the art and future directions," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 253–265, 2021.

[21] Z. Hu, C. Lv, P. Hang, C. Huang, and Y. Xing, "Data-driven estimation of driver attention using calibration-free eye gaze and scene features," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 2, pp. 1800–1808, 2022.

[22] B. Jiang, X. Li, Y. Zeng, and D. Liu, "Human-machine cooperative trajectory planning for semi-autonomous driving based on the understanding of behavioral semantics," *Electronics*, vol. 10, no. 8, p. 946, 2021.

[23] S. Mahmud, X. Lin, and J.-H. Kim, "Interface for human machine interaction for assistant devices: a review," in *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0768–0773, IEEE, 2020.

[24] A. Riener, A. Ferscha, F. Bachmair, P. Hagmüller, A. Lemme, D. Muttenthaler, D. Pühringer, H. Rogner, A. Tappe, and F. Weger, "Standardization of the in-car gesture interaction space," in *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 14–21, 2013.

[25] C. A. Pickering, K. J. Burnham, and M. J. Richardson, "A research study of hand gesture recognition technologies and applications for human vehicle interaction," in *2007 3rd Institution of Engineering and Technology conference on automotive electronics*, pp. 1–15, IET, 2007.

[26] F. Weidner and W. Broll, "Interact with your car: a user-elicited gesture set to inform future in-car user interfaces," in *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*, pp. 1–12, 2019.

[27] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for eeg-based brain–computer interfaces: a 10 year update," *Journal of neural engineering*, vol. 15, no. 3, p. 031005, 2018.

[28] G. Rupp, C. Berka, A. H. Meghdadi, M. S. Karić, M. Casillas, S. Smith, T. Rosenthal, K. McShea, E. Sones, and T. D. Marcotte, "Eeg-based neurocognitive metrics may predict simulated and on-road driving performance in older drivers," *Frontiers in human neuroscience*, vol. 12, p. 532, 2019.

[29] L. Bi, X.-A. Fan, and Y. Liu, "Eeg-based brain-controlled mobile robots: A survey," *IEEE transactions on human-machine systems*, vol. 43, no. 2, pp. 161–176, 2013.

[30] M. Werling, J. Ziegler, S. Kammel, and S. Thrun, "Optimal trajectory generation for dynamic street scenarios in a frenét frame," in *2010 IEEE International Conference on Robotics and Automation*, pp. 987–993, 2010.

[31] N. Deo and M. M. Trivedi, "Multi-modal trajectory prediction of surrounding vehicles with maneuver based lstms," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1179–1184, 2018.

[32] M.-F. Chang, J. W. Lambert, P. Sangkloy, J. Singh, S. Bak, A. Hartnett, D. Wang, P. Carr, S. Lucey, D. Ramanan, and J. Hays, "Argoverse: 3d tracking and forecasting with rich maps," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[33] Y. Rasekhipour, A. Khajepour, S.-K. Chen, and B. Litkouhi, "A potential field-based model predictive path-planning controller for autonomous road vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 5, pp. 1255–1267, 2017.

[34] G. S. Sankar and K. Han, "Adaptive robust game-theoretic decision making strategy for autonomous vehicles in highway," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 14484–14493, 2020.