

Deep Reinforcement Learning for Motion Control Algorithms in Robotics

Qisong Liu *

Faculty of Science, Shijiazhuang University, Hebei, China

* Corresponding Author Email: 1603010720@stu.hrbust.edu.cn

Abstract. This article introduces the concept and practical applications of deep reinforcement learning (DRL), describes the core principles and main algorithms of DRL, and summarizes the advantages of DRL and its differences from traditional reinforcement learning (RL). At the same time, it introduces the broad application prospects of DRL in various fields and lists some examples about the mechanical field. This includes applications for classifying colored objects, agricultural applications for robotic arms to cut fruits, branches, or twigs, a DQN based learning system, a mobile robot that can perform selection and placement operations, A path-planning approach concerning crop picking robotic arms, and a decision algorithm for robot tracking applications, Examples of algorithms for unmanned underwater vehicles and improvements for DRL to solve trajectory tracking control problems in autonomous underwater robots. Finally, the powerful role of DRL in future industrial development and other fields was summarized. Despite the significant progress made in DRL, there are still challenges to be addressed, such as the need for large amounts of training data, the difficulty in designing appropriate reward functions, and the potential for instability during training, which present opportunities for future research and development in this rapidly evolving field.

Keywords: DRL; deep reinforcement learning; robotic arms.

1. Introduction

Robots are widely used to perform a variety of fixed operational tasks in relatively static and simple industrial manufacturing plants. However, in some highly dynamic and complex work environments often encountered in daily life, these operations still pose challenges for robots. However, humans can manipulate and improve in such highly dynamic and complex environments. Because people seem to be able to learn other operational skills by observing the performance and behavior of others (learning from observation), and new knowledge and skills through trial and mistake (discover and learning from exploration). Therefore, DRL has been proposed and fully applied in various industries with the support of various complex algorithms, and in this case. Robot control has been given great play under the expansion of DRL.

Deep reinforcement learning (DRL) is a combination of reinforcement learning (RL) and deep learning (DL), which optimizes reinforcement learning algorithms through deep learning technologies such as neural networks. DRL is a kind of machine learning method. RL refers to learning an agent in an environment in the process of continuous interaction with the environment, so that the agent gets the maximum cumulative reward in the environment. Its core principles include RL, neural networks, value functions, policy functions, and deep learning algorithms. DRL is based on the basic framework of RL, which learns the optimal strategy by observing environmental states, taking actions, receiving rewards, and other processes. Using neural networks to model the mapping between states and actions and representing the agent's behavioral strategies through policy functions, thereby processing high-dimensional, nonlinear state and action spaces. The value function is used to evaluate the value of a state or state action pair, and the optimal strategy is learned and optimized by maximizing the value function. It covers various deep learning algorithms, such as convolutional neural networks, recurrent neural networks, DRL, etc., used to model the mapping between states and actions. Its algorithms mainly include Q-learning, Deep Q-network (DQN), Policy Gradient, Actor Critic, Deep Deterministic Policy Gradient (DDPG), Trust Region Policy Optimization (TRPO),



Trust Region Policy Optimization (TRPO), and so on. Based on these complex sets of algorithms, DRL can handle high-dimensional state and action spaces and automatically extract features to reduce the workload of feature engineering. While adapting to different scenarios and tasks, it can learn more complex strategies to improve performance and efficiency. For dealing with nonlinear and complex relationships and continuous action spaces, DRL is also competent.

The difference between DRL and traditional RL is that traditional RL typically requires manual feature design and then transfer the features to RL algorithms for learning, while DRL can automatically extract features through techniques such as deep neural networks, eliminating the need for manual feature design. And DRL is usually applied to more complex, high-dimensional, and nonlinear tasks, such as game AI, robot control, natural language processing, etc., while traditional RL is usually applied to low dimensional, simple tasks, such as maze problems, inverted pendulum problems, etc. In terms of capability and efficiency, DRL can handle high-dimensional, complex state spaces, and nonlinear relationships through technologies such as deep neural networks, while traditional RL usually can only handle low dimensional, simple state spaces, and basic linear relationships. Meanwhile, DRL can improve training efficiency through techniques such as deep neural networks, allowing for faster learning of better strategies, while traditional RL typically requires longer training time.

2. The application scenarios of deep reinforcement learning

Based on the strong adaptability and self-improvement characteristics of DRL for different scenarios and tasks, its application scenarios are very wide. For example, DRL can be applied to game AI to achieve autonomous learning and optimization of game AI by learning game rules and strategies. For example, AlphaGo learned the strategies and tactics of Go through DRL technology and ultimately defeated human professional players. DRL can also be applied in natural language processing, such as machine translation, text classification, question answering systems, and other tasks. By learning language models and semantic representations, achieve understanding and generation of natural language.

In the field of financial trading, DRL can handle tasks such as stock trading and futures trading. By learning trading strategies and risk control, achieve automation and optimization of financial transactions. DRL can also be applied in intelligent transportation, such as traffic signal control, intelligent driving, public transportation scheduling, and other tasks. By learning traffic rules and traffic flow, achieve the intelligence and optimization of the transportation system.

Even in the field of medical and health, tasks such as disease diagnosis, personalized treatment, drug development, etc. DRL can achieve intelligence and optimization of medical decision-making by learning medical data and knowledge. Based on these application scenarios, DRL can make critical progress in the field of robot control

DRL can be applied in robot control, achieving intelligence of robots by learning their control strategies. For example, DRL can be used to learn tasks such as robot motion planning, path planning, and action selection. For example, the robotic arm in a factory can do a lot with the blessing of DRL

ARC Centre of Excellence for RoboticVision (ACRV) from Queensland University of Technology (QUT) argues that humans can learn in dynamic and complex environments, and improve and operate for a reason. That humans can grow in learning (learning from observation) by observing how others perform [1].

So creating a robot like humans that could learn in the same way and master operational skills became their goal. To ensure that the robot can learn from interpretation proportions as humans, methods that enable continuous learning and the flexibility to handle a range of different operational tasks are required. ACRV believe that DRL (DRL) is a promising candidate for automatic learning. The DRL has many different types of algorithms, of which the more popular example is the DQN, which can match human performance after learning the Atari 2600 game for 38 days

Therefore, allowing DQN can be used in learning robot control is the first step. Based on enabling the robotic arm to learn human skills through visual observation; ACRV proposed a learning system based on the DQN. as shown in Figure 1. The components of this system are collectively divided into three parts: a DQN learner, a 2D robotic arm simulator for Tagg arrival, and a ros-based interface for performing operation commands on the Baxter robot.

ARCV are further evaluated by training agents in a simulated environment and conducting target-reaching experiments in reality. Due to the existence of different levels of noise, offset of image, difference of initial arm posture and difference in link length in the simulation experiment, this leads to some common problems in the robot motion control and manipulation simulation. Finally, the vision-based DRL is implemented in the real-world robot operation.



Fig. 1 Baxter's arm being controlled by a trained deep Q Network (DQN). Synthetic images (on the right) are fed into the DQN to overcome some of the real-world issues encountered, i.e., the differences between training and testing settings [1].

Applied Electronics Department from Fac of Electronics and Telecom, Politecnica Un Applied Electronics Department developed an application for classifying colored objects using robotic arms [2]. It can select cubes of different colors and then classify them in different cups. The color recognition function is realized by using a network camera to identify images. as shown in Figure 2. Mechanical arms are widely used in industry, but most of them are only used for PTP (point-to-point) trajectories, which were previously learned by the robotic arm. They believe that very few robots in these industries can be designed to learn and grow on their own and make the right decisions. In some areas of the future, human work will be largely completely replaced by robots, thus requiring robotic arms that enable self-learning to make decisions. A good example of this is the intelligent robotic arm is its ability to classify objects according to their color.



Fig. 2 Picture of the stand for the robotic arm [2]

Amrita School of Engineering, Amritapuri propose a Kinect based robotic arm control method to address harvesting related issues [3]. Kinect Is a device that can sense human gestures, process these signals, and can also be used to control any robotas, shown in Figure 3 As a revolutionary product from Microsoft for Xbox 360; it marks a breakthrough in robot control in the game industry. A person simply needs to move his body and limbs to enable the Kinect to sense and receive its input and then input the right signal, and these operations can also be copied in the game without the need for any manipulation tools. For Kinect, game users can indeed experience this feeling in a variety of reality simulations in the game without touching or using any hand-held input device. Kinect A device can be a tool for user input through the body or gestures as a vehicle. These input signals will be processed by the host processor and sent to people via the embedded system-based controller and Bluetooth wireless function, and sent to the robotic arm via the signal interface.

In this way, the embedded controller of the robotic arm can achieve the remote control of the arm. Thus, further agricultural applications of robotic arm cutting, fruit picking or pruning can be designed. The positioning and control of the robotic arm in the system can be simpler, and making it easier for trainers to handle the specific control of the robotic arm. They believe that this method can enable the harvesting of crops, pruning or harvesting of agricultural crops in more complex environments.



Fig. 3 Picture of the stand for the robotic arm [3]

School of Electrical and Electronic Engineering, Dun Hussein University, Batu Pahat, Johor, Malaysia, which believe that as the usage of wireless applications increases, there is an increasing demand for a system that can easily connect devices for long-distance data transmission, and therefore Batu Pahat And his team have developed a mobile robot that can perform picking and placing operations, and control it through the use of wireless PS2 controllers [4]. According to the controller specifications, it can rotate forward, reverse, left and right for a certain distance. The development of this robot is based on the Arduino super platform, which will interface with wireless controllers to mobile robotic arms. In order to understand the performance of the robot, analysis has been conducted on speed, distance, load, etc. Finally, the prototype of this robot is expected to overcome issues such as placing or selecting objects far away from the user, and selecting and placing dangerous objects in the fastest and simplest way possible.

A research group from South China Agricultural University proposed a DRL-based robotic arm path planning method to solve the problem of low efficiency and low success rate of DRL in path planning [5]. First, they solved the harvesting path planning problem using the RL method and designed an RL method combining artificial potential field; second, they improved the two DRL algorithms by introducing the long and short term memory (LSTM) structure; finally, they further trained the DRL algorithm using three different unstructured citrus tree environments for the path planning of harvesting mechanical arm. The simulation comparison experiment shows that the RL method and artificial potential field can improve the success rate of manipulator design, improve the convergence rate of the deep deterministic policy gradient (DDPG) algorithm by 57.25%, the soft participant critical (SAC) algorithm by 53.73%, and the success rate of path planning by 23.00% and 9.00% respectively; Compared with the traditional algorithm RRT connection (using LSTM structure, the overall planning path length was shortened by 16.20%, and the success rate of path planning also produced a large gap of 9.67%. This proves that their proposed path planning method can significantly optimize the path planning length and success rate, and provides a reference idea in solving the path planning problem of unstructured harvesting robots.

Another example of DRL in practical mechanical applications is unmanned underwater vehicles (UUVs) [6]. Algorithm has the advantages of simple structure, simple implementation and high reliability. Several researchers have proposed DRL-based algorithms in uuv control. Adaptive neural network application with DRL implements control of uuv trajectory tracking. Developed a DRL method for adaptive low-level control. A model-free UAV depth control DRL algorithm is proposed. Preliminary results show that DRL is feasible in uuv control and can achieve good motor control performance.

Sahar Leisiazar And his team proposed a new approach to robotic tracking applications to address barriers and key challenges of occlusion avoidance [7]. They plan to allow the roadfinding navigation robot to effectively navigate the general route while automatically avoiding collisions and shielding caused by surrounding objects, as shown in Figure 4. Therefore, they designed a more advanced decision algorithm that allows the navigation robot to move and simultaneously generate short-term navigation targets. By combining the DRL method with Monte Carlo tree search, the performance of the navigation robot has been significantly improved in the decision process, and further reduces errors and errors in the process of generating short-term navigation targets. They demonstrate through extensive experiments and evaluations that their proposed method is more efficient and superior compared to existing human-tracking robot methods.

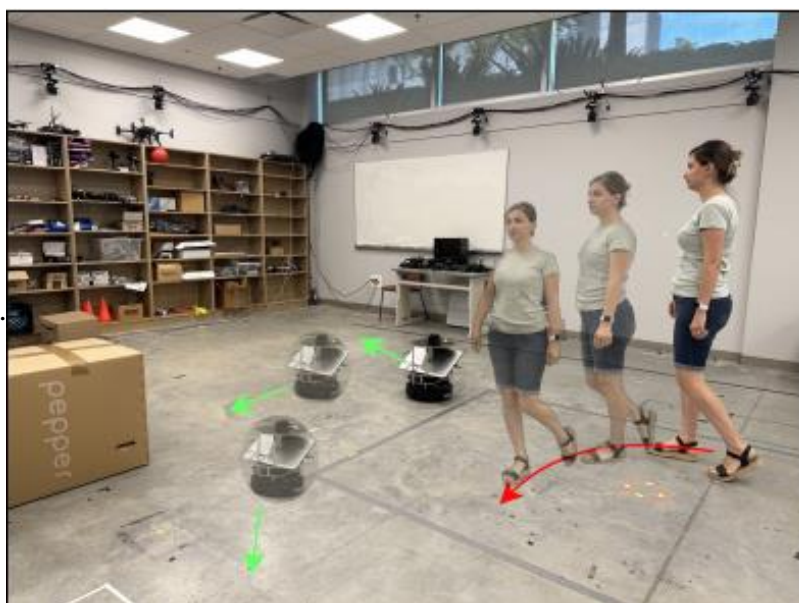


Fig. 4 Road positioning and navigation robots effectively navigate general routes while automatically avoiding collisions and shielding caused by surrounding objects [7].

A research group from South China Normal University, Guangzhou 510631 use and improve DRL (DRL) to solve the problem of trajectory tracking control in autonomous underwater robots (AUV) [8]. DRL in an underwater motion control system consists of two neural networks: one network selects the action, and the other network evaluates whether the selected action is accurate, self-correcting by a deep deterministic policy gradient (DDPG). The two neural networks are composed of multiple fully connected layers. Based on theory and simulation, this algorithm can solve the problem of AUV in complex curves better than traditional PID control (Figure 5).

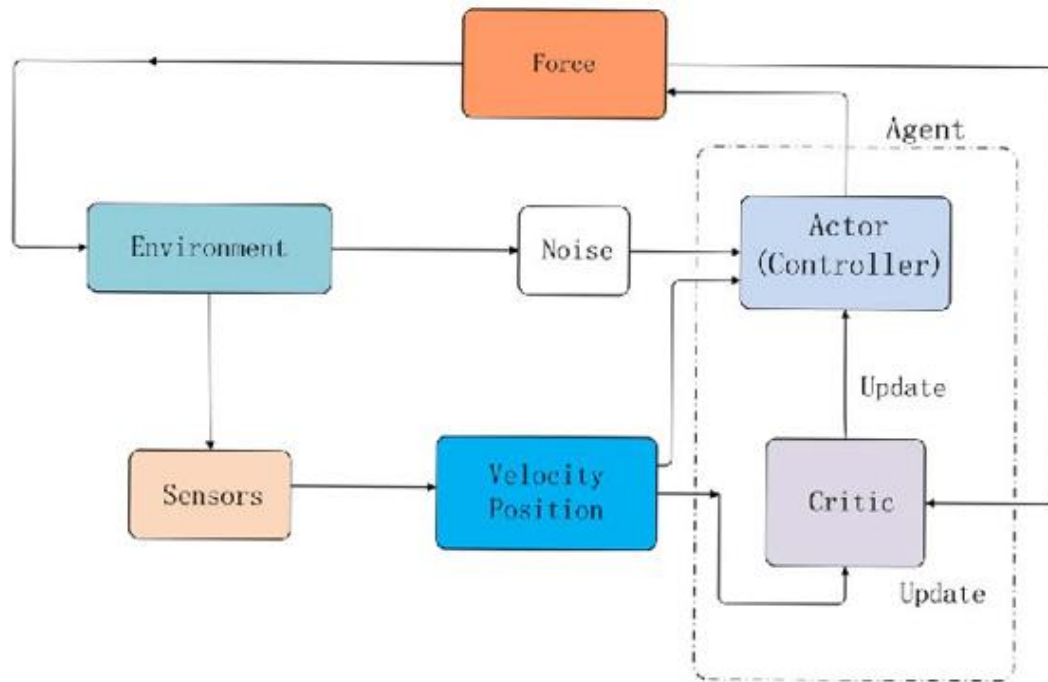


Fig. 5 Overview of the AUVs control system [8]

In a variety of systems for automatic matching applications, scheduling is one of the basic tasks, for example, most of the machine in industrial production in scheduling, as shown in Figure 6, if this can greatly reduce the cost of production and the waste of resources, but in real life, the machine is difficult to find the best scheduling measures, solve scheduling problems often produce many difficulties, and want to combine optimization problem (COP) method to solve the problem due to the limited time is not feasible. However, due to the emergence of DRL, the scheduling problem has made new progress in self-learning and decision-making, and the solution of the mechanical scheduling problem has become possible. Department of Artificial Intelligence and Cybersecurity University of Klagenfurt, Klagenfurt, Austria presents an effective DRL workshop scheduling environment-an important issue in the field. In addition, they designed a meaningful, compact state representation, and a state-of-the-art, concise, dense reward function that is very close to the sparse fabrication-span minimization criterion used by the COP method [9]. They demonstrate that the method has significant advantages over existing DRL methods on classical benchmark examples and is more similar to the state-of-the-art COP methods.

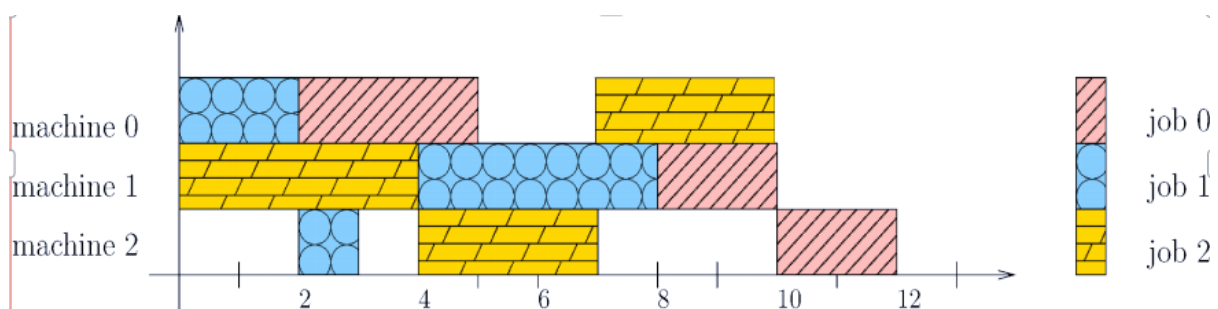


Fig. 6 Scheme for small instances consisting of three tasks and three machines [9].

3. Conclusion

With the continuous development of the times and technology, traditional reinforcement learning has gradually been unable to meet people's needs. More complex and diverse environments and needs, as well as stronger abilities and efficiency, have led to the birth of DRL. For DRL, its neural network, value function, policy function and deep learning algorithm all have broad prospects and technical support. The characteristics of automatic feature extraction make DRL unique in terms of reducing workload. The feature of continuous action space processing creates a gap with the traditional reinforcement learning algorithm. Thanks to the continuous exploration of DRL, natural language processing, game AI, financial transactions, intelligent transportation, medical health, and robot control have all developed to varying degrees. Although there have been many research and application achievements on DRL, DRL still has a huge development space. As described in this paper, DRL has made quite excellent progress in scheduling issues, vision, agriculture, and optimization. It is not only limited to some separate areas while expanding the application field, it also upgrades and integrates its own optimization in the process of being constantly explored and discovered, reducing a lot of workload. Fore predictable in the future, the development of DRL will be further extended, able to demonstrate stronger abilities and efficiency in more fields. The development of DRL will be further extended, whether in natural language processing, game ai, financial transactions, intelligent transportation, health care or robot control, DRL can show great potential in a variety of more complex environment show more strong processing ability and training efficiency, so as to adapt to more scenarios and tasks and complex environments, and meet more and more complex needs. In the process of self-upgrading, this paper will further reduce the proportion of human labor force, and then liberate more positions in more fields.

References

- [1] Zhang Fangyi, Leitner Jürgen, Milford Michael, Upcroft Ben, Corke Peter. Towards Vision-Based Deep Reinforcement Learning for Robotic Motion Control. *Machine Learning*, 2015.
- [2] Szabó R., Lie I. Automated colored object sorting application for robotic arms. In *Proceedings, 10th International Symposium on Electronics and Telecommunications*, Timisoara, Romania, 2012: 95-98.
- [3] Megalingam R. K., Vivek G. V., Bandyopadhyay S., Rahi M. J. Robotic arm design, development and control for agriculture applications. In *Proceedings, 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2017: 1-7.
- [4] Kamaril Yusoff Mohd Ashiq, Samin Reza Ezuan, Kader Ibrahim Babul Salam. *Wireless Mobile Robotic Arm*. *Procedia Engineering*, 2012, 41: 1072-1078.
- [5] Xiong Chunyuan, Xiong Juntao, Yang Zhengang, et al. Path planning method for citrus picking manipulator based on deep reinforcement learning. *Journal of South China Agricultural University*, 2023, 44(3): 473-483.
- [6] Chu Zhenzhong, Sun Bo, Zhu Daqi. Motion control of unmanned underwater vehicles via deep imitation reinforcement learning algorithm. *Volume14, Issue7, July 2020: 764-774*.
- [7] Leisiazar S., Park E. J., Lim A., Chen M. An MCTS-DRL Based Obstacle and Occlusion Avoidance Methodology in Robotic Follow-Ahead Applications. In *Proceedings, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Detroit, MI, USA, 2023: 221-228.
- [8] Yu R., Shi Z., Huang C., Li T., Ma Q. Deep reinforcement learning based optimal trajectory tracking control of autonomous underwater vehicle. In *Proceedings, 36th Chinese Control Conference (CCC)*, Dalian, China, 2017: 4958-4965.
- [9] Tassel Pierre, Gebser Martin, Schekotihin Konstantin. A Reinforcement Learning Environment For Job-Shop Scheduling. *Machine Learning*, 2021.