IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 26, Issue 5, Ser. 3 (Sept. – Oct. 2024), PP 26-28
www.iosrjournals.org

An Overview Of Deep Deterministic Policy Gradient Algorithm And Applications

V.Sumalatha, Dr. Suresh Pabboju
Ph. D Scholar, Dept Of CSE, Osmania University.
Professor Of IT, Director-AEC &COE CBIT Osmania University.

Abstract:
This research summarizes the major developments of the Deep Deterministic Policy Gradient (DDPG) in reinforcement learning Motivated by the ideas, of Deep Q-networks, DDPG has proven capable of confronting more complex problems involving continuous action spaces. The core of DDPG lies in its actor-critic architecture, which enables the learning of highly competitive policies. By leveraging neural network function approximations, it can efficiently operate in large state and action spaces. DDPG has found practical applications across various real-world domains. However, like many model-free reinforcement learning methods, DDPG still faces the challenge of requiring a large number of training steps.
Keywords: *RL, DDPG, DQN, NN.*

Date of Submission: 26-09-2024 Date of Acceptance: 06-10-2024

I. Introduction

An area of artificial intelligence called reinforcement learning (RL) focuses on decision-making through the acquisition of ideal behavior in settings in order to maximize a reward signal [1]. In reinforcement learning, an agent interacts with its surroundings and gets rewards in return, which helps it improve its policy for making decisions [2]. This technique has proven useful in a number of industries, such as robots, autonomous control, anomaly detection, and game development [8].

In order to improve RL's performance in high-dimensional state space, DL and RL are integrated, with Deep Neural Networks (DNNs) representing the agent's decision-making policy [9]. One such integration known as Deep Reinforcement Learning (DRL), has resulted to a substantial advancement in the discipline and has enabled the development of extremely efficient algorithms for making decisions in challenging environments remarks [9, 10]. DDPG, or Deep Deterministic Policy Gradient, is a widely used DRL algorithms [11], which combines the benefits of policy-based and value-based reinforcement learning, using an actor-critic methodology [12].

II. Related Work

DDPG is an actor-critic algorithm; it has two networks: actor and critic. Technically, the actor produces the action to explore.Deep Deterministic Policy Gradient (DDPG) is an algorithm which concurrently learns a Q-function and a policy. It uses off-policy data and the Bellman equation to learn the Q-function, and uses the Q-function to learn the policy.

Deep Deterministic Policy Gradient (DDPG) is an advanced algorithm used in reinforcement learning (RL) to train agents in continuous action spaces. RL is a type of machine learning where an agent learns to make decisions by interacting with an environment to maximize a cumulative reward. Continuous Action Spaces allow for a range of possible actions, such as controlling the speed or angle of a robot arm.

DDPG is based on the actor-critic framework, where two neural networks are used:
Actor Network: This network determines the best action to take given a certain state of the environment [13].
Critic Network: This network evaluates how good the action taken by the actor was by estimating the action's value[13].

DDPG algorithm architecture

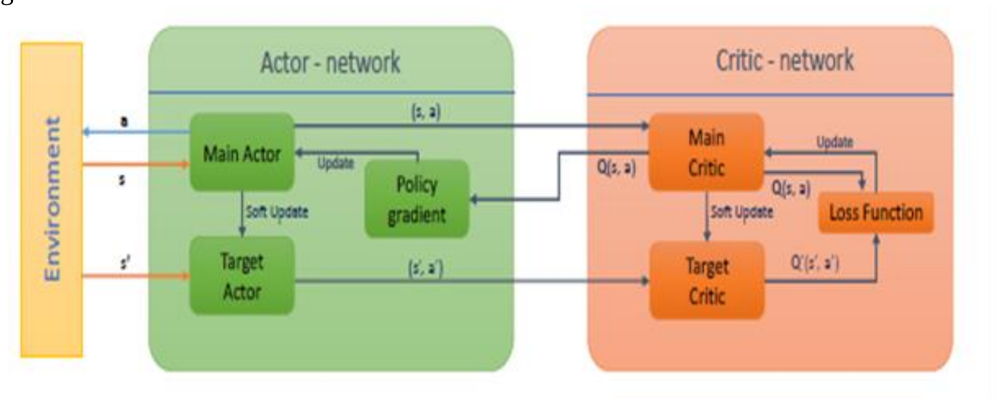

Fig 1: DDPG structure

DDPG is a model-free, off-policy RL algorithm that adopts an actor-critic approach to solving continuous control problems in which the action space is continuous. Figure 1 illustrates the DDPG structure. It was introduced in 2015 by Lillicrap et al. [14] and builds upon the Q-learning and actor-critic algorithms.

DDPG Works on
1. Deterministic Policy: Unlike stochastic policies that select actions based on probabilities, DDPG uses a deterministic policy where a specific action is chosen directly by the actor network.
2. Off-Policy Learning: DDPG uses past experiences stored in a replay buffer to learn. This allows it to learn from data collected at different times and reduces correlation between consecutive experiences.
3. Target Networks: DDPG employs target networks for both the actor and critic. These are slowly updated versions of the original networks, which help stabilize training by reducing the risk of diverging parameters.

III. Applications

DDPG is used in various applications, such as robotic control, autonomous vehicles, and other tasks requiring decision-making in continuous action spaces.

- Robotics: DDPG can be used to train robotic arms for precise control tasks where actions like joint angles must be continuously adjusted.
- Autonomous Vehicles: The algorithm can be used to optimize driving strategies in continuous spaces, such as steering and acceleration.
- Financial Trading: DDPG can be applied to optimize trading strategies where the actions (e.g., buy, sell, hold) can be represented as continuous quantities.

IV. Conclusion

Resource allocation issues in cloud computing and energy management are successfully resolved by DDPG. DDPG can adjust to changing circumstances and optimize resource allocation in real time by drawing on its experience. DDPG is an effective tool for resolving resource allocation issues in dynamic situations because it can learn intricate rules that take into account the nonlinear correlations between the system parameters and the rewards thanks to the use of DNNs.

References

[1] Arulkumaran, K., Deisenroth, M.P., Brundage, M. Bharath, A.A.: Deep Reinforcement Learning: A Brief Survey. Ieee Signal Processing Magazine 34(6),26–38 (2017)37.
[2] Henderson, P., Islam, R., Bachman, P., Pineau, J., Precup, D., Meger, D.: Deep Reinforcement Learning That Matters. In: Proceedings Of The Aaai Conference On Artificial Intelligence, Vol. 32.
[3] Dulac-Arnold, G., Mankowitz, D., Hester, T.: Challenges Of Real-World Reinforcement Learning. Arxiv Preprint Arxiv:1904.12901 (2019).
[4] Rehman, A., Saba, T., Haseeb, K., Alam, T., Lloret, J.: Sustainability Model For The Internet Of Health Things (Ioht) Using Reinforcement Learning With Mobile Edge Secured Services. Sustainability 14(19), 12185 (2022).
[5] Zhao, K., Song, J., Luo, Y., Liu, Y.: Research On Game-Playing Agents Based On Deep Reinforcement Learning. Robotics 11(2), 35 (2022).
[6] Arshad, K., Ali, R.F., Muneer, A., Aziz, I.A., Naseer, S., Khan, N.S., Taib, And S.M.: Deep Reinforcement Learning For Anomaly Detection: A Systematic Review. Ieee Access (2022).
[7] Singh, B., Kumar, R., Singh, V.P.: Reinforcement Learning In Robotic Applications: A Comprehensive Survey. Artificial Intelligence Review, 1–46 (2022).
[8] Shahid, A.A., Piga, D., Braghin, F., Roveda, L.: Continuous Control Actions Learning And Adaptation For Robotic Manipulation Through Reinforcement Learning. Autonomous Robots 46(3), 483–498 (2022).

[9] Arshad, K., Ali, R.F., Muneer, A., Aziz, I.A., Naseer, S., Khan, N.S., Taib, S.M.:Deep Reinforcement Learning For Anomaly Detection: A Systematic Review. Ieee Access (2022).

[10] Kim, M.-S., Eoh, G., Park, T.-H.: Decision Making For Self-Driving Vehicles In Unexpected Environments Using Efficient Reinforcement Learning Methods. Electronics 11(11), 1685 (2022).

[11] Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., Riedmiller, M.: Deter-Ministic Policy Gradient Algorithms. In: International Conference On Machine Learning, Pp. 387–395.

[12] Alturkistani, H., El-Affendi, M.A.: Optimizing Cyber Security Incident Response Decisions Using Deep Reinforcement Learning. International Journal Of Electrical And Computer Engineering 12(6), 6768 (2022).

[13] Ebrahim Hamid Sumiea, Said Jadid Abdulkadir " Deep Deterministic Policy Gradient Algorithm: A Systematic Review" Volume 10, Issue 9, 15 May 2024 Https://Doi.Org/10.1016/J.Heliyon.2024.E30697.

[14] Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., Wierstra, D.: Continuous Control With Deep Reinforcement Learning. Arxiv Preprint Arxiv:1509.02971 (2015)