Federated Reinforcement Learning with Linear Programming for Improving UAV-Enabled Smart Agriculture

A.V. Mayakkannan
Department of ECE, Kings Engineering College, Chennai, Tamil Nadu, India
avmayakkannan@gmail.com (corresponding author)

Chethan Chandra S Basavaraddi
Department of CSE, Faculty of Engineering and Technology, School of Computer Science and Technology, GM University, Davanagere, Bangalore, Karnataka, India
raddi04@yahoo.com

R. Karthick
Department of Management Studies, St. Joseph's College of Engineering, Chennai, Tamil Nadu, India
karthick0528@gmail.com

S. Durga Devi
Department of CSE, Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Chennai, Tamil Nadu, India
sdurgadevi@gmail.com

S. Amudha
Department of CSE, Dr. M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India
amutha.it@drmgrdu.ac.in

R Sasikumar
Department of CSE, K. Ramakrishnan College of Engineering, Tiruchirappalli, Tamil Nadu, India
sasi1986@gmail.com

Bharat Tidke
Symbiosis Institute of Technology, Nagpur Campus, Symbiosis International (Deemed University), Pune, India
batidke@gmail.com

S. Murugan
Department of Biomedical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India
smuresjur@gmail.com

Received: 11 August 2025 | Revised: 6 September 2025 | Accepted: 20 September 2025

ABSTRACT

The incorporation of Unmanned Aerial Vehicles (UAVs) has huge potential to improve crop monitoring, precision farming, and data gathering in smart agriculture. However, optimizing UAV functions on geologically distributed farms poses important challenges related to computational efficiency, energy depletion, and route selection. This paper introduces a Federated Reinforcement Learning with Linear

Programming (FRLP) to address these issues. The proposed system utilizes an FRL algorithm, which is more suitable since every UAV learns about local energy depletion related to every movement path. The FRLP mechanism utilizes Super nodes (SPs) to collect and communicate sensor data to the UAV, and then remove redundant information using Principal Component Analysis (PCA). In this work, a Reinforcement Learning (RL) with Linear Programming (LP) model is utilized to forecast the next state based on a reward function, calculated using SP node energy, queued packets, and link quality. The LP also checks the UAV distance limit, the maximum UAV travels per round, and connectivity. Finally, the UAV decides on an optimal stop point to visit and collect data from the SP nodes. FRL integrated with UAVs offers promising advances in agricultural practices, particularly in optimizing UAV route formation to enhance cultivation efficiency. The simulation results illustrate that the FRLP mechanism reaches a 98.5% success rate and minimizes additional energy utilization.

Keywords-Unmanned aerial vehicles; linear programming; federated reinforcement learning; energy efficiency; agriculture field

I. INTRODUCTION

Smart agricultural systems observe crop fields and collect environmental and soil details through a Wireless Sensor Network (WSN) [1]. Smart agriculture allows precision farming utilizing data collection and automatic decisions [2]. Resources contain water, fertilizers, and insecticides that are applied proficiently for precision farming. Incorporating sensor nodes and UAVs has provided data control to improve agricultural practices, such as specific irrigation, pest control, and monitoring of soil health [3]. However, precision agriculture presents essential challenges, such as delay reduction, bandwidth optimization, and energy utilization [4]. The proposed system uses UAVs to collect data from crop fields through sensor nodes to minimize additional energy [5] and bandwidth utilization in WSNs. In addition, optimal UAV routing is an important factor in smart agriculture [6], as UAVs should reduce flight time and energy utilization while confirming that all SP nodes are visited regularly. This requires solving a complex route optimization issue under UAV limitations. Thus, UAV route planning in precision agriculture minimizes time and enhances service life.

UAV Path Optimization in Smart Agriculture (POSA) proposed a UAV path planning approach for WSNs [7]. The hover-optimized data collection method calculates optimal hover locations in every sensor node's transmission radius, reducing redundant displacement. This mechanism utilizes a cluster head to reduce travel distance. However, this mechanism raises both energy depletion and scalability issues in WSN. In addition, the cluster head approach creates energy holes and utilizes a single UAV, increasing the delay in WSNs. To solve these issues, FRL with LP is used to improve UAV-enabled smart agriculture.

An energy-efficient laser-charged UAV-enabled WSN [8] conveys a joint optimization issue involving power allocation, dynamic charging, and path planning to reduce task completion time and sensor node expiration time. A Deep Reinforcement Learning (DRL) algorithm enables charging scheduling decisions, thus optimizing network performance. A multi-agent Double Deep Q-network (DDQN) mechanism is used to decide the optimal route for all UAVs in complex environments. Multi-agent DRL handles high computational costs when increasing the sensor nodes. In addition, UAVs with reduced energy travel to recharge stations, suspending mission-critical tasks.

FRL combines the strengths of Federated Learning (FL) and Reinforcement Learning (RL) in the challenges of UAV networks [9]. FRL enables UAVs in smart agriculture to learn cooperatively, adjust energetically, and activate efficiently, all while preserving farmer data privacy and considering UAV resource constraints. This makes FRL a powerful approach to intelligent and sustainable agricultural automation. It helps local learning, reduces the transmission overhead of model control, reduces energy utilization, and increases scalability in smart agriculture. Linear Programming (LP) is an optimization method for making the best possible decision in a system with inadequate resources that assists in discovering the optimal solution. Incorporation of FRL and LP allows routing adaptation, improving energy efficiency and scalability, and minimizing delay [10].

Adaptive routing indicates that UAVs can alter their paths based on the status of sensor nodes, environmental fluctuations, and optimal data gathering [11]. Energy efficiency means that UAVs save energy, extending their functional time and minimizing the requirement for frequent recharging. FRL allows the system to scale with the calculation of more UAVs and sensor nodes, despite an important increase in computational complexity. Traditional learning is unusable in agriculture due to privacy, cost, and heterogeneity, while UAV operations demand both adaptive decision-making and strict resource optimization. FRL with LP provides the best of both worlds for justifiable and smart agriculture.

The Federated Reinforcement Learning with Linear Programming (FRLP) mechanism utilizes SPs that collect information from IoT sensors and communicate this information to UAVs. The SP node contains additional resources, such as communication range, energy, and memory. FRL allows multiple UAVs to co-operate in learning optimal routing plans while sustaining data efficiently. UAV separately learns from its local situation and distributes model updates, rather than raw data, to a central server. This method is mostly helpful in situations where data privacy is vital and bandwidth is restricted. The FRLP mechanism applies to an LP method that solves optimization issues related to UAV routing optimization, such as reducing flying time, energy utilization, and timely data collection from sensor nodes. By conveying these intentions as linear constraints and objectives, LP offers an efficient solution to optimize the paths of UAVs.

II. LITERATURE SURVEY

In Hierarchical Federated Learning (HFL), UAVs act as middle aggregators among sensors and central servers [12]. This mechanism is mostly advantageous in agricultural settings where communication may be limited. An unbiased HFL algorithm for UAVs treats the unpredictability of communication by fine-tuning weights during local and global aggregations. The HFL mechanism aims to eliminate bias toward devices with better channel situations, which is vital in heterogeneous agricultural environments. The energy-efficient hybrid low-energy adaptive clustering hierarchy protocol is used to save energy [13].

An adaptive deployment of UAVs adjusts energy depletion, delay, and flexibility to improve network performance [14]. Joint optimization is used for learning configuration, bandwidth allocation, sensor-to-UAV connection, and decision-making in agricultural processes. Combining Internet of Things (IoT) devices with UAVs under an FRL structure leads to an important enhancement in managing resources and energy efficiency [15]. This system applies IoT devices to observe environmental parameters in addition to UAVs to offer aerial imagery, supporting distributed learning that trains a local RL method on every node and associates them into a global model. In UAV-assisted FL [16], a stochastic geometry tool is applied to evaluate the success possibilities of local and global parameter communications to alleviate the influence of unpredictable communication. Including these metrics helps fine-tune the aggregation weights established on the reliability of the communication link. This improves the robustness and performance of the FL models in agricultural applications.

Joint optimization for dynamic data caching and load computation is utilized to reduce the average task handling delay and enhance the UAV cache hit ratio [17]. Cognitive Fish Swarm Optimization (CFWO) is motivated by the combined intelligence and collaboration observed in fish swarms [18]. Each fish in the swarm observes the conditions by collecting data from its surroundings, comprising signal strength, channel availability, and congestion. Then, its cognitive aptitudes are developed to measure different routing and channel options based on precise objectives, namely energy efficiency. RL enables autonomous vehicles to determine ideal routes by fine-tuning to changing situations via trial and error, improving decision-making [19].

III. PROPOSED SYSTEM

The proposed system contains some IoT sensor nodes, UAVs, an edge server, and a cloud server. Figure 1 demonstrates the structure of the FRLP mechanism. UAVs collect data from the field sensors and forward it to the edge server. The edge server collects data from several UAVs and aggregates them. Finally, the edge server forwards the agricultural data to the cloud server. The UAVs travel through the agricultural field and collect data from the sensor nodes at some stopping points. This design aims to avoid sensor node energy hole issues since these nodes require additional energy, memory, and communication range. Thus, it minimizes sensor node energy hole issues and reduces congestion.

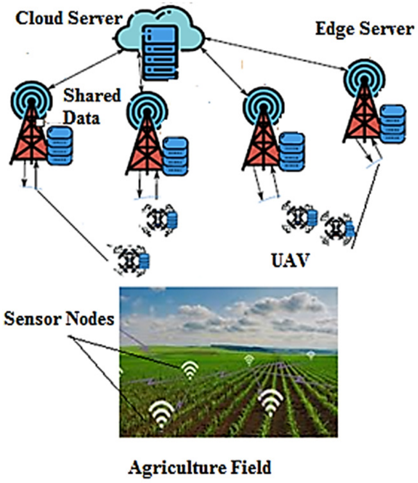

Fig. 1. Structure of the FRLP mechanism.

A. Data Collection

WSNs integrated with IoT sensor nodes have rapidly advanced in agricultural domains. IoT sensors are used to observe environmental features, including soil moisture, soil temperature, and humidity. UAVs capture high-resolution images for soil moisture, crop growth, and pest identification at regular intervals in agricultural fields. PCA is a dimensionality reduction technique for removing repeated and unwanted data, assisting in faster classification, better visualization, and less redundant computation. The objective of the proposed mechanism is to develop an enhanced UAV route optimization.

B. UAV Route Planning

The aim is to find optimal stopping points for collecting sensor data and an optimal route to supervise the field, while avoiding re-visiting stopping points. A stopping point is determined by applying the FRL algorithm to reduce associated costs and visit everyone in a shorter time in the agricultural region. This approach takes advantage of the decentralized nature of FRL to process data locally and reduce communication overhead, while UAVs provide mobility and flexibility in data collection and model aggregation.

C. FRL with LP for UAV Route Formation

UAVs cooperatively construct a learning model based on the FRL concept, using the following three steps:

- Step 1 - Initialization: According to the application, the cloud server recognizes anticipated data types and training hyperparameters, including learning rate and number of epochs. In addition, the cloud server creates an initial global model G^0. Then, data type requirements and training hyperparameters are announced to UAVs (clients). The cloud server chooses both the learning rate and the number of epochs so as not to deplete UAVs' resources.
- Step 2 - Local model training: Every UAV i initiates gathering recent data and update factors of its local model L_j^i, established on the global model G^j, where j is the present iteration index. Every UAV aims to discover optimal parameters that reduce the loss function. The updated parameters are periodically sent to the cloud server.

- Step 3 - Global Model Aggregation: While obtaining the local models from UAVs, the cloud server collects them and transmits back the simplified model factors to the UAVs. The cloud server aims to reduce the average global loss function specified below:

$$Loss(G^j) = \frac{1}{M}\sum_{i=1}^{m} Loss(L_j^i) \quad (1)$$

where M denotes the whole data samples across the UAVs. The local training and aggregation procedures are recurrent till a preferred accuracy is reached or the loss function converges.

The FRLP mechanism utilizes an FL-based multi-agent RL with an LP submodule to optimize route preparation according to linear constraints, such as UAV distance limit, connectivity, and maximum UAV travels per round. The RL process is formalized by the Markov decision process that executes and learns perfect actions using RL and gives rewards (RD). In RL, the value of the Quality (Q)-table helps in determining the highest action for each state (s), where the function of action (a) value $Q(s, a)$ provides the RDs of existing and future actions is implemented at state s. The agent selects an action s, discovers RD, and moves into the next state ns. Next, the Quality Value (QV) is informed as follows.

$$QV(s, a) =$$
$$(1 - \omega)QV(s, a) + \omega\{RD + \gamma * QV(ns, a) \quad (2)$$

where ω indicates the rate of learning, γ defines the future RD discount element. The RD value is calculated as:

$$RD = \gamma(E + PQ + QL) \quad (3)$$

The UAV locally trained RL model is based on SP state attributes such as energy (E), queued packets (PQ), and quality of link (QL). QL represents a measure of how reliably and efficiently a UAV can exchange data with a BS, influenced by factors such as successful packet delivery. PQ denotes the waiting data packets that are stored in a buffer. Actions denote the UAV visiting the next stopping point. However, RL-based UAV paths may impose limitations such as the length of the total path and the number of scheduled SP nodes. To solve these issues, the proposed system utilizes LP, formulating an LP sub-problem as follows.

Decision variables involve a binary value $Y_{i,j} = 1$, if UAV visits node i at time t, and constraints such as UAV distance limit, maximum UAV travels per round, and connectivity.

$$\sum_{i,t} Y_{i,j} \leq \max\_visits \quad (4)$$

UAV distance is linearized through LP relaxation connectivity. The objective is to maximize the sum of routing values weighted by reward values.

This LP is resolved locally by every UAV, altering RL outputs to meet restrictions efficiently. Then the locally trained model is forwarded to the central server through the edge server for global model collection utilizing FL. Figure 2 shows a flowchart of the FRLP mechanism.

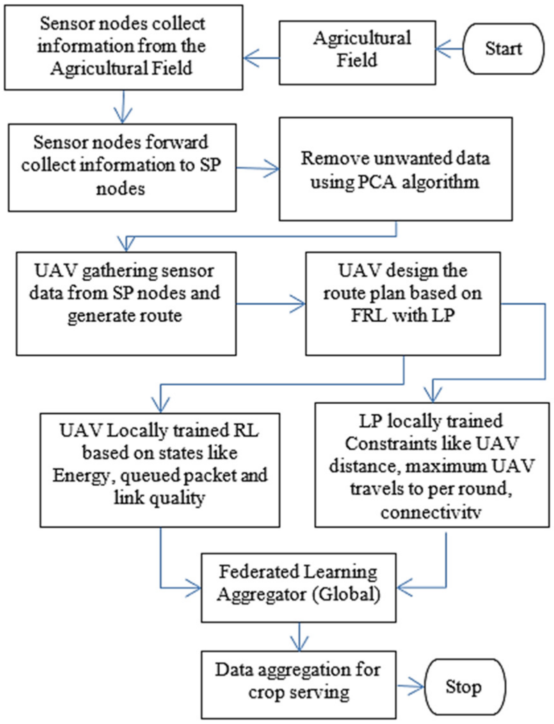

Fig. 2. Flowchart of the FRLP mechanism.

The distinct local models are joined to a global one that embeds combined learning across the network, sustaining the privacy of every local dataset. Weight collection is functional through the central server, which gives greater significance to inconsistent and critical data. The updated model is sent to the edge server after the model update, which enhances decision-making for the next data collection round. Finally, the cloud server collects aggregate data from edge servers and then decides about water irrigation, pest identification, and spraying fertilizer based on sensor information. The outage probability represents the chance that a UAV's link fails to meet the smallest performance thresholds. The proposed system provides a lower outage probability, making it more stable, energy-efficient, and reliable for UAV-assisted smart agriculture functions.

IV. RESULTS AND DISCUSSION

Network Simulator-3 (NS3) is a well-known simulator that was utilized to guide the experiments. The proposed system utilizes a modified NIST dataset, using 150 sensor nodes in a 100×300 m^2 observing region. The simulation setup was run several times and took average values, shown in Figures 3-6. Here, the initial energy for every sensor is 1 J and the simulation topology running time is 200 s. Every sensor node communication range is 30 m, and the data packet size is 512 bytes. Figure 3 shows the energy utilization of DDQN, POSA and FRLP versus UAV rounds in an agricultural field.

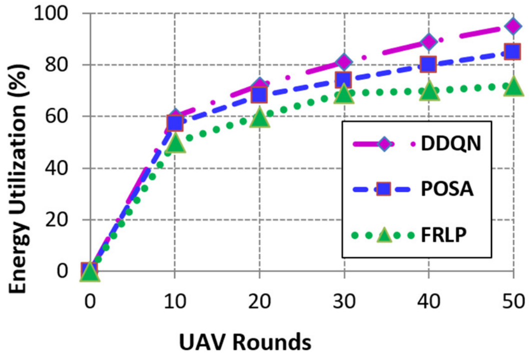

Fig. 3. Energy utilization of DDQN, POSA, and FRLP versus UAV rounds.

The FRL with LP dynamically enhances resource utilization, comprising the reduction of terminated data communication, sensor initiation, and handling UAV paths. The results demonstrate that the proposed FRLP method utilizes UAVs and consumes less energy than DDQN and POSA. The proposed FRLP reduces 20% energy utilization than a POSA and 21% energy utilization than DDQN, demonstrating that FRL with the LP algorithm efficiently reduces needless UAV flying and controls sensor data communication. Figure 4 shows the model convergence in terms of average loss for training rounds in all fields. Loss ratio refers to a key performance metric in wireless communication, which measures the fraction of communicated data packets that fail to effectively reach the receiver.

Initially, a high loss is observed in UAV rounds for FRLP, POSA, and DDQN, which progressively drops as they learn from the data. The FRLP showed less loss compared to POSA and DDQN, suggesting that the UAV performs better data collection from the SP nodes. In FRLP, both SP nodes and UAVs cause the lowest loss values and earlier convergence. The greater loss and slower convergence are due to the more training encounters for sensor nodes and UAVs. The UAVs accurately determine crop growth due to some parameters, including inconsistencies in environmental situations, while the IoT devices exposed deviations in several parameters that delayed model convergence. Figure 5 demonstrates the results of the success ratio for DDQN, POSA, and FRLP versus UAV rounds in the agricultural field.

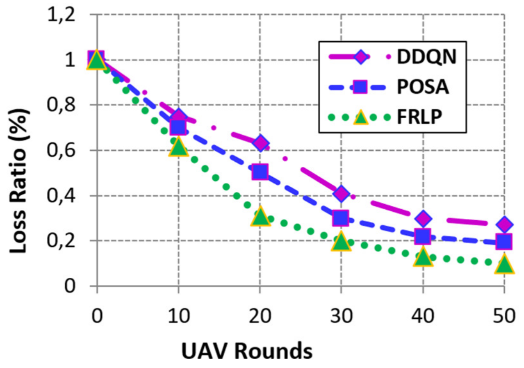

Fig. 4. Loss ratio of DDQN, POSA, and FRLP versus UAV rounds.

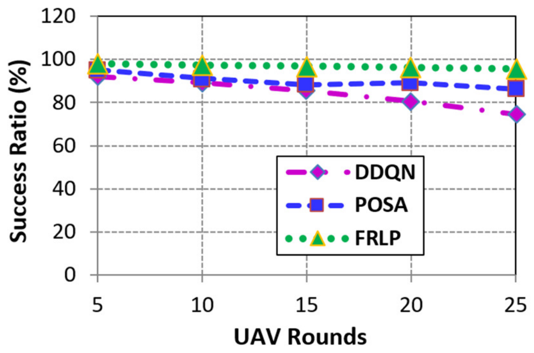

Fig. 5. Success ratio of DDQN, POSA, and FRLP versus UAV rounds.

From this figure, the percentage of the success rate for sensors, SPs, and UAVs ranges from 98.13 to 95.71%, indicating better performance. The small success ratio denotes environmental intrusion, transmission delay, and insufficient task assignments. Concurrently, DDQN and POSA have success rate percentages from 92.14 to 74.42% and 95 to 86% for 25 UAV rounds. The results show that the UAV visiting the SP nodes using FRL with LP enhanced the performance in a smart agriculture environment.

Figure 6 illustrates the delay of the DDQN, POSA, and FRLP mechanisms versus UAV rounds in an agricultural field. The proposed system utilizes SP nodes and UAVs to minimize the delay in collecting data from the agricultural field. In addition, the proposed FRLP system utilizes FRL with LP to reduce UAV flying time. SP nodes remove duplicated data using the PCA algorithm. In the plot, the delays of the DDQN and POSA mechanisms reach 2.3681 and 2.36 s, as these approaches do not build an efficient UAV path for collecting data from the field. In contrast, FRLP reached below 1.5 s in 20 UAV rounds to collect data.

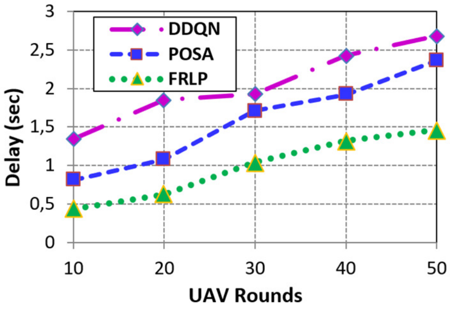

Fig. 6. Delays of DDQN, POSA, and FRLP versus UAV rounds.

V. CONCLUSIONS

UAVs are used in agricultural fields to improve coverage, energy efficiency, and cultivation. This study employed FRL with LP to improve UAV-enabled smart agriculture. The objective is to design a UAV route to minimize time and improve energy efficiency in precision agriculture. The FRLP approach provides promising technology to improve UAV-based data collection in agriculture, merging the cooperative learning capabilities of FRL with the optimization strengths of LP to reduce flight time and energy utilization and offer timely data collection from SP nodes. The simulation results demonstrate that the proposed system minimizes the delay in the UAV route and the UAV energy utilization. Although FRLP improves adaptability, resource efficiency, and privacy, it faces challenges in computation, communication, scalability, convergence speed, and security. Addressing these limitations is essential for real-time deployment in smart agriculture. Future work may focus on real-time adaptation and the use of cryptography algorithms to improve data security.

REFERENCES

[1] Md. N. Mowla, N. Mowla, A. F. M. S. Shah, K. M. Rabie, and T. Shongwe, "Internet of Things and Wireless Sensor Networks for Smart Agriculture Applications: A Survey," *IEEE Access*, vol. 11, pp. 145813–145852, 2023, https://doi.org/10.1109/ACCESS.2023.3346299.

[2] G. Mohyuddin, M. A. Khan, A. Haseeb, S. Mahpara, M. Waseem, and A. M. Saleh, "Evaluation of Machine Learning Approaches for Precision Farming in Smart Agriculture System: A Comprehensive Review," *IEEE Access*, vol. 12, pp. 60155–60184, 2024, https://doi.org/10.1109/ACCESS.2024.3390581.

[3] P. Killeen, C. Lin, F. Li, I. Kiringa, and T. Yeap, "IoT-Based Smart Farming Architecture Using Federated Learning: a Nitrous Oxide Emission Prediction Use Case," *ACM Journal on Computing and Sustainable Societies*, vol. 3, no. 2, Feb. 2025, Art. no. 12, https://doi.org/10.1145/3723039.

[4] M. A. Al-Mashhadani, M. M. Hamdi, and A. S. Mustafa, "Role and challenges of the use of UAV-aided WSN monitoring system in large-scale sectors," in *2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, June 2021, pp. 1–5, https://doi.org/10.1109/HORA52670.2021.9461292.

[5] S. Archana and V. Jayapradha, "Optimized Cluster-Based Communication in MWSN using fuzzy neural Networks and Crow Search Algorithm,", *International Journal of Advances in Signal and Image Sciences*, vol. 11, no. 1, pp. 80–93, June 2025, https://doi.org/10.29284/ijasis.11.1.2025.80-93.

[6] P. K. R. Maddikunta *et al.*, "Unmanned Aerial Vehicles in Smart Agriculture: Applications, Requirements, and Challenges," *IEEE Sensors Journal*, vol. 21, no. 16, pp. 17608–17619, Dec. 2021, https://doi.org/10.1109/JSEN.2021.3049471.

[7] K. Karam, A. Mansour, M. Khaldi, B. Clement, and M. Ammad-Uddin, "UAV Path Optimization for WSN in Smart Agriculture," *IEEE Access*, vol. 13, pp. 87526–87544, 2025, https://doi.org/10.1109/ACCESS.2025.3569642.

[8] M. L. Betalo, S. Leng, A. M. Seid, H. N. Abishu, A. Erbad, and X. Bai, "Dynamic Charging and Path Planning for UAV-Powered Rechargeable WSNs Using Multi-Agent Deep Reinforcement Learning," *IEEE Transactions on Automation Science and Engineering*, vol. 22, pp. 15610–15626, 2025, https://doi.org/10.1109/TASE.2025.3558945.

[9] J. Tursunboev, Y. S. Kang, S. B. Huh, D. W. Lim, J. M. Kang, and H. Jung, "Hierarchical Federated Learning for Edge-Aided Unmanned Aerial Vehicle Networks," *Applied Sciences*, vol. 12, no. 2, Jan. 2022, Art. no. 670, https://doi.org/10.3390/app12020670.

[10] A. O. Hashesh, S. Hashima, R. M. Zaki, M. M. Fouda, K. Hatano, and A. S. T. Eldien, "AI-Enabled UAV Communications: Challenges and Future Directions," *IEEE Access*, vol. 10, pp. 92048–92066, 2022, https://doi.org/10.1109/ACCESS.2022.3202956.

[11] I. Donevski, N. Babu, J. J. Nielsen, P. Popovski, and W. Saad, "Federated Learning With a Drone Orchestrator: Path Planning for Minimized Staleness," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1000–1014, 2021, https://doi.org/10.1109/OJCOMS.2021.3072003.

[12] R. Zhagypar, N. Kouzayha, H. El Sawy, H. Dahrouj, and T. Y. Al-Naffouri, "UAV-Assisted Unbiased Hierarchical Federated Learning: Performance and Convergence Analysis," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 3, pp. 420–447, 2025, https://doi.org/10.1109/TMLCN.2025.3546181.

[13] M. A. Jaleel *et al.*, "An Energy-Efficient Hybrid LEACH Protocol that Enhances the Lifetime of Wireless Sensor Networks," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19364–19369, Feb. 2025, https://doi.org/10.48084/etasr.8458.

[14] X. Yang *et al.*, "Adaptive UAV-Assisted Hierarchical Federated Learning: Optimizing Energy, Latency, and Resilience for Dynamic Smart IoT." arXiv, Mar. 24, 2025, https://doi.org/10.48550/arXiv.2503.06145.

[15] N. Alasbali *et al.*, "IoT-UAV-Enabled Intelligent Resource Management in Low-Carbon Smart Agriculture Using Federated Reinforcement Learning," *IEEE Transactions on Consumer Electronics*, vol. 71, no. 2, pp. 6933–6941, Feb. 2025, https://doi.org/10.1109/TCE.2025.3572552.

[16] M. Akbari, A. Syed, W. S. Kennedy, and M. Erol-Kantarci, "AoI-Aware Energy-Efficient SFC in UAV-Aided Smart Agriculture Using Asynchronous Federated Learning," *IEEE Open Journal of the Communications Society*, vol. 5, pp. 1222–1242, 2024, https://doi.org/10.1109/OJCOMS.2024.3363132.

[17] J. Huang, M. Zhang, J. Wan, Y. Chen, and N. Zhang, "Joint Data Caching and Computation Offloading in UAV-Assisted Internet of Vehicles via Federated Deep Reinforcement Learning," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 11, pp. 17644–17656, Aug. 2024, https://doi.org/10.1109/TVT.2024.3429507.

[18] D. Deepalakshmi and B. Pushpa, "Cognitive Fish Swarm Optimization for Multi-Objective Routing in IoT-based Wireless Sensor Networks utilized in Greenhouse Agriculture," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19472–19477, Feb. 2025, https://doi.org/10.48084/etasr.9203.

[19] S. Arumugam and F. Stomp, "Optimizing Autonomous Vehicle Path Planning Using Reinforcement Learning and Dynamic Mapping," *International Journal of Advances in Signal and Image Sciences*, vol. 10, no. 2, pp. 58–68, Dec. 2024, https://doi.org/10.29284/ijasis.10.2.2024.58-68.