Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 9, No. 1, 2024

Research on Optimization Method of Dredging Robot based on Deep Learning

Guangqiang Lyu [1, a], Yanjing Xie [2, *]

[1] NingboTech University, Ningbo, Zhejiang, 315100, China
[2] Ningbo Guangqiang Robot Technology Co., LTD, Ningbo, Zhejiang, 315000, China
*** Corresponding author:** Yanjing Xie (Email: stoneLgq3@gmail.com), [a] guangqiang@nbt.edu.cn

Abstract: When facing the complex environment, the traditional dredging robot is often limited by the preset rules and algorithms, and it is difficult to adapt to the dynamically changing working environment. in recent years, the rapid development of deep learning has provided new ideas for robot trajectory planning. In this context, this paper proposes an optimization method based on deep learning, combined with deep neural network and reinforcement learning algorithm, aiming to improve the working efficiency and safety of dredging robot in complex environment.

Keywords: Deep Learning; Dredging Robot; Track Planning; Optimization Method.

1. Introduction

Traditional dredging methods mostly rely on manual operation, which is inefficient and has safety risks. The trajectory planning optimization scheme of the desilting robot based on deep learning is a method to deal with the complex challenges brought by the rapid development of underwater desilting operations and deep learning technology, It is the key to improve the desilting efficiency and safety[1].The trajectory planning optimization based on deep learning enables the dredging robot to quickly identify the surrounding environment and intelligently plan the dredging route, which is an important direction for the development of future underwater robot technology.

2. Theoretical Basis and Research Framework

2.1. An Overview of the Core Principles of Deep Learning

The basic principle of deep learning is to imitate the neural network structure of the human brain and build a deep neural network system composed of multi-layer neurons. In essence, it is an important branch of machine learning. Deep neural networks rely on a large number of training data for learning, and the data features can be extracted, predict and classify the data. Hierarchically rich network structure is the core of deep learning, enabling it to learn complex patterns and associations in the data.

In deep learning, the input is the output of each neuron receiving the previous layer, At the same time, the activation function can be used to complete the nonlinear transformation processing, and then output to the next layer[2].In the deep learning model, it updates the parameters of the network according to the error between the predicted value and the true value to capture the hierarchical characteristics of the data, so as to continuously optimize the performance of the model.

2.2. Problem Ition of Trajectory Planning of Dredging Robot

For the definition of the dredging robot trajectory planning problem, in the final analysis is to take measures to plan an optimal path from the starting point to the end of the robot in the state of dredging operation, Thus, the desilting task can be completed safely, efficiently and accurately. This problem can also be carefully divided into the following aspects:

First, Environment perception and modeling: The primary task of the deep learning dredging robot during the operation service is to integrate sonar, sensors and other equipment to perceive and understand the underwater terrain, obstacles and silt distribution information, and thus build an environmental model.

Second, set targets and initialize paths: The target location or area of the robot is planned based on the constructed environment model, and a path from the starting point to the target is initially planned during the dredging operation. Although the constructed path may not be optimal, it provides a starting point for the subsequent optimization.

Third, Trajectory optimization and obstacle avoidance: Based on the preliminary path, deep learning and other technologies are used to adjust the key points of the path to avoid obstacles, reduce energy consumption and improve efficiency. Deep learning models can learn and improve trajectory planning strategies based on historical data and real-time feedback [3]

Fourth, Real-time adjustment and feedback: The robot adjusts the trajectory in real time, based on environmental changes and task progress. For example, when unforeseen obstacles or silt accumulation, the robot needs to be able to rearrange its path.

3. Application of Deep Learning in Trajectory Planning

3.1. Fusion of DNN and RL Trajectory Planning Algorithm

In reinforcement learning, the operation goal of the dredging robot is to select a series of actions to maximize the cumulative reward, which can be realized by finding the strategy π. Policy π maps all the states to the action α. The merits of the strategy is measured by cumulative reward $J(\pi)$:

$$J(\pi) = E_{\tau\sim\pi}[\textstyle\sum_{t=0}^{\infty} \gamma^t r(s_t, \alpha_t)] \tag{1}$$

In formula 1, $J(\pi)$ is that under the πstrategy, The desired cumulative reward of the dredging robot after performing a series of actions; τ Is a trajectory composed of the state and the action; γ Is a discount factor, used to regulate the importance of future rewards;$r(s_t, \alpha_t)$is the immediate reward after performing α_t in state s_t.

In the trajectory planning algorithm that combines DNN with RL, the DNN was used to approximate the valued function $Q(s, \alpha)$, representing the expected return of the function performing action α at a given state s. The output $Q_\theta(s, \alpha)$ of the DNN is the estimated result of the true value $Q(s, \alpha)$,Where θ is the DNN network parameter:

$$Q(s, \alpha) \approx Q_\theta(s, \alpha) \tag{2}$$

3.2. Movement Trajectory Optimization Strategy based on Deep Learning

According to the characteristics of the dynamic model of the dredging robot, based on reinforcement learning, two deep neural networks of Actor network and Critic network are introduced to construct the optimal motion trajectory control model. in Figure 1, The optimal trajectory control model of the dredging robot shows its effect with the surrounding environment.

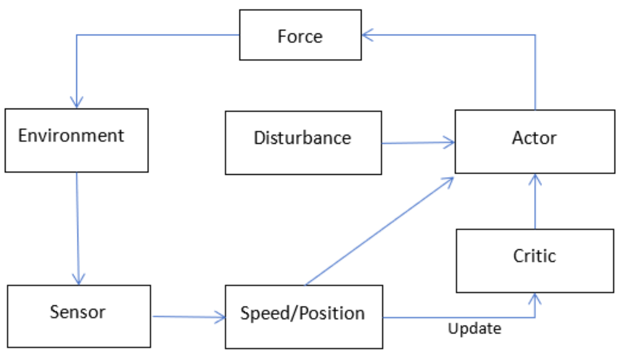

Fig 1. Structural diagram of the control model

As found in the control model structure, its composition is the Actor network, Critic network, Where the actor network $\mu(s_t, \theta)$ replaces the decision function $\mu(s_t)$.When training the Actor network, it is expanded as the controller to obtain the action record a_t of the controller within the t time range. At time t, the dredging robot is as follows:

$$s_t = [v_t, \vartheta_t]^T \tag{3}$$

The propulsion force is expressed as follows:

$$F_t = \mu[v_t, \vartheta_t]^T \tag{4}$$

In the formula, ϑ is buoyancy, μ is the controller.

Use the Critic network $Q(s_t, \mu(s_t, \theta) \mid \omega)$to replace the evaluation function Q, where ω is the Critic network parameter.When training the Critic network, it is expanded as an evaluator, and the output results can be the actor network parameters used to evaluate the training effect of the controller.Deep learning algorithms take different application scenarios as the starting point and foothold, and clarify multiple types of reward signals.The application of deep learning algorithm causes the reward signal to continue to 0, ensuring that the learned strategy tends to the best state. In other words, the closer the actual position of the dredging robot is to the ideal position, the more the reward signal tends to 0 [4]. According to the linear quadratic regulator, the reward signal $r(s_t, \alpha_t)$ can be set as follows:

$$r(s_t, \alpha_t) = -(s_t - \hat{s}_t)^2 A - a_t^2 B \tag{5}$$

In formula 5,$\hat{s}_t$ is the ideal state; A and B are arbitrary positive definite matrices, as necessary for the stable operation of the system. When the dredging robot is running close to the ideal state, the return signal tends to 0.

When tracking and controlling the movement trajectory of the dredging robot, the control system should be in a stable state to avoid the oscillation phenomenon on the premise of maintaining the accuracy standard. Therefore, when training Actor network and Critic network, the discrete degree of reward signal is taken as the criterion to judge the success of training. When performing 100 training the reward signal criterion does not exceed the threshold ϵ_r, it was met:

$$\sqrt{\frac{\sum_{j=1}^{100}\sum_{i=1}^{T}\left[r_j(s_i,a_i)-\overline{r_j}(s_i,a_i)\right]^2}{100T}} < \epsilon_r \tag{6}$$

When satisfying formula 6, where T is the training step length. Adam was used to continuously optimize and update the parameters of both networks to obtain the best scheme.

4. Verification of the Experimental Performance and the Comparative Analysis

Based on the technical background of deep learning depth application, the operation trajectory optimization and tracking of the dredging robot are significantly affected by the training times. in the preliminary experiment, the dredging robot was only trained once or five times, and it was found that it could not accurately track the predetermined movement track.

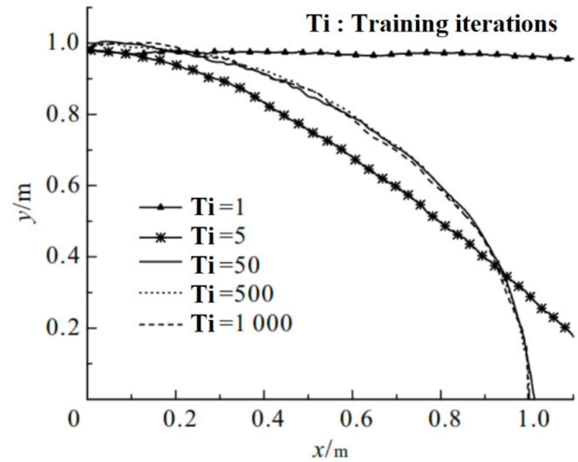

Fig 2. Robot running trajectory for different training times

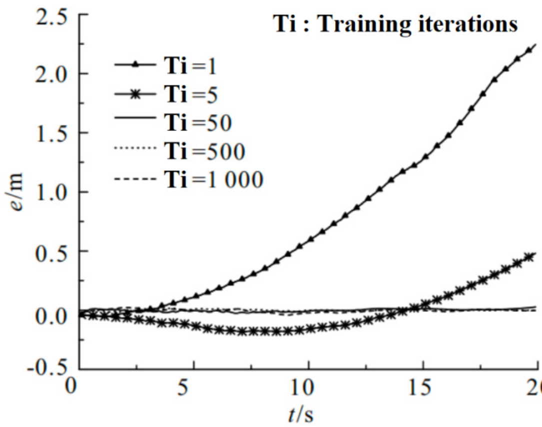

Fig 3. The error occurrence of different training times

After 500 times of training, the dredging robot showed excellent ability when running, indicating that with the increase of training times, the robot can better learn and adapt to the complex underwater environment, to track the predetermined trajectory more accurately. Moreover, with the gradual increase of the operation trajectory training times of the dredging robot, the tracking error gradually decreases, and tends to be stable and close to the zero value.

The changing trend in Fig 3 demonstrates the effectiveness of the deep reinforcement learning algorithm and reveals a positive correlation between training times and trajectory tracking accuracy.

After an in-depth comparison of the effects of the deep reinforcement learning model trained 1,000 times and the PID controllers in trajectory tracking, significant differences were found. Specifically, the trajectory generated by deep reinforcement learning is highly consistent with the target trajectory, showing its excellent learning and adaptability. In contrast, the trajectory produced by the PID controller showed a large deviation.

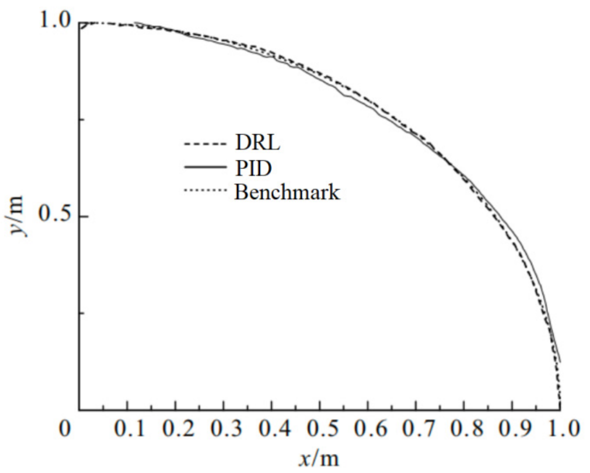

Fig 4. Deep learning and PID controller trajectory comparison

After further analysis of the error situation found that after 1000 training of deep reinforcement learning model can track error stability control within 0.02m, and the maximum error of the PID controller is close to 0.04m, and the error fluctuation is more obvious, shows the deep reinforcement learning in the advantage of track tracking, and provide valuable reference for research and application in related fields.

5. Epilogue

Traditional dredging methods are inefficient and have safety risks. The introduction of deep learning technology significantly improves the efficiency and safety of dredging operations. This paper summarizes the core principles of deep learning, and defines in detail the key links such as environment perception, goal setting, trajectory optimization and real-time adjustment of the dredging robot trajectory planning. At the same time, it also expounds the specific application of deep learning in trajectory planning, the trajectory planning algorithm integrating deep neural network (DNN) and reinforcement learning (RL), as well as the motion trajectory optimization strategy based on deep learning, to realize the intelligent optimization of the motion trajectory of the dredging robot. Through experiments, the effectiveness of deep learning in improving trajectory tracking accuracy and reducing error, demonstrating the significant advantages of deep reinforcement learning over traditional PID controllers.

Acknowledgments

Fund program： Ningbo Key Research and Development Plan and the title of "Key Technology Research and Development of intelligent silting robot for sewage Treatment" (Number2023Z034).

References

[1] Li Min,Zhang Sen,Zeng Xiangguang,WangGang ,Zhang Tongwei, Xie Dijie, Ren Wenzhe,Zhang Tao.Single-leg obstacle-crossing trajectory planning of a quadruped robot based on deep reinforcement learning [J]. Journal of System Simulation, 1-15.

[2] Ying Kaijian. Research on the key technology based on the jumping function of the foot robot [D]. ZHEJIANG UNIVERSITY OF SCIENCE & TECHNOLOGY, 2024.

[3] Li Yingjie,Chen Naijian,Yin Xunrui,Zheng Jiakun,Zhang Shanlin, Li Yingjun.Joint trajectory planning of the handling robot combining deep reinforcement learning with deformed five-time polynomials [J]. Journal of The University of Jinan (Natural Science Edition), 2024,38 (02): 234-241.

[4] Li Zhongwei, Liu Weipeng, Luo Cai.Track-guidance-based mobile robot navigation strategy optimization algorithm [J]. Computer application research,2024,41(05):1456-1461.