

Retraction

Retracted: Deep Reinforcement Learning-Based Path Control and Optimization for Unmanned Ships

Wireless Communications and Mobile Computing

Received 27 June 2023; Accepted 27 June 2023; Published 28 June 2023

Copyright © 2023 Wireless Communications and Mobile Computing. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] D. Wu, Y. Lei, M. He, C. Zhang, and L. Ji, "Deep Reinforcement Learning-Based Path Control and Optimization for Unmanned Ships," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 7135043, 8 pages, 2022.

Research Article

Deep Reinforcement Learning-Based Path Control and Optimization for Unmanned Ships

Di Wu¹, Yin Lei¹, Maoen He², Chunjong Zhang², and Li Ji³

¹College of Electronic and Information Engineering, Tongji University, Shanghai 201804, China

²College of Design and Innovation, Tongji University, Shanghai 200092, China

³School of Architecture and Art, Suzhou Industrial Park Institute of Vocational Technology, Suzhou 215123, China

Correspondence should be addressed to Chunjong Zhang; chunjiongzhang@tongji.edu.cn

Received 6 March 2022; Revised 14 March 2022; Accepted 18 March 2022; Published 11 May 2022

Academic Editor: Mohammad Farukh Hashmi

Copyright © 2022 Di Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Unmanned ship navigates on the water in an autonomous or semiautonomous way, which can be widely used in maritime transportation, intelligence collection, maritime training and testing, reconnaissance, and evidence collection. In this paper, we use deep reinforcement learning to solve the optimization problem in the path planning and management of unmanned ships. Specifically, we take the waiting time (phase and duration) at the corner of the path as the optimization goal to minimize the total travel time of unmanned ships passing through the path. We propose a new reward function, which considers the environment and control delay of unmanned ships at the same time, which can reduce the coordination time between unmanned ships at the same time. In the simulation experiment, through the quantitative and qualitative results of deep reinforcement learning of unmanned ship navigation and path angle waiting, the effectiveness of our solution is verified.

1. Introduction

Unmanned ships are intelligent platforms that rely on ship-board sensors to navigate in an autonomous or semiautonomous manner on the surface of the water and can be widely used in the fields of marine transportation, antimine, and antisubmarine. The unmanned ship is an important node in the networked unmanned system, which will overturn the traditional naval warfare style and give rise to a new marine equipment system and is of great significance to the development of marine resources and the maintenance of national maritime rights and interests [1]. Compared with other unmanned systems, unmanned ships face special challenges such as harsh marine environment (e.g., strong wave and current surges) and special characteristics of unmanned ship motion models (e.g., highly nonlinear model, strong time lag, and time variability) [2].

Path planning [3, 4] is a very important technique in the field of unmanned ships, attracting the attention of countless researchers. Depending on the planning method, there are two different types of path planning: point-to-point and

full-coverage traversal. If the basic information about the entire environment is known to the unmanned vessel during the completion of the task, this is called global path planning, and the main algorithms include greedy algorithms, genetic algorithms, and others [5–7]. If the unmanned vessel can know part of the environmental information during the work process and cannot grasp the full information, this is called local path planning, and the main algorithms include potential field method [8], fuzzy control method [9], and neural network [10, 11].

In recent years, with the rapid progress of high-performance computing, big data, and deep learning technology, the core technology of artificial intelligence software, reinforcement learning algorithms, and their applications have been more widely focused and developed more rapidly [12]. In particular, the combination of reinforcement learning and deep learning has led to several breakthroughs in deep reinforcement learning, and the game between AlphaGo and top human chess players has led to a wider interest in deep reinforcement learning in academia and industry. Not only has reinforcement learning been a great

success in computer gaming, but it is also considered to be the most promising approach to advanced artificial intelligence in areas such as unmanned ship control, inverted pendulum control, and intelligent driving of steam unmanned ships [13–16].

Standard QL algorithms generally give the Q values an initial value and let them start at 0 or a random number, so unmanned ships lack a priori knowledge of the environment, and convergence is slow. During the learning process, it is difficult to weigh long-term and short-term benefits, and it is easy to fall into trap regions when faced with a “symmetric dilemma.” Therefore, the convergence of QL algorithms, balanced exploration, and exploitation, as well as dangerous regions and incomplete access to state-action pairs, has become the focus of reinforcement learning research.

This paper uses reinforcement learning for local path planning of mobile unmanned ships to improve the planning efficiency of optimal paths for unmanned ships. The research objectives are to solve the problems of slow convergence, exploration, and exploitation dilemmas as well as trap regions and incomplete access to state-action pairs of the reinforcement learning algorithm, to speed up the path planning efficiency, and to find the control rules for the optimal path from the starting point to the endpoint.

2. Related Works

2.1. Path Optimization. Based on the situational awareness map, the unmanned vessel navigation planning and navigation consider mission requirements, safety, efficiency, rules, maneuverability, and uncertainty to calculate the key elements of planning and navigation and form intercompatible commands with different granularity, so that the unmanned vessel can be effective while satisfying the navigation safety envelope. Unmanned vessel navigation planning and navigation face special challenges, such as many maritime rules with fuzzy properties, the large time lag in the hull model, high inertia, and large differences between different vessel types. [17] proposes a predictor for estimating the sideslip angle of an unmanned ship to achieve path following based on predicted line of sight. [18] uses an extended state observer for real-time estimation of unmanned ship sideslip angle and combines it with line-of-sight navigation to solve the path-following problem under disturbing conditions. [19] proposed global and local route planning based on Dijkstra and artificial potential field methods. [20] proposed inverse adaptive sliding mode control in the Serret-Frenet coordinate system to solve the path-following problem under model and disturbance uncertainty. [21] implements unmanned ship path-following control based on an improved backstepping method. [22] proposed three PID control methods to solve the linear path-following problem under constant surge disturbance.

[23] used evidence-based reasoning to evaluate the hazards and, based on the evaluation results, used mutual collision avoidance algorithms that satisfy maritime collision avoidance rules to achieve real-time safe obstacle avoidance of unmanned vessels [24, 25]

2.2. Unmanned Boat Control. Unmanned vessel control solves the problems of dynamic positioning, trajectory tracking, and path tracking during navigation, giving the unmanned vessel the control capabilities of an experienced pilot to successfully and stably perform the various maneuvers required for navigation. With the development of control theory, researchers in the marine field can apply the latest control techniques to unmanned ship control. However, the control of unmanned ships faces challenges such as high model nonlinearity and uncertainty, system underdrive, the time lag in the ship itself and in the actuators, saturation characteristics of the actuators, and unpredictable strong external disturbances.

[10, 11] systematically described the progress of research on control of marine electromechanical systems and course keeping of unmanned vessels, respectively. [12] is based on the active method for model identification of unmanned vessels. In [13], model identification of an unmanned boat with integrated propulsion is based on the idea of MMG separation modelling. A robust controller with fast convergence is also proposed based on the idea of multimodel control. The [14] active antidisturbance control law based on compound errors is used to suppress external disturbances. [15] implemented the control of an unmanned ship based on the GPC-PID method and conducted sea trials in the southern Yellow Sea. [16] proposed the robust control with variable delay using smith predictor and extended state observer.

2.3. Unmanned Vessel Cluster Control. Single-ship capabilities are particularly weak in the face of vast and hostile oceans. The path-following control layer uses an inverse step and neural network approach. A graph-theoretic approach is used for the speed and route planning layer. [16] designed a layered control architecture for cluster target tracking, cluster obstacle avoidance, and collision avoidance for members within the cluster. The control architecture is divided into three layers: the cluster strategy layer, the motion planning layer, and the control input layer. [17] uses neural optimization for distributed cluster navigation and a fuzzy approach for approximating the model of an unmanned ship for unmanned ship path maneuver cluster control.

3. Preparation

This section provides an overview of reinforcement learning and the deep reinforcement learning utilised in this paper.

Reinforcement learning is the scheme by which an agent infers the best action rule through its interaction with the environment. MDP is defined by a 4-tuple (S, A, P, R) . Call S the state space and A the action space, and call the respective original $s \in S$ states and $a \in A$ actions. $P : S \times A \times S \rightarrow [0, 1]$ is called the state transfer function and determines the transfer probability to the next state s' when action a is performed in state s . $R : S \times A \times S \rightarrow R$ is the reward function.

Once a strategy has been formulated, the intelligence can interact with the environment as shown in Figure 1. At each moment t , the intelligence in state s_t decides on action a_t

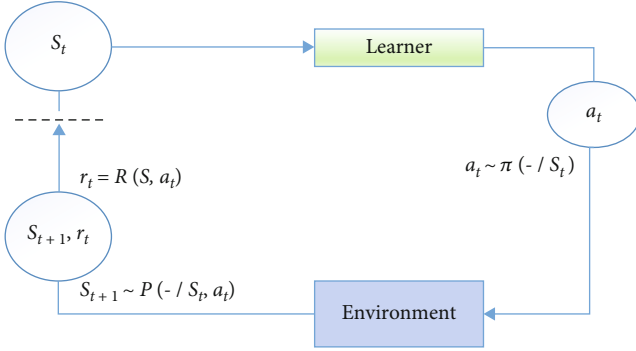


FIGURE 1: Interaction between agent and environment.

according to strategy $\pi(\cdot|s_t)$. The next moment of the intelligence's state $s_{t+1} \sim P_M(\cdot/s_t, a_t)$ and reward $r_t = R(s_t, a_t)$ is then decided according to the state transfer function and the reward function. Repeating this action gives the state and action of the intelligence's history $(s_0, a_0, s_1, a_1, \dots, s_T)$. Subsequently, the state and action that have been repeated for T historical transfers from time 0 are noted as d_t .

Define the value function, which averages the (discounted) reward sum when the action a is selected in state s and thereafter continued according to the strategy π , as given by

$$Q^\pi(s, a) \equiv \lim_{T \rightarrow \infty} \mathbb{E}_{d_T}^\pi \left[\sum_{k=0}^T \gamma^k R(s_k, a_k, s_{k+1}) | s_0 = s, a_0 = a \right], \quad (1)$$

where $\lambda \in [0, 1]$ is the discount rate and $\mathbb{E}_{d_T}^\pi$ represents the average operation about the occurrence mode in policy π . When a policy π, π' meets $Q^\pi(s, a) \geq Q^{\pi'}(s, a)$ in any $s \in S, a \in A$, since strategy π can be expected to bring more rewards to the agent than π' ; $\pi \geq \pi'$ is to strengthen learning and obtain the best method π^* to meet any scheme π and $\pi^* \geq \pi$.

The optimal strategy function is obtained by using its value function Q^* (the optimal value function) set to $\pi^*(a/s) = \delta(a - \arg \max_{a'} Q^*(s, a'))$. The optimal strategy function is the optimal Bernoulli equation:

$$Q^*(s, a) = \mathbb{E}_{s'} \left[R(s, a, s') + \gamma \max_{a'} Q^*(s', a') \right]. \quad (2)$$

The conditions are known to be satisfied and are estimated using the relational formulation above. The representative method is a method called Q-learning, and many experiments have shown that it works well but is difficult to apply to continuous and large state problems if the state space is discrete and the number of states is not huge.

4. Reinforcement Learning for Unmanned Ship Control Mapping

In this subsection, we describe the method proposed in this study for simultaneously controlling the path of movement of an unmanned ship and path corner waiting. We first describe the path corner waiting for optimization

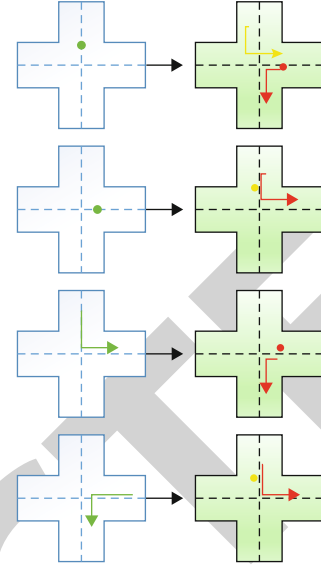


FIGURE 2: Unmanned boat direction change.

problem as a reinforcement learning task by giving details of the state, behaviour, and reward as shown specifically in Figure 2.

4.1. Action Space Definition. The intelligent body must consider the current unmanned vessel signal and the selected action when switching between unmanned vessel signals in order to ensure safety.

When the control object is a beacon and an unmanned ship movement path, the action space related to the path corner waiting as a whole s_{sig} and the space related to the unmanned ship movement path as a whole are defined as A_{veh} . Each action space, e.g., $s_{sig} = \{\text{north-south, east-west}\}$ and $A_{veh} = \{\text{routeA, routeB, routeC}\}$, the north-south (up and down) direction of the signal turns blue, and the east-west (left and right) direction turns red.

Define the action space as the product of the action space s_{sig} for path corner waiting and the action space A_{veh} for each unmanned ship present, with the size of the action space growing according to the number of unmanned ships.

4.2. Path Indicator. We address the onerous task of learning a strategy due to the growth in the number of unmanned vessels in the action space of the intelligence by introducing a virtual machine defined as a path indicator.

For definition, a machine that performs the same route indication for all control object drones present in a defined section is called a path indicator.

In addition, the "virtual" machine, as it is called here, does not necessarily need a physical counterpart, unlike the usual path corner waiting. For example, even if there is no device that can be seen by the pilot as in the case of path corner waiting, it is possible to use a device that is set remotely to allow the unmanned vessel to receive information within a defined section of road and to act as a path indicator by giving the unmanned vessel information on the path it should take.

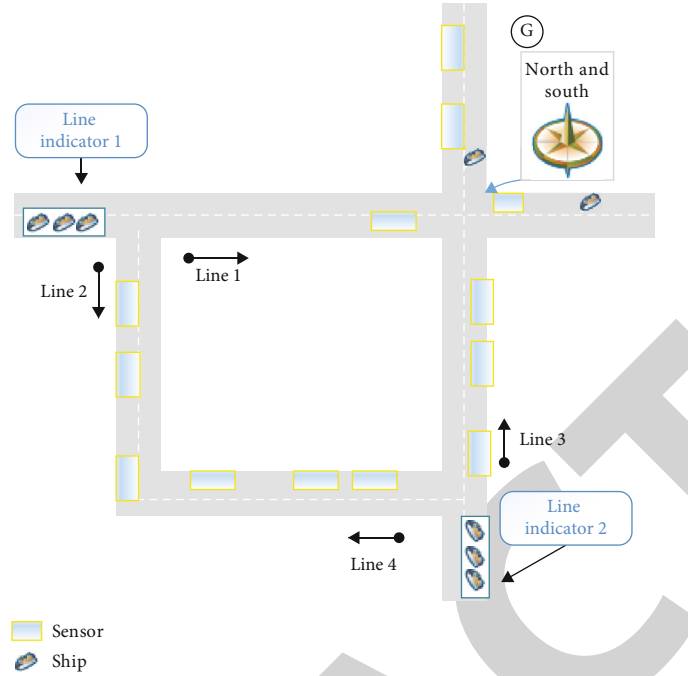


FIGURE 3: Control object area.

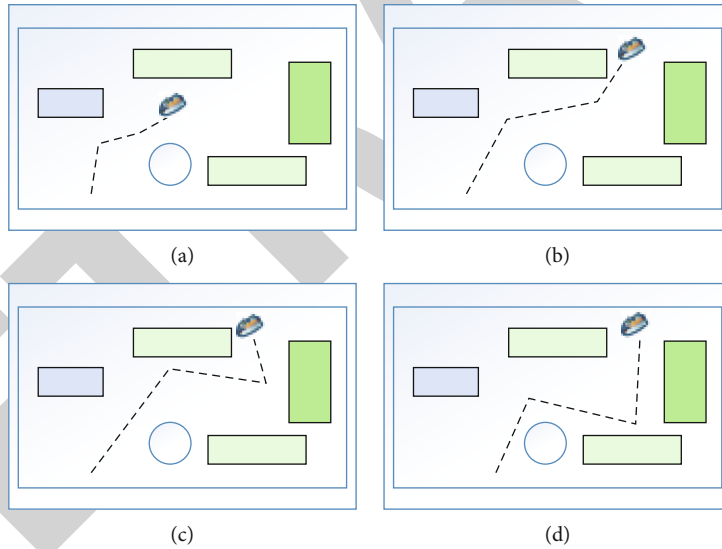


FIGURE 4: Path diagram of the training process.

Based on the above discussion, the action space of the proposed solution is defined as the product of the action space of the path corner wait and the action space A_{veh} of the unmanned vessel where each path indicator is present. In the case of the control object area defined in Figure 3, the action space of path indicator 1 is $A_{veh}^{(1)} = \{\text{indication 1, indication 2}\}$, and the action space of path indicator 2 is $A_{veh}^{(2)} = \{\text{indication 3, indication 4}\}$. The action space is therefore shown below, with a total number of actions of $2 \times 2 \times 2 = 8$. The definition of equation (3) thus allows the problem of simultaneously optimizing path corner waiting and

moving paths to be solved as a practical problem.

$$A = A_{eig} \times \underbrace{A_{veh}^{(1)} \times A_{veh}^{(2)}}_{\text{Number of path indicators}}. \quad (3)$$

5. Improved Reward Function

Deep Policy Gradient (DPG) methods attempt to optimize the policy function iteratively by learning the parameter θ to estimate the gradient of the policy performance. A disadvantage of DPG methods is that they typically lead to high

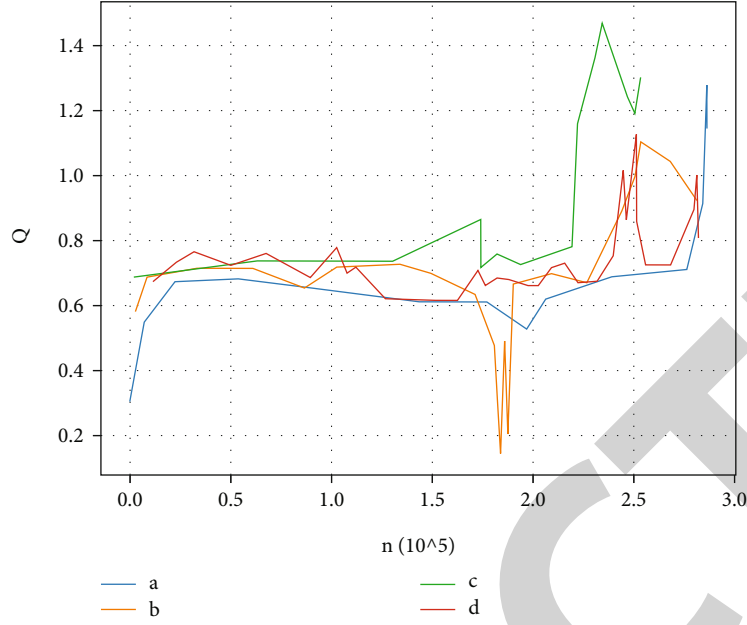


FIGURE 5: Q value.

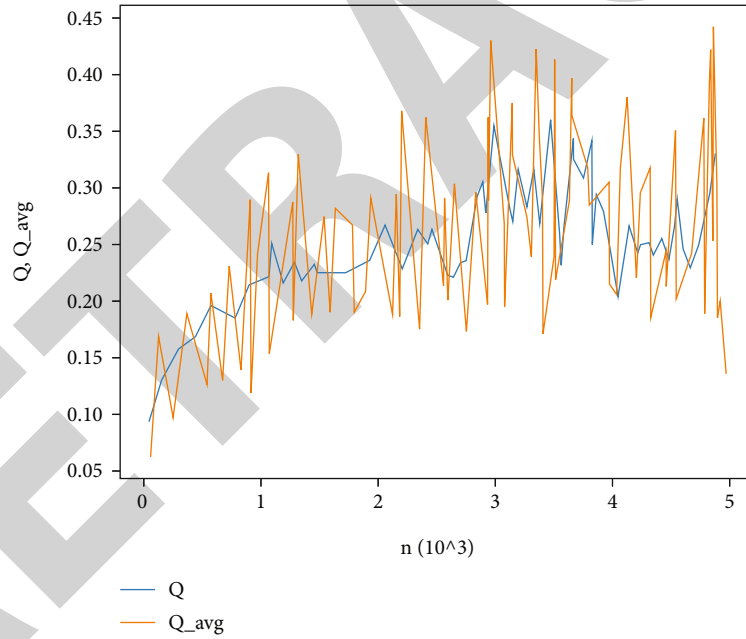


FIGURE 6: Average Q value during training.

variance in the gradient estimation. This problem arises because the trajectories in the gradient estimation are randomly sampled. That is, the derivative of the log and reward variance of a strategy can be very high.

The reward $r_t \in R$ is a scalar value which is the step obtained after each execution of an action by agent, defined as follows:

$$r_t^{(1)} = D_{t-1} - D_t, \quad (4)$$

where D_{t-1} and D_t are the total cumulative delays between the current and previous times.

We can also consider the flow of unmanned vessels in the unmanned ship channel by looking at the occupancy rate, defined as the percentage of the given unmanned ship channel that is full. In addition, we can consider the number of unmanned vessels parked unmanned in the unmanned ship channel as a frustrating phenomenological bonus. Considering the combination of the above parameters, we can positively reward the unmanned boat flow as follows:

$$r_t^{(2)} = D_{t-1} - D_t + \frac{\text{Occupancy}}{\text{Number Of Halting Vehicles} + c}, \quad (5)$$

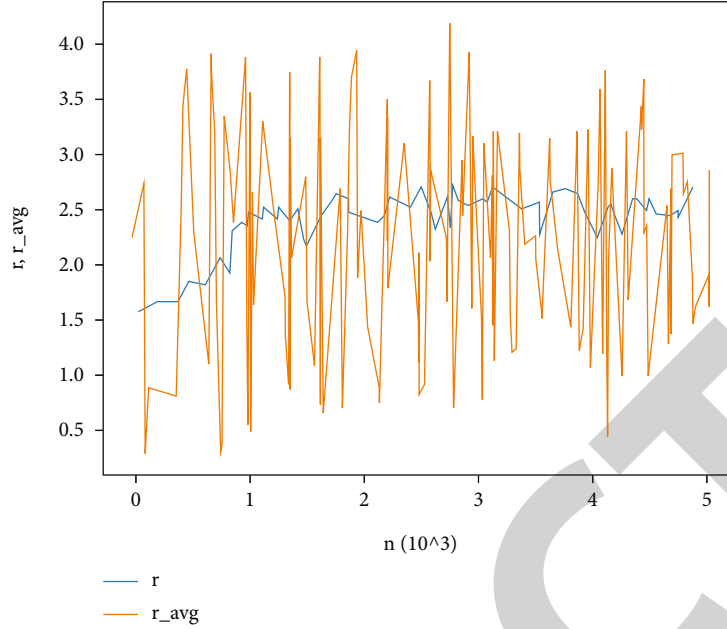


FIGURE 7: Payoff values during training.

where $c \in (0, 1)$ is a constant used to prevent division by zero. Intuitively, if the unmanned boats in the unmanned boat lane do not stop to encourage unmanned boat traffic, we will give a positive reward for a full unmanned boat lane.

Suppose we have a highway that goes 90 miles per hour. Also, suppose that all the unmanned ships in the graph on the left have a speed of 80 miles per hour. Then, the delay for each unmanned boat is $(90 - 80)/90 = 1/9$, since there are 3 unmanned boats in the left graph and the delay at $t = 0$ is $1/3$. For the graph on the right, suppose that another unmanned boat enters the highway at $t = 1$ at a speed of 80 miles per hour. Then, in this case, the delay at time $t = 1$ is $4/9$, and the second reward function is $r_1^{(1)} = 1/3 - 4/9 = -1/9$, assuming that each unmanned boat lane can accommodate 10 unmanned boats and $c = 1$.

6. Experiment

The starting position of the unmanned ship in the simulation scene is $(0 \text{ m}, 0 \text{ m})$, and the target point is $(4.5 \text{ m}, 4.5 \text{ m})$. The scene is surrounded by a wall with some random obstacles inside. The unmanned ship continuously learns and explores the scene according to the reinforcement learning algorithm proposed in this paper, and once the unmanned ship collides with the obstacles or reaches the target point, the whole scene is reset. The parameters of the simulation experiment are set as follows: learning rate $\alpha = 0.2$, $r_{\text{rea}} = 1$, $c_t = -1$, $d_{\text{win}} = 0.25$ discount rate $\gamma = 0.9$, $r_{\text{col}} = -1$, $d_{\text{col}} = 0.05$, and total training times 5,000 [26, 27].

Figure 4 shows the path diagrams during training. Figures 4(a)–4(d) are the results of the 423rd, 1,566th, 3,532nd, and 4,879th training sessions, respectively. Because of the high random probability in the early stage of training, our algorithm in Figure 4(a) does not converge, and the unmanned boat collides with the obstacle.

Figure 5 shows the Q values obtained by the unmanned ship at each step during the above training process. (a) to (d) in the figure correspond to (a) to (d) in Figure 3, respectively, and the unmanned ship selects the action with the largest Q value in the current state to execute each time. In (b), (c), and (d), the unmanned boat successfully avoided the obstacles and reached the end point, especially in the late training period when the unmanned boat was approaching the end point. The expected future payoffs for these states were high, and the selection of these actions resulted in higher payoffs, so the Q values were higher, in line with the results that the reinforcement learning method could produce [28].

Figure 6 shows the average Q value of each selected action during the training process, and it can be seen from Figure 6 that in the early stage of training, the average Q value is low because the random probability is large and the unmanned ship has gained less knowledge and experience, so the number of times the unmanned ship reaches the target point is low; after the training reaches 1,000 times, the number of times the unmanned ship reaches the target point after learning gradually increases, and the number of times it gains positive reward is increasing, so the average Q value gradually increases, i.e., the average cumulative reward value obtained by the unmanned ship for the selected action becomes higher and higher; after that, as the number of training times increases, the number of arriving at the target point becomes more and more, the reward also becomes more and more, the Q value gradually increases, and the algorithm finally converges gradually.

Figure 7 illustrates the cumulative payoffs (r and r_{avg} are the payoffs and average payoffs, respectively) during the training process. The highest payoff value of 1 is obtained when the unmanned ship reaches the end point, the lowest payoff value of -1 is obtained when a collision occurs, and

the corresponding positive payoff is obtained when the unmanned ship moves towards the target point.

7. Conclusions

In this paper, we use deep reinforcement learning to solve the optimization problem in unmanned boat path planning management; specifically, we take the timing of path corner waiting as the optimization objective to minimize the total travel time of an unmanned boat crossing the path. In the experiments, quantitative and qualitative results of deep reinforcement learning on unmanned ship travel and path corner waiting are reported to verify the effectiveness of our solution.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

- [1] J. Woo, C. Yu, and N. Kim, "Deep reinforcement learning-based controller for path following of an unmanned surface vehicle," *Ocean Engineering*, vol. 183, pp. 155–166, 2019.
- [2] N. Wang, Y. Gao, H. Zhao, and C. K. Ahn, "Reinforcement learning-based optimal tracking control of an unknown unmanned surface vehicle," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 7, pp. 3034–3045, 2021.
- [3] H. Xu, N. Wang, H. Zhao, and Z. Zheng, "Deep reinforcement learning-based path planning of underactuated surface vessels," *Cyber-Physical Systems*, vol. 5, no. 1, pp. 1–17, 2019.
- [4] D. H. Chun, M. I. Roh, H. W. Lee, J. Ha, and D. Yu, "Deep reinforcement learning-based collision avoidance for an autonomous ship," *Ocean Engineering*, vol. 234, article 109216, 2021.
- [5] Y. Zhao, X. Qi, Y. Ma, Z. Li, R. Malekian, and M. A. Sotelo, "Path following optimization for an underactuated USV using smoothly-convergent deep reinforcement learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 10, pp. 6208–6220, 2021.
- [6] Y. Cheng, Z. Sun, Y. Huang, and W. Zhang, "Fuzzy categorical deep reinforcement learning of a defensive game for an unmanned surface vessel," *International Journal of Fuzzy Systems*, vol. 21, no. 2, pp. 592–606, 2019.
- [7] B. Du, B. Lin, C. Zhang, B. Dong, and W. Zhang, "Safe deep reinforcement learning-based adaptive control for USV interception mission," *Ocean Engineering*, vol. 246, article 110477, 2022.
- [8] L. Li, D. Wu, Y. Huang, and Z. M. Yuan, "A path planning strategy unified with a COLREGS collision avoidance function based on deep reinforcement learning and artificial potential field," *Applied Ocean Research*, vol. 113, article 102759, 2021.
- [9] D. Wu, C. Zhang, L. Ji, R. Ran, H. Wu, and Y. Xu, "Forest fire recognition based on feature extraction from multi-view images," *Traitement du Signal*, vol. 38, no. 3, pp. 775–783, 2021.
- [10] Y. Zhang, Y. Zhang, and Z. Yu, "Path following control for UAV using deep reinforcement learning approach," *Guidance, Navigation and Control*, vol. 1, no. 1, p. 2150005, 2021.
- [11] J. Xie, R. Zhou, Y. Liu et al., "Reinforcement-learning-based asynchronous formation control scheme for multiple unmanned surface vehicles," *Applied Sciences*, vol. 11, no. 2, p. 546, 2021.
- [12] P. An, Z. Wang, and C. Zhang, "Ensemble unsupervised auto-encoders and Gaussian mixture model for cyberattack detection," *Information Processing & Management*, vol. 59, no. 2, article 102844, 2022.
- [13] Y. Zhao, Y. Ma, and S. Hu, "USV formation and path-following control via deep reinforcement learning with random braking," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 12, pp. 5468–5478, 2021.
- [14] Q. Zhang, W. Pan, and V. Reppa, "Model-reference reinforcement learning control of autonomous surface vehicles," in *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 5291–5296, Jeju, Korea (South), 2020.
- [15] X. Wu, H. Chen, C. Chen et al., "The autonomous navigation and obstacle avoidance for USVs with ANOA deep reinforcement learning method," *Knowledge-Based Systems*, vol. 196, article 105201, 2020.
- [16] N. Wang, Y. Gao, and X. Zhang, "Data-driven performance-prescribed reinforcement learning control of an unmanned surface vehicle," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 12, pp. 5456–5467, 2021.
- [17] A. Hasankhani, Y. Tang, J. VanZwieten, and C. Sultan, "Comparison of deep reinforcement learning and model predictive control for real-time depth optimization of a lifting surface controlled ocean current turbine," in *2021 IEEE Conference on Control Technology and Applications (CCTA)*, pp. 301–308, San Diego, CA, USA, 2021.
- [18] K. Zhu and T. Zhang, "Deep reinforcement learning based mobile robot navigation: a review," *Tsinghua Science and Technology*, vol. 26, no. 5, pp. 674–691, 2021.
- [19] K. Zhu, Z. Huang, and X. Wang, "Tracking control of intelligent ship based on deep reinforcement learning," *Chinese Journal of Ship Research*, vol. 16, no. 1, p. 43, 2021.
- [20] K. Liu, P. Li, C. Liu, L. Xiao, and L. Jia, "UAV-aided anti-jamming maritime communications: a deep reinforcement learning approach," in *2021 13th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–6, Changsha, China, 2021.
- [21] S. Xie, X. Chu, M. Zheng, and C. Liu, "A composite learning method for multi-ship collision avoidance based on reinforcement learning and inverse control," *Neurocomputing*, vol. 411, pp. 375–392, 2020.
- [22] A. Radwan, K. M. S. Huq, S. Mumtaz, K. F. Tsang, and J. Rodriguez, "Low-cost on-demand C-RAN based mobile small-cells," *IEEE Access*, vol. 4, pp. 2331–2339, 2016.
- [23] C. Wang, X. Zhang, L. Cong, J. Li, and J. Zhang, "Research on intelligent collision avoidance decision-making of unmanned ship in unknown environments," *Evolving Systems*, vol. 10, no. 4, pp. 649–658, 2019.
- [24] A. T. Azar, A. Koubaa, N. Ali Mohamed et al., "Drone deep reinforcement learning: a review," *Electronics*, vol. 10, no. 9, p. 999, 2021.

- [25] E. Artusi, F. Chaillan, and A. Napoli, "Path planning for a maritime surface ship based on deep reinforcement learning and weather data," in *OCEANS 2021*, pp. 1–8, San Diego–Porto, 2021.
- [26] K. Duan, S. Fong, and C. P. Chen, "Reinforcement learning based model-free optimized trajectory tracking strategy design for an AUV," *Neurocomputing*, vol. 469, pp. 289–297, 2022.
- [27] X. Lu, C. Zhai, V. Gopalakrishnan, and B. G. Buchanan, "Automatic annotation of protein motif function with gene ontology terms," *BMC Bioinformatics*, vol. 5, no. 1, p. 122, 2004.
- [28] H. L. Chien, Y. H. Chiu, and Y. C. Lee, "Maskless lithography based on oblique scanning of point array with digital distortion correction," *Optics and Lasers in Engineering*, vol. 136, article 106313, 2021.

RETRACTED