Review of Ai-Based Uav Navigation in Gps-Denied Environments

Chenxin Yang[1*]

[1]Department of Mechanical and Aerospace Engineering, Monash University, Melbourne, 3800, Australia

*cyan0119@student.monash.edu

Abstract. Unmanned Aerial Vehicles (UAVs) have increasingly become vital across a wide range of industries in the past decades. However, the conventional GPS-dependent navigation systems face critical limitations in these applications, especially in GPS-denied environments. Recent advances in Artificial Intelligence (AI) offer promising solutions for autonomous navigation without GPS reliance, such as reinforcement learning (RL), deep learning (DL), and deep reinforcement learning (DRL). This study presents a comprehensive review of AI-based navigation systems in these fields, emphasizing trajectory planning, perception, and localization. The main UAV autonomous applications include Search and Rescue (SaR), surveillance, and tracking. AI plays a crucial role in these applications for modeling humans to perform fundamental interactions and decision-making processes. RL methods enable self-learning of navigation policies through environment interaction, while DRL enhances this capability by integrating deep neural networks to manage complex, high-dimensional sensor data. DL approach, notably Convolutional Neural Networks (CNNs) and their derivatives like YOLO and DCNN, further empower UAVs with robust real-time vision-based navigation. Finally, challenges and future research directions are outlined, including lightweight algorithm exploitation, hybrid onboard-cloud computation, and energy-efficient autonomous recharging strategies, aiming to bridge the gap between simulation and real-world deployment for AI-based UAV navigation systems.

Keywords: Unmanned Aerial Vehicles, Navigation, Reinforcement Learning, Deep Learning, Deep Reinforcement Learning

1 Introduction

Unmanned Aerial Vehicles (UAVs), also known as drones, have become an essential and powerful technology in a wide variety of applications across civilian and military fields in recent years. UAVs are capable of tackling certain tasks, carrying payloads, and monitoring a field with significant advantages in accessibility, flexibility, and cost-effectiveness. The main tasks of UAVs in modern society and industries can be classified into four main categories: sensing and inspection, take-off and landing,

A. J. Moshayedi (ed.), *Proceedings of the 2025 2nd International Conference on Electrical Engineering and Intelligent Control (EEIC 2025)*, Advances in Engineering Research 279,
https://doi.org/10.2991/978-94-6463-864-6_51

monitoring and tracking, and search and rescue (SaR). Unmanned Aerial System (UAS), which is the core of a UAV, consists of flight control systems, navigation systems, communication systems, monitoring and sensing. The navigation system is a critical part that plans the trajectory for the UAV to guide the direction and orientation. The capability of autonomous and reliable flight in complex indoor and outdoor environments lies in the effectiveness and efficiency of navigation systems.

The traditional approach of navigation systems relies heavily on the Global Positioning System (GPS), which provides geographic position and time information based on satellites. Although GPS is powerful and advantageous in trajectory planning and obstacle avoidance, it is inherently limited by its dependence on clear line-of-sight communication with satellites and real-time localization. In outdoor environments such as dense urban areas, forests and tunnels, GPS becomes unreliable [1]. The precise localization and motion planning can be tough in indoor missions, especially navigation in 3D multi-story buildings such as SaR in skyscrapers, factories, surveillance, and equipment inspections industries. Meanwhile, the GPS signals could be deliberately jammed or spoofed in military operations. Thus, with the increasing use of UAVs in recent years, the vulnerability of GPS-dependent navigation has enabled a large number of studies on effective and efficient GPS-independent solutions.

As the emergent evolution in Artificial Intelligence (AI), the technology has overturned the traditional approaches in UAVs traditional navigation systems, such as optimization-based approaches and PID controllers. AI has emerged as an effective solution to address the issues in GPS-denied locations by enabling UAVs to localize, perceive, and plan trajectories based on on-board sensors and pre-trained models which involves three main types of methods, reinforcement learning (RL), deep reinforcement learning (DRL), and deep learning (DL). AI-based navigation systems are able to interpret sensor data, predict environmental dynamics, and optimize the decision-making process in real time.

One of the most prominent branches of Machine Learning (ML) in AI is reinforcement learning. RL-based approach UAVs with the RL approach have the capability of learning the optimization of trajectories by the policies and rewards through interaction with their environment. RL modelling the navigation tasks as a Markov Decision Process (MDP), a mathematical framework for maximizing long-term rewards while making decisions in complex and dynamic environments. The self-learning capability of RL enables UAVs to adapt to unknown environments without extensive maps and instructions. In general, RL techniques in UAV navigation applications involve Q-learning, double Q-learning, and multi-agent Q-learning for swarms. A large amount of research has applied these methods to optimize trajectory, tracking targets, avoid obstacles, and improve energy efficiency, especially in scenarios where conventional methods are infeasible due to high-dimensional state spaces and non-linear dynamics.

A more advanced approach, deep reinforcement learning, enhances the abilities of RL by incorporating deep neural networks (DNNs) into the learning processes. DRL is particularly suitable for vision-based navigation due to its capability of handling high-dimensional inputs such as images and LiDAR data. Common approaches for vision-based navigation include Deep Q-Network (DQN), Double Deep Q-Network (DDQN), Proximal Policy Optimization (PPO), Actor-Critic, and an advanced version of Actor-Critic, Asynchronous Advantage Actor Critic (A3C). These approaches have demonstrated remarkable milestones in navigating UAVs in real time, which will be discussed in detail in the following sections, with the relevant studies in these fields. Furthermore, an extension of MDP, Partially Observable MDP, is assessed and evaluated in this study to provide a comprehensive review under the scope of the DRL approach. It accounts for uncertainties utilizing an observation space, with state space and action space, to yield a more precise and optimal decision in unknown circumstances than the MDP method [2].

Another specific class of machine learning is deep learning, which plays a pivotal role in perception and decision-making processes. The deep learning approach is widely applied in vision-based UAS involving CNN and DNN for tasks such as object detection, image classification, and semantic segmentation. UAVs rely solely on onboard cameras and visual information for localization and navigation, including the sensor data from RGB cameras, infrared cameras, LiDAR sensors, and multispectral cameras. DL approaches provide a structure map of the surroundings and estimate the positions of the UAV in real time, integrating with a localization algorithm such as simultaneous localization and mapping (SLAM). Algorithms such as YOLO (You Only Look Once) based on CNNs provide real-time object detection for obstacle avoidance and target tracking in GPS-denied settings. A modification of CNN, DCNN-based approaches such as AlexNet and ResNet have also gained interest in vision-based navigation systems. Moreover, a genetic algorithm (GA) is incorporated to replace manual tuning of hyperparameters with the DCNN in [3].

The incorporation of different AI approaches has illustrated a significant value, for example, a combination of reinforcement learning and visual perception enables dynamic destination path planning. In the context of UAV swarms, decentralized training of navigation policies allows each UAV to learn from its own experiences while contributing to a collective model. Meanwhile, multi-agent Q-learning trained UAVs not only interact with the environment but also avoid collision with other UAVs.

Although promising results have been achieved in the past decades, AI-based navigation systems still face several challenges involving limited computation power on board, high computational power required for training the model, energy consumption, safety verification, and fault handling. In addition, the AI-based navigation requires proof of robustness, interpretability, and generalization for its deployment in safety-critical missions and practicality in a wide range of applications. Future research directions will likely focus on simplification and generalization of algorithms, transfer learning, adaptive control strategies, reducing hardware cost, and optimizing energy efficiency.

2 Machine learning approach

2.1 Reinforcement Learning (RL)

Reinforcement learning, a machine learning framework, provides a self-correcting method for learning and adapting to varying conditions. RL allows UAVs to make decisions by interacting with an environment to achieve the objective, which models the decision-making process based on the Markov Decision Process (MDP). The agent performs certain actions at each time step by assessing the current state using a value function and determining the corresponding action. Next, the agent obtains feedback in the form of rewards and progresses to the next state simultaneously, as shown in Figure 1. Reinforcement learning aims to develop a predefined optimal policy that guides the agent in selecting the most appropriate action for different states to maximize the predicted reward [4].

The policy of an agent's actions is defined by the probability of behavior knowing the current state. The assessment of actions or behaviors has two methods to evaluate: a state-value function to estimate the next-state performance after the action or an action-value function to evaluate the benefit of taking a specific action while in a given state. The latter function values are often referred to as a Q-value and expressed as the summation of the rewards [5]. A Bellman equation, also known as the Q-function (1), calculates the next state action Q-values by introducing a discounted factor δ and a weighting factor α. The agent uses the highest Q values to perform the corresponding action and transfer to the next stage.

$$Q(s_t, a_t) = (1-\alpha)Q(s_t, a_t) + \alpha(R_{t+1} + \delta \max Q(s_{t+1}, a_t)) \tag{1}$$

The reinforcement learning plays an important role in an autonomous system since it has the capability of self-correcting and self-learning. The dynamic target is also possible while implementing RL. As the UAV nears its designated target, the environment grants it progressively larger rewards. This reward structure motivates the UAV to move closer and closer to the target, directing its behavior toward completing the mission effectively and successfully.

The learning targets have two categories, one of which is to directly learn from the control signals. The PID controller parameters for UAV attitude can be optimized by reinforcement learning, which is proposed in [6] to navigate the UAV in an unknown environment by both simulation and real-world implementation. The other approach to control the target learning for the agent is by selecting a specific direction to move, which will be converted to the control signal by the internal controller. Fixed Sparse Representation (FSR) is integrated in RL, a function approximation-based reinforcement learning algorithm designed to handle high-dimensional state spaces and achieve faster convergence in simulation and the real-world environment [7]. In addition, the received signal strength (RSS) method is used for search and rescue (SaR) missions in GPS-denied indoor environments [8]. RSS provides a function of sources and UAV positions as inputs for the Q-learning algorithm. The policy and states are defined by RSS compared to conventional Q-learning methods relying on the GPS, enabling the deployment of highly accurate indoor SaR missions.

Path planning is one of the most critical procedures in UAV navigation. A double Q-learning approach is used to solve the trajectory planning problem in UAV functions as a flying base station [9]. The reward is based on whether the data requested is fulfilled within a specified maximum waiting time. Double Q-learning improves on single Q-learning by using two separate value estimates to reduce overestimation bias. This method separates the selection of the best action and evaluating the best action algorithms. The quality of experience (QoE) for video streaming users utilizing UAVs as mobile Base Stations is improved by utilizing Q-learning for path planning in [10]. This algorithm incorporates a reward function based on a key QoE metric, also referred to as video segment delay, and recharging stations are identified for improved user experiences.

Multi-agent Q learning extends the traditional Q-learning method to train multiple agents where each agents learn from its own policy, environment, and other agents. This algorithm has acquired increasing research for flocking navigation of UAV swarms. A multi-agent Q-learning algorithm is implemented in [11]. The multiple agents received UAV data at each time step, such as position, angle, and speed, while figuring out the next move based on optimal control policy. A two-stage approach is proposed to combine a K-means algorithm for cell partition with a Q-learning-based algorithm where each UAV independently learns its optimal position in 3D space. In addition, a three-step multi-agent RL framework is proposed in [12] to optimize UAV trajectory and power control through leveraging user mobility predictions with proven convergence under mild conditions. The authors implement a decentralized dynamic trajectory planning for navigation in [13] to transfer real-time sensing data to base stations over cellular networks. The authors introduced the sense-and-send protocol and data transmission by Markov chains to deal with the sensing and transmission issues. Simulations of enhanced algorithms prove faster convergence and higher utilities in real-world applications. Lastly, a temporal-difference approach is used to allow a cellular-connected UAV to learn from the state value function of MDP in order to minimize mission duration while maintaining connectivity. The algorithm is enhanced by incorporating linear function approximation in high-dimensional state space, requiring only raw measurements or simulated data, with numerical results indicating a good network connectivity in dense urban regions [14].

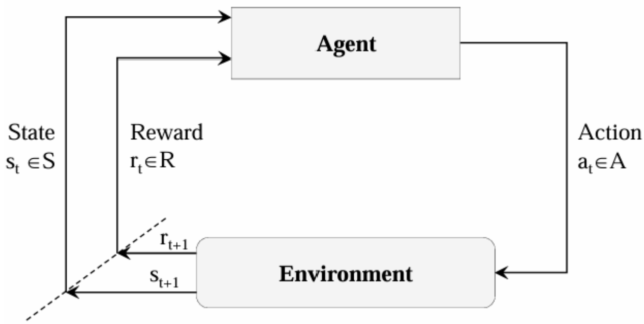

Fig. 1. Reinforcement learning process [15]

2.2 Deep Learning (DL)

DL is mainly applicable for vision-based navigation, and collected data is processed using neural networks (NNs), such as CNN, DNN, fully connected NN (FNN), and deep CNN (DCNN) algorithms. The larger datasets and deeper NN can enhance performance with increasing training time. The parallel processing and CUDA architecture have further improved the computational power in this area. The model is generally pre-trained in a virtual or laboratory environment and adapts the algorithm for new tasks in a similar target environment. The NN can be categorized into three main architectures involving Deep Neural Network (DNN), Recurrent Neural Networks (RNNs), and Convolutional Neural Networks (CNNs) [16]. DNNs, which act as universal function approximators, connect all layers with feedforward neural networks. The second approach allows the network to retain memory of previous inputs through sequential data. An improvement is proposed later called "Long Short-Term Memory (LSTM) networks" to handle the challenges of long-range issues. Finally, CNNs are one of the dominant algorithms for image processing, including convolutional layers for extracting features, pooling layers for simplifying information, and fully connected layers for decision making.

CNN is able to extract hierarchical features from images, making it applicable for UAV navigation, especially in GPS-denied indoor environments. DL is mostly studied using CNN for tasks like obstacle detection and classification, and trajectory planning. The authors proposed a vision-based trail-following method for UAVs utilizing CNNs with integrated obstacle avoidance based on optical flow [17]. The UAV learned from DL to stay centered on a bike trail and maintain stability in case of perturbations. Both simulations and real-world flight tests were implemented, and the performance was evaluated in terms of accuracy, processing time, and GPU usage, confirming its efficiency and robustness. An integration of CNN and Simultaneous localization and mapping (SLAM) methods is used to extract the height of buildings and fuse with the SLAM-generated key points [18]. The data is fed to a model enabling decision-making in UAV directions and paths, indicating the feasibility and efficient operations in complex indoor and outdoor environments. An in-house CNN is implemented to predict forest trail directions using a single monocular image [19]. However, the real-world experiments fell short of simulated performance due to differences in image quality.

Furthermore, YOLO is a real-time object detection model based on CNN that has gained significant popularity, especially in vehicles and UAVs navigation systems. YOLO processes the image in feedforward neural networks, different from conventional object detection approaches, which require multiple passes for one image. The systematic review studied 173 studies from 2019-2024 on deep learning-based computer vision for autonomous UAV navigation, indicating a 39.5% of studies adopting the YOLO model with considerable growth [20]. An innovative DL approach using visual monitoring by integrating YOLO algorithm with navigation data is proposed in [21]. This approach allows UAVs to detect and track targets with a flight test of multirotor UAV pairs, demonstrating the method's high accuracy and robustness. In addition, one of the most practical applications of vision-based UAVs is sensing and inspection, which involves structural health monitoring (SHM), industrial inspections, and wildfire monitoring. YOLOv3 is employed to perform obstacle detection integrated with a fiducial marker-based localization system to navigate UAVs in GPS-denied locations [22]. A real-time damage segmentation method is used for SHM in an indoor or outdoor environment. Moreover, the vision-based UAVs with YOLO method are implemented in SaR operations, such as YOLO v3 with RGB optical sensor approach in [23,24]. Two studies aim to navigate UAVs in GPS-denied locations, where the former study introduced a lightweight and real-time path planning based on YOLO to detect objects. The method was tested under a simulation environment and a constrained real-world environment. However, the result of the real flight test is not as promising as the simulations due to improper localization provided by the stereo camera. The latter study focuses on a navigation system for small coaxial twin-rotor UAVs. The YOLO is implemented as the basic detection algorithm, incorporating a motion planner algorithm for obstacle avoidance and trajectory optimization. A customized simulation and real-world experiments in GPS-denied areas prove the viability and robustness of applications such as indoor SaR missions. Hence, YOLO provides an open-sourced object detection vision-based learning approach that eliminates the need for memory-intensive storage as conventional optimization methods, which is ideal for UAVs designed for lightweight and efficiency.

An improvement of CNN referred to as DCNN, which uses convolutional layers that apply learnable filters to detect hierarchical features, involving genetic algorithm (GA)-based DCNN, AlexNet, and ResNet. Unlike CNN, a DCNN emphasizes deeper layers, which is a state-of-the-art NN in image processing. DCNN-GA algorithm is proposed for indoor navigation without GPS, such as SaR and surveillance. CNN hyperparameters were tuned by a genetic algorithm, and the promising results indicate significant accuracy and efficiency. In addition, the authors proposed a two-stage algorithm using a modification of the AlexNet in [25]. The drone localization is first confirmed by an NN using 3D projections and camera pose estimation. Secondly, camera position and orientation are determined utilizing a second-stage NN, and the designed autopilot controls attitude with a time delay.

2.3 Deep Reinforcement Learning (DRL)

Similar to reinforcement learning, deep reinforcement learning instead of using Q-values in RL but replaces the policy control with a neural network (NN) as shown in Figure 2. This enables the agent to learn directly from data, eliminating the need for manual computation. The classification of DRL involves model-based methods, value-based methods, and policy-based methods. The first method requires a given model for the environment; the agent learns from the model and performs a certain action in the specific environment. Secondly, the policy-based DRL aims to find the optimum policy by neural networks without estimating a value function as the transition step [16]. Policy-based DRL includes deterministic and stochastic methods. Finally, the value-based method allows the agent to find the optimal policy for the value function or the quality function, including on-policy algorithms and off-policy algorithms. Off-policy algorithms include common approaches such as Deep-Q Network (DQN) and dueling DQN (DDQN), where one neural network serves as the target, while the other controls the policy. For example, a Markov decision process (MDP), DDQL with multi-step learning, is used to optimize the trajectory for a cellular-connected UAV [26]. The authors further extend the use of signal measurement to train DQN and generate a radio map for availability simultaneously.

According to Poole and Mackworth, the model-based approach tends to be far more experience-efficient, needing considerably fewer interactions to learn effectively [27]. In contrast, model-free approaches often consume substantial computational resources without yielding similarly effective results. Meanwhile, the authors proposed a memory-enhanced technique in DRL for vision-based UAV navigation by incorporating dynamic goals, penalties and non-sparse rewards [28]. The results are tested in the virtual high-fidelity 3D environment, showing the result of higher success rates with fewer training steps. In addition, a massive MIMO guidance is used to deal with challenges such as UAV motion capture and communication with the ground base in real time [29]. A DQN is designed and trained to make decisions based on signal strengths and has been proven to have better convergence and coverage compared with typical strategies. Furthermore, the dual-UAV system separates the energy transfer and data collection with a control model trained by the Multi-agent DQL (MADQL) algorithm in [30]. The trajectories are optimized to minimize the Age of Information (AoI) and energy efficiency. The proposed approach provides trajectory optimization for multiple UAVs in a dynamic wireless powered communication network (WPCN), with promising results that converge quickly and reliably.

An extension of MDP called a partially observable MDP (POMDP) is able to plan and act under uncertainty, which is ideal for navigating in unknown spaces. POMDP allows the agent to make decisions when the drone cannot provide a comprehensive observation of the environment due to sensor limitations. The authors implement extended double deep Q-learning (EDDQL) in a partially observable environment [30]. The paper used a double-input strategy that integrates raw image data and local maps with location information. The results illustrate that the agent is able to outperform the adapted Deep Q-Network, DDQN and Deep Recurrent Q-Network (DRQN) and navigate UAVs in severe weather conditions of outdoor environments, especially low-visibility areas [31]. A similar algorithm is proposed to optimize coverage path planning (CPP) through an MDP-based DDQN approach [32]. The proposed algorithm generalizes a control policy for varying power constraints, and inputs are fed through a convolutional neural network (CNN), enabling the UAV to make decisions on balancing power consumption while completing target coverage. The proposed method utilized a combination of global MDP and local POMDP [33]. The UAV is navigated for indoor searching by a modular architecture and the TRPO algorithm to train control policies for the local problem. This approach enables the scalable formulation to solve MDP problems and POMDP problems, potentially benefiting applications such as SaR inside a building.

The policy-based algorithm is one of the main categories of DRL, which avoids using memory buffers due to inaccurate estimates based on outdated policies, such as the Proximal Policy Optimization (PPO) algorithm. It is worth noticing that the PPO algorithm is implemented to train models in order to control UAV attitudes with optimal energy consumption and compare different methods of attitude control, respectively [34,35]. Two papers modify and control the signal directly. However, the former paper includes real-world experiments and a reward function to avoid the sudden change in motor rotation speed, thus optimizing energy efficiency. The latter one focuses on inner-loop attitude control for maintaining stability using PPO, comparing their performance to traditional PID controller, Deep Deterministic Gradient Policy (DDGP) approach, and Trust Region Policy Optimization (TRPO) in simulations to assess their suitability for precise and time-critical tasks.

The Actor Critic method is another advanced framework in DRL that combines the benefits of both value-based and policy-based methods. From its name, the actor learns the policy and the critic evaluates the current policy by estimating the value function. Leveraging deep CNN for a saliency algorithm to identify obstacles visually with DRL employing an actor-critic model allows learning from demonstrations and excellent performance in [36]. Meanwhile, Deep Deterministic Policy Gradient (DDPG) is an advanced algorithm for continuous action spaces to produce the Q-values and updates. The agent makes decisions and flies to the static or dynamic targets by incorporating the DDPG algorithm with customized reward functions to inspire shorter paths and avoid collisions [37]. Further improvement is found, where the authors proposed an Asynchronous Advantage Actor-Critic (A3C)-based DRL approach for coordinating UAV swarms to serve as mobile BSs to provide cellular coverage in the target areas [38]. The A3C algorithm tends to optimize the energy efficiency while maintaining stable and maximum coverages. The target network is updated with a modified policy gradient, and each UAV is modelled by DNNs. The simulation results prove that the method outperforms the DDPG-based approach in terms of coverage quality and energy efficiency.

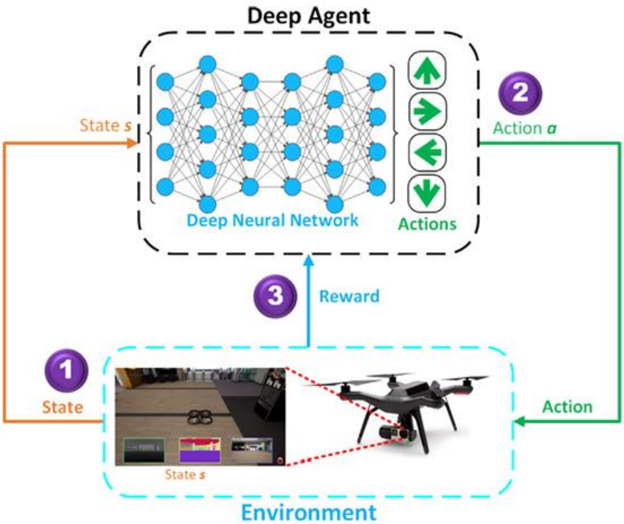

Fig. 2. Deep reinforcement learning process [5]

3 Challenges and Future Trends

3.1 Algorithms

Although AI-based approaches have shown great promise for the next-generation UAV navigation systems, the deployment in real world and technology readiness level still remain a huge gap to be further investigated. For example, RL struggles to generalize the algorithms to dynamic and uncertain real-world scenarios due to the reliance on trial-and-error learning methods. The approach may fail to perform the tasks in complicated and unencountered circumstances. Meanwhile, despite the effectiveness of CNNs in visual perception tasks, the implementation introduces challenges such as long training time and large sets of data.

The emerging technology, such as federated learning (FL), which provides decentralized training across multiple UAVs without raw data exchange and reduces complexity [20]. However, most of the FL approaches are still centralized and indicate unreliability, such as single point failure and sensitivity to environmental perturbations. Furthermore, the decentralized federated learning (DFL) approach is proposed to eliminate these issues with a preliminary simulation, the feasibility and implementation in the real world still require verification [39]. Another open research direction of DFL is to monitor and detect convergence without a central controller, as in the conventional FL approaches, due to different architectures, dynamic topologies, and intermittent communication.

3.2 Energy Consumption

Due to the prolonged limited onboard power for UAVs, persistent research opens the gate to energy efficiency. The power consumption of AI approaches significantly affects the overall performance of a UAV, restricting the range and airborne time. Despite different strategies having been introduced to optimize on-board energy management, such as power-aware task planning, sleep and wake-up schemes, and offline computation methods, these methods are not sufficient and generalizable for long-duration, wide-range missions [40]. In addition, AI-based navigation for planning a charging trajectory by identifying and visiting charging stations automatically remains underexplored.

The future research direction involves developing energy-efficient algorithms and integrating them with smart power management systems to extend endurance and range. AI-based recharging logistics is also prominent in performing autonomous missions, opening gates for autonomous docking and recharging.

3.3 Hardware Limitations

Due to the nature of small size and weight sensitivity of UAVs, the computational power on board has strict constraints, which make the deployment of complex vision-based AI models challenging. The high computational power of CNNs possibly exceeds onboard computation capability, especially for processing high-dimensional data like images and LiDAR sensors. Flight control and trajectory planning rely heavily on low-latency obstacle avoidance and target tracking in dynamic environments. The development of lightweight navigation algorithms and integrating with flight control systems tailored to embedded UAS requires further investigation.

An emerging solution for this problem is to introduce cloud-based or edge computing to free onboard computation. However, the practical limitations of the applicability involve communication latency, signal strengths, generalization, data privacy, and development of compressed onboard algorithms. Thus, the exploration of hybrid computation approaches that balance onboard and cloud computing to generate low-latency models is one of the future directions in overcoming hardware bottlenecks.

4 Conclusion

The navigation systems generally require perception, localization, and trajectory planning, which have been widely researched in recent years, enabling the real-time decision-making processes across various domains. This study presents a comprehensive review of AI-based approaches for GPS-independent UAV navigation systems, focusing on three main types of AI: reinforcement learning (RL), deep reinforcement learning (DRL), and deep learning (DL) techniques.

Although prominent results have been achieved in algorithm development recently, particularly in DRL and CNNs, challenges still remain. One of the most critical challenges is feasibility in real-world environments, due to a large number of studies that have only been verified in simulation environments. This transition is the crucial step to transfer approaches from simulation to practical uses. In addition, a few other obstacles open the gate for future research include the generalization of learning models to real-world environments, high computational demands, energy constraints, and limited onboard hardware capabilities due to unreliable communications, privacy concerns, lack of robustness, and safety considerations. Although the emerging paradigms, such as DFL, offered prominent values in addressing multiple UAV training issues, the real-world implementation is still nascent due to issues like convergence detection and air-to-air communications.

Therefore, future research has pointed to lightweight, energy-efficient generalized and verified learning approaches suitable for UAVs on-board deployment. The robust learning under uncertainty and hybrid frameworks enables the cloud support for onboard systems. Moreover, experimental verification and validation are essential in addressing the AI-based navigation challenges to implement later in the real-world autonomous missions, especially in safety-related applications and military operations.

References

1. Vanegas, F. Gaston, K. J. Roberts, J. and Gonzalez, F.: A Framework for UAV Navigation and Exploration in GPS-Denied Environments, in 2019 IEEE Aerospace Conference, Big Sky, MT, USA: IEEE, Mar. 1–6(2019).

2. Rezwan, S. and Choi, W.: Artificial Intelligence Approaches for UAV Navigation: Recent Advances and Future Challenges, IEEE Access, 10, 26320–26339(2022).
3. Chhikara, P. Tekchandani, R. Kumar, N. Chamola, V. and Guizani, M.: DCNN-GA: A Deep Neural Net Architecture for Navigation of UAV in Indoor Environment, IEEE Internet Things J, 8(6), 4448–4460(2021).
4. Sutton, R. S. and Barto, A. G.: Reinforcement Learning: An Introduction, The MIT Press, 2018.
5. AlMahamid, F. and Grolinger, K.: Autonomous Unmanned Aerial Vehicle navigation using Reinforcement Learning: A systematic review, Engineering Applications of Artificial Intelligence, 115, 105321(2022).
6. Pham, H.X. La, H.M. Feil-Seifer, D. and Nguyen, L.V.: Autonomous UAV Navigation Using Reinforcement Learning, (2018)
7. Pham, H. X. La, H. M. Feil-Seifer, D. and Van Nguyen, L.: Reinforcement Learning for Autonomous UAV Navigation Using Function Approximation, in 2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), Philadelphia, PA: IEEE, Aug. 1–6 (2018).
8. Chowdhury, M. M. U., Erden, F., and Guvenc, I.: RSS-Based Q-Learning for Indoor UAV Navigation, arXiv (2019).
9. Liu, X., Chen, M., and Yin, C.: Optimized Trajectory Design in UAV Based Cellular Networks for 3D Users: A Double Q-Learning Approach, J. Commun. Inf. Netw., 4, no. 1, 24–32, Mar. 2019,
10. Colonnese, S., Cuomo, F., Pagliari, G., and Chiaraviglio, L.: Q-SQUARE: A Q-learning approach to provide a QoE aware UAV flight path in cellular networks, Ad Hoc Networks, 91, 101872, (2019).
11. Liu, X., Liu, Y., and Chen, Y.: Reinforcement Learning in Multiple-UAV Networks: Deployment and Movement Design, IEEE Trans. Veh. Technol., 68(8), 8036–8049(2019).
12. Liu, X., Liu, Y., Chen, Y., and Hanzo, L.: Trajectory Design and Power Control for Multi-UAV Assisted Wireless Networks: A Machine Learning Approach, IEEE Trans. Veh. Technol., 68(8), 7957–7969(2019).
13. Hu, J., Zhang, H., and Song, L.: Reinforcement Learning for Decentralized Trajectory Design in Cellular UAV Networks with Sense-and-Send Protocol, arXiv1809.02934(2018).
14. Zeng, Y. and Xu, X.: Path Design for Cellular-Connected UAV with Reinforcement Learning, arXivARXIV.1905.03440(2019).
15. Azar, A.T., et al.: Drone Deep Reinforcement Learning: A Review, Electronics, 10(9), 999(2021).
16. Fagundes-Junior, L.A., De Carvalho, K.B., Ferreira, R. S., and Brandão, A. S.: Machine Learning for Unmanned Aerial Vehicles Navigation: An Overview, SN COMPUT. SCI., 5(2), 256(2024).
17. Back, S., Cho, G., Oh, J., Tran, X.-T., and Oh, H.: Autonomous UAV Trail Navigation with Obstacle Avoidance Using Deep Neural Networks, J Intell Robot Syst, 100(3–4), 1195–1211(2020).
18. Liu, Y., Xie, K., and Huang, H., VGF-Net: Visual-Geometric Fusion Learning for Simultaneous Drone Navigation and Height Mapping, (2021).
19. Giusti, A. et al.: A Machine Learning Approach to Visual Perception of Forest Trails for Mobile Robots, IEEE Robot. Autom. Lett., 1(2), 661–667(2016).
20. Katkuri, A. V. R. Madan, Khatri, H. N. Abdul-Qawy, A. S. H. and Patnaik, K. S.: Autonomous UAV navigation using deep learning-based computer vision frameworks: A systematic literature review, Array, 23, 100361(2024).

21. Opromolla, R., Inchingolo, G., and Fasano, G.: Airborne Visual Detection and Tracking of Cooperative UAVs Exploiting Deep Learning, Sensors, 19(19), 4332(2019).
22. Waqas, A., Kang, D., and Cha, Y.-J.: Deep learning-based obstacle-avoiding autonomous UAVs with fiducial marker-based localization for structural health monitoring, Structural Health Monitoring, 23(2), 971–990(2024).
23. Tullu, A., Endale, B., Wondosen, A., and Hwang, H.-Y.: Machine Learning Approach to Real-Time 3D Path Planning for Autonomous Navigation of Unmanned Aerial Vehicle, Applied Sciences, 11(10), 4706(2021).
24. Chen, C., Wang, Z., Gong, Z., Cai, P., Zhang, C., and Li, Y.: Autonomous Navigation and Obstacle Avoidance for Small VTOL UAV in Unknown Environments, Symmetry, 14(12), 2608(2022).
25. Kupervasser, O., Kutomanov, H., Levi, O., Pukshansky, V., and Yavich, R.: Using Deep Learning for Visual Navigation of Drone with Respect to 3D Ground Objects, Mathematics, 8(12), 2140(2020).
26. Zeng, Y., Xu, X., Jin, S., and Zhang, R.: Simultaneous Navigation and Radio Mapping for Cellular-Connected UAV With Deep Reinforcement Learning, IEEE Trans. Wireless Commun., 20(7), 4205–4220(2021).
27. Poole, D. L.: Artificial Intelligence, (Cambridge University Press, 1st edn. 2017).
28. Fu, C. et al.: Memory-enhanced deep reinforcement learning for UAV navigation in 3D environment, Neural Comput & Applic, 34(17), 14599–14607(2022).
29. Huang, H., Yang, Y., Wang, H., Ding, Z., Sari, H., and Adachi, F.: Deep Reinforcement Learning for UAV Navigation Through Massive MIMO Technique, IEEE Trans. Veh. Technol., 69(1), 1117–1121(2020).
30. Oubbati, O. S., Atiquzzaman, M., Lakas, A., Baz, A., Alhakami, H., and Alhakami, W.: Multi-UAV-enabled AoI-aware WPCN: A Multi-agent Reinforcement Learning Strategy, in IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Vancouver, BC, Canada: IEEE, May 2021, 1–6.
31. Maciel-Pearson, B. G., Marchegiani, L., Akcay, S.: A. Atapour-Abarghouei, J. Garforth, and T. P. Breckon, Online Deep Reinforcement Learning for Autonomous UAV Navigation and Exploration of Outdoor Environments, 2019, arXiv1912.05684.
32. Theile, M., Bayerlein, H., Nai, R., Gesbert, D., and Caccamo, M.: UAV Coverage Path Planning under Varying Power Constraints using Deep Reinforcement Learning, in 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA: IEEE, Oct. 2020, 1444–1449.
33. Walker, O., Vanegas, F., Gonzalez, F., and Koenig, S.: A Deep Reinforcement Learning Framework for UAV Navigation in Indoor Environments, in 2019 IEEE Aerospace Conference, Big Sky, MT, USA: IEEE, Mar. 2019, 1–14.
34. Agarwal, V. and Tewari, R. R.: Improving Energy Efficiency in UAV Attitude Control using Deep Reinforcement Learning, JSR, 65(03), 209–219 (2021).
35. Koch, W., Mancuso, R., West, R., and Bestavros, A.: Reinforcement Learning for UAV Attitude Control, 2018, arXiv.1804.04154.
36. Ma, Z., Wang, C., Niu, Y., Wang, X., and Shen, L.: A saliency-based reinforcement learning approach for a UAV to avoid flying obstacles, Robotics and Autonomous Systems, 100, 108–118(2018).
37. Bouhamed, O., Ghazzai, H., Besbes, H., and Massoud, Y.: Autonomous UAV Navigation: A DDPG-based Deep Reinforcement Learning Approach, 2020, arXiv2003.10923.
38. Liu, C. H., Ma, X., Gao, X., and Tang, J.: Distributed Energy-Efficient Multi-UAV Navigation for Long-Term Communication Coverage by Deep Reinforcement Learning, IEEE Trans. on Mobile Comput., 19(6), 1274–1285(2020).

39. Qu, Y., et al.: Decentralized Federated Learning for UAV Networks: Architecture, Challenges, and Opportunities, IEEE Network, 35(6), 156–162(2021).
40. Kuru, K.: Planning the Future of Smart Cities With Swarms of Fully Autonomous Unmanned Aerial Vehicles Using a Novel Framework, IEEE Access, 9, 6571–6595(2021).