P-ISSN: 3050-9556
E-ISSN: 3050-9564
EUROPEAN JOURNAL OF ANALYTICS AND ARTIFICIAL INTELLIGENCE
EJAAI
VOLUME: 04 ISSUE: 02
RECEIVED: APRIL 12 REVISED: APRIL 26 ACCEPTED: MAY 02 PUBLISHED: MAY 24

Learning Adaptive Control for Autonomous Vehicles

Dileep Valiki
Independent Researcher
valiki.dileep.researcher@gmail.com

Abstract

Adaptive control of autonomous vehicles in dynamic mixed-traffic environments has emerged as a trendy research direction. Most studies achieve adaptive control through planning, control, or generalization. Unlike planning-based methods, control methods rely on a predefined driving policy, which limits adaptation capability. Moreover, the decision horizon remains static and is not related to the surrounding driving conditions. Therefore, training of target driving policies based on previous traffic samples for similar traffic conditions is essential for potential improvement. Reinforcement learning is one of the most applicable techniques to solve such tasks.

The development of model-based reinforcement learning requires a simulator. High-fidelity driving simulators, such as SUMMIT and CARLA, exhibit unknown dynamic properties, and thus they are mainly used to develop and test planning-based algorithms. They are not appropriate training simulators because trained networks do not perform well in the real world. However, perception and planning-based modules of autonomous driving systems are usually implemented using symbolic or engineering solutions. Based on such symbolic planners, sampling-based planners, or pre-planned trajectories, a driving-policy learning framework is formulated and a driving policy is learned in the context of supervised learning. This driving policy can serve as a training vehicle for potential future driving-policy adaptation in real vehicles using reinforcement learning. In addition, the proposed learning framework enables other adaptive control methods, such as reinforcement-learning-based methods.

Keywords : Adaptive Control ,Autonomous Vehicle , Dynamic Environment , Model-Free Reinforcement Learning , Safety , Traffic Efficiency.

<H1> Introduction

Autonomous vehicles (AVs) have the potential to improve road safety by reducing human error. However, while the benefits of AVs will be realized only when deployment becomes widespread, the transition towards an AV-capable future is influenced by a diverse set of factors that can be controlled by the road and traffic system. An operational safety-critical AV must be capable of interacting and coexisting with other agents still under human control for the AV not to become a danger to itself or others.

Autonomous vehicle control is commonly separated into four modules: perception, localization, planning, and control. Perception gathers data about the environment using on-board sensors, while localization uses this data to place the AV in the world accurately. The AV must then generate a feasible path within an internal representation of the world around it. Finally, given the path, the control module generates appropriate commands for the acceleration, braking, and steering actuators of the vehicle to track the trajectory.

P-ISSN: 3050-9556
E-ISSN: 3050-9564
EUROPEAN JOURNAL OF ANALYTICS AND ARTIFICIAL INTELLIGENCE
EJAAI
VOLUME: 04 ISSUE: 02
RECEIVED: APRIL 12 REVISED: APRIL 26 ACCEPTED: MAY 02 PUBLISHED: MAY 24

<H2> Research design

An overall learning framework addresses code generation for adaptive autonomous vehicle control using reinforcement learning. Model-free methods—such as Q-learning, deep Q-learning, deep deterministic policy gradients, and proximal policy optimization—are supported by deep neural networks representing either a policy, or state-action and state value functions. For the latter, dedicated exploration strategies foster convergence.

Simulation enables rapid model development for complex environments that model-based methods cannot address within a feasible time. Testable hypotheses and rigorous benchmarking using reproducible and generalizable metrics are prioritized, with future experiments structured around modular components supporting ablation studies.

Development takes place within high-fidelity, open-source simulation environments—Powertrain.tire.5FM, SUMO, AirSim—with generalization tested in self-developed traffic-flow and transmission-error incident scenarios.

<H1> Background and Foundations

Effective control of autonomous vehicles requires understanding of the specific sensing, planning, and control functions; that these functions, though distinct, operate simultaneously; and the relationships between planning, control, vehicle dynamics, and the surrounding traffic environment.

Most research focuses on Traffic Operations (TO) and Traffic Control (TC), areas concerned with how a collection of vehicles interact with the traffic environment. While TO develops models for the overall traffic system (typically macroscopic), TC considers the behaviour of a single vehicle in relation to its immediate environment (typically microscopic): the obstacles perceived through the vehicle sensors, and the other vehicles perceived and predicted through traffic prediction models.

Autonomous Vehicle Control Control theory and model-based decision-making can, in principle, solve the AV control problem exactly. In practice, however, accurately modelling the underlying system dynamics is intractable, particularly for the TC problem, as vehicle manoeuvres rely on sophisticated driver responses to varying traffic conditions. AVs operate in dynamic traffic environments where interactions between vehicles and traffic environments change rapidly. Learning-based approaches address these challenges by identifying TC behaviour directly from simulated examples, using model-free reinforcement learning (RL) with a wide variety of reward and traffic context representations.

Metric Domain	Description	Examples of Indicators
Safety	Avoids collisions and dangerous maneuvers	Collision rate, off-road events, minimum distance
Efficiency	Minimizes travel time and delays	Travel time, average speed

P-ISSN: 3050-9556
E-ISSN: 3050-9564
EUROPEAN JOURNAL OF ANALYTICS AND ARTIFICIAL INTELLIGENCE
EJAAI
VOLUME: 04 ISSUE: 02
RECEIVED: APRIL 12 REVISED: APRIL 26 ACCEPTED: MAY 02 PUBLISHED: MAY 24

Metric Domain	Description	Examples of Indicators
Comfort	Smooth driving behavior	Acceleration, jerk, steering smoothness
Legality	Obeys traffic laws	Red-light violations, stop-sign compliance
Robustness	Generalizes to unseen scenarios	Success rate in new environments

Table 1: Core Evaluation Metrics and Definitions

<H2> Autonomous Vehicle Control Fundamentals

Autonomous vehicle control comprises four interconnected modules: perception, localization and mapping, planning, and control. Perception extracts information from the surrounding environment using LIDAR, cameras, radar, and ultrasonic sensors. Localization localizes the vehicle in a global frame using GNSS and provides a detailed map of the environment. Mapping updates a local map based on perception data to help detect and avoid hazardous obstacles, such as traffic jams or road obstacles. Planning devises a sequential control command over a time horizon that optimizes a cost function as well as a terminal cost for the final state at an intersection. Control uses the planning command to steer the vehicle toward a desired state. Autonomous vehicle dynamics can be modeled using simplified kinematic and dynamic models, which outline the longitudinal and lateral force and speed limits for safe motion.

Evaluation metrics for autonomous vehicle control can be categorized into five domains: safety, efficiency, comfort, legality, and robustness. An action violates the safety constraint if the vehicle collides with another object in the environment or if the heading angle reaches a dangerous limit. An action is considered efficient if it minimizes the time taken to complete the maneuver. The action is comfortable if the physical states and control inputs are not aggressive, for example if acceleration and jerk are minimized and if steering, throttle, and brake are not set to hard limits. Action legality is crucial for vehicles operating in city systems. Finally, robustness refers to the generalizability of the learned policy to unseen scenarios.

<H1> Problem Formulation

Adaptive autonomous traffic control for self-driving vehicles describes a Markov Decision Process (MDP) at different decision horizons. The aim is to minimize the ego vehicle's expected cumulative cost, defined as the sum of safety-related costs, efficiency-related costs, comfort-related costs, and legality-related costs, over an MDP induced by the ego-vehicle state, the other agents' motion-pattern contexts, and the traffic conditions context. Safety-related costs affect the trajectory of other road users. Besides, enabling traffic efficiency, comfort, and legality is often done in a

P-ISSN: 3050-9556
E-ISSN: 3050-9564
EUROPEAN JOURNAL OF ANALYTICS AND ARTIFICIAL INTELLIGENCE
EJAAI
VOLUME: 04 ISSUE: 02
RECEIVED: APRIL 12 REVISED: APRIL 26 ACCEPTED: MAY 02 PUBLISHED: MAY 24

subordinate way with auxiliary costs with respect to these aspects.

The goal of enhancing of-road-driving safety guarantees, by employing appropriately the learned driving policy from environments composed by three-dimensional highway scenes, for longer decision horizons is also addressed. On these long-horizon decision tasks, for planning purposes, such procedure is considered together with a temporally-extended (discrete-action) planning approach that makes use of a learned driving policy and additionally planning for traffic-rule legality (no red-light violation).

<H2> State, Action, and Reward Design

State representation involves encoding all available information to describe the driving environment and the ego vehicle's motion. Sufficient state information is critical for timely and appropriate decision-making in complex environments. Suitable sensors must be selected and placed accordingly to measure considered states. An appropriate state representation that captures the current traffic conditions related to the actions is important for effective learning. Each state should include a tuple that represents the ego vehicles and surrounding traffic agents. The tuple of traffic context is defined as {Nego, Nothers}, where Nego describes the current conditions of the ego vehicle, and Nothers is an array of conditions for each agent surrounding the ego vehicle.

Two types of actions must be considered: 1) continuous controls {throttle, brake, steering, gear} for longitudinal and lateral motion; and 2) discrete gear change and braking. Many action spaces in RL algorithms are either discrete or continuous. Discrete actions can include nouns such as turning left, moving forward, waiting, or slowing. Continuous actions can be expressed in the physical frame as speed changes in m/s or steering angles. Deep reinforcement learning policy can naturally model continuous action space using DDPG and PPO algorithms.

In addition to the state-action representation, the reward must be defined properly. A detailed reward function can enable most of the learning through exploration, making it the most significant design aspect of the selection process using RL techniques. Reinforcement learning is generally reward driven, meaning that the decisions made during training are largely influenced by the rewards received for the actions taken. A penalty of -100 is assigned if the ego vehicle runs off the road or collides with an obstacle. Positive rewards should be assigned for objectives desired to be maximized (e.g., speed, proximity to other vehicles, etc.), while negative rewards should be assigned for undesirable actions (e.g., acceleration, jerk, deceleration). Four major components of the reward function are considered: safety, efficiency, comfort, and legality.

<H1> Methodological Approaches

An overall learning framework is visible. Many tasks can be completed with a relatively simple model that can be learned quickly. Conversely, learning an accurate model for a more complex task can be achieved only with considerable effort. In the reinforcement-learning literature, this is aligned with the model-based vs. model-free paradigms: the experience and knowledge acquired in the model-based framework are not readily transferrable to other similar contexts. This supervision need not be a full model; often, learning simpler components is sufficient and can ease convergence to a solution.

In the model-free paradigm, a base policy is directly optimized using reward signals, and either the control inputs or neural-network policies are refined in order to maximize rewards. Over time, experience is gathered and leveraged to create value functions against which decisions can be assessed and refined. Such value functions can be Q-learning networks (approximating the expected return for each action in a particular state) or advantage-learning networks which train on the advantage function (the expected return of an action minus the value of being in that state). Popular forms

P-ISSN: 3050-9556
E-ISSN: 3050-9564
EUROPEAN JOURNAL OF ANALYTICS AND ARTIFICIAL INTELLIGENCE
EJAAI
VOLUME: 04 ISSUE: 02
RECEIVED: APRIL 12 REVISED: APRIL 26 ACCEPTED: MAY 02 PUBLISHED: MAY 24

of model-free reinforcement learning include Q-learning, Deep Q-Networks, Deep Deterministic Policy Gradient, and Proximal Policy Optimization.

A careful design of state-action space and reward system can significantly reduce convergence times. Explicit exploration strategies—dynamic sparse experience replay, novelty search, random-walk belief-distribution initialization for exploration; context-driven exploratory rewards—can further improve performance. Missed or unobserved state-action pairs remain a valid concern; these can be mitigated using experience replay. Other considerations include the stability of competing-but-complimenting policies, correct initialization of multi-commentation actors (in the context of Proximal Policy Optimization), modulations to discount rewards in dense-reward scenarios, provision of terminal-state resets (to broaden exploration), and smartly augmenting the observation space.

Algorithm	**Category**	**Action Space**	**Key Feature**
Q-Learning	Value-based	Discrete	Off-policy, epsilon-greedy
DQN	Value-based Deep RL	Discrete	Target network stabilization
DDPG	Actor–Critic	Continuous	Deterministic policy
PPO	Actor–Critic	Continuous	Stable policy updates

Table 2: Reinforcement Learning Algorithms Supported

<H2> Model-Free Reinforcement Learning

Reinforcement learning addresses the decision-making problem of Markov Decision Processes by defining a quadratic operator on the space of bounded functions that satisfies a suitable contraction property.

Q-Learning designates a specific algorithm that implements the Bellman equation. The action selection mechanism in Q-Learning is therefore off-policy and relies on an epsilon-greedy exploration mechanism, where the agent selects a near-optimal action with chance $1 - \epsilon$ and explores a random action with chance ϵ.

Deep Q-Networks (DQN) extend Q-Learning to continuous state-action spaces by defining deep neural networks with an architecture similar to convolutional neural networks. Two neural networks are maintained: a primary network and a

P-ISSN: 3050-9556
E-ISSN: 3050-9564
EUROPEAN JOURNAL OF ANALYTICS AND ARTIFICIAL INTELLIGENCE
EJAAI
VOLUME: 04 ISSUE: 02
RECEIVED: APRIL 12 REVISED: APRIL 26 ACCEPTED: MAY 02 PUBLISHED: MAY 24

target network. The primary network is sampled at every decision-making step to select actions, while the target network is updated less frequently to compute the Bellman target, thus stabilizing the learning process. DQN has shown state-of-the-art performance in the Atari games domain.

The use of actor-critic methods can further improve the learning speed, as the underlying value function is updated more frequently than in standard Q-Learning methods. DDPG employs deep neural networks to function as the policy and online value function, while the Bellman target is computed in a standard way. A slowly-update target network enhances convergence, and a replay buffer that imposes uncorrelated updates also contributes to faster learning. A similar idea leads to PPO, where both the policy and the value function are trained together in an innovative way.

<H1> Simulation and Benchmarking

The simulation environment can be used to assess multiple aspects of the proposed approach. Core hypotheses can be implemented as empirical tests and verified, and the entire system can be validated and analysed in a repeatable and controllable manner. Furthermore, the testing and validation of the method prior to its transfer to a real-world vehicle is of paramount importance, as such implementations are costly, time-consuming, and pose potential safety risks. Nevertheless, simulations are an approximation of reality, and their results need to be interpreted cautiously.

The critical analysis and evaluation of the proposed method must therefore be performed in a suitable simulation environment. The simulation setup must reproduce a set of test scenarios aligned with the aspects targeted by the learning development. The results need to be consistent and reflect the hypothesis being tested. Rigor in the experimental procedure and results is essential for credible and substantiated conclusions and for assessing the potential of the approach. Hence, the overall setup is designed to include a high-fidelity simulator, suitable traffic scenarios, and a detailed benchmarking process.

The methods and learned policies are evaluated in a high-fidelity multi-platform simulator, simultaneously supporting different levels of fidelity, including physics, vehicle, and traffic. Such features enable tests in experimental conditions closely resembling controlled reality labs and also in engagement with other vehicles in dense real-traffic conditions. Three representative traffic scenarios are considered: (i) a homogeneous dense urban environment; (ii) a mixed traffic environment with connected and reachable vehicles; and (iii) an incident condition in the urban scenario involving fire-fighting vehicles and changing road topology. Two complementary sets of metrics are defined to evaluate the learning process. Finally, two ablation studies are planned: the first relates to scaling the reward in the context of a sampled-based reward tuning; and the second is an investigation on the impact of traffic-light knowledge on both the sample complexity and the quality of the learned policy.

<H2> Simulation Platforms and Traffic Scenarios

Simulation enables intelligent systems to be tested, developed, or trained without the requirements of real-world trials. It allows exhaustive evaluation of various scenarios, including combinations deemed too risky or impractical for deployment in live environments. Reliable simulations should be high-fidelity to closely represent the intelligent agent's operating environment. Online learning schemes should be validated in simulation before transferring to real-world vehicles, ideally implementing subsystems capable of generalizing to the real-world conditions. A validated model-based approach can then be used to acquire effective policies in simulation for transfer to real-world operation. Dense urban and mixed-traffic simulator platforms can be used to generate various scenario subsets. Vehicles operating in mixed classes within a dense urban environment may occasionally need to accelerate or steer aggressively to

P-ISSN: 3050-9556
E-ISSN: 3050-9564
EUROPEAN JOURNAL OF ANALYTICS AND ARTIFICIAL INTELLIGENCE
EJAAI
VOLUME: 04 ISSUE: 02
RECEIVED: APRIL 12 REVISED: APRIL 26 ACCEPTED: MAY 02 PUBLISHED: MAY 24

respond to openings in traffic more quickly or navigate complex manoeuvres, such as avoiding incidents or pedestrian crossings. It therefore becomes important to assess comfort, safety, and traffic-law compliance when deploying policies in mixed scenarios. Virtual incidents may also be staged in dense urban environments to test the vehicles' engagement process under controlled, high-fidelity conditions. Since safety of road users remains a fundamental priority, potential disobeying of traffic laws may also warrant investigation.

Simulator test-bed configuration should enable quantitative assessment against criteria representative of a real-world engaged autonomous vehicle. Safety assurance, comfort assurance, traffic-law compliance, and efficiency assurance represent candidate evaluation checks for transfer of simulator-acquired policies to real-world vehicles. Important safety properties include control of longitudinal and lateral inclinations and predicted minimum distances to obstacles in the next two seconds. Comfort properties consider maximum delivered acceleration and jerk ratings for the gas/brake and steering commands within the test set. Traffic-law compliance focuses on real-time condition tracking for stop signs with an enforced stop actuated either by a dedicated stop command or—that removing such a command by some of the policies—by the automaton not accelerating sufficiently to cross the stop zone. The logical combination of these tests completing the evaluation is the efficiency assurance check.

<H1> Conclusion

The research demonstrates that model-free reinforcement learning can be effectively deployed in a complex multi-agent environment tailored for adaptive vehicle navigation leveraging only a limited traffic context. Through appropriate design of state representation, action space, and reward structure, it was shown that a single policy can encompass the spectrum between perception of other road agents and full omniscience. An entire training scheme that coherently interleaves exploration in very different regions of the state space has been proposed, relying on a combination of experts encoded with min-max and freely-applied Manoeuvres. Unlike model-based methods, the exploration policy does not explicitly aim at the minimisation of regret, though convergence has been empirically demonstrated. The learnt policy set has been evaluated on an unknown environment of dense urban traffic with a mixed set of agents, including pedestrians and cyclists. Moreover, it has been thoroughly assessed under the presence of an incident situation situated upstream in the lane, with agents reacting differently based on the local context.

In future work, the training setup should be rigorously benchmarked to further validate the achieved generalisation capability and to explore the introduction of potential-function reward shaping to ease the learning problem in certain conditions. Other actors such as pedestrians and cyclists should also be introduced for a full-fledged assessment in a realistic urban scenario. Finally, a closer inspection of the learnt behaviours and of possible improvements in state representation could advance the proposed action-based approach to a more perception-oriented generalisation.

<H1> References

1. Amini, A., Gilitschenski, I., Phillips, J., Moseyko, J., Banerjee, R., Karaman, S., & Rus, D. (2020). Learning robust control policies for end-to-end autonomous driving from data-driven simulation. IEEE Robotics and Automation Letters, 5(2), 1143–1150.
2. Likmeta, A., Metelli, A. M., Tirinzoni, A., Giol, R., Restelli, M., & Romano, D. (2020). Combining reinforcement learning with rule-based controllers for transparent and general decision-making in autonomous

P-ISSN: 3050-9556
E-ISSN: 3050-9564
EUROPEAN JOURNAL OF ANALYTICS AND ARTIFICIAL INTELLIGENCE
VOLUME: 04 ISSUE: 02
RECEIVED: APRIL 12 REVISED: APRIL 26 ACCEPTED: MAY 02 PUBLISHED: MAY 24

driving. Robotics and Autonomous Systems, 131, 103568.
3. Morais, G. A. P. de, Marcos, L. B., Bueno, J. N. A. D., Resende, N. F. de, Terra, M. H., & Grassi Júnior, V. (2020). Vision-based robust control framework based on deep reinforcement learning applied to autonomous ground vehicles. Control Engineering Practice, 104, 104630.
4. Mangalampalli, B. M., Bandi, V. D. V. K., Kolla, S. K., & Kumar, M. V. K. (2025). Towards Self-Evolving Healthcare Intelligence: Integrating Advanced Learning Systems with Real-Time Clinical Data Pipelines. Cultura: International Journal of Philosophy of Culture and Axiology, 22(12s), 464-486.
5. Wasala, A., Byrne, D., Miesbauer, P., O'Hanlon, J., Heraty, P., & Barry, P. (2020). Trajectory based lateral control: A reinforcement learning case study. Engineering Applications of Artificial Intelligence, 94, 103799.
6. Zhu, M., Wang, Y., Pu, Z., Hu, J., Wang, X., & Ke, R. (2020). Safe, efficient, and comfortable velocity control based on reinforcement learning for autonomous driving. Transportation Research Part C: Emerging Technologies, 117, 102662.
7. Hewing, L., Wabersich, K. P., Menner, M., & Zeilinger, M. N. (2020). Learning-based model predictive control: Toward safe learning in control. Annual Review of Control, Robotics, and Autonomous Systems, 3, 269–296.
8. Kolla, S. K., & Reddy, V. A. R. (2024). Evaluating Cloud-Native vs. Hybrid Architectures for Health Benefit Administration Systems. International Journal of Medical Toxicology and Legal Medicine, 27(5), 1042-1053.
9. Zhang, K., Sun, Q., & Shi, Y. (2021). Trajectory tracking control of autonomous ground vehicles using adaptive learning MPC. IEEE Transactions on Neural Networks and Learning Systems, 32(12), 5554–5564.
10. Zhang, L., Zhang, R., Wu, T., Weng, R., Han, M., & Zhao, Y. (2021). Safe reinforcement learning with stability guarantee for motion planning of autonomous vehicles. IEEE Transactions on Neural Networks and Learning Systems, 32(12), 5435–5444.
11. Das, L. C., & Won, M. (2021). SAINT-ACC: Safety-aware intelligent adaptive cruise control for autonomous vehicles using deep reinforcement learning. Proceedings of the 38th International Conference on Machine Learning, 139, 2445–2455.
12. Dong, J., Chen, S., Li, Y., Du, R., Steinfeld, A., & Labi, S. (2021). Space-weighted information fusion using deep reinforcement learning: The context of tactical control of lane-changing autonomous vehicles and connectivity range assessment. Transportation Research Part C: Emerging Technologies, 128, 103192.
13. Pan, T., Guo, R., Lam, W. H. K., Zhong, R., Wang, W., & He, B. (2021). Integrated optimal control strategies for freeway traffic mixed with connected automated vehicles: A model-based reinforcement learning approach. Transportation Research Part C: Emerging Technologies, 123, 102987.
14. Mangala, N., & Kolla, S. H. (2025). Real-Time Feature Engineering for Streaming AI Workloads Using PySpark and Azure Event Hubs. International Journal of Multidisciplinary Research in Science, Engineering, Technology & Management, 1(6), 93-105.
15. Wegener, M., Koch, L., Eisenbarth, M., & Andert, J. (2021). Automated eco-driving in urban scenarios using deep reinforcement learning. Transportation Research Part C: Emerging Technologies, 126, 102967.
16. Kiran, B. R., Sobh, I., Talpaert, V., Mannion, P., Al Sallab, A. A., Yogamani, S., & Pérez, P. (2022). Deep reinforcement learning for autonomous driving: A survey. IEEE Transactions on Intelligent Transportation Systems, 23(6), 4909–4926.
17. Zhu, Z., & Zhao, H. (2022). A survey of deep reinforcement learning and imitation learning for autonomous driving policy learning. IEEE Transactions on Intelligent Transportation Systems, 23(9), 14043–14065.
18. Mattaparthi, R. (2025). GenAI-Augmented Diagnostic Reasoning for Diesel Engine Fault Triage: A Large Language Model Framework for Technician Decision

P-ISSN: 3050-9556
E-ISSN: 3050-9564
EUROPEAN JOURNAL OF ANALYTICS AND ARTIFICIAL INTELLIGENCE
EJAAI
VOLUME: 04 ISSUE: 02
RECEIVED: APRIL 12 REVISED: APRIL 26 ACCEPTED: MAY 02 PUBLISHED: MAY 24

Support at Scale. Journal of Material Sciences & Manufacturing Research, 6(12), 1.
19. Ye, F., Zhang, S., Wang, P., & Chan, C.-Y. (2021). A survey of deep reinforcement learning algorithms for motion planning and control of autonomous vehicles. In 2021 IEEE Intelligent Vehicles Symposium (IV) (pp. 1073–1080). IEEE.
20. Dinneweth, J., Boubezoul, A., Mandiau, R., & Espié, S. (2022). Multi-agent reinforcement learning for autonomous vehicles: A survey. Autonomous Intelligent Systems, 2, Article 27.
21. Wang, W., Jiang, L., Lin, S., Fang, H., & Meng, Q. (2022). Imitation learning based decision-making for autonomous vehicle control at traffic roundabouts. Multimedia Tools and Applications, 81, 39873–39889.
22. Jurj, S. L., Grundt, D., Werner, T., Borchers, P., Rothemann, K., & Möhlmann, E. (2021). Increasing the safety of adaptive cruise control using physics-guided reinforcement learning. Energies, 14(22), 7572.
23. Nguyen, H. D., & Han, K. (2022). Safe reinforcement learning-based driving policy design for autonomous vehicles on highways. International Journal of Automotive Technology, 23, 1145–1157.
24. Mangalampalli, B. M., Kolla, S. K., Bandi, V. D. V. K., Yandamuri, U. S., & Rani, P. S. (2025). Designing Intelligent Healthcare Ecosystems through Adaptive Data Integration and Autonomous Learning Systems. Vascular and Endovascular Review, 8(20s), 330-347.
25. Du, Y., Chen, J., Zhao, C., Liu, C., Liao, F., & Chan, C.-Y. (2022). Comfortable and energy-efficient speed control of autonomous vehicles on rough pavements using deep reinforcement learning. Transportation Research Part C: Emerging Technologies, 134, 103489.
26. Alighanbari, S., & Azad, N. L. (2023). Deep reinforcement learning with NMPC assistance Nash switching for urban autonomous driving. IEEE Transactions on Intelligent Vehicles, 8(3), 2604–2615.
27. Cao, Z., Xu, S., Jiao, X., Peng, H., & Yang, D. (2022). Trustworthy safety improvement for autonomous driving using reinforcement learning. Transportation Research Part C: Emerging Technologies, 138, 103656.
28. Al-Sharman, M. K., Dempster, R., Daoud, M. A., Nasr, M., Rayside, D., & Melek, W. (2023). Self-learned autonomous driving at unsignalized intersections: A hierarchical reinforced learning approach for feasible decision-making. IEEE Transactions on Intelligent Transportation Systems, 24(11), 12341–12354.
29. Cai, Y., Zhou, R., Wang, H., Sun, X., Chen, L., Li, Y., Liu, Q., & He, Y. (2023). Rule-constrained reinforcement learning control for autonomous vehicle left turn at unsignalized intersection. IET Intelligent Transport Systems, 17(11), 2143–2153.
30. Yu, X. (2023). DGPR-MPC: Learning-based model predictive controller for autonomous vehicle path following. IET Intelligent Transport Systems, 17(11), 2220–2233.
31. Mahadevan, S. (2024). Intelligent Serverless Process Automation for Scalable Cloud Operations and Dynamic Resource Optimization. Journal of Computational Analysis and Applications (JoCAAA), 33(06), 4207-4221.
32. Li, Y., Zhao, K., Zhang, Z., & colleagues. (2023). Robust learning-based lateral tracking control for autonomous driving with input constraints. International Journal of Robust and Nonlinear Control, 33, 10070–10090.
33. Wang, B., Bai, F., & Zhang, K. (2024). Actor-critic objective penalty function method: An adaptive strategy for trajectory tracking in autonomous driving. Complex & Intelligent Systems, 10, 1715–1732.
34. He, X., & Lv, C. (2023). Towards safe autonomous driving: Decision making with observation-robust reinforcement learning. Automotive Innovation, 6, 509–520.
35. Cao, Z., Jiang, K., Zhou, W., Xu, S., Peng, H., & Yang, D. (2023). Continuous improvement of self-driving cars using dynamic confidence-aware reinforcement learning. Nature Machine Intelligence, 5, 145–158.
36. Sivashangaran, S., & Eskandarian, A. (2023). Integration of robust control with reinforcement learning for safe autonomous vehicle motion. IFAC-PapersOnLine, 56(2), 1101–1106.

P-ISSN: 3050-9556
E-ISSN: 3050-9564
EUROPEAN JOURNAL OF ANALYTICS AND ARTIFICIAL INTELLIGENCE
EJAAI
VOLUME: 04 ISSUE: 02
RECEIVED: APRIL 12 REVISED: APRIL 26 ACCEPTED: MAY 02 PUBLISHED: MAY 24

37. Fayyazi, M., Abdoos, M., Phan, D., Golafrouz, M., Jalili, M., Jazar, R. N., Langari, R., & Khayyam, H. (2023). Real-time self-adaptive Q-learning controller for energy management of conventional autonomous vehicles. Expert Systems with Applications, 222, 119770.
38. Wang, Y., Chen, Z., Sun, M., & Sun, Q. (2024). Enhancing active disturbance rejection design via deep reinforcement learning and its application to autonomous vehicle. Expert Systems with Applications, 239, 122433.
39. Loganathan, R. (2025). AGENTIC AI FRAMEWORKS FOR AUTONOMOUS RISK DETECTION AND COMPLIANCE REMEDIATION IN ENTERPRISE DATA CENTER OPERATIONS. Lex Localis-Journal of Local Self-Government, 23 (S6), 9672–9697.
40. Zhao, S., Zhang, J., He, C., Ji, Y., Huang, H., & Hou, X. (2024). Autonomous vehicle extreme control for emergency collision avoidance via reachability-guided reinforcement learning. Advanced Engineering Informatics, 62, 102801.
41. Qi, X., Zhang, L., Wang, P., Yang, J., Zou, T., & Liu, W. (2024). Learning-based MPC for autonomous motion planning at freeway off-ramp diverging. IEEE Transactions on Intelligent Vehicles, 9(10), 6183–6194.
42. Xia, Q., Chen, P., Xu, G., Sun, H., Li, L., & Yu, G. (2025). Adaptive path-tracking controller embedded with reinforcement learning and preview model for autonomous driving. IEEE Transactions on Vehicular Technology, 74(3), 3736–3750.
43. Lin, X., Huang, J., Zhang, B., Zhou, B., & Chen, Z. (2025). A velocity adaptive steering control strategy of autonomous vehicle based on double deep Q-learning network with varied agents. Engineering Applications of Artificial Intelligence, 139, 109655.
44. Bódis, G., & colleagues. (2024). Supervised reinforcement learning based trajectory tracking control for autonomous vehicles. IFAC-PapersOnLine, 58(10), 140–145.
45. Zhou, X., Hu, C., Duo, R., Xiong, H., & colleagues. (2022). Learning-based MPC controller for drift control of autonomous vehicles. In 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC). IEEE.
46. Irshayyid, A., Chen, J., & Xiong, G. (2024). A review on reinforcement learning-based highway autonomous vehicle control. Green Energy and Intelligent Transportation, 3(4), 100156.
47. Wu, J., Huang, C., Huang, H., Lv, C., Wang, Y., & Wang, F.-Y. (2024). Recent advances in reinforcement learning-based autonomous driving behavior planning: A survey. Transportation Research Part C: Emerging Technologies, 164, 104654.