Exploring the Benefits of Reinforcement Learning for Autonomous Drone Navigation and Control

Ahshanul Haque Department of Computer Science, New Mexico Tech Email: ahshanul.haque@student.nmt.edu Md Naseef-Ur-Rahman Chowdhury Department of Computer Science, New Mexico Tech Email : naseef.chowdhury@student.nmt.edu Mohammad Sahinur Hossen Department of Computer Science, New Mexico Tech Email : mohammad.hossen@student.nmt.edu

-----ABSTRACT------

Drones are now used in a wide range of industries, including delivery services and agriculture. Notwithstanding, controlling robots in powerful conditions can be testing, particularly while performing complex assignments. Conventional strategies for drone mechanization depend on pre-customized directions, restricting their adaptability and versatility. Drones can learn from their interactions with their environment and improve their performance over time with the help of reinforcement learning (RL), which has emerged as a promising method for drone automation in recent years. This paper looks at how RL can be used to automate drones and how it can be used in different industries. In addition, the difficulties of RL-based drone automation and potential directions for future research are discussed in the paper.

Keywords - Reinforcement learning, Drone automation, Machine learning, Navigation, Obstacle avoidance, Object tracking, Safety, Reliability, Training data, Dynamic environments, Decision-making

Date of Submission: May 16, 2023

Date of Acceptance: June 16, 2023

I. INTRODUCTION

Aerial photography and precision agriculture are just two of the many fields where drones are now indispensable tools.

Notwithstanding, controlling robots in powerful conditions can be tested, particularly while performing complex assignments. Pre-programmed instructions are used in traditional drone automation methods, which limit their adaptability [5]. Drones can learn from their interactions with their environment and improve their performance over time with the help of reinforcement learning (RL), which has emerged as a promising method for drone automation in recent years [14].

A type of machine learning called reinforcement learning lets an agent learn by interacting with its surroundings. The agent learns which actions lead to positive outcomes and which lead to negative outcomes by receiving feedback in the form of rewards or punishments based on its actions. With regards to ramble mechanization, RL can be utilized to streamline the robot's way of behaving and dynamic in powerful conditions [7].

In drone automation, RL can be used for object tracking, obstacle avoidance, and navigation. Drones can learn to navigate through complex environments, avoid obstacles, and reach their destinations quickly with the help of RL. Drones can carry out tasks of surveillance and monitoring thanks to the fact that RL can also be used to track moving things like vehicles, animals, or people [3]. There are a few obstacles that must be overcome despite the potential advantages of RL in drone automation. One of the principal challenges is the requirement for a lot of information to prepare RL models. To learn and improve their performance, RL models need a lot of training data, which can be hard to get in the drone automation world. The safety and dependability of RL-based drone automation is another obstacle. Drones should work securely and dependably, especially in conditions with people or different items. Guaranteeing the security and dependability of RL- based drone mechanization requires cautious planning and testing [4].

In general, RL is a revolutionary technology for drone automation. As the interest for drones expansions in different ventures, the expected advantages of RL-based drone mechanization are critical. This paper looks at how RL can be used to automate drones and how it can be used in different industries. Additionally, the paper discusses the difficulties of RL- based drone automation and potential research directions [2][13].

In drone automation, RL has also demonstrated promise in maximizing resource allocation and task scheduling. By considering things like the health of crops, the amount of moisture in the soil, and the weather, RL algorithms can learn to efficiently allocate resources like water or fertilizers in industries like agriculture. Task scheduling for drones can also be aided by RL, which can determine the best order and timing of tasks based on priorities, deadlines, and resources available. In sectors that use drones for multiple tasks, this may result in increased productivity and cost savings. The interpretability of the learned policies is yet another significant obstacle in the way of RL-based drone automation. It is challenging to comprehend the decision- making process and the reasoning behind the drone's actions because RL models frequently function as black boxes. To address this problem, researchers are looking into RL methods and techniques that can be understood by humans and verify the behavior of the drone. Stakeholders can gain insight into the drone's decision-making, increase trust, and facilitate regulatory compliance by displaying or explaining they learned policies.

Integration of multiple drones into collaborative tasks presents a new research challenge as the field of RL-based drone automation develops. Addressing issues like communication, collaboration, and task allocation are necessary for coordinating the actions of multiple drones to achieve collective objectives. Intelligent algorithms that enable drones to work together effectively, whether in synchronized aerial formations, collaborative surveillance, or search-and-rescue missions, can benefit from RL's involvement [29].

In conclusion, RL has enormous potential to transform the automation of drones in a variety of industries. Drones' adaptability and capabilities in dynamic environments can be improved by using RL for everything from object tracking and navigation to resource allocation and task scheduling. To fully reap the benefits of RL-based drone automation, issues with data acquisition, safety, reliability, interpretability, and multi-drone coordination must be resolved. The widespread adoption of RL in the automation of drones in the future will be made possible by ongoing research and development in these areas.

II. REINFORCEMENT LEARNING

A type of machine learning called reinforcement learning lets an agent learn by interacting with its surroundings. The agent learns which actions lead to positive outcomes and which lead to negative outcomes by receiving feedback in the form of rewards or punishments based on its actions. The objective of the specialist is to amplify its combined prize over the long run. With regards to ramble mechanization, RL can be utilized to streamline the robot's way of behaving and dynamic in powerful conditions [6].

Some key terms that describe the basic elements of an RL problem is [8]:

Environment — Physical world in which the agent operates State — Current situation of the agent Reward — Feedback from the environment Policy — Method to map agent's state to actions Value — Future reward that an agent would receive by taking an action in a particular state



Figure 1. Figure showing how RL works

Games provide an excellent framework for understanding reinforcement learning (RL). Consider the classic game Pac-Man, where the objective of the agent (Pac-Man) [1] is to consume all the food in the grid while avoiding the ghosts that roam the board. The interactive environment in which Pac-Man acts is the grid world. The agent receives rewards for consuming food, and punishments for being caught by the ghosts, which results in losing the game. The states in this scenarios correspond to Pac-Man's location in the grid world, while the total cumulative reward is the agent's ultimate success in winning the game [10], [11]. The dilemma known as the Exploration vs. Exploitation trade-off requires the agent to strike a balance between exploring new states and maximizing its overall reward when



Figure 2. RL explained with the simple game Pac-Man [1]

developing an optimal policy. The agent may have to make short-term sacrifices in the interest of long-term gains in order to strike this balance. Therefore, the agent must gather sufficient information to make the best decisions in the future.[12]. Markov Decision Processes (MDPs) are a mathematical framework for describing the RL environment, and MDPs can be used to formulate all RL problems. An MDP consists of a transition model (P (s', s - a)), a set of finite environment states (S), a set of actions (A(s)) in each state, and a real-valued reward function (R(s). Model-free RL methods, on the other hand, are a useful alternative because real-world environments frequently lack prior knowledge of environmental dynamics.[4]. A model-free method known as Q-learning

can be used to create an autonomous Pac-Man agent. Q values, which represent the value of carrying out an action (a) in a particular state (s), are updated as part of the strategy.

III. APPLICATIONS OF RL IN DRONE AUTOMATION

In drone automation, RL can be used for object tracking, obstacle avoidance, and navigation. Drones can learn to navigate through complex environments, avoid obstacles, and reach their destinations quickly with the help of RL. For instance, RL can be used to teach drones how to fly without hitting other objects in crowded areas like cities. Drones are able to carry out tasks of surveillance and monitoring thanks to the fact that RL can also be used to track moving things like vehicles, animals, or people [12]. RL can likewise empower robots to figure out how to perform errands that require complex navigation, for example, bundle conveyance or harvest observing. To ensure that packages are delivered effectively, RL can instruct drones to optimize their routes and delivery schedules. Drones can also learn to monitor crop health with RL, identifying areas that need attention and optimizing resource use, such as fertilizer and water [5]. In reinforcement learning (RL), an agent learns to make decisions by interacting with its environment and receiving feedback in the form of rewards or punishments. Various tasks, including drone control, have been successfully im-lamented using RL.

RL can be used to train an agent to navigate a drone through a complex environment, avoid obstacles, and carry out tasks like inspection or delivery in the context of drone control. The specialist gets tangible contributions from the robot's sensors, like cameras or lidars, and makes moves, for example, changing the robot's speed or heading, to accomplish an objective. Using a deep neural network as the agent, which receives sensory input and produces a set of actions, is one way to use RL for drone control. RL algorithms like Q-learning or policy gradients are used to train the neural network, adjusting the parameters of the network to maximize the expected cumulative reward [15]. The need for extensive training data, which can be costly and time-consuming to collect, is one of the challenges associated with using RL for drone control. One method for relieving this challenge is to utilize recreation conditions, which permit the specialist to prepare in a mimicked climate that intently looks like this present reality climate. RL has numerous practical applications in the automation of drones, including object tracking, obstacle avoidance, and navigation. Drones can learn to master complex environments, deftly avoid obstacles, and arrive at their destinations as quickly as possible by using RL. For example, RL can empower robots to independently fly through thickly populated regions, like urban areas, without crashing into different articles.

Additionally, RL techniques make it easier to track objects, making it possible for drones to effectively carry out surveillance and monitoring tasks, such as tracking individuals, animals, or moving vehicles. Additionally, RL gives drones the ability to carry out tasks like crop monitoring or package delivery that require complex decision-making procedures. RL-trained drones can optimize their routes and delivery schedules to ensure that packages are delivered promptly and effectively [28]. Additionally, RL algorithms make it possible for drones to monitor and evaluate crop health, pointing out areas that require attention and optimizing the use of resources like fertilizer and water. Drones can improve their decisionmaking capabilities by utilizing RL, making them valuable assets in a variety of industries. RL makes use of an agent that interacts with the environment and receives feedback in the form of rewards or penalties when applied to drone control. The agent, which is typically implemented using deep neural networks, processes the sensory information provided by the lidars or cameras on the drone and takes the necessary actions to accomplish goals, such as inspection or delivery. The neural network is trained using RL algorithms like Q-learning or policy gradients by changing its parameters to get the most out of the expected cumulative reward. To teach drones to navigate them autonomously and intelligently surroundings, this method makes use of the power of RL.

The collection of extensive training data, which can be costly and time-consuming, is one significant obstacle when using RL for drone control. A viable solution to this problem is the use of simulation environments, which allow agents to train in virtual environments that closely resemble actual scenarios. Training data can be generated more quickly and efficiently using simulations, thereby accelerating the learning process and decreasing the need for physically collected data. In conclusion, RL has tremendous potential for drone control and is anticipated to become increasingly important in the development of autonomous drone systems. Its applications envelop route, hindrance aversion, object following, as well as dynamic errands like bundle conveyance and harvest observing. Although difficulties exist, for example, information obtaining and reenactment loyalty, the joining of RL strategies carries us nearer to the advancement of exceptionally skilled and smart independent robot frameworks. In general, RL has shown extraordinary potential for drone control, and assuming an undeniably significant part in the improvement of independent robot systems is normal.

IV. COMPARISON RL OVER TRADITIONAL ML ALGORITHMS

The way that traditional machine learning (ML) and reinforcement learning (RL) algorithms learn from data is different. The choice between RL and conventional ML algorithms for drone automation will be determined by the application and the type of data available. A labeled dataset is used to train traditional machine learning algorithms, and the input data is mapped to a set of predefined output categories. When the drone's camera captures visual data that can be labeled and used to train a classifier, this method is typically applied to tasks like image classification or object detection. RL, on the other hand, is used when the drone needs to learn to make decisions based on sensory input and environmental feedback. The robot's activities are not foreordained, and the RL calculation figures out how to choose activities that boosts the aggregate prize over the long haul. RL's ability to deal with situations where the best course of action is unclear or dependent on the context is one of its over conventional machine advantages learning algorithms. For instance, in a drone delivery scenario, the location of obstacles, weather, and other factors may alter the optimal path for the drone. Through trial and error, the drone can adjust to these shifting conditions thanks to RL. However, RL implementation can be more difficult and requires more data than conventional ML algorithms. RL calculations depend on experimentation to realize, which can be tedious and costly in a certifiable setting. To prevent unintended behavior, the rewards or penalties used to train the RL algorithm must also be carefully designed.

In conclusion, when it comes to automating drones, RL and conventional ML algorithms have distinct advantages and disadvantages. Although it requires more data and can be more difficult to implement, RL is ideal for situations in which the drone must learn to make decisions based on sensory input and environmental feedback. When the input data can be labeled and used to train a classifier; traditional ML algorithms are better suited for applications like object detection and image classification. When it comes to automating drones, RL and traditional ML algorithms have distinct advantages and disadvantages in addition to their divergent approaches to learning from data[27]. The decision between the two relies upon the particular application and the idea of accessible information. When trained on labeled datasets, where input data is mapped to predetermined output categories, conventional machine learning algorithms perform admirably. These calculations are normally utilized in errands, for example, picture characterization and article identification, utilizing visual information caught by the robot's camera that can be marked and used to really prepare classifiers.

On the other hand, scenarios in which drones must learn to make decisions based on feedback from their environment. and sensory input are ideal for RL.

The RL algorithm, which learns to select activities that maximize cumulative rewards over time, determines the drone's actions in contrast to conventional ML algorithms. In situations where the best course of action may vary depending on contextual factors like the locations of obstacles, the weather, and other variables, this flexibility is advantageous. The drone is given the ability to adjust to changing conditions and discover the best route through trial and error thanks to RL. Implementing RL can be difficult and require more data than traditional ML algorithms, despite its adaptability. When applied in realworld settings, RL relies on trial-and-error learning, which can be time-consuming and costly. In addition, in order to guarantee effective learning and prevent unintentional behavior, the rewards or penalties used to train the RL algorithm need to be carefully designed.

In conclusion, when it comes to automating drones, RL and conventional ML algorithms each have distinct advantages and disadvantages. Although it requires more data and can be more difficult to implement, RL shines in situations.

where drones need to learn decision-making based on sensory input and environmental feedback. Using labeled input data to effectively train classifiers, traditional ML algorithms, on the other hand, are better suited for tasks like image. classification and object detection.

V. CONCLUSION

In conclusion, the application of reinforcement learning (RL) to the automation of drones has the potential to revolutionize the industry by giving drones the ability to learn from their interactions with their surroundings and continuously enhance their performance. By utilizing RL, robots can embrace complex undertakings in powerful and consistently evolving conditions, including errands like the route through hindrances and following objects of interest. However, for RL-based drone automation to reach its full potential, several obstacles must be overcome despite the promising benefits. The need for a large amount of data to effectively train RL models is one of the main obstacles in implementing RL-based drone automation. Particularly in the context of drone operations, it can be difficult. to gather sufficient data to facilitate learning. Moreover, guaranteeing the security and unwavering quality of RLbased drone computerization represents a critical concern. Drones are supposed to work in conditions that might include the human presence or the presence of different articles, requiring cautious planning, thorough testing, and powerful defenses to relieve possible dangers. Looking forward, future examination endeavors ought to zero in on the advancement of more productive RL calculations customized explicitly for drone robotization. Drone systems' overall performance and capabilities could be improved by combining RL with other machine-learning techniques. In addition, RL-based drone automation's safety and dependability issues will necessitate ongoing advancements in technology, design principles, and regulatory frameworks. All things considered, RL-based drone computerization addresses a significant and game changing mechanical progression. As the interest in drones keeps on developing across different ventures, the possible advantages of coordinating RL into drone frameworks are significant. Drones can continuously improve their performance by being outfitted with the adaptability and learning capabilities that are required to navigate environments that are both complex and dynamic. RLbased drone automation is poised to revolutionize the

industry by providing unprecedented opportunities for increased efficiency, productivity, and innovation across a wide range of applications and industries with continued development and refinement.

References

- [1] https://en.wikipedia.org/wiki/Pac-Man
- [2] Y. Song, M. Steinweg, E. Kaufmann and D. Scaramuzza, "Autonomous Drone Racing with Deep Reinforcement Learning," 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Prague, Czech Republic, 2021,pp.1205-1212,doi: 10.1109/IROS51168.2021.9636053.
- [3] A. Devo, J. Mao, G. Costante and G. Loianno, "Autonomous Single-Image Drone Exploration With Deep Reinforcement Learning and Mixed Reality," in IEEE Robotics and Automation Letters, vol. 7, no. 2, pp. 5031-5038, April 2022, doi: 10.1109/LRA.2022.3154019.
- [4] D. Wang, T. Fan, T. Han and J. Pan, "A Two-Stage Reinforcement Learning Approach for Multi-UAV Collision Avoidance Under Imperfect Sensing," in IEEE Robotics and Automation Letters, vol. 5, no. 2, pp. 3098-3105, April 2020, doi: 10.1109/LRA.2020.2974648.
- [5] V. N. Nguyen, R. Jenssen and D. Roverso, "Intelligent Monitoring and Inspection of Power Line Components Powered by UAVs and Deep Learning," in IEEE Power and Energy Technology Systems Journal, vol. 6, no. 1, pp. 11-21, March 2019, doi: 10.1109/JPETS.2018.2881429.
- [6] W. Wu, M. A. Qurishee, J. Owino, I. Fomunung, M. Onyango and B. Atolagbe, "Coupling Deep Learning and UAV for Infrastructure Condition Assessment Automation," 2018 IEEE International Smart Cities Conference (ISC2), Kansas City, MO, USA, 2018, pp. 1-7, doi: 10.1109/ISC2.2018.8656971.
- [7] N. Imanberdiyev, C. Fu, E. Kayacan and I. -M. Chen, "Autonomous navigation of UAV by using real-time model-based reinforcement learning," 2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV), Phuket, Thailand, 2016, pp. 1-6, doi: 10.1109/ICARCV.2016.7838739.
- [8] C. Wu et al., "UAV Autonomous Target Search Based on Deep Reinforcement Learning in Complex Disaster Scene," in IEEE Access, vol. 7, pp. 117227-117245, 2019, doi: 10.1109/ACCESS.2019.2933002.
- [9] F. Fei, Z. Tu, D. Xu and X. Deng, "Learn-to-Recover: Retrofitting UAVs with Reinforcement Learning-Assisted Flight Control Under Cyber-Physical Attacks," 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 2020, pp. 7358-7364, doi: 10.1109/ICRA40945.2020.9196611.
- [10] A. Gumaei et al., "Deep Learning and Blockchain with Edge Computing for 5G-Enabled Drone Identification and Flight Mode Detection," in IEEE Network, vol. 35, no. 1, pp. 94-100, January/February 2021, doi: 10.1109/MNET.011.2000204.

- [11]E. Bøhn, E. M. Coates, S. Moe and T. A. Johansen, "Deep Reinforcement Learning Attitude Control of Fixed-Wing UAVs Using Proximal Policy optimization," 2019 International Conference on Unmanned Aircraft Systems (ICUAS), Atlanta, GA, USA, 2019, pp. 523-533, doi: 10.1109/ICUAS.2019.8798254.
- [12] H. X. Pham, H. M. La, D. Feil-Seifer and L. Van Nguyen, "Reinforcement Learning for Autonomous UAV Navigation Using Function Approximation," 2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), Philadelphia, PA, USA, 2018, pp. 1-6, doi: 10.1109/SSRR.2018.8468611.
- [13] H. Lu, Y. Li, S. Mu, D. Wang, H. Kim and S. Serikawa, "Motor Anomaly Detection for Unmanned Aerial Vehicles Using Reinforcement Learning," in IEEE Internet of Things Journal, vol. 5, no. 4, pp. 2315-2322, Aug. 2018, doi: 10.1109/JIOT.2017.2737479.
- [14] C. Wang, J. Wang, J. Wang and X. Zhang, "Deep-Reinforcement-Learning-Based Autonomous UAV Navigation With Sparse Rewards," in IEEE Internet of Things Journal, vol. 7, no. 7, pp. 6180-6190, July 2020, doi: 10.1109/JIOT.2020.2973193.
- [15] Subbarayalu, Venkatraman, and Maria Anu Vensuslaus. "An Intrusion Detection System for Drone Swarming Utilizing Timed Probabilistic Automata." Drones 7, no. 4 (2023): 248.
- [16] Moon, S. Papaioannou, C. Laoudias, P. Kolios and S. Kim, "Deep Reinforcement Learning Multi-UAV Trajectory Control for Target Tracking," in IEEE Internet of Things Journal, vol. 8, no. 20, pp. 15441-15455, 15 Oct.15, 2021, doi: 10.1109/JIOT.2021.3073973.
- [17] K. Li, W. Ni and F. Dressler, "LSTM-Characterized Deep Reinforcement Learning for Continuous Flight Control and Resource Allocation in UAV-Assisted Sensor Network," in IEEE Internet of Things Journal, vol. 9, no. 6, pp. 4179-4189, 15 March15, 2022, doi: 10.1109/JIOT.2021.3102831.
- [18] K. Li, W. Ni, E. Tovar and M. Guizani, "Joint Flight Cruise Control and Data Collection in UAV-Aided Internet of Things: An Onboard Deep Reinforcement Learning Approach," in IEEE Internet of Things Journal, vol. 8, no. 12, pp. 9787-9799, 15 June15, 2021, doi: 10.1109/JIOT.2020.3019186.
- [19] M. Z. Anwar, Z. Kaleem and A. Jamalipour, "Machine Learning Inspired Sound-Based Amateur Drone Detection for Public Safety Applications," in IEEE Transactions on Vehicular Technology, vol. 68, no. 3, pp. 2526-2534, March 2019, doi: 10.1109/TVT.2019.2893615.
- [20] D. Chen, Q. Qi, Z. Zhuang, J. Wang, J. Liao and Z. Han, "Mean Field Deep Reinforcement Learning for Fair and Efficient UAV Control," in IEEE Internet of Things Journal, vol. 8, no. 2, pp. 813-828, 15 Jan.15, 2021, doi: 10.1109/JIOT.2020.3008299.
- [21] P. Luong, F. Gagnon, L. -N. Tran and F. Labeau, "Deep Reinforcement Learning-Based Resource

Allocation in Cooperative UAV-Assisted Wireless Networks," in IEEE Transactions on Wireless Communications, vol. 20, no. 11, pp. 7610-7625, Nov. 2021, doi: 10.1109/TWC.2021.3086503.

- [22]L. Li, Q. Cheng, K. Xue, C. Yang and Z. Han, "Downlink Transmit Power Control in Ultra-Dense UAV Network Based on Mean Field Game and Deep Reinforcement Learning," in IEEE Transactions on Vehicular Technology, vol. 69, no. 12, pp. 15594-15605, Dec. 2020, doi: 10.1109/TVT.2020.3043851.
- [23] M. Ezuma, F. Erden, C. K. Anjinappa, O. Ozdemir and I. Guvenc, "Micro-UAV Detection and Classification from RF Fingerprints Using Machine Learning Techniques," 2019 IEEE Aerospace Conference, Big Sky, MT, USA, 2019, pp. 1-13, doi: 10.1109/AERO.2019.8741970.
- [24] J. Yao and N. Ansari, "Wireless Power and Energy Harvesting Control in IoD by Deep Reinforcement Learning," in IEEE Transactions on Green Communications and Networking, vol. 5, no. 2, pp. 980-989, June 2021, doi: 10.1109/TGCN.2021.3049500.
- [25] H. Bou-Ammar, H. Voos and W. Ertel, "Controller design for quadrotor UAVs using reinforcement learning," 2010 IEEE International Conference on Control Applications, Yokohama, Japan, 2010, pp. 2130-2135, doi: 10.1109/CCA.2010.5611206.
- [26] H. Peng and X. Shen, "Multi-Agent Reinforcement Learning Based Resource Management in MEC- and UAV-Assisted Vehicular Networks," in IEEE Journal on Selected Areas in Communications, vol. 39, no. 1, pp. 131-141, Jan. 2021, doi: 10.1109/JSAC.2020.3036962.
- [27] Hossen, Mohammad Sahinur; Islam, Rakibul; Chowdhury, M.NUR.; Haque, Ahshanul; Alahy, Q.E. (2023). Malware Detection in Web Browser Plugins Using API Calls with Permissions, International Journal of Advanced Networking and Applications -IJANA, DOI: 10.35444/IJANA.2023.14603
- [28] Islam, Jahirul; Hasan, Mahadi; Hasan, Md Maruf (2023): Securing the Edge: A Comprehensive Review to Protecting Wireless Sensor Networks and Android Devices from Cyber Attacks. TechRxiv. Preprint.

https://doi.org/10.36227/techrxiv.22776035.v1

[29] Rani, Sangeeta & Dhindsa, Kanwalvir. (2018). Android Malware Detection in Official and Third-Party Application Stores. Int. J. Advanced Networking and Applications. 9. 3506-3509.