



International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Survey on Self-Supervised Machine Learning for Safe Unmanned Aerial Vehicles Navigation

Tankala Sai Pravallika

Student (23341A12B5), IT Department, GMRIT, Rajam

ABSTRACT:

Unmanned Aerial Vehicles (UAVs) have become indispensable in critical applications such as infrastructure inspection, emergency response, and autonomous delivery. Safely navigating dynamic and complex environments continues to present significant challenges. This study introduces a novel two-stage machine learning framework, leveraging advanced technologies in self-supervised learning and deep reinforcement learning (DRL), to improve UAV navigation. In the first stage, a convolutional neural network utilizes self-generated pseudo-labels from the UAV's onboard camera to autonomously learn depth estimation and obstacle segmentation without human annotations. In the second stage, these pretrained features serve as the foundation for a deep Q-network, which is further optimized through reinforcement learning within a simulated 3D environment containing both static and moving obstacles. A lightweight safety filter operates alongside the policy to prevent collision risks and enhance flight reliability. This integrated approach, grounded in state-of-the-art machine learning and reinforcement learning methods, eliminates the need for manual labeling and enables robust deployment in GPS-denied or visually complex settings, offering promising advances toward reliable autonomous aerial navigation.

Keywords: *Self-Supervised Learning, Deep reinforcement Learning, UAV Navigation, Obstacle Detection, Autonomous Flight safety*

1.INTRODUCTION:

Unmanned Aerial Vehicles (UAVs), widely known as drones, are rapidly transforming various sectors such as surveillance, agriculture, infrastructure inspection, and disaster response due to their versatility and ability to operate in challenging environments. Despite their growing usage, achieving safe and reliable autonomous navigation remains a complex problem, particularly in cluttered or GPS-denied environments. Traditional rule-based systems often struggle to adapt to dynamic obstacles and unpredictable conditions, making intelligent decision-making essential. Machine learning offers a powerful solution by enabling UAVs to learn from data and improve performance over time. Among these techniques, self-supervised learning stands out as it allows models to learn meaningful features directly from raw sensor inputs without the need for manually labeled data. When combined with deep reinforcement learning, which trains agents through trial-and-error interactions with the environment, this approach becomes highly effective for developing robust navigation policies. This term paper explores a two-stage learning framework that integrates self-supervised visual learning with deep reinforcement learning to enable UAVs to safely navigate complex 3D 2.

2. LITERATURE SURVEY:

[1] Yanan Tian, Adil Khan, Shabeer Ahmad, Syed Agha Hassnain Mohsan, Faten Khalid Karim, Babar Hayat, Samih M. Mostafa(2025) Proposed an optimized deep learning framework for UAV-assisted wireless networks using Double Exponential Crayfish Optimization Algorithm (DECOA) and Deep Belief Networks (DBN) for energy-efficient resource allocation.

The limitation is Scalability for large UAV networks, handling power limits, and integrating predictive learning with optimization.[2] Hao Li, Zhijun Xie, Xing Jin, Changchun Peng, Cheng Ren, Yunfei Mao, Kefan Zhou, Yingying Li (2025)Developed an adaptive dynamic weighted ensemble (DWE) model with ASO-GSA feature selection for UAV-based water quality monitoring. The limitation is Difficulty in improving both temporal and spatial scalability; accuracy depends on better integration of ML models.[3] Meshari Aljohani, Ravi Mukkamala, Stephan Olariu (2024) Analysed challenges of autonomous strike UAVs in homeland security using blockchain, smart contracts, and ML for secure UAV operations. The limitation is Ensuring secure localization, tamper-proof data storage, and robust ML models for adaptive missions.[4] Hussein Samma, Sami El-Ferik(2024) Proposed enhanced deep reinforcement learning (DQN + self-supervised tuning) for UAV visual navigation in dynamic environments. The limitation is Struggles in highly dynamic obstacle scenarios, slow learning rates, and limited real-world testing.[5] Ala' Abdulmajid Eshmawi, Muhammad Umer, Imran Ashraf, Yongwan Park (2024) Developed a stacked ensemble model (SVM + CNN) to detect GPS spoofing attacks on UAVs without hardware dependency. The limitation is Reliability across diverse real-world spoofing conditions and generalization beyond controlled datasets.[6] Yash Vasant Ahirrao, Rana Pratap Yadav, Sunil Kumar (2024) Designed RF-based UAV detection & identification using ML on RF features (2–6 GHz range).

Achieved 99.6% accuracy in tests. The limitation is Real-time localization accuracy in outdoor environments and lack of diverse UAV RF datasets.[7] Mohammad Hosein Saeedinia, Behzad Hashemi, Ana-Maria Cretu, Shamsodin Taheri (2024) Proposed computational intelligence-based PV parameter extraction for UAV-integrated solar systems. Used MIMO/MISO + ML for real-time predictions. The limitation is Accurate parameter extraction on moving UAVs and modelling effects of UAV 3D movements on PV operation.[8] Dennis Agnew, Alvaro Del Aguila, Janise McNair (2024) Enhanced cybersecurity in SD-UAV networks via queuing analysis + ML to detect zero-day attacks. The limitation is Absence of robust real-time predictive defences and limited testing of zero-day attacks in UAV swarms.[9] Mohamed Amine Ould Rabah, Hamza Drid, Yasmine Medjadba, Mohamed Rahouti (2024) Proposed SDN-UAV architecture using LSTM/Bi-LSTM ensemble + honeypots for DDoS detection & mitigation. The limitation is Single-mode approaches fail against new attack types and real-time adaptive defence's are underdeveloped.[10] Bassem Mokhtar (2024) Introduced collaborative distributed computing with lightweight edge ML models in UAV networks for efficient training & aggregation. The limitation is Efficient ML deployment on constrained UAVs, privacy-preserved learning, and balancing accuracy vs latency.[11] Thinnakon Angkahad, Teerawong Laosuwan, Satith Sangpradid, Narueset Prasertsri, Yannawut Uttarak, Titipong Phoophathong, Joe Nuchthapho (2024) Used UAV photogrammetry + ML for biomass and carbon estimation in forests. Compared UAV vs field surveys. The limitation is Automated biomass estimation is not fully accurate; validation across diverse forestry conditions is lacking.[12] Ahmed Fahim Mostafa, Mohamed Abdel-Kader, Yasser Gadallah, Omar Elayat (2023) Developed ML-based multi-UAV deployment for uplink traffic offloading in cellular networks. Used RL + ARIMA for optimal UAV placement. The limitation is Limited real-time UAV deployment methods for unpredictable uplink loads; high complexity in 3D positioning.[13] Erika Fonseca, Boris Galkin, Ramy Amer, Luiz A. DaSilva, Ivana Dusparic (2023) Proposed RL-based UAV altitude optimization to improve throughput & spectrum efficiency in cellular networks. The limitation is Few solutions for real-time altitude adaptation; lack of evaluation in real-world urban scenarios.[14] Hongyi Zhang, Zhiqiang Qi, Jingya Li, Anders Aronsson, Jan Bosch, Helena Holmström Olsson (2024) Developed DRL-based control of UAV base stations for 5G backhaul in disaster scenarios. Focused on 3D UAV-BS coordination. The limitation is Lack of scalable, decentralized UAV deployment strategies for dynamic and disaster-hit environments.[15] Atefeh Termehchi, Tingnan Bao, Aisha Syed, William Sean Kennedy, Melike Erol-Kantarci (2024) Proposed a goal-oriented DRL + supervised learning framework for UAV-aided THz networks to reduce non-informative interactions. The limitation is Difficulty in reducing redundant UAV-IoT interactions and limited evaluation in highly dynamic IoT scenarios.

3. METHODOLOGY:

This research focuses on developing a self-supervised machine learning framework that enables UAVs to safely navigate complex environments without the need for human-labeled data. The proposed approach consists of four main stages: input, dataset, methodology, and output, as described below.

INPUT DATA

The input for this system consists of multi-modal sensory information collected from the UAV's onboard sensors:

- Visual Data: Captured using a forward-facing RGB camera to identify objects and obstacles.
- Depth Data: Obtained from depth sensors or stereo vision for estimating distances to nearby objects.
- Inertial Data: Includes readings from the Inertial Measurement Unit (IMU) — accelerometer and gyroscope — to measure motion and orientation.
- Telemetry Data: Altitude, speed, and GPS coordinates for flight state estimation and trajectory tracking.

All these inputs are synchronized and preprocessed to create a unified data stream for model training and decision-making.

DATASET USED

The datasets used for this research include both simulated and real-world UAV flight data:

- Simulated Environment Datasets: Generated in platforms such as AirSim, Gazebo, or Microsoft Drone Simulator, providing labeled scenes for obstacle detection and navigation under different conditions (lighting, wind, terrain).
- Real Flight Datasets: UAV flight logs and video feeds collected from real-time flights in outdoor and indoor test zones. These datasets contain sensor readings, depth maps, and collision events automatically labeled through self-supervised learning signals.
- Self-Supervised Data Generation: The UAV generates pseudo-labels like obstacle boundaries, collision warnings, and free-space zones based on its own sensor feedback—removing the need for manual annotation.

METHODOLOGIES FOLLOWED (algorithms and workflows)

To achieve safe and intelligent UAV navigation, this research integrates multiple algorithms that work together within a self-supervised learning and reinforcement learning framework. Each algorithm plays a specific role in data processing, decision-making, optimization, and safety assurance. The detailed workflow and contribution of each are explained below.

The proposed methodology integrates self-supervised machine learning with deep reinforcement learning (DRL) and advanced optimization algorithms to achieve safe autonomous navigation for Unmanned Aerial Vehicles (UAVs). The approach is built upon two primary stages:

- A self-supervised pre-training phase for visual feature learning

- A reinforcement learning phase for autonomous decision-making and flight control.

In self-supervised learning, UAVs use unlabeled data from onboard sensors and cameras to extract meaningful representations without the need for human annotation. These representations form the input foundation for the reinforcement learning model, which learns through interaction with simulated environments to make optimal navigation decisions.

The hybrid use of self-supervised and reinforcement learning ensures adaptability, safety, and energy-efficient performance, even in previously unseen environments such as dense urban landscapes or GPS-denied areas.

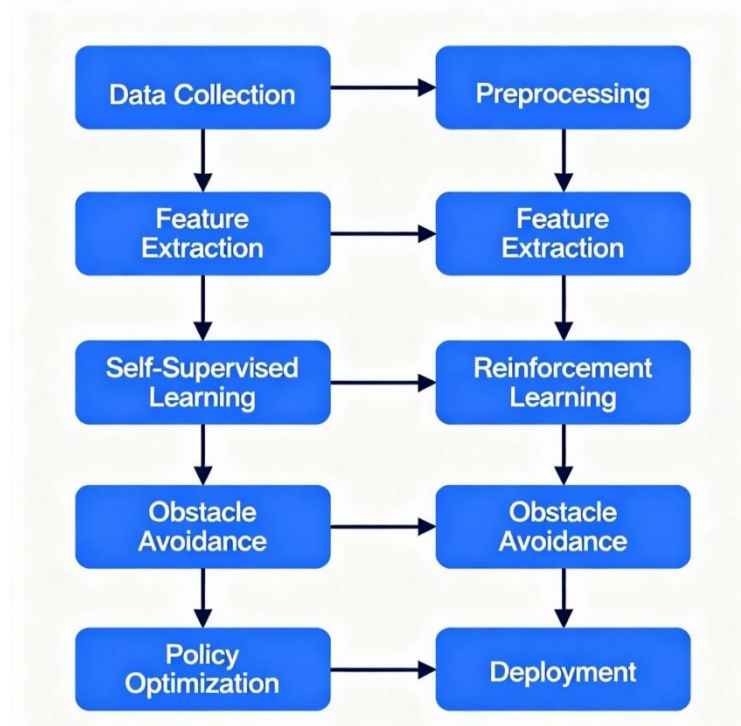


Figure 1

1. Self-Supervised Learning (SSL)

Workflow:

Self-supervised learning enables the UAV to train visual models without human-provided labels. The UAV's onboard camera captures raw images of its surroundings, which are automatically processed to generate pseudo-labels such as obstacle boundaries, free-space areas, and depth maps. These pseudo-labels are derived using techniques like contrastive learning and autoencoding, allowing the CNN to predict spatial and contextual relationships between pixels. The model learns to understand environmental depth and object distance by reconstructing or predicting future frames from its current view.

Role in UAV Safety:

- Helps the UAV perceive its surroundings accurately in real time.
- Eliminates dependency on labeled datasets, making the system adaptable to any new environment.

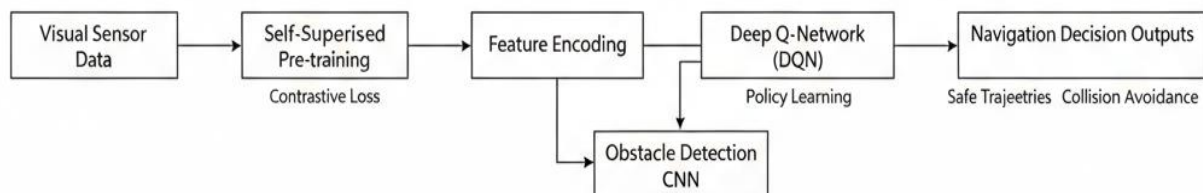


Figure 2

2. Deep Reinforcement Learning (DRL) — Deep Q-Network (DQN)

Workflow:

After the visual encoder (from SSL) extracts scene information, its output features are passed to the Deep Q-Network (DQN). The DQN interacts with a simulated 3D environment that contains both static (buildings, trees) and dynamic obstacles (other drones moving objects). It performs the following steps:

State Observation: Receives environment data (position, velocity, visual frames).

Action Selection: Chooses an action (move forward, turn left, ascend, descend, hover) based on its current policy.

Reward Evaluation: Receives a positive reward for safe movement and negative reward for collision or unsafe proximity.

Policy Update: Continuously improves its strategy through trial-and-error learning using Q-value updates.

Role in UAV Safety:

- Enables autonomous decision-making by learning safe flight actions through experience.
- Minimizes collision risks by optimizing for long-term safety and efficiency.
- Adapts to dynamic environments, maintaining stable flight paths even when new obstacles appear.
- Gradually learns energy-efficient and short routes, improving overall mission performance.

3. Double Exponential Crayfish Optimization Algorithm (DECOA)

Workflow:

DECOA combines the advantages of Double Exponential Smoothing (DES) and the Crayfish Optimization Algorithm (COA) to handle trajectory and resource optimization. The DES component forecasts future UAV positions and energy usage trends. The COA component uses swarm intelligence behavior (inspired by crayfish movement) to optimize UAV flight paths and resource allocation. Through iterative updates, DECOA ensures that UAVs maintain efficient routes while avoiding congested or high-risk regions.

Role in UAV Safety:

- Enhances flight path optimization by selecting safe and energy-efficient routes.
- Reduces the risk of collision by dynamically adjusting paths when obstacles are detected.
- Improves multi-UAV coordination, preventing mid-air conflicts during collaborative missions.

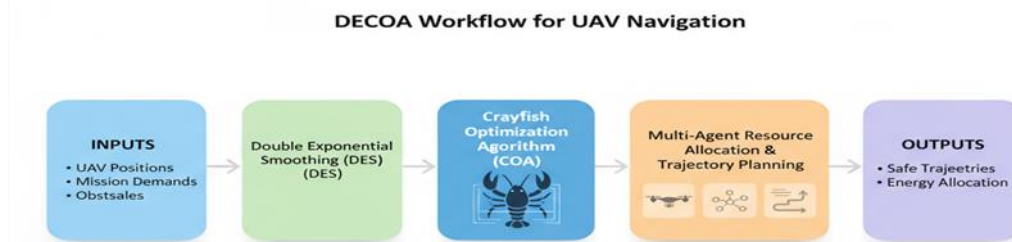


Figure 3

4. Deep Belief Network (DBN)

Workflow:

A DBN is a generative neural network composed of multiple layers of Restricted Boltzmann Machines (RBMs). It processes sensor and environmental data to predict patterns such as wind disturbances, signal loss, or obstacle behavior. The DBN learns hierarchical features — from low-level sensor readings to high-level flight patterns — helping the UAV understand external factors that could affect navigation.

Role in UAV Safety:

- Predicts environmental changes and resource needs in advance.
- Provides context-aware insights to assist the reinforcement learning agent in making safe decisions.
- Increases reliability during long-duration flights by managing power and flight control dynamically.

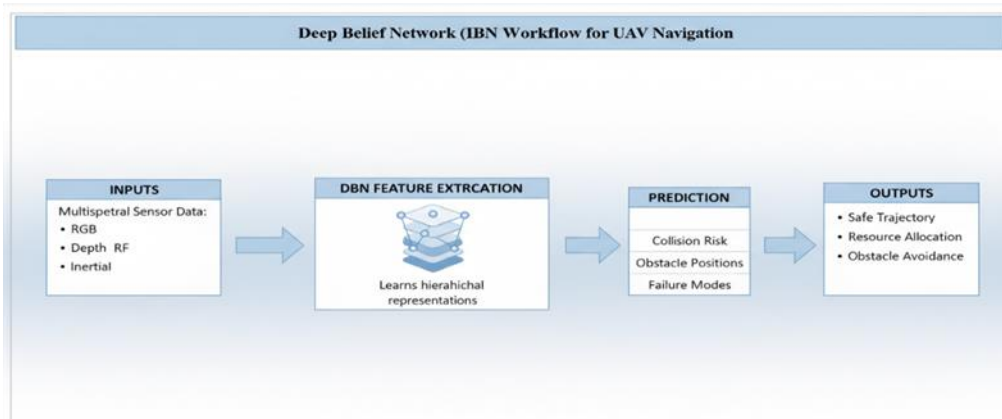


Figure 4

5. Whale Optimization Algorithm (WOA) with Dogleg Trust Region Algorithm (DTRA)

Workflow:

This hybrid algorithm enhances system parameter tuning and model calibration. WOA mimics the bubble-net hunting strategy of whales to explore and exploit the search space efficiently. DTRA refines the solution found by WOA by solving nonlinear optimization problems using trust-region methods, ensuring precise model adjustments. Together, they optimize flight controller parameters such as thrust, angle, and altitude stability.

Role in UAV Safety:

- Ensures stability and smooth control during flight.
- Minimizes the chances of oscillations or sudden drifts that could cause crashes.
- Adapts UAV parameters in real time to maintain balance and reliability under dynamic weather or load conditions.

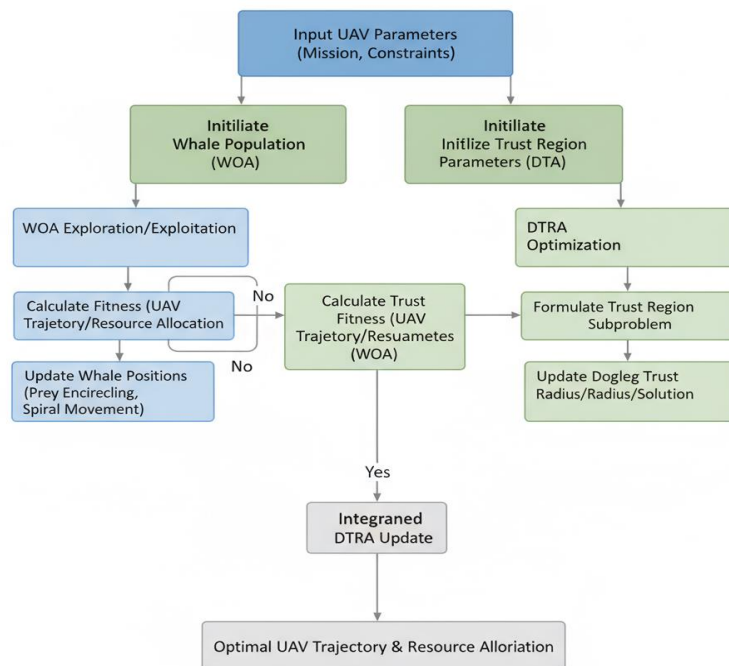


Figure 3

OUTPUT

The output of the proposed framework includes:

- **Autonomous and Collision-Free Navigation:** UAVs c

- an safely maneuver around obstacles in dynamic 3D environments.
- **Real-Time Decision Making:** The model continuously updates its policy for efficient and adaptive flight control.
- **Reduced Human Intervention:** Eliminates the need for manual supervision and labeled training data.
- **Improved Flight Efficiency:** Optimized paths and energy consumption using DECOA and WOA-DTRA.
- **Scalable Multi-UAV Collaboration:** Enables knowledge sharing among UAVs through distributed learning for improved performance in team-based missions.

CONCLUSION:

Safe and autonomous navigation is a critical challenge for UAVs operating in dynamic environments. This research proposes a self-supervised machine learning framework that empowers UAVs to learn from raw sensory data without the need for manual labeling. The integration of self-supervised visual learning and deep reinforcement learning enables UAVs to perceive surroundings, predict obstacles, and make intelligent navigation decisions in real time. The addition of optimization algorithms such as DECOA, DBN, and WOA-DTRA further refines trajectory planning, system stability, and energy efficiency. A lightweight safety filter ensures that the UAV avoids collision-prone maneuvers, improving operational safety. The framework's distributed learning capability allows multiple UAVs to share experiences, enhancing adaptability and scalability. In conclusion, the proposed self-supervised UAV navigation system demonstrates a reliable, data-driven approach for achieving safe, efficient, and autonomous aerial operations, making it suitable for real-world applications such as disaster response, environmental monitoring, and smart logistics.

REFERENCES:

1. Yanan Tian, Adil Khan, Shabeer Ahmad, Syed Agha Hassnain Mohsan, Faten Khalid Karim, Babar Hayat, and Samih M. Mostafa, "An Optimized Deep Learning Framework for Energy-Efficient Resource Allocation in UAV-Assisted Wireless Networks," IEEE Access, Vol. 13, 2025.
2. Hao Li, Zhijun Xie, Xing Jin, Changchun Peng, Cheng Ren, Yunfei Mao, Kefan Zhou, and Yingying Li, "DWE Research on Water Quality Inversion and Scalability in Small Watersheds Using UAV Remote Sensing Data," IEEE Access, Vol. 13, 2025.
3. Meshari Aljohani, Ravi Mukkamala, and Stephan Olariu, "Autonomous Strike UAVs in Support of Homeland Security Missions: Challenges and Preliminary Solutions," IEEE Access, Vol. 12, pp. 90979–90996, June 2024.
4. Hussein Samma and Sami El-Ferik, "Autonomous UAV Visual Navigation Using an Improved Deep Reinforcement Learning," IEEE Access, Vol. 12, pp. 79967–79977, June 2024.
5. Ala Abdulmajid Eshmawi, Muhammad Umer, Imran Ashraf, Yongwan Park, "Enhanced Machine Learning Ensemble Approach for Securing Small Unmanned Aerial Vehicles From GPS Spoofing Attacks," IEEE Access, 2024.
6. Yash Vasant Ahirrao, Rana Pratap Yadav, Sunil Kumar, "RF-Based UAV Detection and Identification Enhanced by Machine Learning Approach," IEEE Access, 2024.
7. Mohammad Hosein Saeedinia, Behzad Hashemi, Ana-Maria Cretu, and Shamsodin Taheri, "A Novel Computational Intelligence-Based Parameter Extraction of UAV-Integrated Photovoltaic System," IEEE Access, Vol. 12, pp. 97664–97677, July 2024. DOI: 10.1109/ACCESS.2024.3428400.
8. Dennis Agnew, Alvaro del Aguila, Janise McNair, "Enhanced Network Metric Prediction for Machine Learning-Based Cyber Security of a Software-Defined UAV Relay Network," IEEE Access, 2024.
9. Mohamed Amine Ould Rabah, Hamza Drid, Yasmine Medjadba, Mohamed Rahouti, "Detection and Mitigation of Distributed Denial of Service Attacks Using Ensemble Learning and Honeypots in a Novel SDN-UAV Network Architecture," IEEE Access, 2024.
10. Bassem Mokhtar, "AI-Enabled Collaborative Distributed Computing in Networked UAVs," IEEE Access, Vol. 13, pp. 96515–96526, July 2024. DOI: 10.1109/ACCESS.2024.3425523.
11. Thinnakon Angkahad, Teerawong Laosuwan, Satith Sangpradid, Narueset Prasertsri, Yannawut Uttaruk, Titipong Phoophathong, Joe Nuchthapho, "An Empirical Analysis of Above-Ground Biomass and Carbon Sequestration Using UAV Photogrammetry and Machine Learning Techniques," IEEE Access, 2024.
12. A. F. Mostafa, M. Abdel-Kader, Y. Gadallah, O. Elayat, "Machine Learning-Based Multi-UAV Deployment for Uplink Traffic Sizing and Offloading in Cellular Networks," IEEE Access, 2023.
13. E. Fonseca, B. Galkin, R. Amer, L. A. DaSilva, I. Dusparic, "Adaptive Height Optimization for Cellular-Connected UAVs: A Deep Reinforcement Learning Approach," IEEE Access, 2023.