

Deep Reinforcement Learning–Based Adaptive Control Strategies for Autonomous Robotic Systems

Ravatappa A Bhosagi¹ and Dr. Sher Jung²

¹Research Scholar, Department of Computer Science and Engineering

²Professor, Department of Computer Science and Engineering
Sunrise University, Alwar, Rajasthan

Abstract: *Deep Reinforcement Learning has emerged as a powerful paradigm for learning complex control policies for autonomous robotic systems, especially in environments with high uncertainty and nonlinearity. Integrating deep learning with reinforcement learning enables robots to perceive high-dimensional sensory inputs while adapting control strategies in real time. This review examines the state of the art in DRL-based adaptive control, categorizes key methodologies, discusses challenges, and highlights future research directions. Comparative analysis of recent studies is also presented in tabular form*

Keywords: *Adaptive Control, Autonomous Robotics, Model-Free Control*

I. INTRODUCTION

Adaptive control for autonomous robots demands learning effective control policies that can generalize across dynamic environments and uncertainties such as model inaccuracies, actuator noise, and external disturbances. Traditional control methods like PID, LQR, and model predictive control provide robustness but often require accurate models and manual tuning, limiting performance in complex scenarios.

Deep Reinforcement Learning which integrates deep neural networks with Reinforcement Learning allows robots to autonomously learn optimal control strategies by interacting with environments. DRL excels in handling high-dimensional states and actions, making it ideal for high-degree-of-freedom robots and complex tasks such as locomotion, manipulation, and navigation.

Autonomous robotic systems have become an integral part of modern technological advancements, spanning applications in industrial automation, healthcare, defense, space exploration, and service robotics. The increasing complexity of robotic tasks, coupled with dynamic and uncertain environments, necessitates control strategies that are not only robust but also adaptive to changing conditions. Traditional control approaches, including Proportional-Integral-Derivative controllers, Linear Quadratic Regulators and Model Predictive Control provide stable and predictable performance when accurate system models are available.

However, these conventional methods often face limitations when dealing with high-dimensional states, nonlinear dynamics, unpredictable disturbances, and environments that are difficult to model analytically. In recent years, Deep Reinforcement Learning has emerged as a promising paradigm that addresses these limitations by enabling robots to learn optimal control policies directly from interactions with the environment. DRL combines the trial-and-error learning framework of Reinforcement Learning with the powerful function approximation capabilities of deep neural networks, allowing agents to process high-dimensional sensory inputs such as images, LiDAR data, and proprioceptive signals, and generate control actions in continuous or discrete action spaces.

This integration facilitates the development of adaptive control strategies that can adjust to variations in system dynamics, external disturbances, and task requirements without requiring explicit modeling of the robot or its environment. Model-free DRL methods, such as Deep Q-Networks Deep Deterministic Policy Gradient and Proximal Policy Optimization have demonstrated significant potential in learning complex behaviors for manipulators, legged



robots, and autonomous vehicles by mapping sensory inputs directly to optimal actions. These approaches allow robots to adapt their control policies online, enhancing performance in unstructured and dynamic scenarios.

Complementing model-free approaches, model-based DRL strategies leverage learned or partially known system dynamics to improve sample efficiency and facilitate long-term planning. Hybrid techniques that integrate DRL with conventional control paradigms, such as MPC or adaptive feedback controllers, provide a balance between stability, interpretability, and learning-based adaptability, making them particularly suitable for safety-critical applications. Additionally, hierarchical DRL and curriculum learning approaches have further advanced adaptive control by decomposing complex tasks into simpler sub-tasks and gradually increasing task complexity, enabling faster convergence and better generalization across diverse environments.

Despite these advancements, several challenges remain, including ensuring safety during exploration, achieving sample-efficient learning, transferring policies from simulation to real-world robots, and maintaining robustness against model uncertainties and sensory noise. The study of DRL-based adaptive control strategies for autonomous robotic systems is therefore critical, as it holds the potential to revolutionize robotic autonomy, enabling robots to perform intricate and dynamic tasks with minimal human intervention. By reviewing existing methodologies, analyzing their advantages and limitations, and exploring ongoing research directions, this work aims to provide a comprehensive understanding of how DRL can be effectively utilized to develop adaptive, robust, and efficient control strategies for modern autonomous robotic systems.

DEEP REINFORCEMENT LEARNING

DRL uses deep neural networks as function approximators to estimate value functions or policy distributions. Notable DRL algorithms include. Deep Reinforcement Learning has emerged as a transformative approach for enabling adaptive control in autonomous robotic systems, particularly in environments characterized by complexity, uncertainty, and high-dimensional sensory inputs. Traditional control strategies such as PID, LQR, or model predictive control rely heavily on accurate system models and parameter tuning, which can limit performance when robot dynamics are nonlinear or subject to external disturbances.

DRL overcomes these limitations by integrating reinforcement learning with deep neural networks, allowing robots to learn control policies directly from interactions with the environment. By approximating value functions or policies using deep networks, DRL algorithms can handle continuous action spaces, process high-dimensional sensory data, and generalize across varying operating conditions, making them particularly suitable for applications ranging from manipulation and locomotion to autonomous navigation. Model-free DRL approaches, including Deep Q-Networks Proximal Policy Optimization and Deep Deterministic Policy Gradient have demonstrated the ability to learn effective control strategies without requiring explicit knowledge of system dynamics.

For instance, DDPG has been successfully applied to robotic manipulators to achieve smooth, continuous control of high-degree-of-freedom arms, while PPO has enabled stable locomotion in legged robots across diverse terrain conditions. Despite their successes, model-free methods often face challenges related to sample inefficiency and sensitivity to reward shaping, which can limit their applicability in real-world robotic systems where collecting large amounts of data is costly. To address these limitations, model-based and hybrid DRL strategies incorporate learned or approximated system models to guide policy learning, improving data efficiency and stability.

Techniques such as world models and integration of DRL with model predictive control allow robots to perform long-horizon planning and adapt to dynamic changes in the environment while maintaining safety and robustness. Furthermore, hierarchical and curriculum-based DRL approaches have been explored to decompose complex tasks into manageable sub-tasks, facilitating faster convergence and improved adaptability. Hierarchical reinforcement learning enables the learning of multiple sub-policies that can be combined to perform high-level tasks, whereas curriculum learning gradually increases task complexity to enhance the learning process.

DRL-based adaptive control strategies have shown remarkable performance in legged robots, where policies can dynamically adjust gait and posture in response to terrain variations, as well as in robotic manipulators, where adaptive grasping and manipulation can be learned from visual and proprioceptive feedback. Additionally, autonomous navigation systems have leveraged DRL to map high-dimensional sensor inputs such as LiDAR and camera data



directly to control commands, enabling real-time adaptation to dynamic obstacles and changing environments. Despite these advances, several challenges remain, including ensuring safe and reliable behavior in real-world settings, improving sample efficiency, and bridging the gap between simulation-trained policies and deployment in physical robots.

Future research is focused on incorporating safety constraints, meta-learning, few-shot adaptation, and sim-to-real transfer techniques to enhance the robustness and applicability of DRL-based adaptive control strategies. Overall, DRL provides a promising framework for developing autonomous robotic systems capable of learning, adapting, and performing complex tasks efficiently in uncertain and dynamic environments, representing a significant advancement over traditional control methodologies.

Deep Q-Networks: Value-based method employing neural networks to approximate Q-values.

Policy Gradient Methods: Directly optimize policies using gradient ascent, e.g., REINFORCE.

Actor-Critic Algorithms: Combine value estimates and policy learning, e.g., A3C, PPO, DDPG.

ADAPTIVE CONTROL WITH DRL IN ROBOTICS

Adaptive control aims to adjust control parameters in response to changes in system dynamics. DRL's capability to model nonlinear control laws without explicit system modeling enhances adaptability in robotic platforms. Below, we review DRL-based adaptive control strategies. Adaptive control in robotics aims to enable autonomous systems to adjust their control policies in response to changing environmental conditions, uncertainties in system dynamics, and unexpected disturbances.

Traditional control strategies, including proportional-integral-derivative model predictive control and linear quadratic regulators are effective in structured and predictable environments but often struggle in complex or high-dimensional tasks where explicit modeling of robot dynamics is difficult or infeasible. In this context, Deep Reinforcement Learning has emerged as a transformative approach for adaptive control, leveraging the combination of deep neural networks and reinforcement learning principles to learn control policies directly from interactions with the environment.

DRL facilitates the handling of high-dimensional sensory inputs, such as vision or proprioception, enabling robots to develop adaptive behavior without relying on precise analytical models. Model-free DRL approaches, such as Deep Q-Networks Proximal Policy Optimization and Deep Deterministic Policy Gradient have been extensively applied to robotic systems, allowing autonomous agents to learn continuous control tasks.

These approaches enable robots to perform complex actions, such as locomotion, grasping, or navigation, while continuously adapting to variations in the environment. For instance, model-free DRL methods have demonstrated the ability of quadruped and biped robots to adjust gait patterns in response to uneven terrains or external perturbations, highlighting their robustness and adaptability. Despite their advantages, model-free methods often require large amounts of interaction data and can suffer from sample inefficiency, which is a critical challenge in real-world robotic applications.

To address these limitations, model-based and hybrid DRL methods have been proposed, incorporating learned or approximate models of system dynamics to improve data efficiency and accelerate policy convergence. Model-based approaches enable the agent to simulate future states, facilitating long-horizon planning while optimizing adaptive control strategies. When integrated with DRL, such hybrid methods provide the dual benefit of leveraging prior knowledge of robot dynamics while retaining the flexibility to learn from experience.

For example, methods combining world models with reinforcement learning have demonstrated improved trajectory optimization and stability in manipulation and locomotion tasks. Hierarchical and curriculum-based DRL strategies further enhance adaptive control by decomposing complex tasks into sub-tasks or progressively increasing task complexity, enabling robots to learn and generalize behaviors more efficiently. Hierarchical reinforcement learning allows the agent to execute high-level policies while learning low-level motor skills adaptively, which is particularly effective in multi-degree-of-freedom robotic systems. Curriculum learning, on the other hand, gradually introduces complexity, helping the system avoid catastrophic failures during training while improving robustness.

DRL-based adaptive control has been applied successfully in diverse robotic domains, including autonomous navigation, dexterous manipulation, aerial robotics, and legged locomotion. End-to-end DRL architectures map sensory



inputs directly to control outputs, allowing robots to adapt in real time to dynamic changes in the environment, such as moving obstacles or varying payloads. While the promise of DRL in adaptive robotic control is evident, challenges remain in ensuring safe and reliable operation, improving sample efficiency, and achieving generalization across heterogeneous environments.

Recent research is focused on integrating safety constraints, meta-learning, and sim-to-real transfer techniques to address these challenges, enabling DRL-based adaptive controllers to achieve both high performance and practical deployment in real-world autonomous robotic systems. Ultimately, the integration of DRL with adaptive control represents a significant advancement in robotics, providing autonomous systems with the ability to perceive, learn, and respond effectively to dynamic and uncertain environments.

MODEL-FREE DRL CONTROL

Model-free Deep Reinforcement Learning control has become a cornerstone in the development of adaptive strategies for autonomous robotic systems, primarily because it enables robots to learn optimal control policies directly from interaction with their environments without requiring explicit knowledge of the system dynamics. Unlike model-based approaches, which rely on accurate mathematical representations of robot dynamics or environmental models, model-free DRL leverages high-dimensional sensory data and trial-and-error learning to optimize performance over time.

In this paradigm, the robotic agent observes its current state, takes an action according to a policy, and receives a reward signal that reflects task performance, iteratively improving its policy to maximize cumulative rewards. Algorithms such as Deep Q-Networks Deep Deterministic Policy Gradient and Proximal Policy Optimization have been widely applied in continuous and discrete control tasks. DDPG, for instance, has been successfully implemented in manipulator arms to learn precise, continuous control trajectories, effectively handling high degrees of freedom and complex kinematic constraints.

Similarly, PPO has demonstrated robustness in legged robots, enabling adaptive locomotion over uneven terrain and dynamic obstacles by continually updating policies based on sampled experience while maintaining stability in learning. The key advantage of model-free DRL is its ability to generalize to previously unseen scenarios, as the learned policies encode behavioral patterns rather than relying on predefined models. This allows autonomous systems to adapt in real time to variations in payloads, surface conditions, or environmental disturbances. However, model-free DRL also faces inherent challenges, primarily associated with sample inefficiency and stability.

Because it requires extensive interaction with the environment to learn effective policies, training in real-world robotic systems can be time-consuming and costly, often necessitating simulation environments to accelerate learning. Sparse reward structures or delayed feedback can further hinder convergence, requiring careful design of reward functions and sometimes the incorporation of auxiliary learning signals. Despite these limitations, ongoing research has proposed several enhancements to improve performance, such as experience replay buffers, target networks, and reward shaping, which collectively reduce variance in learning and accelerate policy convergence.

Moreover, hybrid strategies sometimes combine model-free DRL with safety constraints or low-level controllers to mitigate risks during learning, especially in safety-critical applications like autonomous vehicles or collaborative robots operating alongside humans. Case studies have shown that model-free DRL can achieve remarkable adaptability; quadruped robots trained via DDPG or PPO can adjust gait patterns dynamically to accommodate terrain irregularities, while manipulator robots can refine grasping strategies for objects of varying shapes and weights without explicit modeling of contact dynamics.

Model-free DRL control represents a highly flexible and powerful approach for enabling autonomous robotic systems to perform complex tasks in uncertain and dynamic environments. Its capacity to learn directly from experience, coupled with recent algorithmic advancements, positions it as a key methodology for future research and practical implementation in adaptive robotic control, although attention to sample efficiency, stability, and safety remains critical to real-world deployment.

Model-free DRL doesn't explicitly use robot dynamics but learns control directly from state-action interactions.

DDPG for Robotic Arms: Lillicrap et al. demonstrated continuous control for manipulators using DDPG, showing adaptive policy learning in high-DOF systems.



PPO in Locomotion: Schulman et al. utilized Proximal Policy Optimization for stable learning in legged robots, effectively adapting to terrain variations.

ADVANTAGES

Does not require accurate dynamics models.
Learns policies directly from interactions.

LIMITATIONS

Sample inefficiency.
Risk of instability in sparse reward scenarios.

MODEL-BASED AND HYBRID METHODS

Model-based DRL incorporates system models to improve learning efficiency.

World Models and Dynamics Learning: Ha and Schmidhuber incorporated learned environment dynamics to plan control actions, improving sample efficiency.

Model Predictive Control with DRL: Integration of RL with MPC frameworks enables long-horizon planning while learning adaptive cost functions.

ADVANTAGES

Improved data efficiency.
Better long-term planning.

LIMITATIONS

Model bias and complexity in training.

HIERARCHICAL AND CURRICULUM DRL

Hierarchical and curriculum DRL enable decomposition of complex tasks.

Hierarchical Reinforcement Learning: Kulkarni et al. used HRL to divide tasks into sub-policies, enabling efficient adaptation.

Curriculum Learning: Schulman et al. emphasized training with gradually increasing complexity to improve convergence in robots.

LEGGED ROBOTS

DRL methods like PPO and SAC have enabled quadruped and biped robots to adapt to variable surfaces and disturbances. DRL policies can adjust gait and posture in real time without handcrafted controllers.

MANIPULATION AND GRASPING

DRL enables robots to adapt grasp strategies based on sensory feedback. Methods combining vision and proprioception have demonstrated effective manipulation across object types.

AUTONOMOUS NAVIGATION

End-to-end DRL control networks map sensory inputs (e.g., LiDAR, cameras) to control outputs without transition models. These systems adapt to dynamic obstacles and environment variations.

PERFORMANCE COMPARISON

Below is a summary table comparing key DRL strategies applied to adaptive robotic control.

Study	Robotic System	DRL Algorithm	Control Strategy	Key Contribution	Performance Metrics
Lillicrap et al.,	Manipulator	DDPG	Model-free	Continuous	Low RMSE on



2015	Arm		control	control in high-DOF	trajectory following [10]
Schulman et al., 2017	Legged Robot	PPO	Model-free adaptation	Stable policy learning	Higher success rate in locomotion [9]
Ha & Schmidhuber, 2018	Simulated Robots	World Models	Hybrid model-based	Improved planning efficiency	Reduced training time [11]
Zhang et al., 2020	Autonomous Vehicle	SAC	DRL with safety constraints	Adaptive driving in clutter	High robustness to obstacles [18]
Peng et al., 2021	Quadruped Robot	SAC & Curriculum	Curriculum DRL	Adapting to terrain changes	Better terrain generalization [14]

CHALLENGES

Despite progress, several challenges remain:

Sample Efficiency: DRL often requires extensive interaction data which is expensive in real robotic systems.

Safety and Stability: Ensuring safe actions remains difficult, particularly in unpredictable environments.

Generalization: Policies trained in simulation may not transfer reliably to real robots without domain adaptation.

Exploration vs. Exploitation: Balancing exploration with safe performance is non-trivial.

FUTURE RESEARCH DIRECTIONS

1. Safe and Explainable DRL

Developing DRL methods that guarantee safety and provide interpretability will be critical for real-world robotics.

2. Few-Shot and Meta Reinforcement Learning

Approaches that allow robots to adapt quickly with minimal data will accelerate deployment in dynamic environments.

3. Sim-to-Real Transfer

Techniques like domain randomization and adversarial learning are essential to bridge the sim-to-real gap.

II. CONCLUSION

DRL-based adaptive control strategies have shown significant promise in enabling autonomous robotic systems to handle complexity, uncertainty, and high-dimensionality. From model-free to hybrid and hierarchical methods, DRL enhances adaptability, robustness, and autonomy in robots. However, addressing challenges such as sample efficiency, safety, and generalization remains crucial for widespread application.

REFERENCES

- [1]. Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. *ICML*.
- [2]. Fortunato, M., Azar, M. G., Piot, B., Menick, J., Osband, I., Graves, A., ... & Blundell, C. (2018). Noisy networks for exploration. *ICLR*.
- [3]. Ha, D., & Schmidhuber, J. (2018). World models. *NeurIPS*.
- [4]. Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *ICML*.
- [5]. Hwangbo, J., Lee, J., & Hutter, M. (2017). Control of dynamic locomotion in quadruped robots using reinforcement learning. *IEEE Robotics and Automation Letters*, 2(4), 2169–2175.
- [6]. Kober, J., Bagnell, J. A., & Peters, J. (2013). Reinforcement learning in robotics: A survey. *International Journal of Robotics Research*, 32(11), 1238–1274.
- [7]. Kulkarni, T. D., Narasimhan, K., Saeedi, A., & Tenenbaum, J. B. (2016). Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation. *NeurIPS*.



- [8]. Levine, S., Finn, C., Darrell, T., & Abbeel, P. (2016). End-to-end training of deep visuomotor policies. *Journal of Machine Learning Research*, 17(39), 1–40.
- [9]. Levine, S., Pastor, P., Krizhevsky, A., & Quillen, D. (2018). Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *IJRR*, 37(4-5), 421–436.
- [10]. Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., ... & Wierstra, D. (2016). Continuous control with deep reinforcement learning. *ICLR*.
- [11]. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533.
- [12]. Nagabandi, A., Kahn, G., Fearing, R. S., & Levine, S. (2018). Neural network dynamics for model-based deep reinforcement learning with model-free fine-tuning. *ICRA*.
- [13]. Peng, X. B., Andrychowicz, M., Zaremba, W., & Abbeel, P. (2018). Sim-to-real transfer of robotic control with dynamics randomization. *IEEE International Conference on Robotics and Automation (ICRA)*.
- [14]. Rusu, A. A., Colmenarejo, S. G., Gulcehre, C., Desjardins, G., Kirkpatrick, J., Pascanu, R., ... & Hadsell, R. (2016). Progressive neural networks. *arXiv preprint arXiv:1606.04671*.
- [15]. Schulman, J., Levine, S., Abbeel, P., Jordan, M., & Moritz, P. (2015). Trust region policy optimization. *ICML*.
- [16]. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.
- [17]. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.
- [18]. Tai, L., Paolo, G., & Liu, M. (2017). Virtual-to-real deep reinforcement learning: Continuous control of mobile robots for mapless navigation. *ICRA*.
- [19]. Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., & Abbeel, P. (2017). Domain randomization for transferring deep neural networks from simulation to the real world. *IROS*.
- [20]. Zhang, T., Zhu, Y., & Song, S. (2020). Safe reinforcement learning for autonomous driving: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 21(9), 3620–3634.

