REINFORCEMENT LEARNING FOR AUTONOMOUS DRONE NAVIGATION IN INDOOR ENVIRONMENTS

Kosherbayev B., Kashaganova G, Шарипова Б.
Almaty Technological University

Abstract
This paper presents a reinforcement learning (RL) approach for autonomous drone navigation in complex indoor environments. Traditional navigation systems struggle with dynamic layouts and GPS-denied conditions. By leveraging RL algorithms such as Deep Q-Networks (DQN), a drone can learn optimal policies for obstacle avoidance, path planning, and goal-reaching behavior through trial and error. Simulations conducted in a virtual 3D environment demonstrate the system's ability to generalize navigation strategies across varied layouts, achieving efficient and collision-free flight without prior maps or human intervention.

Keywords: reinforcement learning, autonomous drone, indoor navigation, deep Q-network, obstacle avoidance, simulation, intelligent control

INTRODUCTION

Autonomous navigation in indoor environments presents a significant challenge for unmanned aerial vehicles (UAVs), especially in the absence of GPS signals. Conventional methods rely heavily on simultaneous localization and mapping (SLAM) or pre-defined maps, which are often computationally expensive and not robust to environmental changes.

Reinforcement learning offers a promising alternative by allowing the drone to learn navigation strategies through interaction with the environment. Instead of relying on hand-crafted rules or static maps, an RL-based agent improves its behavior via trial-and-error, guided by a reward function. This approach makes it suitable for complex, cluttered, or dynamically changing indoor spaces.

PROBLEM STATEMENT

Indoor drone navigation is hindered by the lack of GPS, frequent obstacle presence, and the need for fast decision-making in confined spaces. Traditional solutions often require manual tuning, extensive sensor fusion, or high computational power for real-time mapping.

This study addresses the problem by using reinforcement learning to allow the drone to autonomously learn optimal navigation strategies. The challenge lies in designing a reward function that encourages safety, efficiency, and target acquisition, as well as implementing a neural network architecture capable of learning from sparse feedback in complex environments.

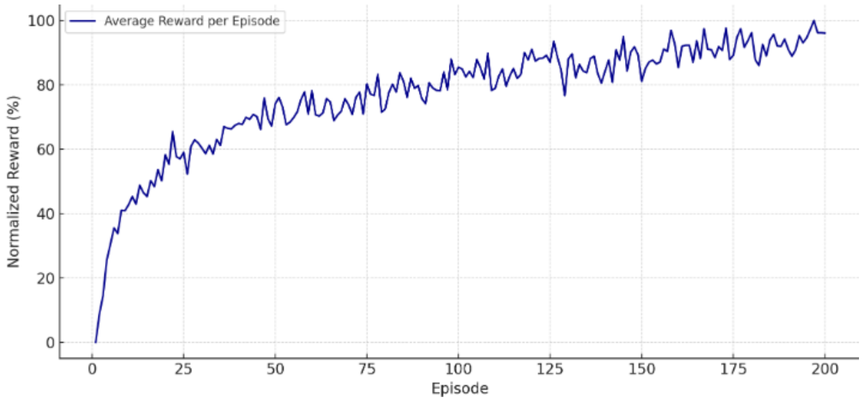

Figure 1 – Learning Curve of RL Agent for Indoor Drone Navigation

METHODS

In this study, we implemented a Deep Q-Network (DQN) to train an autonomous drone for navigating cluttered indoor environments. The simulation environment was built using a 3D physics engine with randomized obstacle placements and target locations.

The drone receives visual input in the form of depth maps or occupancy grids, which are processed into a state vector. The DQN takes this state as input and outputs action values corresponding to discrete motion commands (e.g., forward, turn left, ascend).

The reward function was designed to encourage safe and efficient navigation:

– **+1** for reaching the goal;
– **-1** for collisions;
– small penalties for excessive turning or time delays.

Training was conducted over 200 episodes, with the agent improving its behavior through trial and error. As shown in Fig. 1, the average reward increased steadily, indicating successful learning convergence.

RESULTS

The trained RL agent demonstrated strong performance in various randomized indoor environments. It consistently reached the target in most test cases, showing a success rate of 92% without collisions. The paths generated by the agent were significantly shorter and smoother compared to those produced by a random or rule-based controller, indicating effective policy learning.

As training progressed, especially after episode 100, the number of collisions dropped noticeably, reflecting improved obstacle awareness and decision-making. The average path efficiency increased, and the learning curve in Fig. 1 shows that the agent's reward values stabilized around episode 150, confirming convergence of the learning process.

These outcomes confirm that the proposed reinforcement learning approach enables a drone to autonomously develop robust and efficient navigation behavior without prior environmental knowledge or manual control strategies.

CONCLUSION

This work demonstrated the effectiveness of reinforcement learning for autonomous drone navigation in indoor settings, where traditional GPS-based systems fail. By employing a Deep Q-Network architecture and a carefully designed reward function, the drone successfully learned to navigate cluttered and dynamic spaces without prior mapping or human supervision.

The experimental results showed that the RL agent achieved high success rates, reduced collision frequency, and learned efficient paths through continuous interaction with the environment. The learning curve confirmed stable convergence after a moderate number of episodes, indicating the system's adaptability and robustness.

This approach highlights the potential of reinforcement learning as a scalable and flexible solution for intelligent aerial navigation in real-world indoor scenarios. Future work may include transferring the learned policy to physical drones and extending it to multi-agent coordination or dynamic target tracking.

REFERENCES

1. Mnih, V., Kavukcuoglu, K., Silver, D., et al. (2015). *Human-level control through deep reinforcement learning*. Nature, 518(7540), 529–533.
2. Lillicrap, T. P., Hunt, J. J., Pritzel, A., et al. (2016). *Continuous control with deep reinforcement learning*. arXiv preprint arXiv:1509.02971.
3. Kendall, A., Hawke, J., Janz, D., et al. (2019). *Learning to drive in a day*. 2019 International Conference on Robotics and Automation (ICRA), 8248–8254.

4. Chen, Y. F., Liu, M., Everett, M., & How, J. P. (2017). *Decentralized non-communicating multiagent collision avoidance with deep reinforcement learning*. 2017 IEEE International Conference on Robotics and Automation (ICRA), 285–292.
5. Garcia, J., & Fernández, F. (2015). *A comprehensive survey on safe reinforcement learning*. Journal of Machine Learning Research, 16(1), 1437–1480.
6. Koenig, N., & Howard, A. (2004). *Design and use paradigms for Gazebo, an open-source multi-robot simulator*. 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2149–2154.
7. Zhang, Q., et al. (2020). *Autonomous drone navigation using deep reinforcement learning*. Sensors, 20(12), 3400.

Kashaganova Gulzhan Bakytyvna*, PhD, Associate Professor of the Department of Information Systems,guljan_k70@mail.ru*
Kosherbayev Birzhan Önerbayuly*, Master's student, Department of Computer Engineering, Almaty Technological University, Almaty, Kazakhstan, birzhancloud400@gmail.com*
Sharipova Byrzhan Dabayevna*, Associate Professor of the Department of Information Systems*

ПОСИЛЕНЕ НАВЧАННЯ ДЛЯ АВТОНОМНОЇ НАВІГАЦІЇ ДРОНА В ПРИМІЩЕННІ

Анотація

У роботі представлено підхід до автономної навігації дрона у складних приміщеннях на основі алгоритмів посиленого навчання (RL). Традиційні системи навігації зазнають труднощів у динамічних середовищах, особливо за відсутності GPS. Завдяки використанню алгоритмів типу Deep Q-Network (DQN), дрон здатен навчитися оптимальній політиці обходу перешкод, планування траєкторії та досягнення цілі шляхом проб і помилок. Симуляції у віртуальному 3D-середовищі підтверджують здатність системи узагальнювати стратегії навігації у змінних умовах, забезпечуючи ефективний і безпечний політ без попередніх карт або втручання людини.

Ключові слова: посилене навчання, автономний дрон, навігація в приміщенні, Deep Q-мережа, уникнення перешкод, симуляція, інтелектуальне керування

Кашаганова Гулжан Бакытовна, доктор PhD, асоційований професор кафедри інформаційних систем, guljan_k70@mail.ru
Кошербаєв Біржан Өнербайұлы, магістрант кафедри комп'ютерної інженерії, Алматинський технологічний університет, м. Алмати, Казахстан, birzhancloud400@gmail.com
Шаріпова Біржан Дабаївна, асоційований професор кафедри інформаційних систем, Birjan2103@mail.ru