Autonomous Navigation and Obstacle Avoidance of UAVs Based on Deep Reinforcement Learning

Xinyuan Wang*

Department of Transportation, Nanjing University of Aeronautics and Astronautics, Nanjing, China

**Corresponding author*

Abstract: This paper studies the autonomous navigation and obstacle avoidance technology of UAVs based on deep reinforcement learning, analyzes the limitations of traditional navigation and obstacle avoidance methods, and proposes a solution to optimize the path planning and obstacle avoidance of UAVs in complex environments through deep reinforcement learning. Deep reinforcement learning can perceive the surrounding environment through autonomous learning, adjust the flight path in real time, avoid static and dynamic obstacles, and significantly improve the autonomy and task execution efficiency of UAVs. This paper systematically explores the core concepts and algorithms of deep reinforcement learning, as well as its application in UAV navigation and obstacle avoidance. In particular, an in-depth analysis is conducted on environmental perception, convergence and stability of dynamic obstacle avoidance, and optimization of algorithm performance, emphasizing the importance of improving the real-time performance of the algorithm in autonomous flight of UAVs, and demonstrating the technical advantages and broad application prospects of deep reinforcement learning in UAV navigation tasks.

Keywords: UAV Navigation; Deep Reinforcement Learning; Obstacle Avoidance; Path Planning; Autonomous Flight

1. Introduction

1.1 Research Background and Importance

With the rapid development of unmanned aerial vehicle technology in recent years, applications for UAVs have been incessantly developed within logistics, agriculture, disaster relief, and national defense. However, achieving self-acting navigation and obstacle avoidance in UAVs still faces great challenges, especially within dynamic and complex feature environments. Conventional rule-based navigation and obstacle avoidance methods usually perform quite poorly in these environments and cannot cope with the unpredictability and complexity of such an environment [1]. As a frontier in artificial intelligence, DRL has recently shown great potential in achieving efficient autonomous navigation and obstacle avoidance for UAVs through autonomously learning and optimizing strategies [2].

Traditional navigation solutions, including simultaneous localization and mapping algorithms, all depend on high-precision sensors and a large amount of computing resources to realize real-time updating of environmental maps, which is greatly limited in practical use [3]. On the contrary, DRL may reduce dependence on sophisticated sensor systems through mapping the sensor data of UAVs into control signals directly and thus have the independent navigation and obstacle avoidance of complex three-dimensional space realized. For this unique solution to the problem of independent navigation and obstacle avoidance of UAVs, DRL is exceptionally capable of conquering the challenge brought by a complex and dynamic environment. The successful application of DRL in UAVs not only promotes the development of automation technology but also offers new possibilities for the realization of unmanned and autonomous systems. Through DRL, drones can perceive the surrounding environment in an increasingly smart way, optimize paths, and avoid obstacles to improve the efficiency of finishing a task [4]. Therefore, a study on drone navigation and obstacle avoidance based on DRL will be academically valuable and have a wide range of applications [5].

1.2 Research Objectives

This study aims to explore the application of deep reinforcement learning in drone autonomous navigation and obstacle avoidance, and propose a drone navigation and obstacle avoidance strategy that can adapt to complex environments. Specifically, this study will solve the following key problems: first, how to improve the path planning ability of drones through environmental perception technology; second, how to ensure the convergence and stability of deep reinforcement learning algorithms in dynamic obstacle environments; and finally, how to optimize the performance of deep reinforcement learning algorithms to improve their efficiency in real-time navigation tasks. This study will not only verify the effectiveness of the algorithm through simulation, but also explore the future development direction of deep reinforcement learning in drone navigation, aiming to provide a new technical path for the intelligent and automated development of drones.

2. Current Status of Drone Autonomous Navigation and Obstacle Avoidance Technology

2.1 Traditional Methods of Drone Autonomous Navigation

In the early days, drone autonomous navigation technology mainly depends on some traditional navigation algorithms and sensor technologies. One of the classic methods represents navigation based on GPS and INS. These approaches provide very accurate positioning and heading information by fusing GPS signals and IMU data, and are widely applied in UAV navigation in outdoor open space [6]. However, GPS navigation in a complex environment with features such as an urban canyon, forest, and underground has its limitation, thus it usually needs to be combined with other sensors for assistance [7]. These methods, while doing so, LiDAR-and vision-based navigation methods have gradually become the core of navigation in UAVs. LiDAR generates a three-dimensional map of the environment using laser emission by measuring the reflection time, while a visual sensor would use image processing algorithms like SLAM, enabling UAVs to navigate through unknown environments. While these technologies are in wide use, there are also high computing costs and hardware device limitations, especially in the need for real-time processing of massive amounts of data to ensure flying safety for UAVs [8].

2.2 Classification and Comparison of Obstacle Avoidance Technologies

Obstacle avoidance technology is one of the key technologies in the autonomous navigation of UAVs, which can be divided into two categories: reactive obstacle avoidance approaches and planning-based obstacle avoidance methods. Normally, the reactive obstacle avoidance methods would utilize sensors for the purpose of acquiring real-time information from the surrounding environment and thus avoiding obstacles simply based on certain rule or strategy [9]. This kind of method is fast and suitable for scenarios with limited computing resources. But in a complex dynamic environment, its performance is very poor and it cannot predict the movement trajectory of obstacles effectively [10].

Obstacle avoidance with path planning, on the other hand, avoids an obstacle by first planning a safe path in advance. This class of methods usually depends on map building technologies such as SLAM and couples global path-planning algorithms (Dijkstra's algorithm, for instance, or A* algorithm) for the purpose of navigation and obstacle avoidance [11]. Such an approach may design more optimized obstacle avoidance routes in complex environments, but it is at a very high computational cost. The updating of maps in real time-especially in large-scale three-dimensional environments-can result in long computational delays. Obstacle avoidance technology based on deep reinforcement learning has received much attention recently due to the development of deep learning technology. Such technology has prominent advantages, in that it could make drones learn independently how to avoid obstacles and cope with sophisticated situations in dynamic environments [12].

2.3 Limitations and Challenges of Current Technologies

While the Navigation and obstacle avoidance in existing technologies have gone a long way, there are still so many limitations. Firstly, navigation and obstacle avoidance technologies based on traditional sensors perform very poor, including in computing resources and energy consumption, especially when flying drones are limited in size and weight and hardly portable with high-performance sensors [13]. Delays in real-time planning due to high-velocity flying could result in collision-risk conditions. Second, most current obstacle avoidance methods lack the ability to handle dynamic obstacles. Third, for training in an environment with complex conditions, most deep reinforcement learning obstacle avoidance methods require large computational resources and time, and learned policies cannot be quickly adapted to changes in the environment for real-time applications [14]. The other serious challenge is related to the lack of current environmental perception along with navigation stability, especially when the drones have to fly for an extended duration in the three-dimensional complex environment; a system cannot be robust enough [15].

3. Overview and Application of Deep Reinforcement Learning

3.1 Basic Concepts and Algorithms of Deep Reinforcement Learning

Deep reinforcement learning (DRL) is a technology that combines reinforcement learning and deep learning, aiming to solve tasks with high-dimensional input space and complex decision-making processes. The core idea of reinforcement learning is to learn the best strategy to maximize the cumulative reward through the interaction between the agent and the environment in a continuous trial and error process. DRL uses deep neural networks to represent the policy function or value function, enabling it to process complex high-dimensional input data such as images or sensor data.

In DRL, reinforcement learning problems are usually modeled as a Markov decision process (MDP), whose core elements include state space S, action space A, transition probability P and reward function R. At time t, the agent is in state $s_t \in S$, chooses an action $a_t \in A$, and transfers to the next state s_{t+1} according to the transition probability $P(s_{t+1} | s_t, a_t)$, and obtains an immediate reward $r_t = R(s_t, a_t)$. The goal of the agent is to find the optimal strategy π^* to maximize the cumulative discounted reward G_t:

$$G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \tag{1}$$

$$Q(s_t, a_t) = r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) \tag{2}$$

Among them, $\gamma \in [0,1]$ is a discount factor used to balance short-term rewards and long-term benefits. The key algorithms of deep reinforcement learning include deep Q network (DQN), policy gradient method and actor-critic method. In DQN, the Q value function $Q(s,a)$ represents the expected cumulative reward after selecting action a in state S. Approximate $Q(s,a)$ through a deep neural network, and continuously update the Q value through the Bellman equation

In addition, the policy gradient method directly optimizes the policy $\pi(a| s)$ by performing gradient ascent on the policy parameters θ to maximize the cumulative reward. The update formula of the policy gradient is:

$$\nabla_\theta J(\theta) = \mathbb{E}_{\pi_\theta}\left[\nabla_\theta \log \pi_\theta(a | s) Q^\pi(s,a)\right] \tag{3}$$

This formula indicates that the policy is updated based on the parameter gradient of the policy π by sampling experience to maximize the performance of the policy $J(\theta)$.

3.2 Analysis of the Application of Deep Reinforcement Learning in Navigation and Obstacle Avoidance

Deep reinforcement learning is increasingly used in autonomous navigation and obstacle avoidance of drones. Drones need to navigate autonomously in complex environments and avoid static or dynamic obstacles, which places high demands on the real-time, robustness and computing performance of the algorithm. Traditional path planning algorithms often require precise environmental modeling and are difficult to cope with dynamically changing environments. Deep reinforcement learning can adapt to environmental changes and update strategies in real time through autonomous learning and online adjustment.

For example, a DRL-based drone can

autonomously adjust its flight path in a dynamic environment to avoid collisions and minimize flight time. In an environment with dynamic obstacles, the DRL algorithm can not only learn how to avoid static obstacles, but also predict and avoid moving targets by continuously adjusting its strategy.

Figure 1 shows the application process of deep reinforcement learning in autonomous navigation and obstacle avoidance of drones [12]. The simulation environment on the left trains an agent (using a deep neural network) to perform navigation tasks in a simulated environment. The agent adjusts its strategy to maximize the reward by observing obstacles and position feedback in the environment. The right side shows the application of the trained strategy to the navigation task of an actual drone in the real world.

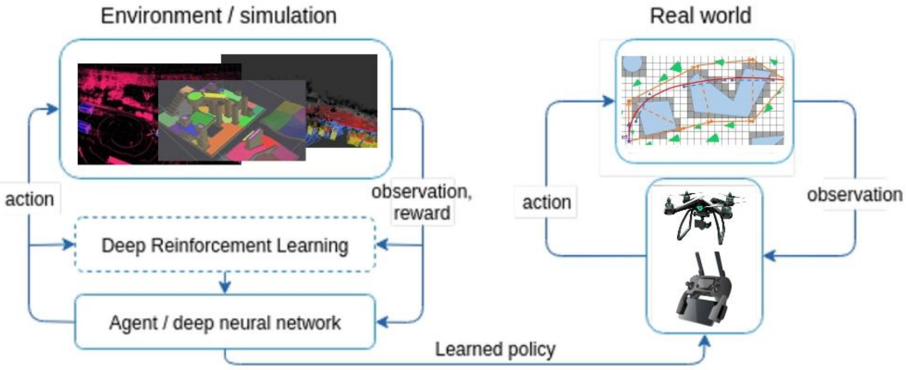

Figure 1. Application Process of Deep Reinforcement Learning in Autonomous Navigation and Obstacle Avoidance of Drones

4. Key Issues and Challenges of Deep Reinforcement Learning for Drone Navigation

4.1 Environmental Perception and Path Planning

In the process of autonomous navigation of drones, environmental perception and path planning are the key to achieving accurate navigation and safe flight. Environmental perception relies on the drone's sensors, such as lidar, cameras, or ultrasonic sensors, which are used to detect surrounding obstacles and path features. However, sensor data is usually uncertain or noisy, and how to deal with these imprecise data is a challenge. At the same time, deep reinforcement learning needs to use these perception information for path planning to ensure that the drone can find the optimal path in a complex environment. In a high-dimensional continuous state space, path planning must not only avoid collisions, but also consider task efficiency, such as the shortest path or the least energy consumption. Therefore, it is crucial to design a perception and planning system that can respond to environmental changes in real time.

4.2 Convergence and Stability in Dynamic Obstacle Avoidance

Dynamic obstacle avoidance is an important task in autonomous navigation of drones, especially when there are moving obstacles in the environment. A major challenge of deep reinforcement learning in this case is how to ensure that the algorithm converges to a stable strategy in a dynamic environment. Because the position and speed of obstacles are constantly changing, traditional path planning algorithms often cannot be updated in time, resulting in navigation failure. Deep reinforcement learning copes with these dynamic changes by constantly adjusting strategies, but slow convergence or unstable strategies may lead to collisions or path deviations during training. Therefore, ensuring the rapid convergence and stability of the algorithm in a dynamic environment is an important issue for UAV navigation, which requires more efficient exploration strategies and more accurate value function estimation.

4.3 Algorithm Performance Optimization and Real-Time Issues

UAV navigation tasks require algorithms to have high real-time performance and be able to make decisions quickly in complex three-dimensional environments. However, the computational complexity of deep reinforcement learning algorithms is high, especially when dealing with large-scale state and action spaces, which may cause decision delays and affect flight safety. To this end, optimizing the performance of the algorithm to improve its application in real-time tasks is a key challenge. The optimization directions include accelerating the network's reasoning process, reducing unnecessary computational burdens, and increasing decision speed through parallel computing. In addition, designing lightweight deep neural network models and reducing training and reasoning time are also effective means to improve the real-time performance of UAV navigation algorithms.

5. Conclusion

This paper systematically studies and discusses the application of deep reinforcement learning in autonomous navigation and obstacle avoidance of UAVs. Through the analysis of existing UAV navigation and obstacle avoidance technologies, it can be found that traditional sensor- and rule-based navigation methods have many limitations in dynamic and complex environments and are difficult to cope with the challenges of real-time changes. The introduction of deep reinforcement learning technology enables UAVs to perceive the environment, plan paths, and effectively avoid obstacles through autonomous learning, showing great technical advantages. Combined with the algorithmic characteristics of deep reinforcement learning, UAVs can show stronger adaptability and autonomous decision-making capabilities in dynamic and high-dimensional environments, and realize full-process intelligent navigation from perception to execution.

However, although deep reinforcement learning provides a new solution for UAV autonomous navigation, this paper also analyzes its challenges in environmental perception, dynamic obstacle avoidance, convergence and stability, and real-time algorithm problems. In order to solve these problems, future research needs to further explore the improvement of algorithm efficiency, optimization of training process, enhancement of environmental adaptability, and improvement of real-time decision-making capabilities. Overall, deep reinforcement learning has broad application prospects in UAV autonomous navigation, but there are still many key technical issues that need to be solved to ensure its reliability and practicality in actual complex scenarios.

References

[1] Xue, Y., Chen, W.: A UAV Navigation Approach Based on Deep Reinforcement Learning in Large Cluttered 3D Environments. IEEE Transactions on Vehicular Technology 72, pp. 3001-3014 (2023).

[2] Roghair, J., Ko, K., Asli, A. E. N., Jannesari, A.: A Vision Based Deep Reinforcement Learning Algorithm for UAV Obstacle Avoidance. ArXiv abs/2103.06403 (2021).

[3] Wang, C., Wang, J., Shen, Y., Zhang, X.: Autonomous Navigation of UAVs in Large-Scale Complex Environments: A Deep Reinforcement Learning Approach. IEEE Transactions on Vehicular Technology 68, pp. 2124-2136 (2019).

[4] Joshi, B., Kapur, D., Kandath, H.: Sim-to-Real Deep Reinforcement Learning based Obstacle Avoidance for UAVs under Measurement Uncertainty. ArXiv abs/2303.07243 (2023).

[5] [5] Grando, R. B., Jesus, J. C., Drews-Jr, P. L. J.: Deep Reinforcement Learning for Mapless Navigation of Unmanned Aerial Vehicles. 2020 Latin American Robotics Symposium (LARS), 2020 Brazilian Symposium on Robotics (SBR) and 2020 Workshop on Robotics in Education (WRE), pp. 1-6 (2020).

[6] Lin, H., Peng, X.: Autonomous Quadrotor Navigation With Vision Based Obstacle Avoidance and Path Planning. IEEE Access 9, pp. 102450-102459 (2021).

[7] Lu, Y., Xue, Z., Xia, G., Zhang, L.: A survey on vision-based UAV navigation. Geo-spatial Information Science 21, pp. 21 - 32 (2018).

[8] Back, S., Cho, G., Oh, J., Tran, X.-T., Oh, H.: Autonomous UAV Trail Navigation with Obstacle Avoidance Using Deep Neural Networks. Journal of Intelligent & Robotic Systems 100, pp. 1195 - 1211 (2020).

[9] Pant, S., Lee, S.: Obstacle Avoidance Method for UAVs using Polar Grid. Journal of Korea Multimedia Society 23, pp. 1088-1098 (2020).

[10] Wei, J., Li, S.: A Method for

Collision-free UAV Navigation around Moving Obstacles over an Uneven Terrain. 2023 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 1-6 (2023).

[11] Guo, T., Jiang, N., Li, B., Zhu, X., Wang, Y., Du, W.: UAV navigation in high dynamic environments: A deep reinforcement learning approach. Chinese Journal of Aeronautics (2020).

[12] Dobrevski, Matej & Skočaj, Danijel. (2021). Deep reinforcement learning for map-less goal-driven robot navigation. International Journal of Advanced Robotic Systems. 18. 172988142199262. 10.1177/1729881421992621.