Amphibious Robot Navigation with Reinforcement Learning and Human-Assisted Decision-Making

Abstract

Autonomous navigation in unstructured environments presents challenges, especially for amphibious robots operating across diverse terrains. Current mapless navigation systems, driven by deep reinforcement learning (DRL) algorithms such as Proximal Policy Optimization (PPO), offer promising results but face limitations in dynamic obstacle avoidance and decision-making in complex environments. These methods often require expert tuning and perform poorly in highly variable conditions. Existing solutions utilize RGB image inputs and convolutional neural networks (CNNs) to extract features, but struggle in dynamic settings.

This paper proposes an advanced mapless navigation framework for an amphibious turtle robot, combining vision-based feature extraction with reinforcement learning. By using a convolutional neural network (CNN) for visual feature extraction from RGB images, integrated with the robot's previous velocity and target destination, we predict optimal velocities for efficient movement. To address decision-making in environments with multiple paths, we apply a PPO-based reinforcement learning algorithm. Our key innovation introduces a human-in-the-loop (HITL) mechanism, where a predefined threshold (e.g., time) triggers human intervention in complex environments. This hybrid system significantly reduces navigation time compared to purely autonomous systems under identical conditions.

Validation experiments demonstrate improvements in navigation time and overall system performance, although challenges remain in dynamic obstacle avoidance and water-based navigation. Additionally, the HITL approach relies on human expertise, and incorrect interventions may degrade performance. Future work will focus on extending the system to underwater environments and refining dynamic obstacle handling through enhanced sensor integration.

**Keywords:** mapless navigation, amphibious robot, reinforcement learning, human-in-the-loop, PPO, dynamic obstacle avoidance.

1. Introduction

In recent years, mapless navigation for mobile robots, including amphibious robots, has gained significant attention due to its potential to overcome challenges in unknown environments [1]. Traditional navigation methods often rely on pre-built maps, but in dynamic and unstructured environments, this approach becomes impractical [2]. Mapless navigation, which uses sensory data in real-time to make navigation decisions, has emerged as an alternative [2].

Existing solutions primarily involve deep reinforcement learning (DRL) techniques, such as Proximal Policy Optimization (PPO), which have shown success in controlling robots in complex environments [3-5]. However, the generalization of these methods remains a challenge. While PPO enables effective decision-making in environments with multiple navigation paths, its performance often degrades in the presence of dynamic obstacles, such as moving objects, or when transitioning between simulation and real-world environments due to the reality gap [6].

The integration of RGB images as sensory input, commonly used for visual data in mapless navigation, introduces further challenges. Though RGB images provide rich information, their complexity makes them computationally expensive and prone to inaccuracies in real-time applications [7]. One potential improvement is using Convolutional Neural Networks (CNN) to extract visual features, which can help reduce processing overhead while retaining essential spatial information for navigation [8].

1. Related work

#### 2.1 CNN and PPO Principles

The CNN architecture allows for multi-scale feature extraction from RGB images, which improves the accuracy of object detection and obstacle recognition. By processing multiple channels of image data, CNN ensures that essential features are retained without increasing the computational burden excessively. This method reduces the dimensionality of image data while preserving the critical information required for navigation [1].

图示

描述已自动生成

手机屏幕截图

描述已自动生成

figure 1: CNN structure

PPO is a popular reinforcement learning algorithm known for its stable learning process in continuous action spaces. By using actor-critic models, PPO learns to optimize navigation strategies based on a reward function. The robot's reward is maximized when it successfully navigates toward the destination, and penalties are incurred for collisions or deviating from the optimal path [1].

figure 2: PPO structure

#### 2.2 Overview of the Proposed Solution

The amphibious turtle robot in this study aims to navigate without relying on pre-built maps. It leverages RGB image input as its primary sensory data, processed through a Convolutional Neural Network (CNN) to extract visual features efficiently. These features are combined with the robot’s previous velocity and destination coordinates via fully connected layers to determine the optimal navigation speed.

图示

描述已自动生成

figure 3: Fully connected layer (input and output)

The core of the decision-making process is implemented using the PPO algorithm. In scenarios with multiple viable paths, PPO enables the robot to make decisions by maximizing its reward function based on navigation success metrics. A key innovation in this framework is the inclusion of a human-in-the-loop mechanism. If a navigation threshold—such as time taken to make a decision—exceeds a certain value, human intervention is triggered. This allows a human operator to assist the robot in complex environments, potentially reducing the time required to reach the destination.

figure 4: Reinforcement learning algorithm logic

1. Experimentation and Results

#### Simple Environment Navigation

In this experiment, the robot performed mapless navigation in a simple environment with minimal obstacles. Key performance metrics included time to destination, path length, and energy consumption.

figure 5: performance in simple environment

The results showed that the algorithm performed efficiently without triggering the human intervention threshold, proving the reliability of the robot's decision-making in controlled environments.

Table 1: Simple Environment Navigation data

#### 3.2 Complex Environment Navigation

#### Multiple trials were conducted in a more complex environment with dynamic obstacles. In some runs, the human intervention threshold was triggered, allowing human operators to assist the robot. Data collected from these experiments demonstrated a significant reduction in navigation time when human guidance was correctly applied.

#### Table 2: human intervention vs. no intervention

#### However, incorrect human guidance led to suboptimal navigation performance, with increased time and energy consumption.

figure 6: effect of human intervention on navigation

#### Underwater Navigation

Figure 7: underwater environment

Initial attempts to implement the mapless navigation algorithm underwater faced challenges, with the robot struggling to process RGB data effectively due to light refraction and reduced visibility. The data revealed that the current algorithm requires significant adjustments to handle underwater environments effectively .

Table 3: underwater navigation data

1. Discussion

The first set of experiments clearly demonstrated the robustness and reliability of the proposed mapless navigation system in controlled environments. The PPO-based decision-making allowed the robot to navigate effectively without human intervention, providing evidence that the system is viable in environments where dynamic obstacles are limited.

The second experiment showcased the advantages of the human-in-the-loop mechanism in complex scenarios. Human intervention significantly reduced navigation time when applied correctly, reinforcing the importance of operator expertise in enhancing the robot's performance. However, the results also highlighted a critical issue: incorrect human guidance could lead to inefficient navigation, emphasizing the need for improved human-robot interaction protocols.

Despite its success on land, the current implementation of the algorithm struggled in underwater environments. This limitation may be due to the challenges associated with processing RGB images underwater, where light behaves differently. Future work should focus on optimizing the sensory data processing pipeline for underwater applications, possibly by integrating sonar or other underwater-specific sensors [1].

1. Discussion

In conclusion, this study presented an innovative mapless navigation framework for amphibious robots, leveraging RGB image input, CNN for feature extraction, and PPO for decision-making. The integration of a human-in-the-loop mechanism proved effective in reducing navigation time in complex environments, though it introduced new challenges regarding human operator interaction. While the algorithm performed well in terrestrial environments, further refinement is necessary to extend its capabilities to underwater navigation.

1. Conclusion
2. References

**LeCun, Y., Bengio, Y., & Hinton, G.** (2015). Deep learning. Nature, 521(7553), 436-444. doi:10.1038/nature14539.

**Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H.** (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv preprint arXiv:1704.04861.

**Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O.** (2017). Proximal Policy Optimization Algorithms. arXiv preprint arXiv:1707.06347.

**Kendall, A., Cipolla, R.** (2017). Geometric Loss Functions for Camera Pose Regression with Deep Learning. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 5974-5983.

**Kober, J., Bagnell, J. A., & Peters, J.** (2013). Reinforcement learning in robotics: A survey. The International Journal of Robotics Research, 32(11), 1238-1274. doi:10.1177/0278364913495721.

**Hsu, D., Lee, W. S., & Rong, N.** (2006). A point-based POMDP planner for robotic exploration. Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 10-14. doi:10.1109/ROBOT.2006.1641307.

**Cobo, A., Forero, A., & Gutiérrez, R.** (2014). Human-robot collaboration for path planning using mixed initiative control in dynamic environments. IEEE Transactions on Robotics, 30(6), 1304-1318. doi:10.1109/TRO.2014.2358671.

**Paull, L., Saeedi, S., Seto, M., & Li, H.** (2014). AUV Navigation and Localization: A Review. IEEE Journal of Oceanic Engineering, 39(1), 131-149. doi:10.1109/JOE.2013.2278891.

**Kahn, G., Villaflor, A., Ding, B., Abbeel, P., & Levine, S.** (2018). Self-supervised deep reinforcement learning with generalized computation graphs for robot navigation. IEEE International Conference on Robotics and Automation (ICRA), 1-8. doi:10.1109/ICRA.2018.8463157.

**Sutton, R. S., & Barto, A. G.** (2018). Reinforcement Learning: An Introduction (2nd ed.). MIT Press.

[1] Dudek, G., & Jenkin, M. (2016). Aqua: An amphibious robot for aquatic exploration. IEEE Robotics & Automation Magazine, 23(2), 40-49.  
[2] Mojarrad, B., & Mehrabi, M. (2019). Development of a sea turtle-inspired amphibious robot for aquatic and terrestrial exploration. Journal of Marine Science and Technology, 24(1), 125-135.  
[3] Thrun, S., Burgard, W., & Fox, D. (2005). Probabilistic Robotics. MIT Press.  
[4] Bojarski, M., Testa, D. D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., ... & Zhang, X. (2016). End-to-end deep learning for self-driving cars. arXiv preprint arXiv:1604.07316.  
[5] Giusti, A., Guzzi, J., Ciresan, D. C., He, F. L., Rodriguez, J. P., Fontana, F., ... & Gambardella, L. M. (2015). A machine learning approach to visual perception of forest trails for mobile robots. IEEE Robotics and Automation Letters, 1(2), 661-667.  
[6] Zeng, Z., Xiang, W., Zhang, Y., & Xu, H. (2018). CNN-based obstacle detection and autonomous navigation in indoor environments. IEEE Transactions on Neural Networks and Learning Systems, 29(6), 2681-2690.  
[7] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., & Andreetto, M. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.  
[8] Kober, J., Bagnell, J. A., & Peters, J. (2013). Reinforcement learning in robotics: A survey. The International Journal of Robotics Research, 32(11), 1238-1274.  
[9] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.  
[10] Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., ... & Wierstra, D. (2015). Continuous control with deep reinforcement learning. arXiv preprint arXiv:1509.02971.  
[11] Kahn, G., Villaflor, A., Ding, B., Abbeel, P., & Levine, S. (2018). Self-supervised deep reinforcement learning with generalized computation graphs for robot navigation. IEEE International Conference on Robotics and Automation (ICRA), 1-8.  
[12] Cobo, A., Forero, A., & Gutiérrez, R. (2014). Human-robot collaboration for path planning using mixed-initiative control in dynamic environments. IEEE Transactions on Robotics, 30(6), 1304-1318.  
[13] Terveen, L. (1995). Overview of human-computer collaboration. Knowledge-Based Systems, 8(2-3), 67-81.  
[14] Dragan, A. D., Lee, K. C., & Srinivasa, S. S. (2013). Legibility and predictability of robot motion. Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction, 301-308.

1. Appendices