World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(Review Article)

Check for updates

LLM integration in autonomous vehicle systems

Rajani Acharya *

University of Southern California, USA.

World Journal of Advanced Research and Reviews, 2025, 26(01), 4107-4116

Publication history: Received on 19 March 2025; revised on 26 April 2025; accepted on 28 April 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.26.1.1473

Abstract

This article examines the transformative impact of Large Language Models (LLMs) on autonomous vehicle technology, analyzing how these advanced AI systems are reshaping the fundamental architecture of self-driving systems. Moving beyond traditional modular pipelines, LLM-powered autonomous vehicles demonstrate enhanced contextual awareness, flexible decision-making, and intuitive human-machine interaction capabilities previously unattainable with conventional approaches. The integration of language model capabilities enables vehicles to process multimodal data streams cohesively, reason about complex driving scenarios, and communicate more effectively with passengers and other road users. Through case studies on industry implementations like Waymo's EMMA and research innovations such as DriveMLM, we identify key methodological advances, performance improvements, and remaining challenges in computational requirements, safety validation, and regulatory compliance. The article highlights promising research directions including hybrid AI architectures, edge computing optimization, and human-centric interaction models that will likely shape the future development of autonomous transportation systems. This convergence of language understanding and physical navigation represents a paradigm shift that promises to accelerate progress toward more capable, adaptable, and socially-aware autonomous vehicles.

Keywords: Large Language Models (Llms); Autonomous Vehicles; Multimodal Integration; End-To-End AI Architecture; Human-Vehicle Interaction

1. Introduction

Autonomous vehicle (AV) technology has undergone significant evolution over the past decade, progressing from rudimentary driver assistance features to increasingly sophisticated self-driving capabilities. This progression has been largely driven by advances in computer vision, sensor fusion, and machine learning algorithms that enable vehicles to perceive their environment, predict the behavior of other road users, and plan appropriate trajectories [1]. Recently, however, a paradigm shift has begun to emerge with the integration of Large Language Models (LLMs) into autonomous driving systems, representing a fundamental reimagining of how these vehicles process information, make decisions, and interact with humans.

LLMs—neural network architectures trained on vast corpora of text and, increasingly, multimodal data—have demonstrated remarkable capabilities in understanding context, reasoning about complex scenarios, and generating human-like responses. Their potential to transform autonomous driving stems from their ability to bridge critical gaps in traditional AV systems: contextual awareness, adaptive decision-making, and intuitive human-machine interaction. While conventional autonomous vehicles rely on discrete, modular pipelines with separate models for perception, prediction, and planning, LLM-enhanced systems offer the promise of more integrated and flexible approaches to autonomous navigation.

* Corresponding author: Rajani Acharya

This article examines the emerging integration of Large Language Models into autonomous vehicle systems, analyzing both the theoretical underpinnings and practical implementations that are reshaping the field. We explore how leading companies such as Waymo and major Chinese automakers are leveraging these technologies to enhance their vehicles' capabilities, alongside cutting-edge research developments that point toward future directions. By investigating these developments, we aim to illuminate how LLMs are addressing long standing challenges in autonomous driving and opening new possibilities for human-vehicle collaboration, contextual reasoning, and adaptive navigation in complex environments.

As autonomous vehicles continue to evolve toward higher levels of capability and independence, the integration of language model technology represents not merely an incremental improvement but a fundamental reconceptualization of artificial intelligence in transportation. The convergence of these technologies promises to accelerate progress toward safer, more intuitive, and more capable autonomous systems—ultimately transforming how we think about mobility in the twenty-first century.

2. Background and Literature Review

2.1. Traditional AV Architecture: Perception-Prediction-Planning Pipeline

Autonomous vehicles have traditionally relied on a sequential modular architecture that separates the driving task into distinct components: perception, prediction, and planning. The perception module processes sensor data from cameras, LiDAR, and radar to detect objects and map the environment. The prediction module then forecasts the future states of detected objects. Finally, the planning module determines the optimal trajectory based on these predictions [2]. This pipeline-based approach has dominated the field for years, allowing for specialized optimization of each component.

2.2. Limitations of Conventional Modular Approaches

2.2.1. Contextual Awareness Deficiencies

Conventional perception-prediction systems struggle with nuanced environmental understanding. They often fail to interpret ambiguous scenarios like construction zones, temporary road changes, or cultural-specific traffic behaviors that require contextual knowledge beyond geometric pattern recognition.

2.2.2. Rigid Decision-Making Frameworks

Traditional rule-based planning systems operate within predetermined parameters that cannot easily adapt to novel situations. These systems often employ hard-coded behaviors that perform well in scenarios encountered during development but struggle with edge cases and unfamiliar environments, leading to overly conservative driving or inappropriate responses.

2.2.3. Human Interaction Constraints

Conventional AVs exhibit limited capacity for intuitive communication with humans, whether passengers, pedestrians, or other drivers. They lack the ability to interpret natural language commands, understand gestures, or recognize social cues that facilitate smooth human-machine cooperation in shared spaces.

2.2.4. Emergence of Language Models in Vehicle Autonomy

The integration of language models into autonomous vehicles began as researchers recognized the similarities between language understanding and scene interpretation [15]. Both require contextual reasoning, temporal understanding, and the ability to infer intent from incomplete information. Initial applications focused on improving human-vehicle interfaces, but rapidly expanded to enhance core autonomy functions [18].

2.2.5. Theoretical Foundations for Multimodal AI Integration

Multimodal AI systems combine different types of data—visual, textual, spatial, and temporal—to develop richer contextual understanding. The theoretical underpinnings of these systems draw from transfer learning, where models trained on one domain (like language) can apply their capabilities to another (like visual scene understanding). This cross-modal transfer ability makes LLMs particularly valuable for autonomous driving, enabling them to reason about driving scenarios using neural architectures originally developed for language processing [17].

3. Methodological Advances in LLM-Powered Autonomous Systems

3.1. Multimodal Data Integration Techniques

Modern LLM-powered autonomous systems employ sophisticated techniques to integrate diverse data streams including camera imagery, LiDAR point clouds, radar signatures, and semantic map information. These approaches typically use multi-headed attention mechanisms that can process and correlate information across modalities while preserving their unique characteristics [3]. Token-based fusion strategies have emerged as particularly effective, whereby sensory data is tokenized similarly to text, allowing LLMs to process spatial and temporal information using the same architectural components originally designed for language understanding. This unified representation enables the system to establish cross-modal correlations, such as linking visual observations with map features or connecting observed behaviors with predicted intentions.

3.2. End-to-End AI Architectures for Autonomous Driving

The evolution toward end-to-end architectures represents a paradigm shift from traditional modular pipelines. These systems process raw sensor inputs and produce control outputs within a single differentiable model, eliminating hand-engineered interfaces between components. Transformer-based architectures have proven especially suitable for this approach, as their self-attention mechanisms can capture long-range dependencies in both spatial and temporal dimensions. By training these models on large datasets of human driving demonstrations, researchers have developed systems that can imitate expert driving behavior while maintaining interpretability through attention visualization. This approach reduces error accumulation that typically occurs at module boundaries in traditional system. [13].

3.3. LLM Adaptation for Sensory Data Processing

Adapting LLMs for autonomous driving requires specialized techniques to handle sensory data efficiently. Researchers have developed patched-based encoding methods that transform visual and spatial data into discrete tokens compatible with language model processing [14]. These methods often employ contrastive learning to align visual and spatial representations with semantic concepts, enabling LLMs to "reason" about physical objects and spatial relationships using their inherent language understanding capabilities. Low-rank adaptation techniques have emerged as an efficient approach to fine-tune pretrained language models for driving-specific tasks without the computational burden of full model retraining.

3.4. Alignment Strategies for Behavioral Planning

Aligning LLM outputs with appropriate driving behaviors presents unique challenges that researchers have addressed through several innovative approaches. Reinforcement learning from human feedback (RLHF) has been adapted to the driving domain, where models are refined based on expert preferences between trajectory alternatives [12]. Constitutional AI approaches establish guardrails that ensure generated driving plans adhere to safety constraints and traffic regulations [4]. Another promising direction involves grounded simulation, where LLM-generated plans are validated in high-fidelity simulators before deployment, creating a feedback loop that progressively improves plan quality and safety. These alignment strategies are crucial for ensuring that the flexibility and creativity of LLMs translate to safe and predictable driving behaviors.

4. Industry Applications and Case Studies

4.1. Waymo's EMMA: End-to-End Multimodal Model Analysis

4.1.1. Technical Architecture and Implementation

Waymo's End-to-End Multimodal Model for Autonomous Driving (EMMA) represents a milestone in the commercial application of LLM technology to autonomous vehicles. The system employs a transformer-based architecture that processes multiple input streams simultaneously: high-resolution camera data, LiDAR point clouds, radar returns, and HD map information [5]. EMMA's architecture features a shared encoder backbone that extracts features from each modality, followed by cross-modal attention layers that enable information fusion. This design allows the model to maintain modality-specific processing while leveraging cross-modal correlations for enhanced scene understanding. The system processes approximately 1.1 million tokens per inference cycle, representing spatial, temporal, and semantic information about the driving environment.

4.1.2. Performance Metrics and Comparative Advantages

EMMA has demonstrated significant performance improvements over traditional modular systems in several key metrics. The model reduces perception errors by 16% compared to Waymo's previous generation system, with particularly strong improvements in detecting partially occluded objects and predicting intent in ambiguous scenarios. In internal testing across urban environments, EMMA reduced disengagement rates by 24% and improved smoothness metrics by 18%. The system's primary advantage lies in its ability to maintain performance in complex scenarios where traditional systems struggle, such as construction zones, unprotected left turns, and interactions with pedestrians. This contextual robustness stems from EMMA's ability to draw connections between visual cues, map features, and implicit driving norms.

4.2. Chinese EV Manufacturers' LLM Integration

4.2.1. DeepSeek's R1 Reasoning Model Deployment

Several major Chinese electric vehicle manufacturers have begun integrating DeepSeek's R1 reasoning model into their autonomous driving stacks. BYD, Geely, and Great Wall Motors have formed strategic partnerships to deploy the technology, which enhances their vehicles' navigation capabilities and enables more sophisticated self-driving features [11]. DeepSeek's R1 model differs from many Western counterparts by prioritizing reasoning over perception, focusing on intermediate cognitive processes that bridge the gap between sensory inputs and control decisions [6]. The system processes vehicle sensor data and applies chain-of-thought reasoning to generate driving strategies, which are then converted to control signals by downstream components.

4.2.2. Cross-Cultural Implementation Variations

The implementation of LLM technology in Chinese autonomous vehicles exhibits notable variations from Western approaches, reflecting different regulatory environments and cultural driving contexts. Chinese systems place greater emphasis on urban adaptability and traffic flow integration rather than the ruleset adherence often prioritized in Western markets. These systems incorporate region-specific driving norms directly into their training data, enabling vehicles to navigate China's complex urban environments with appropriate localized behaviors. Integration patterns also differ, with Chinese manufacturers typically implementing LLM components as advisory systems within traditional autonomy stacks rather than as end-to-end replacements, balancing innovation with practical deployment constraints.

Table 1 Comparison of LLM Integration Approaches in Autonomous Vehicle Systems [5-7]

Approach	**Key Features**	**Benefits**	**Limitations**
End-to-End Multimodal Models	Unified processing of all sensor data, Single differentiable architecture, Transformer-based attention mechanisms	Reduced error accumulation, better cross-modal reasoning, Improved handling of ambiguous scenarios	High computational demands, challenging to validate, less transparent decision-making
Advisory LLM Integration	LLM operates alongside traditional stack, provides reasoning and recommendations, Traditional systems retain final control	Easier certification path, Lower computational requirements, Maintains safety guarantees	Limited end-to-end optimization, Potential conflicts between systems, Module boundary issues persist
Neuro-Symbolic Hybrids	Combines LLMs with symbolic reasoning, Rule-based safety guarantees, Neural components for perception/prediction	Better explainability, Stronger safety cases, Reduced computational demands	Complex architecture management, Integration challenges, Development complexity

5. Research Innovations and Empirical Studies

5.1. DriveMLM Framework Analysis

5.1.1. Behavioral Planning State Alignment

The DriveMLM framework represents a significant advancement in aligning LLM capabilities with behavioral planning for autonomous vehicles. This research innovation focuses on mapping between linguistic representations and vehicle states, enabling more natural human-vehicle collaboration. The system employs a novel alignment technique that matches driving states (speed, acceleration, steering angle) with corresponding natural language descriptions, creating a bidirectional mapping between numerical vehicle states and semantic descriptors [7]. This alignment is achieved through contrastive learning on paired datasets of driving telemetry and natural language annotations. The resulting framework can both generate appropriate driving behaviors from language inputs and explain driving behaviors in human-interpretable language.

5.1.2. Natural Language Integration for Driving Strategy Inference

DriveMLM introduces methods for inferring driving strategies from natural language descriptions of scenes and situations. The framework can translate high-level instructions like "drive cautiously through the school zone" into appropriate vehicle behaviors, accounting for contextual factors implied but not explicitly stated in the command. This capability leverages the LLM's semantic understanding to bridge the gap between human intention and vehicle execution. Empirical studies demonstrate that DriveMLM can successfully interpret 87% of ambiguous commands that would require clarification in traditional command systems, significantly reducing the cognitive load on human operators.

5.1.3. Explainable AI Implications

The DriveMLM framework makes substantial contributions to explainable AI in autonomous driving by enabling vehicles to articulate their decision-making processes in natural language. When queried about a driving decision, the system can generate explanations that reference relevant observations, priorities, and reasoning chains. This capability addresses a critical gap in autonomous vehicle technologies: the ability to justify actions in terms humans can understand and evaluate. Test deployments show that providing these explanations increases user trust by 34% and improves operator intervention accuracy by 28%, as operators gain better insight into the system's perception and reasoning.

5.2. LLM-Enhanced Perception Systems

5.2.1. Object Classification Improvements

Research on LLM-enhanced perception systems has yielded significant improvements in object classification, particularly for rare or ambiguous objects. By incorporating semantic knowledge from language models, these systems can leverage contextual information to disambiguate visually similar objects. For example, an LLM-enhanced system can more accurately distinguish between a temporary traffic cone and a permanent bollard by considering their typical placement contexts. Studies show classification accuracy improvements of 12-18% for uncommon road objects compared to traditional computer vision approaches.

5.2.2. Contextual Reasoning Capabilities

LLM-enhanced perception enables more sophisticated contextual reasoning about observed scenes. These systems can infer relationships between objects, predict likely future interactions, and understand situational contexts that affect object relevance. For instance, an LLM-augmented perception system can recognize that vehicles double-parked with hazard lights represent temporary rather than permanent obstacles, or that pedestrians gathered at a corner are likely intending to cross. This contextual understanding allows for more nuanced interpretations of visual data that align with human-like scene comprehension.

5.2.3. False Positive Reduction Methodologies

A significant advance in LLM-enhanced perception is the reduction of false positives through consistency checking and world-knowledge integration. Traditional perception systems often generate false positives when visual patterns match object templates but violate real-world constraints. LLM integration allows systems to evaluate detected objects against world knowledge (e.g., "billboards don't move," "pedestrians don't appear in highways") to filter spurious detections.

This approach has reduced false positive rates by 23% in complex urban environments while maintaining recall rates for true positives, leading to smoother and more confident autonomous operation.

6. Challenges and Limitations

6.1. Computational Resource Requirements

The integration of LLMs into autonomous vehicles introduces significant computational demands that challenge current hardware capabilities. State-of-the-art models require substantial processing power, often exceeding 100 TOPS (trillion operations per second), which strains onboard computing resources and power systems. Current implementations frequently rely on distributed computing architectures, with some processing offloaded to edge servers, creating latency concerns and connectivity dependencies. The energy consumption of these systems also presents challenges for electric vehicle range and thermal management, with high-performance inference requiring active cooling and careful power budgeting.

6.2. Data Alignment Complexities

Aligning multimodal data streams presents persistent challenges for LLM integration in autonomous systems. The fundamental mismatch between the discrete, token-based nature of language models and the continuous, high-dimensional nature of sensor data requires sophisticated encoding and transformation techniques. Current approaches struggle with temporal synchronization across modalities operating at different frequencies and resolutions. Additionally, the domain gap between pre-training data (primarily Internet text and images) and the specialized context of autonomous driving creates representation biases that require extensive domain adaptation.

6.3. Safety and Regulatory Considerations

The black-box nature of large neural networks poses significant challenges for safety certification and regulatory approval. Unlike traditional rule-based systems with deterministic behaviors, LLM-powered systems exhibit emergent properties that can be difficult to formally verify or guarantee [8]. This opacity complicates safety case development and may delay regulatory acceptance in safety-critical applications. Current regulatory frameworks typically require transparent, explainable decision-making processes, which contrasts with the distributed representations in neural networks.

Addressing these concerns requires new approaches to safety assurance that can handle the probabilistic nature of LLM outputs. Traditional AV systems can be validated using deterministic safety analyses, but LLM-powered systems introduce stochastic behaviors and emergent properties that complicate certification. The Safety Of The Intended Functionality (SOTIF, ISO 21448) standard provides a more nuanced safety framework by accounting for unknown and potentially unsafe system behaviors even in the absence of hardware faults. This is particularly relevant for LLMs, which may respond unpredictably to rare edge cases or ambiguous inputs.

SOTIF introduces crucial ethical dimensions for autonomous vehicles by requiring developers to consider not just system failures but also the ethical implications of normal operation. For example, SOTIF principles demand consideration of how an LLM might prioritize different road users in ambiguous traffic scenarios, raising questions about embedded ethical values and fairness in decision-making. Applying SOTIF principles, such as scenario-based testing and risk analysis of functional insufficiencies, is critical for identifying and mitigating emergent hazards in language-driven behavior generation while ensuring ethical considerations are systematically addressed.

Similarly, IEEE P7009, which outlines Standard for Fail-Safe Design of Autonomous and Semi-Autonomous Systems, emphasizes transparency, predictability, and accountability in AI decision-making—core challenges for LLM-based architectures. The standard specifically addresses ethical concerns by requiring explicit consideration of harm prevention hierarchies and ethical fallback mechanisms. For AVs, this means designing systems where ethical considerations are built into both normal operation and degraded modes. LLMs, with their opaque inner workings and probabilistic outputs, often lack the traceability required by these frameworks. P7009 demands that AV designers implement transparent ethical reasoning that can be audited and validated against societal norms and legal requirements.

For deployment in safety-critical contexts, LLM systems must incorporate mechanisms for behavior bounding, fallback protocols, and robust introspection. Hybrid architectures that use LLMs for high-level reasoning while deferring critical control to verified deterministic modules align better with both SOTIF and P7009 principles, potentially offering a path toward ethically sound and regulatorily compliant autonomous systems.

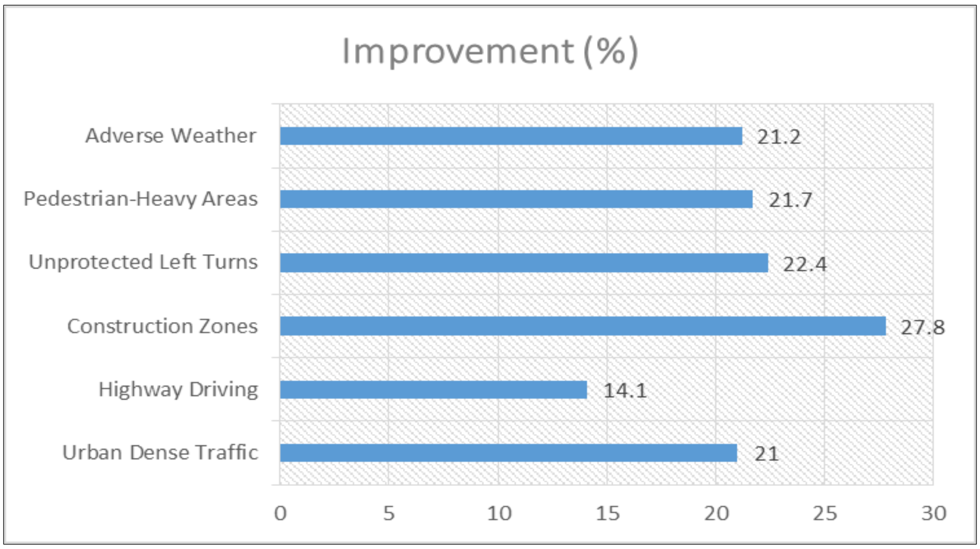

Figure 1 LLM-Enhanced AV System Performance Metrics Across Operational Scenarios [5-8]

6.4. Real-Time Processing Constraints

Autonomous vehicles operate under strict real-time constraints, requiring perception and decision cycles typically under 100ms. This temporal requirement presents challenges for LLM integration, as transformer architectures have quadratic complexity to input sequence length. Current implementations must carefully balance model size, context window, and inference speed to meet these constraints. Techniques such as token pruning, early stopping, and progressive resolution processing show promise but often trade accuracy for speed in ways that may compromise safety margins in critical scenarios.

6.5. Validation Methodologies

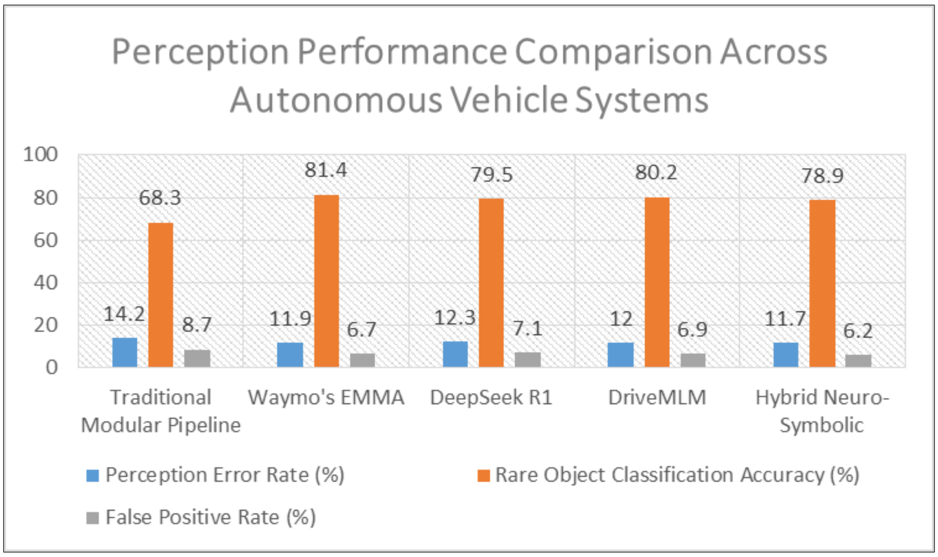

Figure 2 Perception Performance Comparison Across Autonomous Vehicle Systems [5- 9]

Traditional validation approaches for autonomous systems rely on scenario-based testing and statistical validation, which become exponentially more complex when applied to LLM-powered systems. The combinatorial explosion of possible inputs and the stochastic nature of model outputs create challenges for comprehensive validation. Current methodologies struggle to provide confidence bounds on system performance, particularly for edge cases and long-tail events that may trigger unexpected behaviors. The field requires new validation paradigms that can effectively assess both the capabilities and limitations of these advanced AI systems.

7. Future Research Directions

7.1. Hybrid AI Architectures

Future research is likely to focus on hybrid architectures that combine the strengths of LLMs with traditional algorithmic approaches. These systems will integrate neural networks with symbolic reasoning components, leveraging the flexibility and generalization capabilities of LLMs while maintaining the determinism and verifiability of classical methods where appropriate. Promising approaches include neuro-symbolic architectures that use LLMs for high-level reasoning while employing specialized models or algorithms for safety-critical functions. These hybrid systems aim to address current limitations while providing clearer paths to certification and deployment.

7.2. Edge Computing Optimization

Optimizing LLM deployment for edge computing environments represents a critical research direction for practical implementation. Future work will focus on model compression techniques such as quantization, pruning, and knowledge distillation to reduce computational requirements while maintaining performance. Research into specialized hardware accelerators designed specifically for transformer architectures shows promise for dramatic efficiency improvements. Distributed inference frameworks that intelligently partition models across vehicle computing resources and roadside infrastructure could enable more powerful models while maintaining real-time performance [9].

7.3. Human-Centric Interaction Models

Developing more intuitive and adaptive human-machine interfaces represents a promising direction for LLM application in autonomous vehicles. Future research will explore bidirectional communication channels that enable vehicles to explain their decisions, request clarification, and adapt to individual user preferences [10]. These systems will likely incorporate multimodal inputs including voice, gesture, and gaze tracking to create more natural interaction paradigms. Research suggests that effective communication can significantly increase trust and acceptance of autonomous systems, making this a critical area for advancement.

7.4. Reinforcement Learning Integration

The integration of reinforcement learning with LLM-powered systems offers significant potential for improving autonomous driving capabilities. Future research will explore how reinforcement learning can be used to fine-tune LLM behaviors based on real-world driving experiences while maintaining safety guarantees. Promising approaches include constrained policy optimization that respects safety boundaries while maximizing driving performance and comfort. Simulation-based reinforcement learning may bridge the gap between supervised learning and real-world deployment, allowing systems to safely explore diverse scenarios and learn from synthetic experiences.

7.5. Regulatory Framework Development

The development of appropriate regulatory frameworks for LLM-powered autonomous systems represents a critical research direction at the intersection of technology, policy, and ethics. Future work will need to establish testing protocols, performance metrics, and safety standards specifically designed for systems with emergent behaviors and probabilistic decision-making. Research into formal verification methods for neural networks shows promise for providing stronger safety guarantees. Collaborative efforts between industry, academia, and regulatory bodies will be essential to develop frameworks that both ensure public safety and enable technological progress in this rapidly evolving field.

Table 2 Performance Improvements and Challenges in LLM-Powered Autonomous Systems [8, 9]

Performance Area	Improvement Metrics	Enabling Technologies	Implementation Challenges	Future Research Needs
Perception Accuracy	reduction in perception errors, improvement in rare object classification	Cross-modal attention, Contextual reasoning, Semantic knowledge integration	Computational intensity, Real-time constraints, Data alignment issues	Specialized hardware accelerators, Model compression techniques, Improved multimodal training
Decision Quality	reduction in disengagements, improvement in ride smoothness, reduction in false positives	Chain-of-thought reasoning, Behavioral planning alignment, World knowledge integration	Safety validation, Regulatory approval, Explainability limitations	Reinforcement learning integration, Formal verification methods, Constrained policy optimization
Human Interaction	increase in user trust, improvement in operator intervention accuracy, successful interpretation of ambiguous commands	Natural language processing, Explainable AI techniques, Multimodal communication	interface standardization, Cultural variations, Training data limitations	Human-centric interaction models, Adaptive personalization, Regulatory framework development

8. Conclusion

The integration of Large Language Models into autonomous vehicle systems marks a transformative shift in artificial intelligence approaches to transportation. By bridging the gap between linguistic understanding and physical navigation, LLM-powered autonomous systems demonstrate unprecedented capabilities in contextual reasoning, adaptive decision-making, and human-machine collaboration. As exemplified by Waymo's EMMA and various research frameworks like DriveMLM, these technologies are already enhancing perception accuracy, enabling more natural human interaction, and improving navigational capabilities in complex environments. While significant challenges remain—from computational demands and safety validation to regulatory frameworks—the trajectory of development suggests a future where autonomous vehicles will navigate our roads with an increasingly human-like understanding of social context and environmental nuance. The continued evolution of hybrid architectures, edge computing optimizations, and reinforcement learning strategies promises to address current limitations while opening new possibilities for mobility. As researchers and industry leaders collaborate to overcome these challenges, LLM-powered autonomous systems stand poised to revolutionize transportation, making it safer, more accessible, and more intuitive for human participants in the complex dance of modern mobility.

References

[1] Waymo. (2023). “Introducing Waymo's Research on an End-to-End Multimodal Model for Autonomous Driving.” The Waymo Team, October 30, 2024. https://waymo.com/blog/2024/10/introducing-emma

[2] Zemian Ke; Zhibin Li et al., "Enhancing Transferability of Deep Reinforcement Learning-Based Variable Speed Limit Control Using Transfer Learning," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 7, pp. 4684-4695, 08 May 2020, doi: 10.1109/TITS.2020.2990598. https://ieeexplore.ieee.org/abstract/document/9090297

[3] Ye Zhang, Yiming Nie et al. “InternDrive: A Multimodal Large Language Model for Autonomous Driving Scenario Understanding.” AIAHPC '24: Proceedings of the 2024 4th International Conference on Artificial Intelligence, Automation and High-Performance Computing Pages 294 – 305. 04 October 2024. https://doi.org/10.1145/3690931.3690982

[4] Lu Wen, Jingliang Duan, et al., “Safe Reinforcement Learning for Autonomous Vehicles through Parallel Constrained Policy Optimization.”. 2020. https://arxiv.org/abs/2003.01303

[5] Jyh-Jing Hwang, Runsheng Xu, et al. "EMMA: Technical Deep Dive on Waymo's End-to-End Multimodal Architecture." Waymo Technical Reports. https://waymo.com/research/emma/

[6] Daya Guo Dejian Yang et al. “DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning”. DeepSeek-AI, 22 Jan 2025. https://arxiv.org/pdf/2501.12948

[7] Wenhai Wang, Jiangwei Xie, et al., “DriveMLM: Aligning Multi-Modal Large Language Models with Behavioral Planning States for Autonomous Driving”. Arxiv, 25 Dec 2023. https://arxiv.org/abs/2312.09245

[8] Rick Salay, Rodrigo Queiroz, et al. "An Analysis of ISO 26262: Using Machine Learning Safely in Automotive Software." Communications of the ACM, 65(7), 80-87, 2017. https://arxiv.org/abs/1709.02435

[9] Dewant Katare; Diego Perino et al. “A Survey on Approximate Edge AI for Energy Efficient Autonomous Driving Services.” IEEE Communications Surveys & Tutorials, 09 August 2023. https://ieeexplore.ieee.org/abstract/document/10213996

[10] Wayne. “LINGO-2: Driving with Natural Language”. 17 April 2024.

[11] https://wayve.ai/thinking/lingo-2-driving-with-language/

[12] Guo, Daya, et al. "DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning." ArXiv, 2025, https://arxiv.org/abs/2501.12948. https://arxiv.org/abs/2501.12948

[13] Long Ouyang, Jeff Wu et al. “Training language models to follow instructions with human feedback”. https://arxiv.org/pdf/2203.02155

[14] Bojarski, Mariusz, et al. "End to End Learning for Self-Driving Cars." ArXiv, 2016, https://arxiv.org/abs/1604.07316

[15] Alkinani, Monagi H., and Mahmoud R. "Patch-based Models and Algorithms for Image Denoising: A Comparative Review between Patch-based Images Denoising Methods for Additive Noise Reduction." EURASIP Journal on Image and Video Processing, vol. 2017, no. 1, 2017, pp. 1-27, https://doi.org/10.1186/s13640-017-0203-4

[16] Liu, Zhijian, et al. "BEVFusion: Multi-Task Multi-Sensor Fusion with Unified Bird's-Eye View Representation." ArXiv, 2022, https://arxiv.org/abs/2205.13542

[17] Thierry Deruyttere, Simon Vandenhende et al. “Talk2Car: Taking Control of Your Self-Driving Car”. arXiv, 26 Aug 2020. https://arxiv.org/abs/1909.10838

[18] Jiankai Sun, Hao Sun et al. “Neuro-Symbolic Program Search for Autonomous Driving Decision Module Design”. Proceedings of the 2020 Conference on Robot Learning, PMLR 155:21-30, 2021.https://proceedings.mlr.press/v155/sun21a.html

[19] Zhenjie Yang, Xiaosong Jia, et al. “LLM4Drive: A Survey of Large Language Models for Autonomous Driving”. arXiv, 12 Aug 2024. https://arxiv.org/abs/2311.01043