Reinforcement Learning-Based Optimal Path Planning for Mobile Robot with Obstacles Avoidance

Mohammed Abdallah*, Manar Lashin*, Ahmed El-Awamry*, Walaa Ibrahim Gabr *
[1]Electrical Eng. Dept., Benha Faculty of Engineering, Benha University, Benha, Egypt
Emails: {mohamed.abdallah@bhit.bu.edu.eg, manar.lashin@bhit.bu.edu.eg, ahmedelawamry@bhit.bu.edu.eg, walaa_gabr@bhit.bu.edu.eg}

Abstract

This paper presents a deep reinforcement learning (RL) approach for training mobile robots to navigate complex environments using the Twin Delayed Deep Deterministic Policy Gradient (TD3) method, which is known for its stability in continuous control tasks. The robot model simulates real-world bicycle kinematics with nonholonomic constraints and tackles three key navigation tasks: point tracking with obstacle avoidance, linear path following, and circular path tracking.The study focuses on enhancing tasks like point tracking, linear path following, and circular path tracking, aiming to reduce the distance to the goal, minimize tracking errors, and lower control effort over time. This approach replaces traditional methods and significantly improves upon them, enabling the system to reach targets even at points it hasn't been trained on before, thereby boosting efficiency and adaptability. Synthetic environments with obstacles are created using the MATLAB® Reinforcement Learning toolbox for realistic simulations. The system employs an actor-critic neural network that processes occupancy map data and outputs continuous velocity commands. Evaluations show the approach's effectiveness in teaching robots collision-free navigation, achieving human-level competency in complex environments through iterative learning. This work demonstrates the potential of model-free deep RL for real-world mobile robot navigation. ***Keywords— Deep Reinforcement learning (DRL), mobile robots, environment, TD3, MATLAB®/Reinforcement Learning toolbox.***

I. Introduction

In recent years, there has been significant growth in the field of robotics, with robots now being commonly deployed across various domains, including social interactions, agriculture, logistics, manufacturing, medicine, and education [1]–[4].For autonomous navigation, mobile robots need effective path planning, motion control, and obstacle avoidance. Traditional path-planning algorithms, such as A*, Dijkstra's , RRT, and PRM [5]–[8], have limitations including slow processing times, high computational requirements, and challenges in complex environments. These limitations highlight the need for improved methods in robotic path planning. Recently, researchers have turned to biologically inspired algorithms like Artificial Neural Networks (ANNs), Genetic Algorithms (GAs), Particle Swarm Optimization (PSO), and Reinforcement Learning (RL) to address the limitations of traditional path-planning methods [9]–[12]. These algorithms consider various optimization criteria such as path length, cost-effectiveness, and collision avoidance. This study tackles these challenges by integrating innovative methodologies to enhance path length, efficiency, robustness, and obstacle avoidance, aiming to advance path-planning techniques and improve navigation systems.

The literature review examines reinforcement learning (RL) techniques for mobile robot path planning and motion control [13]. RL is a method where robots are trained through rewards and penalties to develop desirable behaviors. Key components of RL include an agent, environment, reward function, and policy [14]. RL algorithms are categorized into model-free and model-based methods [15]. Model-free algorithms, like Q-Learning [16], SARSA [17] and temporal-difference methods [18], rely on agent-environment interactions to improve the policy or value function. Model-based algorithms use an environment model for training. The Q-Learning algorithm, using the Q-function, helps robots learn optimal actions based on feedback. Several prior studies have applied deep learning techniques for mobile robot navigation in environments without predefined targets, successfully using deep reinforcement learning (DRL) [19]–[21]. However, the effect of including targets remains underexplored.For example, Nakamura [22] used a Deep Q-Network for mobile robot path planning in a narrow 2m x 3m road environment with a width of 0.35m. A target was included, but it took 15,000 episodes to reach the goal in 39.78 seconds and 20,000 episodes to reduce time to 35.8 seconds. These times seem long, given the simplicity of the environment. Quiroga et al. [23]compared DDPG and Deep Q-Learning to traditional control algorithms for mobile robot navigation between a start point of (-0.4,0) and a target at (0.8,0) with no obstacles. The deep learning techniques outperformed traditional methods but still took significant time, even increasing in a static obstacle environment. Furthermore, the environments used were quite simple. These findings suggest that DRL holds promise for target-based navigation, but training times can be long, motivating further research into the effect of targets on learning performance. This paper evaluates the Twin Delayed DDPG (TD3) algorithm for mobile robot navigation in static obstacle environments. TD3 improves on DDPG by using two critic networks to reduce overestimation bias and delaying policy updates for better performance [24], [25]. The study trains a robot to navigate between set points while following a predefined trajectory.

REFERENCES

[1] R. Gautron, O.-A. Maillard, P. Preux, M. Corbeels, and R. Sabbadin, "Reinforcement learning for crop management support: Review, prospects and challenges," *Computers and Electronics in Agriculture*, vol. 200, p. 107182, 2022.

[2] M. Hu, J. Zhang, L. Matkovic, T. Liu, and X. Yang, "Reinforcement learning in medical image analysis: Concepts, applications, challenges, and future directions," *Journal of Applied Clinical Medical Physics*, vol. 24, no. 2, p. e13898, 2023.

[3] C. Li, P. Zheng, Y. Yin, B. Wang, and L. Wang, "Deep reinforcement learning in smart manufacturing: A review and prospects," *CIRP Journal of Manufacturing Science and Technology*, vol. 40, pp. 75–101, 2023.

[4] Y. Yan, A. H. Chow, C. P. Ho, Y.-H. Kuo, Q. Wu, and C. Ying, "Reinforcement learning for logistics and supply chain management: Methodologies, state of the art, and future opportunities," *Transportation Research Part E: Logistics and Transportation Review*, vol. 162, p. 102712, 2022.

[5] H. Wang, S. Lou, J. Jing, Y. Wang, W. Liu, and T. Liu, "The ebs-a* algorithm: An improved a* algorithm for path planning," *PloS one*, vol. 17, no. 2, p. e0263841, 2022.

[6] Y. Huang, "Research on the improvement of dijkstra algorithm in the shortest path calculation," in *2017 4th International Conference on Machinery, Materials and Computer (MACMC 2017)*. Atlantis Press, 2018, pp. 745–749.

[7] N. Makariye, "Towards shortest path computation using dijkstra algorithm," in *2017 International Conference on IoT and Application (ICIOT)*. IEEE, 2017, pp. 1–3.

[8] J. Xu, Z. Tian, W. He, and Y. Huang, "A fast path planning algorithm fusing prm and p-bi-rrt," in *2020 11th International Conference on Prognostics and System Health Management (PHM-2020 Jinan)*. IEEE, 2020, pp. 503–508.

[9] Z. Ullah, Z. Xu, L. Zhang, L. Zhang, and W. Ullah, "Rl and ann based modular path planning controller for resource-constrained robots in the indoor complex dynamic environment," *IEEE Access*, vol. 6, pp. 74 557–74 568, 2018.

[10] C. Lamini, S. Benhlima, and A. Elbekri, "Genetic algorithm based approach for autonomous mobile robot path planning," *Procedia Computer Science*, vol. 127, pp. 180–189, 2018.

[11] B. Song, Z. Wang, and L. Zou, "An improved pso algorithm for smooth path planning of mobile robots using continuous high-degree bezier curve," *Applied Soft Computing*, vol. 100, p. 106960, 2021.

[12] G. Farias, G. Garcia, G. Montenegro, E. Fabregas, S. Dormido-Canto, and S. Dormido, "Reinforcement learning for position control problem of a mobile robot," *IEEE Access*, vol. 8, pp. 152 941–152 951, 2020.

[13] B. Singh, R. Kumar, and V. P. Singh, "Reinforcement learning in robotic applications: a comprehensive survey," *Artificial Intelligence Review*, pp. 1–46, 2022.

[14] F. Quiroga, G. Hermosilla, G. Farias, E. Fabregas, and G. Montenegro, "Position control of a mobile robot through deep reinforcement learning," *Applied Sciences*, vol. 12, no. 14, 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/14/7194

[15] A. Kamalova, S. G. Lee, and S. H. Kwon, "Occupancy reward-driven exploration with deep reinforcement learning for mobile robot system," *Applied Sciences*, vol. 12, no. 18, 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/18/9249

[16] A. Maoudj and A. Hentout, "Optimal path planning approach based on q-learning algorithm for mobile robots," *Applied Soft Computing*, vol. 97, p. 106796, 2020.

[17] V. N. Sichkar, "Reinforcement learning algorithms in global path planning for mobile robot," in *2019 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*. IEEE, 2019, pp. 1–5.

[18] Z. Miao, Y. Xia, F. Zhou, and X. Yuan, "Fault diagnosis of wheeled robot based on prior knowledge and spatial-temporal difference graph convolutional network," *IEEE Transactions on Industrial Informatics*, 2022.

[19] M.-F. R. Lee and S. H. Yusuf, "Mobile robot navigation using deep reinforcement learning," *Processes*, vol. 10, no. 12, p. 2748, 2022.

[20] M. Luong and C. Pham, "Incremental learning for autonomous navigation of mobile robots based on deep reinforcement learning," *Journal of Intelligent & Robotic Systems*, vol. 101, no. 1, p. 1, 2021.

[21] P. Wenzel, T. Schön, L. Leal-Taixé, and D. Cremers, "Vision-based mobile robotics obstacle avoidance with deep reinforcement learning," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 14 360–14 366.

[22] T. Nakamura, M. Kobayashi, and N. Motoi, "Path planning for mobile robot considering turnabouts on narrow road by deep q-network," *IEEE Access*, vol. 11, pp. 19 111–19 121, 2023.

[23] F. Quiroga, G. Hermosilla, G. Farias, E. Fabregas, and G. Montenegro, "Position control of a mobile robot through deep reinforcement learning," *Applied Sciences*, vol. 12, no. 14, p. 7194, 2022.

[24] N. Abo Mosali, S. S. Shamsudin, O. Alfandi, R. Omar, and N. Al-Fadhali, "Twin delayed deep deterministic policy gradient-based target tracking for unmanned aerial vehicle with achievement rewarding and multistage training," *IEEE Access*, vol. 10, pp. 23 545–23 559, 2022.

[25] R. B. Grando, J. C. de Jesus, V. A. Kich, A. H. Kolling, and P. L. J. Drews-Jr, "Double critic deep reinforcement learning for mapless 3d navigation of unmanned aerial vehicles," *Journal of Intelligent & Robotic Systems*, vol. 104, no. 2, p. 29, 2022.

[26] S. Erfani, A. Jafari, and A. Hajiahmad, "Comparison of two data fusion methods for localization of wheeled mobile robot in farm conditions," *Artificial Intelligence in Agriculture*, vol. 1, pp. 48–55, 2019.

[27] M. Kim, D.-K. Han, J.-H. Park, and J.-S. Kim, "Motion planning of robot manipulators for a smoother path using a twin delayed deep deterministic policy gradient with hindsight experience replay," *Applied Sciences*, vol. 10, no. 2, p. 575, 2020.

[28] R. Fernandez-Fernandez, M. Aggravi, P. R. Giordano, J. G. Victores, and C. Pacchierotti, "Neural style transfer with twin-delayed ddpg for shared control of robotic manipulators," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 4073–4079.

[29] J. Wu, Q. M. J. Wu, S. Chen, F. Pourpanah, and D. Huang, "A-td3: An adaptive asynchronous twin delayed deep deterministic for continuous action spaces," *IEEE Access*, vol. 10, pp. 128 077–128 089, 2022.

[30] Z. Zhang, X. Li, J. An, W. Man, and G. Zhang, "Model-free attitude control of spacecraft based on pid-guide td3 algorithm," *International Journal of Aerospace Engineering*, vol. 2020, pp. 1–13, 2020.