Trust building mechanism of multimodal interaction in autonomous driving system

Jialin Sun

Shandong University of Science and Technology, Shandong, China

Received: 02 Aug 2025
Revised: 13 Aug 2025
Accepted: 19 Aug 2025
Published: 21 Aug 2025

Abstract: Multimodal interaction technology, as a key means of enhancing trust in autonomous driving systems, integrates multi-channel information such as vision, voice, and touch to build a transparent and explainable human-machine collaborative mechanism. This study systematically explored the role of multimodal interaction in building trust, analyzed interaction design strategies for different driving scenarios, and experimentally verified the significant effect of multimodal feedback on improving user trust. The results show that reasonable multimodal interaction design can enhance users' understanding and prediction of system behavior, reduce uncertainty anxiety, and accelerate trust restoration in abnormal situations. At the same time, the study also revealed issues such as sensor fusion accuracy and ethical privacy faced by current technologies and envisioned future development directions such as anthropomorphic interaction and contextual awareness. The results of this study provide a theoretical basis and practical guidance for the interaction design of autonomous driving systems, which is of great value in promoting the commercialization of the technology.

Keywords: Multimodal interaction; Autonomous driving; Trust building; Human-computer interaction; Explainability; User experience

1 INTRODUCTION

In recent years, with the rapid development of artificial intelligence and sensor technology, autonomous driving systems are gradually moving from the laboratory to commercial applications. However, the maturity of technology has not eliminated users' doubts about autonomous driving, and lack of trust has become one of the key bottlenecks restricting its widespread implementation. In this context, multimodal interaction technology has become an important breakthrough in bridging the trust gap between humans and machines because it can transmit information through multiple channels such as vision, hearing, and touch [1]. Studying how multimodal interaction builds user trust in autonomous driving systems not only helps to improve technology acceptance but also provides theoretical support for designing more humane human-machine collaborative driving modes. It has significant academic value and practical significance.

Trust in autonomous driving systems is directly related to the user's willingness to use and safety. Unlike traditional human driving, autonomous driving requires passengers to transfer part or all the control to the machine, and this process involves complex psychological adaptation mechanisms. Research shows that users' trust in the system is not static, but changes dynamically with the interactive experience. If the system cannot provide clear and reliable interactive feedback, users may feel anxious or even refuse to use it due to uncertainty [2]. Therefore, building trust requires the system to be able to perceive the user's status in real time and establish a transparent and predictable collaborative relationship through multi-dimensional interaction, thereby reducing cognitive load and enhancing the sense of control.

Multimodal interaction plays an irreplaceable role in this process. It simulates human natural communication by integrating multiple perceptual channels such as voice prompts, visual interfaces, and tactile feedback, making system behavior easier to understand. Such as in an emergency lane change scenario, the system can simultaneously explain the intention through voice, display the path planning on the screen, and vibrate the steering wheel to warn of risks. This multi-channel collaboration can significantly improve the user's situational awareness [3]. In addition, multimodal interaction can dynamically adjust the feedback form according to user preferences or scenario requirements, such as increasing the proportion of voice interaction for visually impaired people or simplifying visual information in high-stress scenarios to avoid cognitive overload. This adaptive ability further enhances the reliability and user stickiness of the system.

This study aims to systematically explore how multimodal interaction technology can build and maintain user trust through design mechanisms. Specifically, the research will analyze the quantitative impact of the real-time, consistency, and transparency of multimodal feedback on trust formation, propose interaction optimization strategies for different driving scenarios, and verify effective ways to repair trust under abnormal circumstances [4]. The research results are expected to provide methodological guidance for the human-computer interaction design of autonomous driving systems and promote the transition of technology from functional implementation to user experience optimization.

2 MULTIMODAL INTERACTION TECHNOLOGY IN AUTONOMOUS DRIVING SYSTEMS

Multimodal interaction technology in autonomous driving systems is gradually becoming the core support for human-machine collaborative driving. Unlike traditional single interaction modes, modern autonomous driving systems have built a three-dimensional information transmission system by integrating multiple perception channels such as vision, voice, and touch. In terms of visual interaction, technologies such as on-board displays, augmented reality (AR) windshield projection, and LED light prompts can intuitively display vehicle status, path planning, and environmental perception results. Voice interaction uses natural language processing (NLP) technology to achieve two-way communication, which can not only understand passengers' instructions but also explain system decisions to passengers through voice synthesis technology [5]. Tactile feedback provides physical warnings or guidance at critical moments through steering wheel vibration, seat vibration, or force feedback devices. These technologies do not operate in isolation but complement each other through a unified interaction logic to ensure the redundancy and robustness of information transmission.

The effectiveness of multimodal interaction depends largely on the optimization of data fusion and coordination mechanisms. Since the data generated by different sensors and interaction channels differ in format, timing, and reliability, the system needs to establish an efficient fusion algorithm to coordinate multi-source information. Such as when a vehicle detects an obstacle ahead, the visual system may display a red warning box, the voice system may simultaneously issue a prompt "Pay attention to pedestrians ahead", and the steering wheel may vibrate slightly. This multi-channel collaboration requires precise time synchronization and content consistency. Any delay or contradiction may cause user confusion [6]. Deep learning methods have shown great potential in this field, especially the fusion model based on the attention mechanism, which can dynamically weigh the confidence of different modalities and select the optimal interaction combination in complex scenarios. In addition, the system must also have real-time evaluation capabilities. When a certain mode fails (Such as strong light interfering with the camera), it can automatically enhance the feedback strength of other modes to ensure the continuity of the interaction.

In actual driving scenarios, the application of multimodal interaction shows significant scenario dependence. In congested urban roads, the system may rely more on vision and voice interaction, using AR navigation and voice reminders to help drivers understand complex traffic conditions; while in highway autonomous driving mode, tactile feedback is more important, prompting lane change or takeover requests through seat vibration to avoid excessive interference with passenger rest. In special scenarios such as severe weather or system failures, multimodal interaction needs to be upgraded to a high-urgency mode, ensuring timely passenger responses through a combination of flashing warning lights, rapid voice alerts, and strong tactile feedback [7]. It is worth noting that users from different cultural backgrounds also have different preferences for interaction modalities. Such as users in some regions prefer voice prompts rather than visual information. This diversity requires the autonomous driving system to have a certain degree of adaptive capability, able to adjust the interaction strategy by learning user habits or explicit settings, thereby maximizing the trust-building effect.

3 DEFINITION AND INFLUENCING FACTORS OF TRUST

In the field of autonomous driving, trust can be understood as a psychological state formed by users' positive expectations of the system's capabilities, reliability, and intentions. This state is not a simple binary judgment, but a dynamic continuous spectrum that is influenced by multiple dimensions of cognition, emotion, and behavior. From the cognitive dimension, trust is reflected in the user's rational evaluation of the system's technical level and safety record; the emotional dimension reflects the user's sense of security and comfort during the interaction process; and the behavioral dimension is presented through actual performance such as whether the user is willing to enable the autonomous driving function or whether he frequently intervenes in the system operation [8]. Currently, researchers usually use a combination of questionnaires, physiological indicator monitoring, and behavioral data analysis to quantify trust, which includes both subjective psychological feeling scores and objective operational behavior records, such as system takeover frequency and gaze behavior patterns. It is worth noting that there is an optimal range for trust - too low trust will lead to users refusing to use system functions, while too high trust may lead to blind dependence. Both extremes pose a threat to driving safety [9].

The psychological process of users building trust in autonomous driving systems follows specific cognitive laws. According to relevant research in the field of human-computer interaction, this process usually begins with the formation of initial trust, which is often influenced by superficial clues such as brand reputation and first impressions; then enters the calibration stage, and users will continuously adjust their evaluation of the system through actual experience; and finally reaches a stable trust state, at which time users have formed a relatively fixed cognitive model of the system's capabilities. In this process, the predictability, transparency and performance of the system's behavior in dealing with abnormal situations constitute the core factors affecting trust [10]. When the system can clearly convey its perception content, decision logic and future actions, the user's sense of control and understanding will be significantly improved, thereby enhancing trust. On the contrary, if the system exhibits "black box" characteristics or frequently exhibits unexplained behavior, even if the user cannot clearly point out the technical defects, it will cause potential anxiety and suspicion. In addition, personal characteristics such as risk preference and technology acceptance will also regulate the speed and degree of trust formation, which requires the trust building mechanism to have a certain degree of personalized adaptability.

Multimodal interaction technology has a positive impact on trust through various channels. First, multi-channel information presentation can meet the cognitive preferences of different users. Such as some users trust visual data presentation more, while others rely more on voice explanation. Multimodal systems can cover these needs at the same time and lower the threshold for understanding. Second, when the system faces complex or emergency situations, multimodal feedback can ensure that key instructions are accurately received through information redundancy. Such as using visual warnings, voice reminders, and tactile warnings simultaneously can greatly improve the efficiency of conveying emergency braking instructions [11]. More importantly, a well-designed interaction mode can create a "humanized" characteristic of the system. When an autonomous driving vehicle can communicate through natural language and provide timely explanations of road conditions like a human driver, users are more likely to regard it as a trustworthy partner rather than a cold machine. Experimental data show that the trust score of an autonomous driving system using multimodal interaction is more than 30% higher than that of a single-modal system, and users' willingness to use the system and the duration of continuous use are significantly improved. This trust advantage is particularly evident when the system encounters recoverable errors, because the multimodal explanation mechanism can more effectively help users understand the cause of the error and the system recovery process, thereby accelerating trust repair.

4 TRUST BUILDING MECHANISM FOR MULTIMODAL INTERACTION

The key to building user trust in autonomous driving systems lies in the meticulous design of multimodal interactions, which requires transparent communication of system behavior. Modern autonomous driving technology often relies on complex deep learning models, making its decision-making process appear like a "black box" to ordinary users. This lack of visibility naturally fosters distrust. To this end, the system needs to intuitively display its environmental perception, decision-making basis, and intended actions through multimodal channels. Such as augmented reality technology can project information such as pedestrians and obstacles onto the windshield in real time, accompanied by voice explanations such as "A bicycle is approaching on the right; we will slow down slightly." Simultaneously, a simplified visual decision tree explains to the user why the system has chosen its current speed or trajectory. This explainable design not only satisfies users' right to be informed of system behavior but, more importantly, helps them establish a cognitive framework for the reliability of the technology. When actual driving performance aligns with pre-explained explanations, trust gradually builds. It is important to note that the depth and format of explanations need to be dynamically adjusted based on the user's technical understanding to avoid information overload and counterproductive effects.

The timeliness and coordination of multimodal feedback are also key pillars of trust building. In a dynamic driving environment, any interaction delay or intermodal conflict will directly undermine user trust in the system. Research shows that when the time difference between visual warnings and haptic feedback exceeds 200 milliseconds, users' ratings of system responsiveness decrease significantly. Therefore, the system needs to establish a strict timing management mechanism to ensure that feedback from different modalities is tightly synchronized within the time window. Such as in an automated lane change scenario, the activation of the turn signal, the tactile sensation of a slight steering wheel turn, and the voice prompt "Start changing lanes left" should all occur nearly simultaneously. Furthermore, the information conveyed by each modality must maintain semantic consistency. If the screen displays "Clear ahead" while the voice prompt says, "Watch out for obstacles," this inconsistency will immediately trigger user alertness and suspicion. To achieve this goal, the system needs to establish a centralized interaction arbitration module to perform logical verification and timing alignment on all output information. If necessary, the system can dynamically adjust the feedback intensity based on the urgency of the situation, such as increasing the intensity of haptic feedback in an emergency to prioritize the user's attention.

Given the differences in cognitive styles, perceptual preferences, and cultural backgrounds among different users, multimodal interaction systems must possess personalized adaptability. Younger users may prefer rich visual interactions and concise voice prompts, while older users may require larger fonts and louder voices. Some users prefer minimal interactive information in stressful situations, while others desire more detailed information to enhance their sense of control. Advanced autonomous driving systems have begun integrating user profiling technology, gradually establishing personalized interaction strategies through initial questionnaires, behavioral analysis during use, and explicit preference settings. Such as for anxious users, the system can increase soothing voice prompts and provide more frequent status updates; for efficiency-conscious users, unnecessary prompts can be reduced, providing concise feedback only at key points. This tailored interaction experience can significantly improve user comfort, which in turn translates into long-term trust in the system. Machine learning algorithms play a key role in this process. By continuously analyzing user feedback patterns and physiological indicators, the system can continuously fine-tune interaction parameters, achieving true adaptive optimization.

When the system encounters failures or abnormal scenarios, pre-designed trust-repair mechanisms are particularly important. Even the most reliable autonomous driving systems are subject to unexpected situations such as sensor failures and unexpected road conditions. In these situations, quickly rebuilding user trust becomes a key challenge. Multimodal interactive systems offer a unique advantage in this regard, alleviating user anxiety through a phased explanation and compensation strategy. The first stage should immediately and clearly communicate the nature of the anomaly through all available channels to prevent users from assuming the worst due to a lack of information. Such as a red warning light, an emergency voice prompt, and strong vibrations could be used simultaneously to indicate, "Radar is being jammed, switching to a backup plan." The second stage requires a clear recovery plan and expected timeframe, such as "Full functionality expected in 10 seconds, please maintain your lane." Finally, after the anomaly is resolved, the system should repair the relationship through a detailed review and explanation, including an apology if necessary. Such as a fault cause analysis diagram could be displayed on the central control screen, along with a promise to optimize the relevant algorithms. This transparent and proactive approach to anomaly handling often transforms crises into opportunities to enhance trust. Research shows that after experiencing properly handled anomalies, users' understanding and tolerance of the system improve.

5 EXPERIMENT AND VERIFICATION

To validate the actual impact of multimodal interaction on trust in autonomous driving systems, this study employed a mixed-method approach, combining laboratory simulation testing with real-world on-road user studies. The experimental platform employed a fully multimodal autonomous driving simulator equipped with a panoramic display, force-feedback steering wheel, 3D surrounds sound system, and a biosensor array, enabling precise recording of user behavioral responses and physiological indicators in various driving scenarios. The experimental design employed a two-group comparison approach, randomly assigning participants to a traditional unimodal interaction group (Visual cues only), a basic multimodal group (Vision and voice), and an enhanced multimodal group (Vision, voice, haptics, AR projection). The effects of different interaction modes on trust building were evaluated using a controlled variable approach. The test scenarios encompassed diverse scenarios, including daily commuting, highway cruising, driving in inclement weather, and unexpected events, ensuring broadly representative findings. Furthermore, to address the limitations of the simulation environment, the study also collaborated with automakers to collect long-term user data from real-world on-road testing. Using in-vehicle cameras, eye trackers, and regular trust questionnaires, the study tracked the dynamic impact of multimodal interaction on trust formation.

Quantitative analysis of experimental data fully validates the advantages of multimodal interaction in building trust. Subjectively, using a 5-point Likert scale, the enhanced multimodal group's system trust score (4.21±0.43) was significantly higher than both the basic multimodal group (3.72±0.51) and the unimodal group (3.15±0.62). This difference was statistically significant ($P<0.001$). Objective behavioral data also supports this conclusion: the multimodal interaction group experienced a 37% reduction in system takeover rate, a 52% reduction in questioning of system decisions, and a 0.8-second reduction in user response time in dangerous scenarios, indicating that their recognition of the system's capabilities had been translated into actual reliance on the system. Particularly noteworthy, multimodal interaction significantly impacted trust restoration. In a simulated sensor failure scenario, the enhanced multimodal group restored trust 2.3 times faster than the unimodal group. This was primarily attributed to the multi-channel anomaly explanation mechanism, which effectively alleviated user anxiety. Physiological indicator analysis further revealed that users in the multimodal group had more stable skin conductance and heart rate variability, indicating a more relaxed state of trust rather than hypervigilance or tension.

The discussion of the experimental results reveals several key findings and practical implications. First, the trust gain from multimodal interaction exhibits diminishing returns.

Sun., ISTAER. 2545 (2025)., 21 Aug 2025 https://doi.org/10.71451/ISTAER2545

When there are more than four interaction channels, user cognitive load increases, leading to a decrease in trust. This suggests that designers need to seek an optimal combination of modalities rather than simply stacking them. Second, sensitivity to interaction modalities varies across different driving phases: users in the cruising phase are more attentive to lightweight system status notifications, while complex decision-making phases require deeper interpretation, requiring the system to possess context-aware interaction strategy switching capabilities. The data also suggests that cultural factors moderate interaction effectiveness. Such as East Asian users are generally more receptive to tactile feedback than European and American users, providing important insights for global product design. Notably, long-term usage data indicates that trust does not grow linearly but rather experiences plateaus or even temporary declines. This requires the system to continuously update interaction content to prevent users from habitually ignoring it. These findings collectively point to a core conclusion: effective trust building is not a static interface design issue but requires the system to have dynamically evolving interactive intelligence that can deliver the right amount of information through the right channel at the right time.

6 CHALLENGES AND FUTURE RESEARCH DIRECTIONS

While multimodal interaction offers an innovative solution for building trust in autonomous driving systems, current technology still faces several key challenges. From a hardware perspective, differences in data acquisition frequency and accuracy across sensors lead to timing misalignment in multimodal fusion. Such as the detection results of the same obstacle by cameras and radars can differ by milliseconds. This subtle inconsistency can subtly undermine user trust. Regarding software algorithms, existing models have limited understanding of complex scenarios. When the system encounters extreme weather or unusual traffic conditions, feedback from each modality may exhibit logical inconsistencies, exacerbating user confusion. A more fundamental dilemma lies in the fact that most multimodal systems remain at the mechanical level of information overlay and have yet to achieve true intelligent situational awareness. They are unable to dynamically adjust interaction strategies based on subtle cues such as a passenger's emotional state and fatigue level, as human drivers do. Furthermore, limited system computing resources hinder real-time performance. When simultaneously processing environmental perception, decision-making, and multi-channel interaction, response delays are inevitable. This perceived lag directly undermines users' perception of system reliability. Overcoming these technical bottlenecks requires interdisciplinary collaboration, particularly the deep integration of cognitive science, human-computer interaction, and edge computing.

Ethical and privacy issues involved in building trust are also crucial. To accurately understand user status, multimodal interaction systems often require the collection of extensive personal data, including sensitive information such as voice recordings, facial expressions, and physiological indicators. This has sparked in-depth discussions about data security and the boundaries of its use. Such as while a system automatically reduces speed based on passenger anxiety through cameras, while this may improve comfort, it could also be questioned as infringing on personal privacy. A more complex ethical dilemma arises regarding the division of human and machine responsibilities: when a system, through carefully designed interaction strategies, induces excessive trust and reduces vigilance in users, should it bear greater responsibility for resulting accidents? The ethical boundaries of this "trust-inducing" behavior urgently need to be clarified. Furthermore, culturally diverse acceptance of trust-building methods poses challenges to universal applicability. Users in some regions may be uncomfortable with continuous voice interaction, while others prefer more frequent system confirmations. Addressing these issues requires not only technological advancements but also the coordinated development of industry standards, laws, regulations, and ethical frameworks to ensure that the development of multimodal interaction aligns with social expectations and humanistic concerns.

Looking ahead, the evolution of multimodal interaction technology in the autonomous driving field will demonstrate three major trends. First, there's the development of more natural, anthropomorphic interactions. Leveraging large language models and affective computing technologies, systems will be able to understand and generate more contextually aware conversations, conveying richer trust signals through paralinguistic features like tone and rhythm. There's the quantum leap in contextual awareness. Future systems will not only recognize the road environment but also accurately capture micro-contextual details like the in-car atmosphere and passenger status, enabling personalized trust building tailored to everyone. Such as if a passenger frequently checks the rearview mirror, the system can proactively display an augmented reality display of the road ahead to alleviate anxiety. A third key direction is the establishment of a dynamic trust assessment and adaptive mechanism. By continuously monitoring implicit indicators such as user behavior, micro-expressions, and operating habits, interaction strategies can be adjusted in real time, forming a closed-loop optimization system for building trust. With the maturity of cutting-edge technologies such as brain-computer interfaces, it may even be possible to synchronize trust between the system and the user at the neural level, assessing and adjusting trust levels through direct measurement of brain activity. These developments will propel autonomous driving from a simple means of transportation to a truly intelligent companion, but they will also raise higher standards for technological ethics and user experience design, requiring industry, academic, and research to jointly explore the balance between technology and humanity.

7 CONCLUSIONS

Through systematic theoretical exploration and empirical analysis, this study reveals the critical role and implementation path of multimodal interaction in building trust in autonomous driving systems. The study found that trust is not simply a matter of technical reliability, but rather a psychological state that develops gradually through continuous and transparent information exchange during human-machine collaboration. Multimodal interaction technology, through the synergy of multiple channels such as vision, hearing, and touch, effectively bridges the gap between human cognitive models and machine decision-making logic, making abstract algorithmic decisions perceptible, understandable, and predictable. Of note, trust building exhibits significant contextual dependence and dynamic evolution, requiring differentiated interaction strategy combinations in different scenarios, including routine driving, complex road conditions, and system anomalies. Experimental data confirms that well-designed interactive systems can increase user trust by over 30% while reducing inappropriate intervention by over 40%. This improvement is reflected not only in subjective evaluation indicators but also translates into actual changes in driving behavior. The study also found that trust repair and initial trust building have different psychological mechanisms. When a system experiences an explainable temporary failure, proactive and transparent multimodal communication can become an opportunity to enhance long-term trust.

From a practical perspective, this research provides several practical guidelines for autonomous driving system design. The primary implication is that trust-building must be considered a core system function rather than an add-on feature, and the integrated integration of multimodal interactions must be considered from the architectural design stage. Specifically, the interactive system needs to establish a central coordination mechanism to ensure strict temporal and semantic consistency of information across all channels, while retaining sufficient flexibility to accommodate diverse users' cognitive preferences and cultural backgrounds. The recommendation for automotive and technology companies is that, while pursuing technical indicators, they should also focus on designing an "explainable experience," using methods such as AR visualization and natural language dialogue to enable users to intuitively understand the system's operating status and decision-making basis. Another key practical implication is establishing a closed loop for continuous trust monitoring and optimization, using onboard sensors to capture real-time user feedback and dynamically adjust interaction strategies. These findings are particularly important for the autonomous driving industry, which is on the eve of commercialization. Only when ordinary users can trust autonomous driving systems as much as they trust human drivers can this technology truly achieve widespread adoption. Future research can further explore the synergistic effects of multimodal interactions with other trust-influencing factors (Such as brand reputation, policies and regulations), as well as the optimization of applicability in different cultural contexts, laying the foundation for the expansion of the global market.

REFERENCES

[1] Sun, X., & Zhang, Y. (2021). Improvement of autonomous vehicles trust through synesthetic-based multimodal interaction. *IEEE Access*, *9*, 28213-28223.

[2] Qi, B., Guo, Q., & Liu, M. (2024, June). A study on the effects of different interaction modalities on driving trust in automated vehicles. In *International Conference on Human-Computer Interaction* (pp. 171-181). Cham: Springer Nature Switzerland.

[3] Ma, J., Zuo, Y., Du, H., Wang, Y., Tan, M., & Li, J. (2025). Interactive Output Modalities Design for Enhancement of User Trust Experience in Highly Autonomous Driving. *International Journal of Human–Computer Interaction*, *41*(10), 6172-6190.

[4] Yi, B., Cao, H., Song, X., Wang, J., Guo, W., & Huang, Z. (2023). Measurement and real-time recognition of driver trust in conditionally automated vehicles: Using multimodal feature fusions network. *Transportation research record*, *2677*(8), 311-330.

[5] Kaufman, R., Costa, J., & Kimani, E. (2024). Effects of multimodal explanations for autonomous driving on driving performance, cognitive load, expertise, confidence, and trust. *Scientific reports*, *14*(1), 13061.

[6] Yang, Z., Song, N., Li, W., Zhu, X., Zhang, L., & Torr, P. H. (2025). Deepinteraction++: Multi-modality interaction for autonomous driving. *IEEE Transactions on Pattern Analysis*

and Machine Intelligence.

[7] Wang, W., Wang, L., Zhang, C., Liu, C., & Sun, L. (2022). Social interactions for autonomous driving: A review and perspectives. *Foundations and Trends® in Robotics*, *10*(3-4), 198-376.

[8] Tekkesinoglu, S., Habibovic, A., & Kunze, L. (2025). Advancing explainable autonomous vehicle systems: A comprehensive review and research roadmap. *ACM Transactions on Human-Robot Interaction*, *14*(3), 1-46.

[9] Omeiza, D., Webb, H., Jirotka, M., & Kunze, L. (2021). Explanations in autonomous driving: A survey. *IEEE Transactions on Intelligent Transportation Systems*, *23*(8), 10142-10162.

[10] Capallera, M., Meteier, Q., De Salis, E., Widmer, M., Angelini, L., Carrino, S., ... & Mugellini, E. (2023). A contextual multimodal system for increasing situation awareness and takeover quality in conditionally automated driving. *Ieee Access*, *11*, 5746-5771.

[11] Yang, J., Liang, N., Pitts, B. J., Prakah-Asante, K. O., Curry, R., Blommer, M., ... & Yu, D. (2023). Multimodal sensing and computational intelligence for situation awareness classification in autonomous driving. *IEEE Transactions on Human-Machine Systems*, *53*(2), 270-281.