Acta Scientific COMPUTER SCIENCES

Volume 5 Issue 5 May 2023

Review Article

Advances in Reinforcement Learning: A Comprehensive Review of Real-World Applications in Industry

K Ussenko* and VI Goncharov

Division for Automation and Robotics, School of Computer Science and Robotics, Tomsk Polytechnic University, Russian Federation

***Corresponding Author:** K Ussenko, Division for Automation and Robotics, School of Computer Science and Robotics, Tomsk Polytechnic University, Russian Federation.

Received: March 15, 2023

Published: April 11, 2023

Abstract

This paper investigates the current feasibility of utilizing reinforcement learning algorithms in the industrial sector. Although many studies have showcased the success of these algorithms in simulations or on isolated real-world objects, there is a paucity of research examining their wider implementation in real-world systems. In this study, we identify the obstacles that must be surmounted to fully leverage the potential benefits of reinforcement learning algorithms in practical applications. Moreover, we present a thorough overview of existing literature aimed at tackling these challenges.

Keywords: Reinforcement Learning; Deep Learning; Sim-to-real; Engineering; Artificial Intelligence; Control; Robotics; Autonomous Control

Abbreviations

RL: Reinforcement Learning; DQN: Deep Q-Network; DDQN: Double Deep Q-Network; DDPG: Deep Deterministic Policy Gradient; MRAC-RL: Model Reference Adaptive Control and Reinforcement Learning; SCM: Supply Chain Management; IoT: Internet of Things; AIoT: Autonomous Internet of Things

Introduction

Reinforcement learning (RL) has emerged as a powerful class of algorithms for optimization and control tasks. RL gained widespread attention after the pioneering work by R. Sutton and A. Barto [1]. which spurred a surge of research in this field [2]. Notably, Q-learning and related algorithms such as DQN [3], DDQN [4] have become popular choices for RL applications. DQN and DDQN are among the earliest algorithms that use artificial neural networks as universal function approximators for Q-values. These models have demonstrated impressive performance on various control tasks, including the Cartpole problem [5].

An important development in RL was the introduction of actor-critic models, which separate the Q-function into two components: an actor responsible for policy selection, and a critic responsible for estimating action values. Actor-critic models [6] have brought RL closer to achieving human-level performance in Atari games [3].

However, applying RL algorithms to real-world tasks presents significant challenges [7]. Real-world systems often require extensive computational resources, formalization of the Markov decision process, and a controlled environment. Moreover, learning in the real world can cause wear and tear on the control object, potentially leading to its failure. These challenges have motivated researchers to develop methods for improving the performance and robustness of RL algorithms in real-world applications, either through algorithmic or statistical means.

In this overview, we aim to collect and synthesize the most significant studies on RL in real-world applications, particularly in industrial and control domains. By doing so, we hope to provide a comprehensive understanding of the current challenges and opportunities in applying RL to complex real-world problems.

Materials and Methods

The initial attempts at applying RL algorithms to real-world objects faced numerous challenges, as highlighted by [7]. These challenges are sticking point in resolving tasks with RL algorithms. Authors of this paper already tried overview the problematic of RL real-world application from systematic point of view. However, it still needs broader studies in the way of resolving this problematic.

Some of the earliest successful works in this area focused on modifying specific Actor-Critic models, such as those used in [8-10]. For example, [9] trained a Minitaur robot to walk on a flat surface, but the resulting algorithm was also able to perform well on more complex and curved surfaces. Similarly, [10] RL algorithms succeed in manipulating dynamic and deformable objects.

For more complex tasks with higher levels of disturbance and complexity, a combined approach such as that used by [11]. This approach utilizes a PI regulator for low-level control and RL algorithms for trajectory creation.

Another approach involves a fully hybrid use of RL and classical algorithms. However, in these works, the task is typically decomposed, and RL is only responsible for one part of the task [11].

These approaches seek to reduce the impact of the stochastic component of RL algorithms (exploration) to minimize control error and increase the deterministic component (exploitation) for greater robustness.

One major limitation of existing approaches is their narrow applicability to specific control objects and laboratory conditions. To extend their use in real-world applications, researchers are now exploring more complex tasks and entire systems where RL algorithms can be integrated.

To further improve the performance and applicability of RL algorithms, researchers are also exploring the integration of other techniques and methods, such as imitation learning, meta-learning, and hierarchical reinforcement learning. These approaches seek to enhance the sample efficiency, stability, and generalization capabilities of RL algorithms, making them more suitable for real-world applications.

One potential solution to improve the performance of reinforcement learning algorithms is to explicitly incorporate the control object model into the RL algorithm, as demonstrated in [12]. This category of algorithms is known as model-based reinforcement learning algorithms. In contrast, algorithms that do not include a model of the control object in their structure are referred to as model-free reinforcement learning algorithms. While Model-Based Reinforcement Learning models are generally more sample-efficient, their performance can be sensitive to the model used [13]. In cases where the model is simple and contains inaccuracies and uncertainties, model-free algorithms may be more appropriate. In addition, model-based algorithms can suffer from systematic errors, which can also negatively impact their performance [14]. However, in cases where the model is without significant uncertainties and disturbances, a model-based approach may be more appropriate. This was demonstrated by [15], where a model-based RL algorithm was successfully applied to the dynamic control of soft robotic manipulators.

When it comes to Model-free RL algorithms, there are concerns about their sample efficiency and safety [16]. To address these issues, one popular approach is to train RL algorithms on simulations and then transfer the learned model to the real environment, as described in [16]. However, due to the simulation to reality gap, the performance of the transferred model can drastically drop on real objects [17], a problem commonly known as the sim-to-real problem.

While there is no perfect solution to this problem, some researchers attempt domain adaptation [18] or domain randomization [19] to mitigate the gap between simulation and reality. Other researchers add disturbances to the data to increase the model's robustness [20], while some apply inverse dynamic modeling [21,22] to reduce the simulation to reality gap.

This is a summary of the techniques used to reduce the gap between simulation and real-world reality, as described in table 1. The summary provides information about the methods used, a

Method	References	Description	Application
Domain adaptation	[23]	Domain adaptation by transferring image labels in the source domain to images in the target domain	A 'hook loop' manipulation task
Domain randomization	[19]	Randomize visual appearance in simulations for policy training and sim-to-real policy transfer	Grasp objects in a cluttered real-world environment
	[24]	Randomize both visual input and physical parameters in simulations for sim-to-real transfer	Learning dexterous in-hand manipulation
Inverse dynamics model	[21]	Sim-to-real transfer by adapting the action selected by the simulation policy with a learned inverse dynamics model in the real world	Back-and-forth swing of a robot arm
	[22]	Sim-to-real policy transfer by modifying the simulation environment to be equivalent to the real world	Bipedal robot walking
Continual and multitask learning (PNNs)	[25]	Sim-to-real policy transfer via PNNs with raw pixels as input	Dynamic conveyor task
Meta-reinforcement learning	[26]	Meta-train a global dynamics model for fast online adaptation in dynamic environments	Track desired trajectories
	[27]	Meta-train a policy with model-free reinforcement learning for sim-to-real domain adaptation	Shoot a hockey puck to a target location

Table 1: Summary of main sim-to-real deep reinforcement learning policy transfer methods and their applications.

brief description of the work, and the application tasks to which they have been applied.

Overall, these techniques are used to bridge the gap between the simulation and real world, allowing models trained in simulation to generalize better to real-world scenarios.

These solutions, of course, facilitate the process of pre-trained model transfer, but do not solve the problems themselves. Moreover, there is currently no comprehensive study of the factors influencing the process of algorithm transfer and its final performance.

In various studies, these proposed solutions have been widely employed to investigate the feasibility of using reinforcement learning algorithms in practical settings. In this article, we will explore the different algorithm modifications used in research that applies reinforcement learning to industrial and robotics objects, the Internet of Things, and supply chain management.

By examining the types of modifications and techniques that have been used, we can better understand the adaptability of reinforcement learning algorithms in addressing real-world challenges. With this knowledge, researchers and practitioners can better tailor reinforcement learning algorithms to suit the specific needs of their respective industries, leading to more effective solutions for a range of applications.

Figure 1: Types of proposed solutions to real-world challenges.

Reinforcement learning in robotics and autonomous control

Many types of problems in autonomous control and robotics may be model as reinforcement learning problem [14]. Trial and error mechanism of RL architecture helps a robot and control system to autonomously learn an optimal behavior by interacting with the environment [28,29].

Some of the RL applications to solve robotic problems by using model-based algorithms include [30,31]. These works were proposed to enable a robot for the penalty kick, navigation task, vision-based mobile robot docking task and the task of obstacle avoidance respectively.

One approach to RL is to use model-based algorithms to enable robots to perform tasks such as penalty kicks, navigation, vision-based mobile robot docking, and obstacle avoidance. The Brainstorner Tribots introduced in [32] won the Robocup 2006 Midsize league and can learn various skills such as penalty shots, defenses, dribbling, interception, kicking, and motor speed and position control.

Soft robots have also gained significant attention in the industrial sector, and deep RL has been widely used for manipulation tasks such as reaching, door opening, picking and dropping [15,33].

Deep RL has also been successfully used in soft robotic navigation to assist control systems and robots in performing tasks autonomously, such as autonomous driving. Some important studies related to the use of deep RL methods for navigation are [34,35].

Reinforcement learning in internet of things

The integration of autonomous control systems and the Internet of Things (IoT) has led to the development of autonomous IoT (AIoT) systems. Reinforcement learning algorithms have introduced ambient intelligence into AIoT systems by providing solutions to closed-loop tasks involving processing sensory data to make control decisions [36]. Although AIoT is a relatively new trend, several recent works have explored the application of RL in autonomous IoT systems. For example, [37] proposes an IoT-enabled mobile robot for monitoring plant health using Q-learning and a CNN-based method. Q-learning has also been used for energy consumption optimization and delay in [38], while [39] discuss AIoT systems for energy storage management, energy trading processes, and other Q-learning applications in AIoT, respectively.

RL actor-critic methods have been widely used in AIoT systems to learn stochastic policies for continuous state or action space problems and, in a few cases, for discrete state problems. Additionally, the actor-critic method can be used to train deep RL models with fewer computational resources and samples. Some examples of using RL actor-critic methods in AIoT systems are [40].

Reinforcement learning in supply chain management

According to the supply chain driver classification [41], inventory management problems are the most common application in SCM, followed by information and transportation problems. The models that address inventory management problems usually use the RL agent to orchestrate the material flow between multiple sites in the supply chain. Information is a broader class that includes applications that aim to increase information availability, such as forecasting, collaboration, or risk management. In the transportation class, RL is used to address vehicle routing or scheduling problems. The prevalence of RL models in planning tasks with short decision horizons is significantly higher than in long-term decision-making because they mostly have a limited scope, precise inputs and objectives, and a requirement for fast decision making. On the contrary, long-term decisions taken on the managerial level require valid reasoning.

Most models take advantage of Q-learning [42]. This fact can be explained by the maturity of the long- standing technique capable of learning stochastic transitions and reward problems. Besides, Q-learning belongs to the model-free class, implying that an RL agent can operate directly by sampling and learning the expected rewards. However, more recent publications use such deep learning techniques as DQN [43], A2C [44] and PPO [45]. By incorporating deep artificial neural networks, RL agents gain the capabilities to make decisions from large-scale and potentially unstructured input data.

Results and Discussion

Despite efforts to address the challenges of scalability, robustness, and interpretability in reinforcement learning, the majority of these attempts have been either statistical in nature, such as manipulating the resulting sample, or algorithmic, such as modifying reward functions or adding extra coefficients. Some studies have utilized a combination of both statistical and algorithmic methods. Another studies used transfer techniques to reduce sim-to-real gap. While these approaches have shown some promise in specific contexts, further research is needed to identify more comprehensive and effective solutions to these challenges.

Method	References	Description
Algorithmic modifications	[8,9,15,22,25,30,31,46]	Algorithmic modifications involve changes to the basic RL algorithm, such as modifying the reward function, exploration strategy, and action selection mechanism.
Statistical modifications	[7,29,32,35,37,39,42]	Statistical modifications involve changes to the learning algorithm's statistical assumptions or objectives, changes in observations, input vectors, etc.
Transfer techniques	[21-24,26,47,48]	Transfer techniques involve leveraging knowledge learned from one task to accelerate learning in a related task. One example is using transfer learning to reuse a pre-trained neural network to perform a similar task with minimal retraining. Another example is using domain adaptation to transfer knowledge between different environments with similar but not identical dynamics.
Algorithmic and Statistical modifications	[11,49]	Combining statistical and algorithmic methods.
Combination of all methods	[2,50]	Combining all methods involves integrating multiple solutions to address different challenges faced by RL agents.

Table 2: Summary of real-world application of reinforcement learning algorithms and type of proposed solutions.

Conclusion

Based on our analysis, it appears that while there are numerous algorithms available for controlling specific objects in certain conditions, there has been limited success in generalizing the approach to develop control systems for real-world applications using reinforcement learning. Further research is needed to overcome the challenges of scalability, robustness, and interpretability, which are essential for practical implementation in industrial settings. Nonetheless, we remain optimistic about the potential of reinforcement learning algorithms and believe that with continued investigation and development, they will become increasingly valuable tools for industrial control systems in the future.

Bibliography

1. Sutton Richard S and Andrew G Barto. "Reinforcement Learning: An Introduction". The MIT Press, (1998).
2. Nian Rui., *et al*. "A Review On Reinforcement Learning: Introduction and Applications in Industrial Process Control". *Computers and Chemical Engineering* 139 (2020): 106886.
3. Mnih, Volodymyr., *et al*. "Human-Level Control through Deep Reinforcement Learning". *Nature* 518.7540 (2015): 529-533.
4. van Hasselt Hado., *et al*. "Deep Reinforcement Learning with Double Q-Learning". arXiv:1509.06461, arXiv, 8 Dec. (2015).
5. Barto Andrew G., *et al*. "Neuronlike Adaptive Elements That Can Solve Difficult Learning Control Problems". *IEEE Transactions on Systems, Man, and Cybernetics* SMC-13.5 (1983): 834-846.
6. Lillicrap Timothy P., *et al*. "Continuous Control with Deep Reinforcement Learning". 1, arXiv:1509.02971, arXiv (2015).
7. Dulac-Arnold Gabriel., *et al*. "Challenges of Real-World Reinforcement Learning: Definitions, Benchmarks and Analysis". *Machine Learning* 110.9 (2021): 2419-2468.
8. Hwangbo Jemin., *et al*. "Control of a Quadrotor With Reinforcement Learning". *IEEE Robotics and Automation Letters* 2.4 (2017): 2096-2103.
9. Haarnoja Tuomas., *et al*. "Soft Actor-Critic Algorithms and Applications". arXiv:1812.05905, arXiv, 29 Jan. (2019).
10. Singh Avi., *et al*. "End-to-End Robotic Reinforcement Learning without Reward Engineering". arXiv:1904.07854, arXiv, 15 May (2019).
11. McClement Daniel G., *et al*. "Meta-Reinforcement Learning for the Tuning of PI Controllers: An Offline Approach". *Journal of Process Control* 118 (2022): 139-152.
12. Hui Jonathan. "RL — Model-Based Reinforcement Learning". *Medium* (2019).
13. Moerland Thomas M., *et al*. "Model-Based Reinforcement Learning: A Survey". arXiv:2006.16712, arXiv, 31 Mar. (2022).
14. Sünderhauf Niko., *et al*. "The Limits and Potentials of Deep Learning for Robotics". *The International Journal of Robotics Research* 37.4-5 (2018): 405-420.

15. Thuruthel Thomas George., *et al*. "Model-Based Reinforcement Learning for Closed-Loop Dynamic Control of Soft Robotic Manipulators". *IEEE Transactions on Robotics* 35.1 (2019): 124-134.

16. Ju Hao., *et al*. "Transferring Policy of Deep Reinforcement Learning from Simulation to Reality for Robotics". *Nature Machine Intelligence* 4.12 (2022): 1077-1087.

17. Zhao Wenshuai., *et al*. "Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: A Survey". *2020 IEEE Symposium Series on Computational Intelligence (SSCI)* (2020): 737-744.

18. Yang Chenhao., *et al*. "Relative Camera Pose Estimation Using Synthetic Data with Domain Adaptation via Cycle-Consistent Adversarial Networks". *Journal of Intelligent and Robotic Systems* 102.4 (2021): 79.

19. Tobin Josh., *et al*. "Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World". *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2017): 23-30.

20. Mankowitz Daniel J., *et al*. "Situational Awareness by Risk-Conscious Skills". arXiv:1610.02847, arXiv, 10 Oct. (2016).

21. Christiano Paul., *et al*. "Transfer from Simulation to Real World through Learning Deep Inverse Dynamics Model". arXiv:1610.03518, arXiv, 11 Oct. (2016).

22. Hanna Josiah P., *et al*. "Grounded Action Transformation for Sim-to-Real Reinforcement Learning". *Machine Learning* 110.9 (2021): 2469-2499.

23. Tzeng Eric., *et al*. "Adapting Deep Visuomotor Representations with Weak Pairwise Constraints". Algorithmic Foundations of Robotics XII: Proceedings of the Twelfth Workshop on the Algorithmic Foundations of Robotics, edited by Ken Goldberg *et al*., Springer International Publishing, (2020): 688-703.

24. Andrychowicz OpenAI: Marcin., *et al*. "Learning Dexterous In-Hand Manipulation". *The International Journal of Robotics Research* 39.1 (2020): 3-20.

25. Rusu Andrei A., *et al*. "Sim-to-Real Robot Learning from Pixels with Progressive Nets". Proceedings of the 1st Annual Conference on Robot Learning, PMLR, (2017): 262-270.

26. Nagabandi Anusha., *et al*. "Learning to Adapt in Dynamic, Real-World Environments Through Meta-Reinforcement Learning". arXiv:1803.11347, arXiv, 27 Feb. (2019).

27. Arndt Karol., *et al*. "Meta Reinforcement Learning for Sim-to-Real Domain Adaptation". 2020 IEEE International Conference on Robotics and Automation (ICRA) (2020): 2725-2731.

28. Wang Pin., *et al*. "A Reinforcement Learning Based Approach for Automated Lane Change Maneuvers". 2018 IEEE Intelligent Vehicles Symposium (IV), (2018): 1379-1384.

29. Gu Shixiang., *et al*. "Deep Reinforcement Learning for Robotic Manipulation with Asynchronous Off-Policy Updates". 2017 IEEE International Conference on Robotics and Automation (ICRA), (2017): 3389-3396.

30. Bakker B., *et al*. "Quasi-Online Reinforcement Learning for Robots". Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006., (2006): 2997-3002.

31. Hester Todd., *et al*. "RTMBA: A Real-Time Model-Based Reinforcement Learning Architecture for Robot Control". 2012 IEEE International Conference on Robotics and Automation, (2012): 85-90.

32. Riedmiller Martin., *et al*. "Reinforcement Learning for Robot Soccer". *Autonomous Robots* 27.1 (2009): 55-73.

33. Fang Guoxin., *et al*. "Efficient Jacobian-Based Inverse Kinematics With Sim-to-Real Transfer of Soft Robots by Learning". *IEEE/ASME Transactions on Mechatronics* 27.6 (2022): 5296-5306.

34. Chaplot Devendra Singh., *et al*. "*Active Neural Localization*". arXiv:1801.08214, arXiv, 24 Jan. (2018).

35. Wu Chunxue., *et al*. "UAV Autonomous Target Search Based on Deep Reinforcement Learning in Complex Disaster Scene". *IEEE Access* 7 (2019): 117227-117245.

36. Lei Lei., *et al*. "Deep Reinforcement Learning for Autonomous Internet of Things: Model, Applications and Challenges". *IEEE Communications Surveys and Tutorials* 22.3 (2020): 1722-1760.

37. Saravanan M., *et al*. "IoT Enabled Indoor Autonomous Mobile Robot Using CNN and Q-Learning". 2019 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), (2019): 7-13.

38. Liu Xiaolan., *et al*. "Resource Allocation for Edge Computing in IoT Networks via Reinforcement Learning". ICC 2019 - 2019 IEEE International Conference on Communications (ICC), (2019): 1-6.

39. Foruzan Elham., *et al*. "Reinforcement Learning Approach for Optimal Distributed Energy Management in a Microgrid". *IEEE Transactions on Power Systems* 33.5 (2018): 5749-5758.

40. Zhang Yameng., *et al*. "A Deep Reinforcement Learning Approach for Online Computation Offloading in Mobile Edge Computing". 2020 IEEE/ACM 28th International Symposium on Quality of Service (IWQoS), (2020): 1-10.

41. Rolf Benjamin., *et al*. "A Review on Reinforcement Learning Algorithms and Applications in Supply Chain Management". *International Journal of Production Research* (2022): 1-29.

42. Wang Fang and Lin Lin. 'Spare Parts Supply Chain Network Modeling Based on a Novel Scale-Free Network and Replenishment Path Optimization with Q Learning". *Computers and Industrial Engineering* 157 (2021): 107312.

43. Li Xijun., *et al*. "Learning to Optimize Industry-Scale Dynamic Pickup and Delivery Problems". 2021 IEEE 37th International Conference on Data Engineering (ICDE), (2021): 2511-2522.

44. Zhu Zheng., *et al*. "A Mean-Field Markov Decision Process Model for Spatial-Temporal Subsidies in Ride-Sourcing Markets". *Transportation Research Part B: Methodological* 150 (2021): 540-565.

45. Kegenbekov Zhandos and Ilya Jackson. 'Adaptive Supply Chain: Demand-Supply Synchronization Using Deep Reinforcement Learning". *Algorithms* 14.8 (2021): 240.

46. Zhao Zhijia., *et al*. "Reinforcement Learning Control for a 2-DOF Helicopter With State Constraints: Theory and Experiments". *IEEE Transactions on Automation Science and Engineering* (2022): 1-11.

47. Yang Tao., *et al*. "A Soft Artificial Muscle Driven Robot with Reinforcement Learning". *Scientific Reports* 8.1 (2018): 14518.

48. Zhu Wei., *et al*. "A Survey of Sim-to-Real Transfer Techniques Applied to Reinforcement Learning for Bioinspired Robots". *IEEE Transactions on Neural Networks and Learning Systems* (2021): 1-16.

49. Ha Sehoon. "Quadrupedal Robots Trot into the Wild". *Science Robotics* 5.47 (2020): eabe5218.

50. Naug Avisek., *et al*. "Deep Reinforcement Learning Control for Non-Stationary Building Energy Management". *Energy and Buildings* 277 (2022): 112584.