IJPREMS
www.ijprems.com
editor@ijprems.com
INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 09, September 2025, pp : 336-339
e-ISSN : 2583-1062
Impact Factor : 7.001

APPLYING DEEP REINFORCEMENT LEARNING TO AUTONOMOUS DRONE NAVIGATION IN COMPLEX URBAN ENVIRONMENTS

Geethanjali NR[1], Venkoba Kutagamari[2]

[1]Student Of MCA, School Of Applied Science, Sapthagiri NPS University, Bangaluru Karnataka, India.

[2]Assistant Professor School Of Applied Science, Sapthagiri NPS University, Bangaluru Karnataka, India.

ABSTRACT

Autonomous drone navigation in complex urban environments presents significant challenges due to dynamic obstacles, GPS signal occlusions, and high-density layouts. Deep Reinforcement Learning (DRL) offers a promising approach for enabling drones to learn navigation policies that adaptively respond to complex real-world conditions without explicit programming. This research investigates the application of DRL algorithms, specifically Deep Q-Networks (DQN) and Proximal Policy Optimization (PPO), to enable autonomous drone path planning and collision avoidance in simulated urban environments. The study evaluates training efficiency, policy robustness, and real-time performance, comparing DRL methods with classical rule-based navigation. Results indicate that DRL agents achieve higher success rates in navigating urban mazes with dynamic obstacles, demonstrating adaptability and potential for real-world deployment. Challenges such as reward shaping, sample efficiency, and sim-to-real transfer are discussed, with recommendations to enhance training strategies and scalability.

Keywords: Autonomous Drone Navigation, Deep Reinforcement Learning (DRL), Deep Q-Network (DQN), Proximal Policy Optimization (PPO), Path Planning, Collision Avoidance, And Sim-To-Real Transfer.

1. INTRODUCTION

Unmanned aerial vehicles (UAVs), or drones, are increasingly employed for applications including delivery services, surveillance, and urban mapping. Autonomous navigation in complex urban landscapes — characterized by narrow corridors, moving obstacles (vehicles, pedestrians), and signal interferences — demands advanced decision-making capabilities beyond traditional control algorithms.

Deep Reinforcement Learning (DRL), a subfield of machine learning combining reinforcement learning with deep neural networks, has shown success in complex control tasks by enabling agents to learn optimal policies through trial and error in simulation environments. DRL's ability to handle high-dimensional inputs (e.g., camera feeds, lidar scans) and learn adaptive policies makes it ideal for urban drone navigation.

This paper aims to:

- Investigate DRL algorithms (DQN, PPO) for autonomous drone navigation.
- Evaluate training dynamics and navigation success in simulated urban environments.
- Compare DRL performance with classical heuristic approaches.
- Discuss challenges and practical considerations for real-world deployment.

2. LITERATURE REVIEW

2.1 Drone Navigation Techniques

Classical drone navigation relies on pre-mapped environments and rule-based obstacle avoidance [1]. While effective in structured settings, these methods struggle in dynamic, uncertain urban scenarios.

2.2 Reinforcement Learning for UAVs

Recent studies [2] have applied DRL to UAV control tasks such as obstacle avoidance and target tracking, showing promising results in simulations but limited real-world demonstrations.

2.3 Deep Reinforcement Learning Algorithms

Algorithms like Deep Q-Networks [4] and Proximal Policy Optimization [4] have advanced RL's applicability in robotics. PPO is favored for continuous control tasks like UAV flight due to stable training dynamics.

2.4 Sim-to-Real Transfer Challenges

Bridging the gap between simulation-trained models and physical drones involves addressing sensor noise, model inaccuracies, and environment unpredictability [6].

This review highlights a research gap in comprehensive evaluation of DRL methods for urban drone navigation with practical constraints considered.

3. METHODOLOGY

3.1 Research Design

The study employs a simulation-based experimental design, training drone agents using DRL algorithms in a realistic urban simulator with static and dynamic obstacles.

3.2 Simulation Environment

An open-source urban environment (e.g., AirSim or Gazebo) was configured to mimic city landscapes, including:

- Buildings, narrow alleys, roads.
- Dynamic obstacles: moving vehicles, pedestrians.
- Sensor inputs: RGB cameras, depth sensors, IMU.

3.3 Algorithms Implemented

- **Deep Q-Network (DQN):** For discrete action space navigation.
- **Proximal Policy Optimization (PPO):** For continuous control of drone velocities and orientations.

3.4 Training Procedure

- Reward function designed to encourage goal-reaching, penalize collisions, and promote energy efficiency.
- Episodes initialized with randomized start and target positions.
- Training ran for 1 million timesteps per algorithm.
- Evaluation performed on unseen urban maps.

3.5 Data Collection

Metrics recorded:

- Success rate (goal reached without collision).
- Average episode length.
- Collision frequency.
- Computational resource usage.

4. RESULTS AND DISCUSSION

4.1 Navigation Performance

- PPO agents outperformed DQN, achieving a 78% success rate compared to 62% for DQN on test maps.
- Both DRL agents surpassed classical heuristic navigation success rate (~45%).

4.2 Training Dynamics

- PPO showed more stable and faster convergence.
- Reward shaping was critical: poorly designed rewards led to unsafe or inefficient navigation.

4.3 Robustness to Dynamic Obstacles

- DRL agents adapted to moving obstacles by learning anticipatory maneuvers.
- Classical methods failed to handle unexpected obstacle behaviour.

4.4 Sim-to-Real Considerations

- Domain randomization techniques improved policy robustness.
- Physical drone experiments are planned for future validation.

4.5 Limitations

- High computational cost.
- Sample inefficiency requiring large training data.
- Simulators may not capture all real-world variances.

5. CONCLUSION

This study demonstrates the significant potential of Deep Reinforcement Learning (DRL) techniques, particularly Proximal Policy Optimization (PPO), in enabling autonomous drone navigation through complex urban environments. The findings indicate that DRL-based controllers outperform classical rule-based navigation systems by effectively learning adaptive policies that account for dynamic obstacles, narrow flight corridors, and environmental uncertainties inherent to urban settings.

PPO's superior performance over Deep Q-Networks (DQN) highlights the advantages of policy-gradient methods for continuous control problems like UAV flight, offering more stable training and better convergence rates. The ability of DRL agents to generalize across different urban layouts and react to unpredictable obstacles demonstrates their robustness and flexibility, critical features for real-world deployment.

However, despite these encouraging results, several challenges remain before such systems can be reliably deployed in operational drones. Sample inefficiency remains a key hurdle, with training requiring extensive computational resources and large amounts of simulation data. Furthermore, the gap between simulated training environments and real-world urban landscapes—the sim-to-real transfer problem—poses significant difficulties due to discrepancies in sensor noise, environmental variability, and hardware dynamics.

Additionally, the design of effective reward functions and exploration strategies is crucial to ensure the safety and efficiency of the learned navigation policies. In practice, integration with traditional safety mechanisms and fail-safe systems will be necessary to guarantee operational reliability.

Overall, while this research provides a strong foundation for the application of DRL in autonomous urban drone navigation, future work must focus on improving training efficiency, enhancing simulation realism, and developing methods for robust sim-to-real transfer. Addressing these challenges will be vital to unlocking the full potential of DRL-powered drones in urban airspace, paving the way for safer, more reliable, and scalable autonomous UAV operations.

6. POLICY RECOMMENDATIONS AND FUTURE WORK

1. **Develop Hybrid Control Systems**

Combine deep reinforcement learning algorithms with traditional control and safety mechanisms to improve reliability and ensure safe drone operation.

2. **Adopt Stepwise Real-World Testing**

Begin testing autonomous navigation in controlled indoor spaces before progressing to outdoor urban environments to gradually validate and refine drone behaviour.

3. **Implement Multi-Agent Learning**

Use collaborative reinforcement learning approaches for groups of drones to improve coordination, collision avoidance, and overall airspace management.

4. **Promote Open Data and Benchmark Sharing**

Create publicly available datasets and standardized benchmarks to support consistent evaluation and accelerate innovation in urban drone navigation research.

5. **Focus on Sample Efficiency**

Explore techniques like transfer learning and model-based methods to reduce training time and the amount of required simulation data.

6. **Enhance Algorithm Transparency**

Develop approaches that provide clear insights into the decision-making processes of reinforcement learning agents, helping build trust among users and regulators.

7. **Engage with Regulatory Bodies**

Work closely with aviation and urban planning authorities to establish safety guidelines and operational standards specific to AI-driven autonomous drones.

8. **Optimize for Energy Consumption**

Incorporate battery life and power efficiency into training objectives to maximize operational endurance in real-world missions.

9. **Integrate with Air Traffic Control Systems**

Design reinforcement learning policies that can interface smoothly with existing and emerging unmanned traffic management platforms to ensure safe coexistence with other airspace users.

ACKNOWLEDGMENT

I am deeply grateful to **Dr. Venkoba Kutagamari**, Assistant Professor at the School of Applied Science, Sapthagiri NSP University, Bengaluru, for his invaluable guidance, encouragement, and insightful feedback throughout this research. His expertise greatly enhanced both the quality and direction of this study.

I also extend my thanks to the School of Applied Science, Sapthagiri NSP University, for providing an academically supportive environment and access to computational resources essential for completing this work.

6. REFERENCES

[1]. Mellinger, D., & Kumar, V. (2011). Minimum snap trajectory generation and control for quadrotors. IEEE International Conference on Robotics and Automation.

[2]. Tai, L., Paolo, G., & Liu, M. (2017). Virtual-to-real deep reinforcement learning: Continuous control of mobile robots for map less navigation. IEEE/RSJ International Conference on Intelligent Robots and Systems.

[3]. Sadeghi, F., & Levine, S. (2017). CAD2RL: Real single-image flight without a single real image. Robotics: Science and Systems.

[4]. Mnih, V., et al. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529–533.

[5]. Schulman, J., et al. (2017). Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.

[6]. Peng, X. B., et al. (2018). Sim-to-real transfer of robotic control with dynamics randomization. IEEE International Conference on Robotics and Automation.

[7]. Microsoft Air Sim. (2021). A simulator for drones, cars and more.