International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering
Vol. 9, No. 3, 2022, pp. 21-25.

ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

Reinforcement Learning Model-oriented Autonomous Drone Navigation and its Applications

Tanya Singh [a], Aditya Nair [b]

[a] *Multilingual NLP Lab, Punjab Institute of Artificial Intelligence, India.*
[b] *Multilingual NLP Lab, Punjab Institute of Artificial Intelligence, India.*

Abstract

Mobile robots, including unmanned aerial vehicles (drones), are utilized for surveillance, tracking, and data collection in structures, infrastructure, and various surroundings. The significance of precise and comprehensive monitoring is widely recognized for the early identification of issues and preventing their escalation. This necessitates the development of versatile, autonomous, and robust decision-making mobile robots. Those systems must possess the capability to learn by integrating data from numerous sources. Until relatively recently, they were specialized in tasks. This study presents a generic navigation method that utilizes data from onboard sensors to direct the drone to the issue location. In dangerous and safety-critical scenarios, the precise and swift identification of issues is essential. The research employs the proximal strategy optimization Deep Reinforcement Learning (DRL) system with incremental curricula learning and long-term memory neural networks to develop a versatile and flexible navigation system. The study assesses several configurations compared to a heuristic method to illustrate its precision and efficacy. The study evaluates the assurance of the drone's security by analyzing its performance using its navigation system in real-world situations.

Keywords: Drones, Navigation, Reinforcement Learning, Drones.

DOI: 10.71086/IAJSE/V9I3/IAJSE0924
Received: 14 June 2022/Revised: 18 July 2022/Accepted: 24 August 2022/Published: 30 September 2022

Introduction

Autonomous Systems (AS) [1] can execute designated tasks independently, exemplified by robots operating without human intervention, self-driving vehicles, and delivery robots. Artificial intelligence is infiltrating various sectors to enhance operational efficiency and mitigate costs and risks associated with human involvement.

An Unmanned Aerial Vehicle (UAV) [2] is an aircraft devoid of a human pilot, sometimes called a drone for short. Independent UAVs are garnering heightened interest due to their varied applications, including sending packages, reacting to traffic accidents to assist those injured, monitoring military targets, and aiding in rescue efforts.

Drones are generally outfitted with cameras and various sensors that gather data from their surroundings, facilitating autonomous navigation. Navigation instruction is usually performed in a virtual 3D world due to the limited computational capacity and electricity available for Drones and the high costs associated with replacing parts after crashes.

Various Reinforcement Learning (RL) [3] techniques instruct drones in independent environmental navigation. RL can address diverse challenges wherein the agent functions as a human specialist. The agent engages with the surroundings by evaluating its condition, executing an action, and obtaining an incentive. Drone sensors and camera systems acquire environmental data for the representation of states. The agent analyzes the recorded information and generates a reaction that dictates the Drones' trajectory or regulates the rotors' thrust, as depicted in Fig. 1.

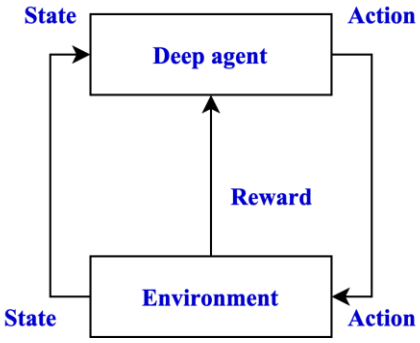


Fig. 1: UVA Training

Related Works

Diverse Deep RL (DRL) [4] techniques and methodologies were employed to address multiple challenges in automated Drone navigation, including drone control, obstacle evasion, route planning, and swarming behavior. The DRL agent functioned as an authority on all these issues, choosing the optimal action that optimizes the reward to attain the intended goal. The specified objective and the employed methodology typically dictate the inputs and results of the DRL algorithms.

The design of an RL agent for Drone navigation, illustrated in Fig. 2, demonstrates several Drone input devices employed to collect the state analyzed by the RL agent. The agent generates action values that represent either the movement data of the Drones or the waypoint data to which the Drones must relocate. Upon executing the action within the surroundings, the agent obtains the subsequent state and the corresponding reward depending on the action [5]. The function of reward is formulated to produce rewards aligned with the intended purpose, utilizing diverse environmental data. The design of the agent (depicted as the 'Agent' boxes) is shaped by reinforcement learning computations, with its elements and internal mechanisms differing across different approaches.

Table 1 encapsulates the utilization of reinforcement learning in various drone navigation goals, whereas the subsequent subsections elaborate on these navigation objectives in greater detail. This table indicates that most studies concentrated on two Drone navigation goals: (1) avoiding obstacles utilizing diverse drone sensor technologies, including webcams, and (2) routing planning to determine the optimal or quickest route.

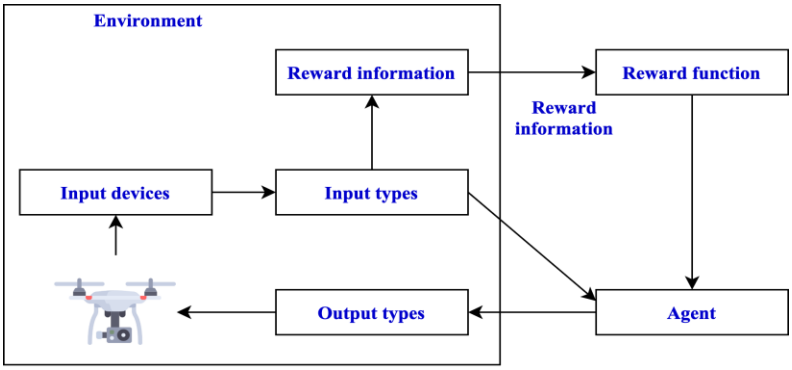


Fig. 2: Drone Navigation

Drones Navigation Systems and Software for Simulation

Drones Navigation Frameworks

The software framework is a conceptual framework similar to a blueprint that directs the understanding and development of software by delineating various functionalities and their interconnections [6]. RL can be regarded as an independent framework by design. The research focused exclusively on drone navigation systems that enhance classical navigation by integrating sensors or camera information. The following sections elaborate on the frameworks in greater detail.

- **Energy-efficient Drones Navigation Systems**

Drones have restricted flying duration and mostly rely on batteries for operation. Designing aircraft routes and recharging stopovers is essential for reaching locations. Energy-aware drone navigational frameworks seek to facilitate avoiding obstacles while accounting for the Drone's battery life.

The study established a structure utilizing the Deep Deterministic Policy Gradient (DDPG) algorithms to direct the Drones to a designated place while maintaining communication with ground stations, enabling them to refuel their batteries when they fall below a certain threshold [7].

Monitoring the power source level, rotor circumstance, and sensor information to strategize the route and implement appropriate adjustments for battery recharging.

- **Structures for Path Planning**

Path planning involves identifying the most efficient trajectory that fulfills the flying aim, such as locating the shortest, quickest, or safest path. Multiple systems have used a modular route planning approach, wherein every component performs a specific function while sharing information with other modules to teach decision-making rules and identify the ideal path.

The study established a four-tier structure wherein each tier produces a collection of objective and constraint variables. The roles are designed to support the lower layer while considering the aims and limits of the topmost layer, with the primary objective of creating trajectories.

Alternative models proposed stage-based learning to select actions based on current environmental interactions at the preferred stage.

The research suggests acquiring a set of motion basics offline and employing them online to formulate rapid maneuvers that facilitate effortless changes in two ways: near-hover movements, which generate motion plans for stable maneuver execution, and swift actions to manage abrupt supplies effectively.

In a collaborative environment, the study proposed a partnership between Unmanned Ground Vehicles (UGVs) [8] and Droness to achieve destination objectives since battery limitations constrain Droness from accessing far areas. At the same time, UGVs are hindered by their inability to ascend to elevated altitudes.

- **Flocking Structures**

Drones flocking systems possess capabilities that extend beyond navigation while preserving flight formation. The research introduced a reinforcement learning-based spatial scheduling method for autonomous Drones. The

technology allows Droness to independently organize their timetables to maximize coverage of the essential pre-scheduled activities in time and space within an established region and timeframe. The study forecasted the trajectories of drones and changing impediments within the airspace to formulate efficient and secure pathways.

- **Vision-Centric Frameworks**

It relies on Drones cameras for navigation, utilizing the images captured to enhance navigational capabilities. Frameworks can be used to improve the agent's CNN design for integrating data from multiple sensors, implement Long-Short Term Memory (LSTM) to preserve navigation decisions, employ Recurrent Neural Networks (RNN) to record Drone states across various time intervals, or prepare pictures of to enrich environmental information.

- **Frameworks for Transfer Learning**

Drones undergo training in target environments before completing flight missions, with training conducted in either virtual or real-world settings. The Drones necessitate retraining when exposed to new surroundings or transitioning from online instruction due to variations in terrain and object configurations or textures. Drone instruction necessitates considerable time and is resource-intensive, whereas actual Droness possesses restricted hardware capabilities. When Drones are deployed in novel surroundings, transfer learning systems diminish training duration by reutilizing neural network weights acquired from prior settings and retraianing only specific components of the agent's neural network.

The study presents an algorithm-hardware co-design in which the Drones learn within a computer simulation. Upon deployment in a real-world setting, the agent retrieves weights from rooted Non-Volatile Memory (NVM) and assesses new behaviors, training exclusively the final layers of the Convolutional Neural Network (CNN) whose weights are retained in on-die Static Random Access Memory (SRAM).

- **Simulative Software**

The research group employed many evaluation methodologies for unmanned drone navigation utilizing reinforcement learning. Simulation software is extensively used instead of actual Droness for evaluation purposes, owing to the high costs associated with the hardware (drone) and the expenses incurred from repairs necessitated by Drones crashes. The objective is not to compare simulation software but to inform the study community of the most frequently utilized evaluation techniques.

RL Agent: denotes the reinforcement learning method employed, encompassing all computations necessary to create rewards, process states, and determine the ideal action. The RL Agent communicates directly with the Drones Flight simulation to transmit and receive Drones operations and states.

Drones Flight Simulation: tasked with simulating drones' motions and interactions within a three-dimensional setting, including capturing pictures via the drone's camera and reporting collisions with different barriers. Exemplary Drones flight simulations include Robotic Operating Systems (ROS) and Windows AirSim.

The 3D Graphics System: offers a 3D graphics environment integrated with a physics engine that simulates gravity and motion like the actual world. Instances of 3D graphics engines include Gazebo and the Unreal Engine.

ROS is utilized alongside Gazebo due to interoperability and support concerns, whereas AirSim employs Unreal Engine for simulation. The three components are not consistently available, mainly if the program used for simulation includes internal components or extensions that deliver the necessary capabilities, such as MATLAB.

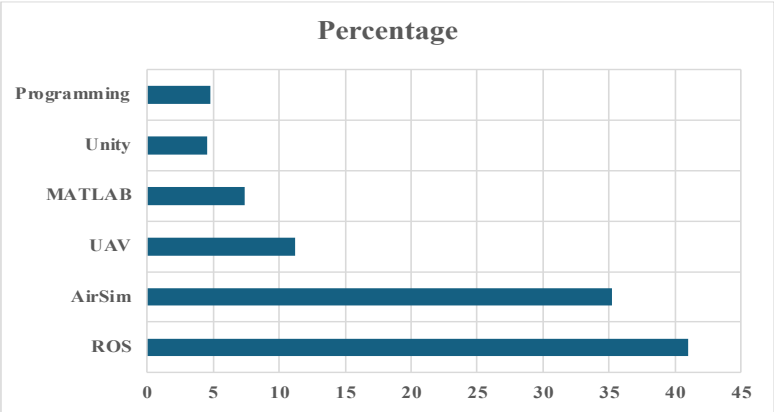


Fig. 3: Drones Software Usage

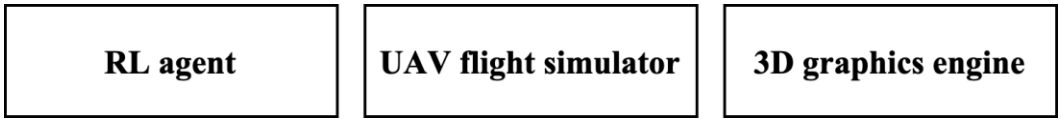


Fig. 4: Drones Simulation Components

Conclusion

Detecting anomalies in surroundings, buildings, and equipment is essential for early problem identification and prevention of escalation. This paper presents an anomaly-detecting drone. A sensor module is installed on a drone. The research devised a DRL-driven navigation system that utilizes data from these devices to direct the drone to the precise location of the issue. An outlier locator is essential for swift identification in safety-critical or dangerous scenarios, enabling individuals to reduce their exposure to danger. The drone can be operated by a human or navigate independently.

Regardless of whether a person or a robot operates it, the mapping algorithm serves as a guide, allowing the pilot to concentrate on safely maneuvering the drone. The system utilizes the Policy Optimization (PPO) DRL framework. It uses progressive curriculum learning to enhance training, with an LSTM recurrence layer that enables the agent to retain its previous locations and retreat when encountering obstacles. DRL methods facilitate experience-driven training for real-world challenges, rendering them optimal for the objective. The research has intentionally designed the algorithm to be versatile, flexible, and capable of functioning in intricate and dynamic contexts. In the future, the research will integrate AI with a wireless sensor package affixed to a drone to roam the Unity 3-D world utilizing actual data from sensors instead of artificially generated information for qualitative evaluation.

References

Chen, H., Wen, Y., Zhu, M., Huang, Y., Xiao, C., Wei, T., & Hahn, A. (2021). From automation system to autonomous system: An architecture perspective. *Journal of Marine Science and Engineering*, *9*(6), 645. https://doi.org/10.3390/jmse9060645

Otto, A., Agatz, N., Campbell, J., Golden, B., & Pesch, E. (2018). Optimization approaches for civil applications of unmanned aerial vehicles (UAVs) or aerial drones: A survey. *Networks*, *72*(4), 411-458. https://doi.org/10.1002/net.21818

Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine*, *34*(6), 26-38. https://doi.org/10.1109/MSP.2017.2743240

Garnier, P., Viquerat, J., Rabault, J., Larcher, A., Kuhnle, A., & Hachem, E. (2021). A review on deep reinforcement learning for fluid mechanics. *Computers & Fluids*, *225*, 104973. https://doi.org/10.1016/j.compfluid.2021.104973

Nguyen, T. T., Nguyen, N. D., & Nahavandi, S. (2020). Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications. *IEEE transactions on cybernetics*, *50*(9), 3826-3839. https://doi.org/10.1109/TCYB.2020.2977374

Sokol, R., Albanese, M., Chew, A., Early, J., Grossman, E., Roll, D., ... & Schuman-Olivier, Z. (2019). Building a Group-Based Opioid Treatment (GBOT) blueprint: a qualitative study delineating GBOT implementation. *Addiction science & clinical practice*, *14*, 1-17. https://doi.org/10.1186/s13722-019-0176-y

Qiu, C., Hu, Y., Chen, Y., & Zeng, B. (2019). Deep deterministic policy gradient (DDPG)-based energy harvesting wireless communications. *IEEE Internet of Things Journal*, *6*(5), 8577-8588. https://doi.org/10.1109/JIOT.2019.2921159

Liu, Q., Li, Z., Yuan, S., Zhu, Y., & Li, X. (2021). Review on vehicle detection technology for unmanned ground vehicles. *Sensors*, *21*(4), 1354. https://doi.org/10.3390/s21041354. https://doi.org/10.3390/s21041354