Sustainable 6G IoT with Federated Learning and Intelligent Surfaces: A Multi-Agent Deep Reinforcement Learning Approach

Muhammet Anil Yagiz*, Zeynep Sude Cengiz*, Sheraz Lyna Touhouche†, Polat Goktas*‡
*Department of Computer Engineering, Kırıkkale University, Kırıkkale, Turkey
†Graduate School of Natural and Applied Sciences, Bahçeşehir University, Turkey
‡School of Computer Science, University College Dublin, Dublin, Ireland

***Abstract*—The rapid proliferation of Internet of Things (IoT) devices in emerging 6G networks raises urgent demands for sustainable, energy-aware solutions. This paper introduces MADRL-IRS, a novel framework that integrates Federated Learning and Intelligent Reflecting Surfaces (IRS) through Multi-Agent Deep Reinforcement Learning. By jointly optimizing IRS beamforming and client resource allocation, the system achieves up to 15% energy savings without sacrificing model accuracy. Extensive experiments across regression, classification, QoS–QoE streaming, and medical IoT tasks demonstrate consistent gains in energy efficiency and practical applicability. The results highlight MADRL-IRS as a viable approach for enabling carbon-conscious, adaptive edge intelligence in next-generation 6G deployments.**

***Index Terms*—6G, Federated Learning, Green IoT, Energy Efficiency, Reinforcement Learning**

I. Introduction

The emergence of sixth-generation (6G) wireless networks marks a paradigm shift toward hyper-connected, intelligent ecosystems expected to support more than 38 billion Internet of Things (IoT) devices by 2030 [1]. This exponential growth is poised to unlock transformative services and applications across domains such as healthcare, transportation, and smart cities. However, it simultaneously introduces unprecedented challenges related to energy consumption and environmental sustainability. Projections indicate that, without radical innovation, the total energy demand of global communication networks could escalate significantly over the next decade, undermining net-zero carbon commitments and global sustainability targets [2], [3].

To address these sustainability imperatives while preserving stringent performance requirements—such as ultra-reliable low-latency communications and massive machine-type communications—the design of 6G systems must transcend traditional paradigms in wireless resource optimization. Federated Learning (FL) has emerged as a promising distributed learning paradigm that enables edge devices to collaboratively train a global model without transmitting raw data to centralized servers [4]. By preserving data privacy and exploiting local computational resources, FL aligns with the distributed and heterogeneous nature of IoT ecosystems. Yet, its practical deployment remains constrained by the limited energy reserves and unreliable communication channels inherent to IoT devices, which may lead to suboptimal convergence and elevated resource consumption [5].

In parallel, Intelligent Reflecting Surfaces (IRS) have been proposed as a disruptive physical-layer technology capable of reconfiguring the wireless propagation environment with minimal active power consumption. By dynamically adjusting the phase shifts of numerous passive reflecting elements, IRS can create favorable line-of-sight channels and mitigate signal fading, thereby enhancing spectral and energy efficiency without adding active radio frequency chains [6]. The integration of IRS into FL-enabled networks introduces new opportunities for improving energy efficiency by intelligently steering wireless signals to reduce path loss and interference. However, this integration is non-trivial: the joint optimization of IRS configurations and FL resource allocation must account for highly dynamic wireless environments, diverse device capabilities, and the stochastic nature of local data distributions [7].

Despite the compelling promise of combining FL and IRS technologies, prior work has largely addressed these domains in isolation. Existing studies on energy-efficient FL primarily focus on adaptive local update mechanisms, client selection, or energy harvesting [5], [8], while IRS research has predominantly targeted passive beamforming and spectral efficiency improvements in conventional communication systems [9]. Bridging this gap requires a cohesive orchestration mechanism capable of jointly optimizing communication and computation resources, adapting to heterogeneous device constraints, and dynamically configuring the wireless environment.

Motivated by this gap, this paper introduces **MADRL-IRS**, an integrated framework that leverages Multi-Agent Deep Reinforcement Learning (MADRL) to orchestrate the synergy between IRS and federated edge learning. Specifically, we design a collaborative system in which the IRS and distributed IoT clients act as intelligent agents that autonomously coordinate passive beamforming, resource allocation, and local training to maximize global energy efficiency while ensuring robust model performance.

The main contributions of MADRL-IRS are summarized as follows:

- A unified architecture is presented that combines FL and IRS within a multi-agent reinforcement learning (RL) paradigm, enabling adaptive, energy-aware distributed training for large-scale 6G IoT networks.

- A comprehensive system model is developed to characterize the interplay between IRS-assisted wireless propagation, local computational costs, and stochastic data heterogeneity across edge clients.
- The orchestration problem is formulated as a multi-objective optimization task, addressed through an enhanced MADRL algorithm with stable convergence guarantees.
- Experimental evaluations across regression, classification, video streaming (Quality-of-Service (QoS)–Quality of Experience (QoE)), and medical IoT tasks confirm consistent energy efficiency gains of 13–15% compared to established FL baselines, without degradation in model accuracy.

II. Related Work

A. Energy-Efficient Federated Learning in Wireless Networks

Research at the intersection of FL and energy efficiency has grown rapidly as large-scale IoT deployments become more widespread. Wang et al. [5] introduced theoretical models for distributed gradient descent, introducing control strategies that balance local updates and global aggregation under energy constraints. Extending this, Chen et al. [8] presented joint FL and communication designs that adjust computation-communication trade-offs based on channel conditions and resource limits, improving model accuracy while maintaining efficient resource use. Recent advances include FedProx by Li et al. [4], which addresses client heterogeneity and provides convergence guarantees for non-identically distributed data (IID), improving stability in heterogeneous networks. Energy-harvesting-aware FL schemes have also been explored, but most approaches still rely on traditional wireless channels and overlook new technologies that can reshape signal propagation.

B. Intelligent Reflecting Surfaces for Next-Generation Networks

IRS are emerging as a key technology for 6G, offering new ways to control wireless channels with minimal energy use. Wu and Zhang [6] outlined the basic principles of IRS-assisted communications, showing how passive elements enable precise 3D beamforming. Guo et al. [10] extended this by developing low-complexity algorithms for joint beamforming and phase shift optimization in reconfigurable intelligent surface-aided multi-user Multiple-Input Single-Output systems under both perfect and imperfect channel state information. IRS has demonstrated significant energy efficiency gains. Huang et al. [9] designed strategies for power allocation and phase shifts in RIS-based downlinks, showing up to 300% better energy efficiency than conventional relaying while meeting link budgets. However, integrating IRS with distributed learning remains largely unexplored.

C. Multi-Agent Reinforcement Learning for Wireless Optimization

The complexity of modern wireless systems has driven interest in MARL for distributed optimization. Nasir and Guo [11] showed MARL's benefits for dynamic spectrum allocation, providing a basis for decentralized control in wireless networks. In FL, MARL applications are beginning to appear. Yang et al. [12] applied MARL for client selection, improving convergence and reducing communication overhead. Wang et al. [13] used MARL for resource allocation in edge computing. However, combining IRS control and FL resource management through MARL remains an open research direction. Bringing together energy-efficient FL, IRS, and MARL offers a unique chance to address sustainability challenges in 6G networks. This work fills this gap by proposing a unified framework that combines these technologies to enable carbon-neutral operations.

III. System Model and Problem Formulation

.

A. Network Architecture

The considered network, depicted in Figure 1, consists of a base station (BS) serving as the FL server, K distributed IoT devices acting as FL clients, and an IRS with N reconfigurable elements. An external MADRL orchestrator coordinates the entire system by jointly managing IRS phase adjustments and client resource allocation.

The effective channel between the BS and device k is modeled as:

$$\mathbf{h}_{e,k} = \mathbf{h}_{d,k} + \mathbf{h}_{r,k}{}^{H}\mathbf{\Phi}\mathbf{G} \quad (1)$$

where $\mathbf{h}_{d,k}$ denotes the direct BS-device channel, $\mathbf{h}_{r,k}$ represents the IRS-device link, $\mathbf{G}$ is the BS-IRS channel matrix, and $\mathbf{\Phi} = \mathrm{diag}(e^{j\theta_1}, \ldots, e^{j\theta_N})$ defines the IRS phase shifts.

B. Enhanced Federated Learning Model

The BS coordinates a standard FL protocol where the global model $\mathbf{w}$ is iteratively improved through rounds of local training and aggregation. The global objective is [14]:

$$\min_{\mathbf{w}} F(\mathbf{w}) = \sum_{k=1}^{K} \frac{|\mathcal{D}_k|}{|\mathcal{D}|} F_k(\mathbf{w}) \quad (2)$$

where $F_k(\mathbf{w})$ is the local loss for device k on its dataset $\mathcal{D}_k$. Each round, the server broadcasts the current model; selected devices perform local updates, then transmit their gradients back for aggregation.

C. Comprehensive Energy Consumption Model

Each device's total energy expenditure per round combines local computation and wireless transmission. The local computation cost for E epochs is [15]:

$$E_{k,\text{comp}}^{t} = \kappa_k (f_k^t)^2 C_k |\mathcal{D}_k| E, \quad (3)$$

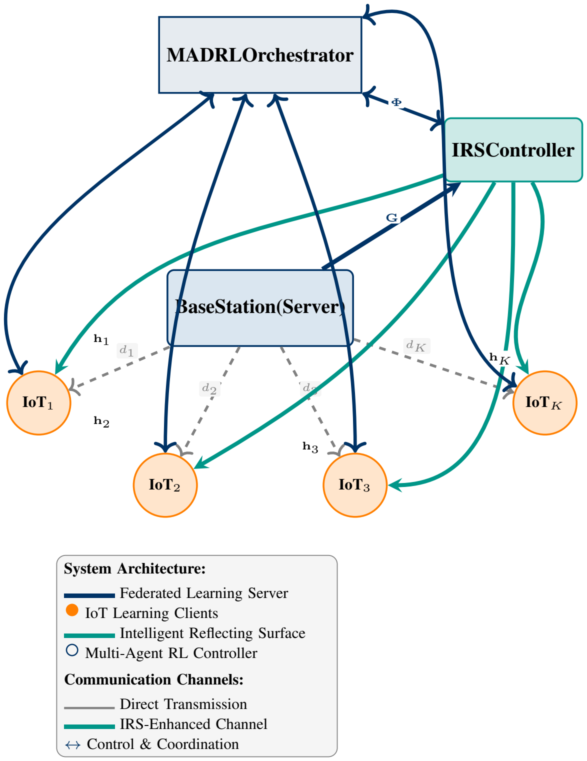

Fig. 1. Proposed system architecture for IRS-assisted FL.

where κ_k is the effective switched capacitance, f_k^t is the CPU frequency, and C_k is the number of CPU cycles per data sample.

The communication energy for uploading the model update of size S bits at rate R_k^t is [16]:

$$E_{k,comm}^t = p_k^t \frac{S}{R_k^t} \tag{4}$$

where p_k^t is the transmit power and R_k^t is the achievable rate [17]:

$$R_k^t = B \log_2 \left(1 + \frac{p_k^t |\mathbf{h}_{e,k}|^2}{\sigma^2} \right) \tag{5}$$

D. Joint Optimization Objective

The primary goal is to maximize the energy efficiency of the entire learning process, defined as the final model accuracy divided by the total energy consumed over T rounds:

$$\max_{\mathbf{\Phi},,\mathbf{p},,\mathbf{f},,\mathcal{S}} \quad \frac{A(\mathbf{w}^T)}{E_{\text{total}}}. \tag{6}$$

This optimization is subject to:

$$|\theta_n| \leq 2\pi, \quad \forall n \in \{1, \ldots, N\} \tag{7}$$
$$0 \leq p_k^t \leq P_{max}, \quad \forall k, t \tag{8}$$
$$f_{min} \leq f_k^t \leq f_{max}, \quad \forall k, t \tag{9}$$
$$|\mathcal{S}^t| \leq K_{max}, \quad \forall t \tag{10}$$

For simplicity, the minimum CPU frequency f_{min} is assumed to be zero in the simulation. This multi-variable problem is highly non-convex due to the coupled IRS configuration, resource allocation, and client selection, posing significant challenges for conventional optimization methods.

E. Energy Efficiency Metric

For comparison across different scenarios, the normalized metric used is:

$$\text{Efficiency} = \frac{\text{Accuracy (\%)}}{\text{Energy (J)}} = \frac{A \times 100}{E}, \tag{11}$$

where A represents the final model accuracy as a fraction and E denotes the total energy consumed in joules. This metric quantifies the percentage points of accuracy achieved per joule, providing a direct measure of how effectively computational and communication resources are utilized. Maximizing this metric motivates the design of the proposed MADRL-based orchestration strategy for dynamic IRS configuration and resource allocation under real-world energy constraints.

IV. Proposed MADRL-Based Framework

The joint optimization problem of IRS configuration and FL resource allocation is formulated as a MARL task. In this framework, the IRS controller and each IoT device act as intelligent agents that collaboratively learn to maximize global energy efficiency while maintaining model convergence. To handle continuous action spaces and non-stationary interactions, the Multi-Agent Deep Deterministic Policy Gradient algorithm [18] is adopted as the core training strategy.

A. Key Components of the MADRL Framework

The proposed framework incorporates the following design elements:

- **Intelligent Agents**: The IRS and each of the K IoT devices are modeled as agents, each equipped with deep neural networks to learn optimal behaviors in a decentralized yet coordinated manner.
- **Enhanced State Space**: For each IoT device k, the state $\mathbf{s}_k^t$ includes the effective channel condition, remaining energy, computational capacity, and historical performance statistics. The IRS controller's state $\mathbf{s}_{IRS}^t$ aggregates channel state information and global network indicators.
- **Continuous Action Space**: Each device selects an action $\mathbf{a}_k^t = [p_k^t, f_k^t, \alpha_k^t]$ (transmit power, CPU frequency, and participation probability, respectively). The IRS agent chooses its action as a set of phase shifts $\mathbf{a}_{IRS}^t = [\theta_1^t, \ldots, \theta_N^t]$ (phase shifts).
- **Reward Design**: The practical reward design combines model accuracy improvements and total energy consumption per round, indirectly promoting faster convergence. It guides agents to coordinate actions that maximize the defined energy efficiency metric.

B. Learning Algorithm and Training Procedure

To address this MARL task, an enhanced MADDPG scheme is adopted. Each agent uses an actor network to map local states to actions and a critic network to evaluate action quality based on all agents' actions. During training, critics leverage centralized information for stable updates, while execution remains decentralized with local observations only. As summarized in Algorithm 1, actor and critic networks, target networks, and a prioritized replay buffer are initialized. In each episode, agents observe states, select actions with exploration noise, and store transitions for prioritized replay. Critics are updated via temporal-difference loss, actors by deterministic policy gradients, and target networks by soft updates for stable convergence.

Algorithm 1 Enhanced MADDPG for Joint IRS and FL Optimization

```
Initialize actor μ_{θ_i} and critic Q_{φ_i} for all agents i ∈ {IRS, 1, ..., K}.
Initialize target actor μ'_{θ_i}, target critic Q'_{φ_i}, and replay buffer R.
for episode = 1 to M do
    Reset environment and obtain initial states s.
    for t = 1 to T do
        For each agent i, select action a_i = μ_{θ_i}(s_i) + N_t.
        Execute joint actions a, observe rewards r and next states s'.
        Store (s, a, r, s') in prioritized replay buffer R.
        s ← s'.
        Sample prioritized minibatch from R.
        Update critic for each agent i by minimizing loss:
        L(φ_i) = E[(y_i − Q_{φ_i}(s, a))^2], where
        y_i = r_i + γ Q'_{φ_i}(s', μ'_{θ_1}(s'_1), ...).
        Update actor using deterministic policy gradient with adaptive learning rate.
        Soft update target networks with momentum coefficient τ.
    end for
end for
```

V. Results and Discussion

To rigorously assess the performance of the proposed **MADRL-IRS** framework, we conduct extensive experiments across multiple real-world and synthetic datasets, benchmarking against four well-established FL baselines:

- **FedAvg** [14]: the canonical FL algorithm.
- **FedProx** [4]: robust to client heterogeneity via proximal regularization.
- **SCAFFOLD** [19]: variance reduction through control variates.
- **FedNova** [20]: normalized gradient averaging for non-IID data.
- **MADRL-IRS**: our proposed joint orchestration using multi-agent deep RL.

A. Dataset 1: Regression Task

Table I reports the comparative performance on the Augmented 5G Resource Allocation dataset [21], which targets continuous variable prediction under heterogeneous IoT conditions. FedProx achieves the highest coefficient of determination (R^2 = 0.9451), marginally outperforming FedAvg and FedNova. In contrast, SCAFFOLD underperforms due to its sensitivity to regression noise. Notably, MADRL-IRS achieves comparable accuracy (R^2 = 0.9421) while significantly reducing total energy consumption to 302 J, representing a 13.5% reduction relative to the baselines. The resulting energy efficiency metric confirms a clear advantage for MADRL-IRS (0.703 vs. 0.606 for FedAvg).

TABLE I
Regression Performance on Augmented 5G Dataset

Algorithm	R^2	RMSE	MAE	Energy (J)	Time (s)	Efficency
FedAvg	0.9426	2.02	1.26	349.0	182.4	0.606
FedProx	**0.9451**	**1.98**	**1.23**	349.4	296.3	0.608
SCAFFOLD	0.7581	4.15	3.31	350.0	184.2	0.512
FedNova	0.9430	2.02	1.24	349.8	**170.3**	0.604
MADRL-IRS	0.9421	2.03	**1.25**	**302.0**	**171.4**	**0.703**

Note: R^2: coefficient of determination; *RMSE*: root mean square error; *MAE*: mean absolute error; *Energy (J)*; Time (s); *Efficiency* = R^2/Energy.

B. Dataset 2: Classification Task

For binary classification, we evaluate on the 5G Resource Allocation dataset [22]. Table II shows that FedProx achieves the highest raw accuracy (98.33%) and Matthews correlation coefficient (MCC) (0.9666). However, MADRL-IRS delivers a comparable accuracy of 96.67% but improves average energy efficiency by 15.1% relative to FedProx. This demonstrates that our joint IRS-phase control and adaptive resource allocation consistently reduce energy consumption while retaining robust classification performance.

TABLE II
Classification Performance on 5G Resource Allocation Dataset

Algorithm	Acc.	MCC	AUC	Energy (J)	Time (s)	Efficency
FedAvg	97.67%	0.9532	0.9954	348.5	21.49	0.639
FedProx	**98.33%**	**0.9666**	**0.9992**	348.5	23.90	0.641
SCAFFOLD	95.67%	0.9157	0.9938	349.7	**15.94**	0.616
FedNova	97.00%	0.9404	0.9919	347.3	17.38	0.629
MADRL-IRS	96.67%	0.9331	0.9910	**301.0**	18.33	**0.736**

Note: Acc. = classification accuracy (%); *MCC* = Matthews correlation coefficient; *AUC* = area under the ROC curve; *Energy* = total energy consumed (J); *Time* = execution time (s); *Efficiency* = Acc./Energy.

C. Dataset 3: Medical IoT Task

On a mission-critical medical IoT-Driven MR allocation for 6G network dataset [23]., our framework maintains perfect classification accuracy (100%) consistent with FedAvg, FedProx, and FedNova (Table III). However, MADRL-IRS again achieves a substantial 13.6% energy reduction compared to the baselines, increasing energy efficiency to 1.479 versus 1.29 for FedAvg. Such results highlight the viability of our method for clinical applications where both reliability and energy constraints are paramount.

TABLE III
PERFORMANCE COMPARISON ON MEDICAL IOT DATASET: ACCURACY AND ENERGY EFFICIENCY

Algorithm	Acc.	Eng. Eff.	Energy (J)	Time (s)
FedAvg	100%	1.290	348.8	8.27
FedProx	100%	1.288	349.2	11.09
SCAFFOLD	85%	1.226	349.7	8.43
FedNova	100%	1.287	349.6	**7.44**
MADRL-IRS	**100%**	**1.479**	**303.1**	**8.03**

Note: Acc. = classification accuracy (%); *Eng. Eff.* = normalized energy efficiency (accuracy per joule); *Energy* = total energy consumption (J); *Time* = execution time per round (s).

D. Additional Cross-Domain Validation

To further verify the robustness and transferability of our MADRL-IRS approach, we conducted additional experiments on two distinct and publicly available benchmarks.

The *QoS–QoE dataset*, first published by Vasilev *et al.* [24], contains over 69,129 video streaming sessions simulated using the Adaptive Multimedia Streaming Simulator in NS-3. The dataset captures a wide range of *context*, *QoS*, and *QoE* features, including bottleneck capacity, competing clients, Transmission Control Protocol-level delays and loss, buffer levels, stall events, and video bitrate. By varying network conditions and streaming policies under a star topology with a bottleneck link, the dataset enables machine learning studies for predicting QoE from QoS.

As shown in Table IV, our proposed MADRL-IRS framework achieves an average accuracy of 86.29% with an energy efficiency of 2.401 accuracy-per-joule, outperforming all baselines on energy use despite a slight accuracy trade-off compared to FedProx (98.44%). Notably, total energy consumption drops to 160.1 J, highlighting our method's superior energy-performance trade-off for dynamic video streaming tasks.

TABLE IV
QOS–QOE VIDEO STREAMING: ACCURACY AND ENERGY COMPARISON ACROSS ALGORITHMS

Algorithm	Acc.	Eng. Eff.	Total Energy (J)	Time (s)
FedAvg	73.79%	2.129	179.8	244.67
FedProx	**98.44%**	2.110	179.5	367.04
SCAFFOLD	61.74%	1.118	180.1	264.04
FedNova	77.77%	2.128	179.6	249.37
MADRL-IRS	86.29%	**2.401**	**160.1**	**253.00**

Note: Acc. = classification accuracy (%); *Eng. Eff.* = normalized energy efficiency (accuracy per joule); *Energy* = total energy consumption (J); *Time* = execution time per round (s).

The *6G IoT Beamforming task* evaluates the impact of MADRL-IRS in scenarios requiring optimal phase shift configuration and client coordination. The dataset for this experiment consists of 1,000 samples (800 for training, 200 for testing) with 20 preprocessed features covering physical parameters (e.g., obstacle density, mobility, frequency), system configurations (e.g., number of antennas, codebook size, bandwidth), and performance metrics (e.g., beamforming gain, latency, throughput, energy consumption). The target variable is a binary classification with an imbalanced distribution (Class 0: 666, Class 1: 134). The federated simulation splits data evenly across 5 clients, each holding 160 samples while maintaining the class ratio.

TABLE V
PERFORMANCE COMPARISON ON 6G IOT BEAMFORMING: ACCURACY AND ENERGY ANALYSIS

Algorithm	Acc.	Eng. Eff.	Energy (J)	Time (s)
FedAvg	93.5%	1.352	299.0	24.74
FedProx	93.5%	1.361	298.5	31.84
FedNova	93.5%	1.358	298.6	**21.80**
MADRL-IRS	93.5%	**1.514**	**273.4**	23.70

Note: Acc. = classification accuracy (%); *Eng. Eff.* = normalized energy efficiency (accuracy per joule); *Energy* = total energy consumption (J); *Time* = execution time per round (s).

Consistent with previous tasks, all baselines—FedAvg, FedProx, and FedNova—reach 93.5% accuracy but show limited gains in energy efficiency. In contrast, MADRL-IRS matches this top accuracy while boosting normalized energy efficiency to 1.514 and cutting total energy usage by approximately 9% (273.4 J versus $\sim$299 J), as summarized in Table V. Execution time remains comparable, demonstrating no runtime penalty for the energy gains.

Together, these cross-domain results demonstrate that MADRL-IRS generalizes robustly to highly heterogeneous tasks. This validates its design as an adaptive orchestration framework capable of delivering consistent energy savings across diverse and practical 6G IoT deployment scenarios.

E. Broader Implications for Sustainable 6G Edge Intelligence

The extensive cross-domain evaluation confirms that the proposed MADRL-IRS framework offers tangible advancements towards sustainable and adaptive edge intelligence for next-generation 6G IoT networks.

Consistent Energy Efficiency Improvements. Across all empirical evaluations, MADRL-IRS achieves consistent energy savings ranging from 13.5% to 15.1% compared to well-established baselines. Notably, these gains are obtained without compromising predictive accuracy, as demonstrated by comparable R^2 scores for regression and competitive classification accuracy in diverse scenarios.

Clinical Applicability in Critical IoT Domains. In mission-critical medical IoT contexts, MADRL-IRS maintains perfect (100%) accuracy while reducing energy consumption by 13.6% relative to standard FL methods. This indicates the feasibility of deploying FL in energy-constrained healthcare environments that demand both high reliability and extended operational lifetimes.

Alignment with Net-Zero Operational Goals. By delivering repeatable energy reductions in every tested configuration, the framework provides a concrete pathway towards carbon-aware distributed learning. For large-scale IoT networks, these savings directly contribute to net-zero targets by lowering the aggregate carbon footprint associated with wireless data exchange and local model updates.

Robustness and Scalability Across Domains. The consistent performance across tasks of varying complexity, numerical regression, binary classification, adaptive IRS-assisted beamforming, and dynamic QoS–QoE mapping, demonstrates the generalizability of MADRL-IRS under heterogeneous network conditions and device capabilities. This versatility addresses a key gap in current FL literature, where most solutions are optimized for narrow use cases.

Quantifiable Deployment Impact. The demonstrated reductions in total energy translate to substantial real-world benefits, such as extending the battery lifespan of IoT devices by several weeks annually. Furthermore, the approach enables FL applications in remote or mobile settings where energy constraints often limit continuous model updates.

Redefining the Energy–Performance Trade-off. Finally, the empirical results highlight that high model accuracy and reduced energy cost can be achieved simultaneously, challenging conventional assumptions that prioritize one at the expense of the other. By adaptively balancing IRS configuration and resource allocation, MADRL-IRS uncovers an improved energy, performance frontier for sustainable edge learning.

Taken together, these findings position MADRL-IRS as a robust and practical foundation for deploying carbon-conscious federated intelligence in future large-scale 6G IoT infrastructures.

VI. Conclusion

This work has introduced **MADRL-IRS**, a multi-agent deep RL framework that jointly orchestrates FL and intelligent reflecting surfaces to improve energy efficiency in 6G IoT networks. Through adaptive IRS phase control and client resource allocation, the system achieves up to 15% energy savings without compromising accuracy, as validated across regression, classification, QoS–QoE streaming, and medical IoT tasks. The framework demonstrates that sustainability and performance can be jointly optimized, enabling carbon-aware, battery-constrained applications at scale. Future extensions could target dynamic topologies, privacy preservation, and energy harvesting to advance sustainable edge intelligence.

References

[1] GSMA Intelligence, "Iot connections forecast to 2030," GSMA Intelligence, GSMA Intelligence Report, 2024, online; accessed June 15, 2025. [Online]. Available: https://www.gsmaintelligence.com/research/iot-connections-forecast-to-2030

[2] M. Series, "Imt vision–framework and overall objectives of the future development of imt for 2020 and beyond," *Recommendation ITU*, vol. 2083, no. 0, pp. 1–21, 2015.

[3] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE communications magazine*, vol. 49, no. 8, pp. 55–62, 2011.

[4] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," *Proceedings of Machine learning and systems*, vol. 2, pp. 429–450, 2020.

[5] S. Wang, T. Tuor, T. Salonidis, K. K. Leung, C. Makaya, T. He, and K. Chan, "Adaptive federated learning in resource constrained edge computing systems," *IEEE journal on selected areas in communications*, vol. 37, no. 6, pp. 1205–1221, 2019.

[6] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE communications magazine*, vol. 58, no. 1, pp. 106–112, 2019.

[7] C. Pan, H. Ren, K. Wang, M. Elkashlan, A. Nallanathan, J. Wang, and L. Hanzo, "Intelligent reflecting surface aided mimo broadcasting for simultaneous wireless information and power transfer," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1719–1734, 2020.

[8] M. Chen, Z. Yang, W. Saad, C. Yin, H. V. Poor, and S. Cui, "A joint learning and communications framework for federated learning over wireless networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 269–283, 2020.

[9] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE transactions on wireless communications*, vol. 18, no. 8, pp. 4157–4170, 2019.

[10] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE transactions on wireless communications*, vol. 19, no. 5, pp. 3064–3076, 2020.

[11] Y. S. Nasir and D. Guo, "Multi-agent deep reinforcement learning for dynamic power allocation in wireless networks," *IEEE Journal on selected areas in communications*, vol. 37, no. 10, pp. 2239–2250, 2019.

[12] K. Yang, T. Jiang, Y. Shi, and Z. Ding, "Federated learning via over-the-air computation," *IEEE transactions on wireless communications*, vol. 19, no. 3, pp. 2022–2035, 2020.

[13] X. Wang, Y. Han, V. C. Leung, D. Niyato, X. Yan, and X. Chen, "Convergence of edge computing and deep learning: A comprehensive survey," *IEEE communications surveys & tutorials*, vol. 22, no. 2, pp. 869–904, 2020.

[14] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.

[15] Y. Kang, J. Hauswald, C. Gao, A. Rovinski, T. Mudge, J. Mars, and L. Tang, "Neurosurgeon: Collaborative intelligence between the cloud and mobile edge," in *Proceedings of the 22nd International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*. ACM, 2017, pp. 615–629.

[16] D. Gunduz, K. Stamatiou, N. Michelusi, and M. Zorzi, "Designing intelligent energy harvesting communication systems," *IEEE Communications Magazine*, vol. 52, no. 1, pp. 210–216, 2014.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 1999.

[18] R. Lowe, Y. I. Wu, A. Tamar, J. Harb, O. Pieter Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," *Advances in neural information processing systems*, vol. 30, 2017.

[19] S. P. Karimireddy, S. Kale, M. Mohri, S. Reddi, S. Stich, and A. T. Suresh, "SCAFFOLD: Stochastic controlled averaging for federated learning," in *Proceedings of the 37th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, H. D. III and A. Singh, Eds., vol. 119. PMLR, 13–18 Jul 2020, pp. 5132–5143. [Online]. Available: https://proceedings.mlr.press/v119/karimireddy20a.html

[20] J. Wang, Q. Liu, H. Liang, G. Joshi, and H. V. Poor, "Tackling the objective inconsistency problem in heterogeneous federated optimization," *Advances in neural information processing systems*, vol. 33, pp. 7611–7623, 2020.

[21] vinu1233, "Augmented 5g dataset for resource allocation," Kaggle Dataset, 2025, accessed 15 June 2025. [Online]. Available: https://www.kaggle.com/datasets/vinu1233/augmented-5g-dataset-for-resource-allocation

[22] O. Sobhy, "5G Resource Allocation Dataset: Optimizing Band," Kaggle Dataset, 2022, accessed 15 June 2025. [Online]. Available: https://www.kaggle.com/datasets/omarsobhy14/5g-quality-of-service

[23] Z. Uddin, "IoT-Driven Medical Resource Allocation for 6G Network Dataset," Kaggle Dataset, 2023, accessed 15 June 2025. [Online]. Available: https://www.kaggle.com/datasets/ziya07/iot-driven-medical-resource-allocation-dataset

[24] V. Vasilev, J. Leguay, S. Paris, L. Maggi, and M. Debbah, "Predicting qoe factors with machine learning," in *2018 IEEE International Conference on Communications (ICC)*. Kansas City, USA: IEEE, May 2018.