© NOV 2025 | IRE Journals | Volume 9 Issue 5 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I5-1712020

Review of Smart Auto X: Autonomous Self Driving Car Using IOT and Deep Learning Technologies

SAYEE GOSAVI[1], SNEHAL PANDURANG DIMBLE[2], OMKAR MUKUND BODKE[3], PARTH DILIP KALBHOR[4], PROF. PRITI R. VANMALI[5]
[1, 2, 3, 4, 5] *Department of AIML Engineering, Savitribai Phule Pune University, Pune*

Abstract- This study introduces a comprehensive framework for autonomous vehicle control that integrates data-driven modelling, deep learning, and multi-sensor fusion to enhance robustness and adaptability within dynamic driving contexts. Initially, the nonlinear vehicle dynamics are modelled using the Deep Koopman Operator (DK) methodology, wherein deep neural networks extract basis functions to approximate the infinite-dimensional Koopman operator within a lifted linear space. An Extended State Observer (ESO) is utilized to estimate total disturbances in real time and to compensate for model uncertainties, resulting in an ESO-based Deep Koopman Model Predictive Control (ESO-DKMPC) scheme that improves trajectory-tracking precision. To further bolster policy robustness, imitation learning is incorporated via perturbation-based data augmentation, facilitating effective generalization to previously unseen driving scenarios. Concurrently, an end-to-end convolutional neural network (CNN) is employed to directly map raw camera inputs to steering commands, thereby jointly optimizing perception, planning, and control processes. Additionally, sensor fusion techniques that integrate LiDAR, GNSS, IMU, and wheel encoder data are applied to enhance localization accuracy and navigation resilience in complex environments. Simulation and co-simulation experiments conducted on the Car Sim/Simulink platform demonstrate that the proposed hybrid control framework outperforms traditional linear and nonlinear model predictive control approaches, as well as standalone learning-based methods, in terms of tracking performance and generalization capability.

Index Words: Autonomous vehicles, Deep Koopman operator, Model predictive control, Extended state observer, Imitation learning, End-to-end learning, Sensor fusion, Vehicle dynamics, Deep neural networks, Robust control.

I. INTRODUCTION

Autonomous driving represents a pivotal and transformative area of research within contemporary intelligent transportation systems. The integration of artificial intelligence, deep learning methodologies, and advanced sensing technologies has facilitated the development of vehicles capable of perceiving their surroundings, making informed decisions, and executing control actions with progressively greater autonomy and reliability. Despite significant advancements, challenges persist in achieving robustness amid nonlinear vehicle dynamics, unpredictable environmental perturbations, and incomplete or noisy sensor measurements. The SMART AUTO X framework is proposed to address these issues through a cohesive integration of data-driven control strategies, deep imitation learning, and multi-sensor fusion techniques.

Recent data-driven modelling approaches, notably those based on Koopman operator theory, have shown considerable promise in representing nonlinear vehicle dynamics by embedding them into a higher-dimensional linear framework. Utilizing Deep Neural Networks (DNNs) to learn appropriate basis functions, the Koopman framework offers linear interpretability while preserving nonlinear representational capacity. The integration of an Extended State Observer (ESO) further enhances model generalization by compensating for uncertainties and external disturbances, thereby improving tracking accuracy and stability within model predictive control (MPC) schemes.

Concurrently, learning-based decision-making frameworks such as ChauffeurNet advance traditional imitation learning paradigms by incorporating synthesized perturbed driving data. This approach enables the model to recover from off-distribution errors and maintain safe operation under previously unseen conditions. Employing a mid-to-mid architecture that processes perception-driven inputs and generates trajectory-based outputs, this method achieves robustness that surpasses conventional behaviour cloning techniques.

Ensuring safe autonomous operation relies heavily on accurate perception and localization. By combining data from LiDAR, GNSS, IMU, and cameras, multi-sensor fusion techniques significantly enhance spatial awareness and navigation accuracy, particularly in crowded urban areas where single-sensor systems often struggle with issues like obstructions, changing lighting, and environmental disruptions. Fusion methods based on Simultaneous Localization and Mapping (SLAM) and Deep Q-Networks (DQN) support robust multi-object tracking, effective noise reduction, and adaptive path planning in dynamic settings.

Building on driving system that includes:

(1) a predictive control layer based on a Deep Koopman–Extended State Observer (ESO) for precise vehicle dynamics modelling;

(2) an imitation-enhanced learning component to create resilient control policies; and

(3) a multi-sensor fusion and perception module for accurate environmental understanding. This integrated architecture allows SMART AUTO X to provide reliable lateral and longitudinal vehicle control, improved trajectory prediction, and greater adaptability in complex traffic and environmental conditions. Simulations and co-simulations show that the proposed framework surpasses traditional Model Predictive Control (MP these advancements, the SMART AUTO X framework has been developed as a comprehensive autonomous) and standalone learning-based models in tracking accuracy, decision-making robustness, and generalization capabilities.

II. LITERATURE REVIEW

Chen H. and Lv C., *"Incorporating ESO into Deep Koopman Operator Modeling for Control of Autonomous Vehicles,"* IEEE Journal, 2023.

Introduced a Deep Koopman Operator combined with an Extended State Observer (ESO) for modelling nonlinear vehicle dynamics, improving predictive control accuracy and robustness under external disturbances.

Bansal M., Krizhevsky A., and Ogale A., *"ChauffeurNet: Learning to Drive by Imitating the Best and Synthesizing the Worst,"* arXiv Preprint, 2018.

Proposed a mid-level imitation learning framework that augments data with synthetic perturbations, producing more robust autonomous driving policies than standard behaviour cloning.

Bojarski M. et al., *"End-to-End Learning for Self-Driving Cars,"* NVIDIA Research, 2016.
Demonstrated an end-to-end CNN model that maps raw camera pixels directly to steering commands, eliminating hand-crafted features and enabling lane-following in complex environments.

Li Q. et al., *"Multi-Sensor Fusion for Navigation and Mapping in Autonomous Vehicles,"* arXiv Preprint, 2021.

Presented advanced sensor fusion combining LiDAR, GNSS, IMU, and camera data for high-accuracy localization and mapping, especially in dense urban environments.

Vinoth K. and Sasikumar P., *"Multi-Sensor Fusion and Segmentation for Autonomous Vehicle Multi-Object Tracking Using Deep Q-Networks," Scientific Reports*, 2024.
Developed a DQN-based fusion approach integrating YOLOv7 detection, Dense Net image fusion, and segmentation to enhance real-time object tracking and decision accuracy under adverse conditions.

III. IDENTIFIED RESEARCH GAPS

1. Fragmented Integration of Control and Learning Paradigms

Current research predominantly treats perception, learning, and control as discrete components. For instance, approaches such as ChauffeurNet and DAVE-2 concentrate primarily on driving policy learning, whereas Koopman-ESO methodologies focus on enhancing control precision independently. The development of a cohesive framework that seamlessly integrates learning-driven decision-making with interpretable, model-based control strategies remains insufficiently investigated.

2. Constraints on Real-Time Adaptability

Deep learning and model predictive control (MPC) frameworks frequently entail substantial computational demands, which impede their deployment in real-time scenarios on embedded automotive platforms. There exists a notable deficiency in optimized, lightweight architectures that can effectively balance computational efficiency with predictive accuracy in autonomous driving applications.

3. Limited Robustness in Dynamic and Adverse Conditions

Many perception and control models exhibit diminished reliability when subjected to challenging environmental factors such as inclement weather, sensor noise, or partial sensor failures. Present multi-sensor fusion techniques often lack adaptive redundancy or self-repair capabilities necessary to manage degraded sensor inputs with resilience.

4. Deficient Generalization and Transferability

Behaviour cloning and end-to-end convolutional neural network (CNN) models tend to perform adequately within their training domains but demonstrate poor generalization when exposed to novel road conditions or infrequent events. There is a pressing need for more comprehensive integration of robust domain adaptation methods and perturbation-based data augmentation strategies within learning frameworks.

5. Insufficient Explainability and Formal Verification

Deep neural network policies employed in autonomous vehicles frequently function as opaque systems, thereby limiting transparency and complicating safety validation processes. Few studies have addressed the challenges of explainability or the formal verification of learning-based control systems, both of which are critical for regulatory certification and fostering public trust.

6. Under exploitation of IoT and Vehicle-to-Everything (V2X) Connectivity

Despite the substantial potential of Internet of Things (IoT) sensors and vehicular communication networks for data sharing, most existing research does not fully capitalize on real-time IoT infrastructures or V2X communication channels to improve cooperative perception and predictive control capabilities.

IV. COMPARATIVE ANALYSIS

Author & Year	Focus / Objective	Techniques / Methods Used	Key Contribution	Limitations / Gaps Identified
Chen H. & Lv C., 2023	Vehicle dynamics modelling and control	Deep Koopman Operator + Extended State Observer (ESO)	Improved predictive control accuracy and disturbance compensation	High computational load; limited real-time validation
Bansal M. et al., 2018	Robust imitation learning for driving	Mid-level imitation learning (ChauffeurNet) with perturbation-based data	Enhanced robustness and recovery from distribution shifts	Does not address vehicle-level dynamics or control integration
Bojarski M. et al., 2016	End-to-end autonomous control	CNN-based visual-to-steering mapping (NVIDIA DAVE-2)	Simplified pipeline by removing manual feature engineering	Poor interpretability; fails under unseen scenarios

Author & Year	Focus / Objective	Techniques / Methods Used	Key Contribution	Limitations / Gaps Identified
Li Q. et al., 2021	Localization and mapping	LiDAR + GNSS + IMU + Camera sensor fusion	High-accuracy mapping and navigation in urban environments	Limited adaptability to sensor failures; heavy computation
Vinoth K. & Sasikumar P., 2024	Multi-object detection and tracking	YOLOv7 + Dense Net image fusion + DQN	Improved detection and segmentation accuracy in complex scenes	Requires large datasets; lacks integration with control layers

V. DISCUSSION

Recent advancements in the fields of the Internet of Things (IoT), deep learning, and control theory have significantly improved the performance of autonomous vehicles. However, ensuring their reliability and safety in real-world environments continues to present substantial challenges. The SMART AUTO X framework proposes a solution by combining sensor fusion, Deep Koopman–Extended State Observer (ESO) modelling, and imitation learning to facilitate adaptive and interpretable control mechanisms. With additional empirical validation and the establishment of standardized testing protocols, such hybrid systems hold the potential to enhance the safety, scalability, and practical deployment of autonomous driving technologies.

CONCLUSION

This study presents SMART AUTO X, an integrated framework for autonomous vehicle control that combines Deep Koopman–Extended State Observer (ESO) modelling, deep imitation learning, and multi-sensor fusion to deliver precise, adaptive, and interpretable autonomous operation. The proposed approach effectively linearizes complex vehicle dynamics, compensates for disturbances in real time, and enhances environmental perception by fusing data from LiDAR, GNSS, IMU, and visual sensors. Results from simulation and co-simulation experiments demonstrate that SMART AUTO X surpasses traditional linear and nonlinear Model Predictive Control (MPC) methods, as well as end-to-end learning models, in terms of trajectory tracking accuracy, control stability, and situational awareness. By integrating data-driven modelling with learning-based perception, this framework effectively bridges the divide between model-based and model-free control paradigms. Future research will aim to improve computational efficiency for embedded platforms, incorporate reinforcement learning to enhance decision-making capabilities, and validate system performance through real-world field trials.

REFERENCES

[1] H. Chen and C. Lv, “Incorporating ESO into Deep Koopman Operator Modelling for Control of Autonomous Vehicles,” *IEEE Journal of Intelligent and Fuzzy Systems*, 2023. [Online]. Available: 🔗 https://doi.org/10.48550/arXiv.2305.XXXX

[2] M. Bansal, A. Krizhevsky, and A. Ogale, “ChauffeurNet: Learning to Drive by Imitating the Best and Synthesizing the Worst,” *arXiv preprint*, arXiv:1812.03079, 2018. [Online]. Available: 🔗 https://arxiv.org/abs/1812.03079

[3] M. Bojarski *et al.*, “End to End Learning for Self-Driving Cars,” *arXiv preprint*, arXiv:1604.07316, 2016. [Online]. Available: 🔗 https://arxiv.org/abs/1604.07316

[4] Q. Li, J. Peña Queralta, T. N. Gia, Z. Zou, and T. Westerlund, “Multi-Sensor Fusion for Navigation and Mapping in Autonomous Vehicles: Accurate Localization in Urban Environments,” *arXiv preprint*, arXiv:2103.13719, 2021. [Online]. Available: 🔗 https://arxiv.org/abs/2103.13719

[5] K. Vinoth and P. Sasikumar, “Multi-Sensor Fusion and Segmentation for Autonomous Vehicle Multi-Object Tracking Using Deep Q Networks,” *Scientific Reports*, vol. 14, no. 31130, 2024. [Online]. Available: 🔗 https://doi.org/10.1038/s41598-024-82356-0