applied sciences

MDPI

Review

Federated Reinforcement Learning in IoT: Applications, Opportunities and Open Challenges

Euclides Carlos Pinto Neto, Somayeh Sadeghi *, Xichen Zhang and Sajjad Dadkhah

Canadian Institute for Cybersecurity (CIC), University of New Brunswick (UNB), Fredericton, NB E3B 5A3, Canada; e.neto@unb.ca (E.C.P.N.); xichen.zhang@unb.ca (X.Z.); sdadkhah@unb.ca (S.D.)

* Correspondence: s.sadeghi@unb.ca

Abstract: The internet of things (IoT) represents a disruptive concept that has been changing society in several ways. There have been several successful applications of IoT in the industry. For example, in transportation systems, the novel internet of vehicles (IoV) concept has enabled new research directions and automation solutions. Moreover, reinforcement learning (RL), federated learning (FL), and federated reinforcement learning (FRL) have demonstrated remarkable success in solving complex problems in different applications. In recent years, new solutions have been developed based on this combined framework (i.e., federated reinforcement learning). Conversely, there is a lack of analysis concerning IoT applications and a standard view of challenges and future directions of the current FRL landscape. Thereupon, the main goal of this research is to present a literature review of federated reinforcement learning (FRL) applications in IoT from multiple perspectives. We focus on analyzing applications in multiple areas (e.g., security, sustainability and efficiency, vehicular solutions, and industrial services) to highlight existing solutions, their characteristics, and research gaps. Additionally, we identify key short- and long-term challenges leading to new opportunities in the field. This research intends to picture the current FRL ecosystem in IoT to foster the development of new solutions based on existing challenges.

Keywords: internet of things (IoT); federated reinforcement learning (FRL); reinforcement learning (RL); federated learning (FL); survey

Citation: Pinto Neto, E.C.; Sadeghi, S.; Zhang, X.; Dadkhah, S. Federated Reinforcement Learning in IoT: Applications, Opportunities and Open Challenges. *Appl. Sci.* **2023**, *13*, 6497. https://doi.org/10.3390/app13116497

Academic Editor: Dimitris Mourtzis

Received: 5 April 2023
Revised: 29 April 2023
Accepted: 23 May 2023
Published: 26 May 2023

1. Introduction

The internet of things (IoT) represents a disruptive concept that has changed society in several ways. This paradigm connects businesses and optimizes operational factors in a variety of industries [1]. The recent increase in IoT services has shed light on the valuable resources it brings to operations and to society in general [2]. In fact, new technologies are expected to be developed in the near future, and new solutions are currently under development in different sectors.

There have been several successful applications of IoT in the industry. For example, in transportation systems, the novel internet of vehicles (IoV) concept has enabled new research directions and automation solutions [3]. Similarly, logistics have been supported by new IoT-based solutions [4]. Finally, there are other successful IoT applications and more opportunities to develop new solutions in the next few years [5,6].

With the increasing adoption of IoT, new cybersecurity threats have been engineered to exploit the vulnerabilities of such devices [7,8]. The lack of standards regarding vulnerability documentation [9]; the variety of devices, models, and brands [10]; and the simplicity of IoT architecture (both in terms of software and hardware) [11] harden the mitigation of such threats. Although current vulnerabilities can be addressed in future models in IoT, new vulnerabilities can be discovered. Thus, security solutions are paramount for the success of IoT operations.

Furthermore, privacy concerns have become more relevant in the past few years [12]. The interaction with such devices can disclose private information in different forms, e.g., financial transactions [13] and transportation states [14]. The need for privacy-preserving solutions is critical in several applications. In fact, strategies to enable different IoT systems to interact while preserving confidential information can be used in the development of solutions in multiple domains.

Moreover, reinforcement learning (RL) has demonstrated remarkable success in solving complex problems in different scenarios [15,16]. Additionally, federated learning (FL) has enabled global statistical models to be developed based on distributed remote systems in a way to reduce local error [17]. The combination of these two concepts, namely, federated reinforcement learning (FRL), focuses on enabling joint and privacy-preserving learning in sequential decision-making problems [18].

In recent years, new solutions have been developed based on this combined framework [19,20]. These solutions enable different entities to work collaboratively to achieve faster convergence and more robust results [21,22]. The same applies to IoT systems, where different systems can improve internal operations based on the experiences collected from global systems.

Conversely, although some works focus on analyzing FRL contributions [23], there is a lack of analysis concerning IoT applications. Additionally, an analysis of FRL applications subdomains in the IoT context is necessary to shed light on short- and long-time research directions. There is a lack of understanding regarding challenges and future directions analysis of the current FRL landscape in different IoT applications. Finally, the description of open challenges is important to foster the development of new solutions based on the current issues faced in the intersection between FRL and IoT.

Thereupon, the main goal of this research is to present a literature review of federated reinforcement learning (FRL) applications in IoT from multiple perspectives. We focus on analyzing applications in multiple areas (e.g., security, sustainability and efficiency, vehicular solutions, and industrial services) to highlight existing solutions, their characteristics, and research gaps. This is due to the fact that several solutions (both in terms of software and devices) are under development in these sectors, and this presence is expected to be even more significant in the next few years. Additionally, we identify key short- and long-term challenges leading to new opportunities in the field. This research intends to picture the current FRL ecosystem in IoT to foster the development of new solutions based on existing challenges. In this context, the main contributions of this research are as follows:

- A comprehensive review of efforts regarding FRL-based solutions for IoT, and their main contributions, methods, resources, and future directions;
- An analysis of timely solutions divided into categories concerning problems faced, methods used, and immediate directions tailored to each domain;
- An extensive list of short- and long-term open challenges regarding the proposal of new IoT solutions supported by federated reinforcement learning (FRL).

This paper is organized as follows: Section 2 presents the background of this research, in which concepts related to reinforcement learning (RL), federated learning (FL), and federated reinforcement learning (FRL) are depicted. After that, Section 3 presents several FRL applications in IoT solutions. The areas considered include security, sustainability and efficiency, and vehicular and industrial solutions . Finally, Sections 4 and 5 present open challenges concerning FRL applications in IoT and the conclusion of this research.

2. Background

This Section overviews critical topics for a better understanding of the FRL analysis in IoT applications. First, we present the reinforcement learning (RL) characteristics. After that, we define federated learning (FL) and federated reinforcement learning (FRL).

2.1. Reinforcement Learning (RL)

Reinforcement learning (RL) is the result of the fusion between the trial-and-error "law-of-effect" tradition, optimal control theory, the secondary reinforcement tradition, and the use of different stimulus [24].

Reinforcement learning techniques have demonstrated their efficacy in various important applications [25]. When applying reinforcement learning (RL) in complex applications, it is common to use generalizing function approximators such as neural networks, decision-trees, or instance-based methods [26]. RL studies how systems can learn and predict the consequences of environmental interactions [27]; FL relies on an agent interacting with these environments, learning an optimal policy in many fields [28], making a series of decisions over time, aiming to achieve goals that may be delayed while also managing uncertainty and randomness. It focuses on making decisions quickly rather than relying on lengthy analysis or higher-level reasoning [29]. To solve problems related to reinforcement learning, there are two primary approaches. The first involves exploring different behaviors to discover one that is effective in a given environment. The second approach is to utilize dynamic programming methods and statistical techniques to gauge the usefulness of actions taken in specific states of the world [30].

In the past few years, there have been several RL contributions across multiple fields. Furthermore, there are several opportunities to use RL in new applications and develop new RL approaches [15,16].

There are many important RL components. A *state* represents the configuration of the environment for a given task [31,32], while actions are functions RL agents can execute to change or interact with the environment [33,34]. These actions can generate new states. In fact, the reward function produces a score for an action executed for a given state [35] (the output—i.e., the reward—can be positive and negative. Although the term "rewards" may lead to an understanding that the outcome is always positive, negative signals are also provided to inhibit non-optimized decisions.). Furthermore, a policy represents an association between actions and states from the reward standpoint [36,37]. Figure 1 illustrates the general RL process model.

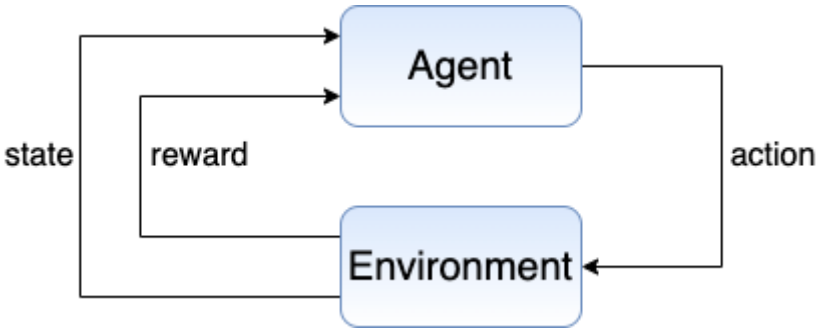

Figure 1. Reinforcement learning (RL) process [38,39].

Moreover, several RL techniques have been proposed in the last decade. These models are becoming more complex and solving challenging problems with high performance. Although several efforts have been made in this direction, some of them can be considered applications for different IoT systems.

Q-learning [40,41] has been used in IoT monitoring [42] and resource allocation [43]. Deep Q-learning [44] and double deep Q-learning (DDQN) [45] have supported transmission scheduling in IoT [46,47]. Proximal policy optimization (PPO) [48] and advantage actor–critic (A2C) [49,50] were adopted for computation offloading [51,52], while the deep deterministic policy gradient (DDPG) [53] can be adopted for intrusion detection [54] in green IoT. Finally, the asynchronous advantage actor–critic (A3C) [55] has been adopted in service placement [56].

2.2. Federated Learning (FL)

Federated learning (FL) entails the process of learning a shared model from distributed sources on various client systems in order to reduce prediction error [17]. Kairouz et al., define FL as "a machine learning setting where multiple entities (clients) collaborate in solving a machine learning problem, under the coordination of a central server or service provider. Each client's raw data is stored locally and not exchanged or transferred; instead, focused updates intended for immediate aggregation are used to achieve the learning objective" [57].

A common training procedure comprises client identification, broadcast, local computation, aggregation, and local and global updates [57]. All of these phases are pivotal for training and defining a federated model, and they can be performed in different ways.

For instance, the aggregation phases include the mixture of various factors (or weights) executed by a centralized system as illustrated in Figure 2. This entity is in charge of considering the various inputs provided by the different clients. An approach used in several efforts is the FedAvg [58,59], which is based on merging clients' weights considering local updates.

FL has been successfully used in multiple scenarios [60–63] and presents pivotal advantages (e.g., privacy). In reference [60], the authors suggest a federated learning (PEFL) method for IAI that is both efficient and provides enhanced privacy. PEFL is designed to be non-interactive and capable of protecting sensitive data from disclosure, even if multiple entities cooperate to breach it. The authors of [61] introduce a novel architecture that enhances data privacy through security. They propose a privacy-preserving federated learning mechanism, incorporating a two-phase approach involving intelligent data transformation and collaborative data leakage detection to mitigate privacy risks. By combining blockchain technology with on-device learning, this [62] study proposes a novel approach to federated learning. This offers a promising solution for securing machine learning while preserving privacy. Another scenario proposed in [63] is a framework for privacy-preserving federated learning in smart agriculture. The framework employs a deep privacy encoding method to protect the data privacy of each participant in the federated learning process.

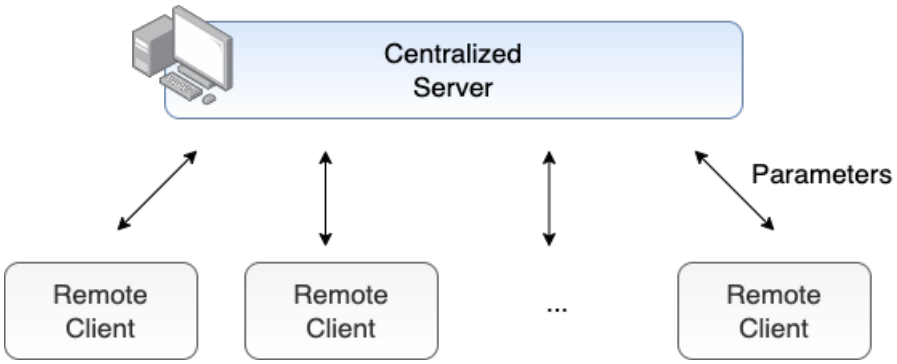

Figure 2. Federated learning (FL) overview [64,65].

There are different categories of FR models. Although new approaches are under development by the scientific community, some of the most popular classes are as follows:

- **Vertical FL:** This strategy refers to scenarios where data samples present in each client's environment share the same target while presenting different features [66]. In other words, clients may have different features referring to the same target [67,68];
- **Horizontal FL:** Refers to the use of the same feature space by different clients while considering different data samples [66]. This refers to a more structured way of distributed learning as the model weights can be combined due to the reference to the same features [69];

- **Cross-device FL:** This refers to the utilization of multiple devices (e.g., IoT devices) to train a global model in a likely massive distributed dataset [70]. In other words, the number of training clients can be extremely large [71,72];
- **Cross-silo FL:** This approach is based on the consideration of entities (e.g., companies or organizations) as training clients in different industrial sectors (e.g., transportation) [73]. Compared to cross-device FL, the number of clients tends to be smaller. Furthermore, each entity participates in the entire training process [74,75].

There are new federated learning (FL) efforts under active development, e.g., trustworthy federated learning [76,77], and federated anomaly detection [78–80]. Finally, new paradigms in artificial intelligence (AI) as well as in reinforcement learning (RL) are under development and evaluation, e.g., data-centric AI and its applications to IoT [81].

2.3. Federated Reinforcement Learning (FRL)

RL is capable of solving a variety of complex problems with high performance. However, there are some challenges in its applications to practical scenarios, e.g., sampling is a decisive factor in the agent's experience, which entails that learning efficiency relies on sample efficiency [23]. In fact, sample efficiency refers to the manifestation of the actual decision-making challenge in reinforcement learning theory [82]. Although some efforts have focused on distributed RL [83,84], protecting agents' privacy represents an issue to be faced. Moreover, the simulation-reality gap is another challenge once simulated environments can present some limitations compared to real environments [23].

Moreover, federated reinforcement learning (FRL) aims to enable agents to jointly learn how to solve a given RL task performing uniformly well across multiple environments [18]. FRL focuses on privacy-preserving sequential decision-making, in which the sample, feature, and label are not used and the environment, state, and action are included [23]. In fact, FRL extends RL capabilities regarding distributed learning, new sampling techniques, generalization, and privacy. Conversely, it also entails new challenges to be faced (e.g., poisoning threats [85]). Figure 3 illustrates the federated reinforcement learning (FRL) training process.

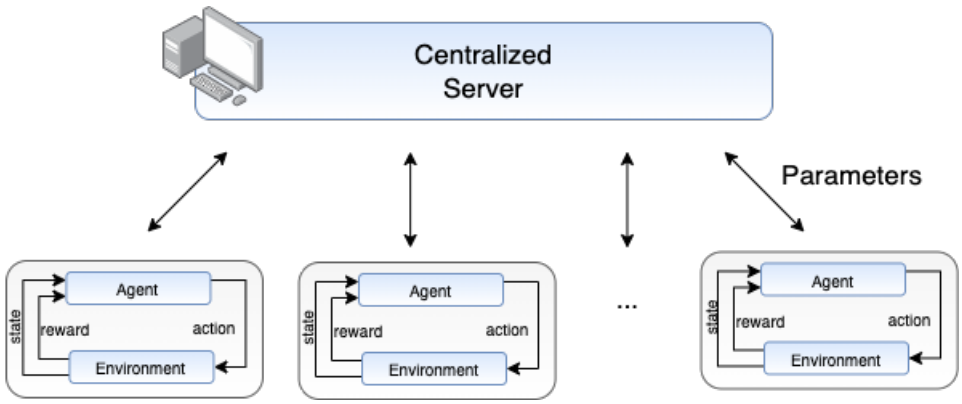

Figure 3. Federated reinforcement learning (FRL) overview [86,87].

FRL is a recent topic that has been successfully adopted in several works. For example, there has been FRL application to 5G [88], autonomous driving [89], robotics [90], healthcare [91], and transportation [92]. These examples demonstrate how FRL can tackle several problems faced nowadays and foster the development of new approaches related to this combined framework. Moreover, the use of deep reinforcement learning (DRL) in a federated environment is referred to as federated deep reinforcement learning (FDRL). Finally, FRL can be applied to IoT solutions in several different ways and environments [91,93].

3. Methodology

This Section presents an in-depth analysis of FRL efforts in different IoT applications. First, we consider efforts toward IoT security. Secondly, we consider FRL-based sustainabil-

ity and efficiency solutions for IoT operations. After that, vehicular solutions are depicted. Finally, we focus on industrial IoT (IIoT) applications.

3.1. FRL Applications in the IoT in Terms of Security

The authors of [94] propose a novel approach for dynamic spectrum access (DSA) in the context of the IoT using federated deep reinforcement learning (FDRL). They present a set of techniques to enhance the efficiency of dynamic spectrum access in IoT environments. The complete system can be categorized into three stages: the first involves training a local double deep Q-learning network (DDQN) model, the second consists of aggregating the central model, and the third involves releasing the model parameters. Their findings indicate that the reinforcement learning algorithm has a significantly higher success rate than random channel access, particularly after several iterations. Incorporating FL into the approach results in even faster convergence. The proposed scheme aims to address the security of terminal data and personal privacy data in a context where multiple devices communicate with each other. Additionally, it maintains IoT users' confidentiality since the FDRL method only requires the uploading of model parameters to edge servers.

In reference [95], the authors propose a framework that merges blockchain and federated learning to improve the protection and confidentiality of acquired model characteristics. The structure comprises three components: local training, blockchain for parameter verification, and global aggregation. The edge servers maintain the blockchain. The deep reinforcement learning (DRL) method they use has three crucial elements: the main network, the target network, and replay memory. The primary and target networks each have two deep neural networks (DNNs): the actor DNN that links system states to actions, and the critic DNN that evaluates the effectiveness of policies and directs actions toward policy gradient direction. The objective network has a comparable arrangement to the primary network, and it generates desired outputs that help train the main critic DNN. The proposed approach suggests using blockchain to preserve learning parameters and validate their accuracy, which can improve the security and quality of the learning process.

The authors of [88], similar to [95], present a new framework for channel resource allocation in 5G/B5G networks using federated reinforcement learning (FRL). They suggest utilizing FRL to enhance incumbent technologies' security in beyond 5G networks. The authors suggest using the FRL model for faster learning convergence and combining RL and FL methods in the proposed framework. The framework involves both local and global learning phases. The results of the study show that the FRL framework assembles much more quickly than the standard RL approach, indicating a significant performance improvement. The experimental results confirm that the presented technique improves the ability of WiFi networks in regard to throughput by selecting the optimal channel access parameters through collaborative learning.

In reference [96], the authors introduce a decision-making system called Devote, which presents a solution to address security challenges in fog-based IoT environments. Devote utilizes a great algorithm to prioritize data services according to their importance while assuming the accessibility of resources at the fog node (FN). To deal with the playful nature of the IoT domain, the authors process IoT data in a way to adapt to changing conditions over time. In addition, they present a technique based on an online secretary approach for selecting the appropriate candidate FN for data offloading. The results demonstrate that Devote achieves lower service delay than other systems and a user satisfaction rate of 88.4%.

The authors of [97] present a new approach to introduce a solution to address security concerns regarding privacy-preserving offloading in IoT environments enabled by the cloud. The proposed method combines the concepts of context awareness and privacy preservation to make offloading decisions more securely and efficiently. The authors use federated deep reinforcement learning (fDRL) to train an offloading policy that can adapt to IoT environment changes while considering privacy requirements. This approach allows for better control over the trade-off between offloading performance and privacy.

The experimental results show that the C-fDRL approach outperforms traditional offloading methods in terms of both efficiency and privacy protection. The authors conclude that using fDRL in IoT offloading can deliver a more secure and efficient answer while preserving the privacy of sensitive data.

Miao et al. [98] present a new method called FL2S, designed to provide secure data sharing in IoT systems. The proposed approach utilizes federated deep reinforcement learning (FDRL) to create a policy that can adapt to IoT environment changes. This policy is designed to enhance data-sharing security in IoT systems. The proposed approach, FL2S, employs fDRL to create a data-sharing policy that adapts to the context of each IoT device and the requirements of each sharing scenario. The approach provides a more flexible and secure way of managing data sharing in IoT systems, considering privacy and security risks. Experimental results demonstrate that the fDRL-based approach outperforms traditional methods regarding both efficiency and security. The authors conclude that using fDRL in IoT data sharing can lead to a more secure and efficient solution while preserving the privacy of sensitive data.

Zheng et al. [99] suggest a new method to solve security issues related to resource allocation in privacy-preserving EdgeIoT, which involves using DRL to train a policy to make online resource allocation decisions based on the context of each edge device. They suggest a novel FL-enabled twin-delayed deep deterministic policy gradient (FL-DLT3) framework to balance privacy preservation and resource utilization. This approach is designed to optimize resource allocation by learning from experience, considering the context of each edge device, and balancing privacy preservation with the efficient use of resources. The proposed approach can potentially enhance the efficiency and security of resource allocation by using deep reinforcement learning, federated learning, differential privacy, and access control mechanisms. According to the numerical results, for the better prediction of time-varying data size, bandwidth, channel gain, and remaining energy of IoT devices, the authors developed a new state characterization layer based on LSTMs in FL-DLT3. The FL-DLT3 that has been suggested achieves rapid convergence, taking fewer than 100 iterations. By comparing it to the current leading benchmark, FL shows a 51.8% improvement in accuracy-to-energy consumption.

Anwar et al. [100] present a new framework based on multi-task federated RL for learning multiple tasks while protecting against adversarial attacks. The suggested framework utilizes federated RL to allow agents to learn from one another while defending against adversaries. The multi-task RL formulation ensures that the agents can learn multiple tasks while assessing the trade-off between task performance and adversary defense. This framework can enhance learning efficiency in multi-agent systems in the presence of adversaries. The effectiveness of the proposed method is assessed through simulated experiments, revealing that the agents can effectively acquire knowledge for multiple assignments and protect themselves from adversarial attacks. The authors proposed a new attack method called AdAMInG that considers the aggregation operator used in federated RL. Further, they proposed modifying the conventional federated RL algorithm called ComA-FedRL to manage the issue of adversaries in the multi-task federated RL problem. The solution incorporates FRL, differential privacy, and Byzantine-resilient aggregation mechanisms to ensure the privacy and security of local models and detect malicious behavior.

The author in [20] proposes a federated RL-based clinical decision system for edge computing environments with resource constraints and privacy preservation concerns. The system uses multiple agents to collaborate and utilize each other's experiences to make clinical decisions collaboratively and optimizes decision-making policies using reinforcement learning. The authors proposed a new algorithm called double deep Q-network (DDQN) within a fully decentralized federated framework (FDFF) made possible by an integrated system known as SMEC. This algorithm ensures the privacy and security of patient data and detects malicious behavior by providing a reliable way to develop a treatment policy in real-time using multiple distributed electronic medical records (EMRs). To ensure the confidentiality of EMRs, additively homomorphic encryption is employed.

Security issues related to traffic monitoring in SDN-based IoT networks are a big problem. To overcome these issues, Nguyen et al. [101] suggest a method for monitoring IoT network traffic based on SDN using FDRL. The proposed approach utilizes a control algorithm based on double deep Q-network (DDQN) and a flow-rule matching field to supervise a specific IoT edge, facilitating effective traffic monitoring. The authors suggest utilizing federated RL to enhance the learning version of the DDQN algorithm mentioned earlier and enable the SDN controllers to collaborate and make traffic monitoring decisions based on shared data. Their results demonstrate that the system can allow SDN controllers to monitor network traffic and make informed decisions efficiently. The study found that using a federated DDQN approach can decrease learning loss and learning cycles by 66% and 40%, compared to the traditional approach, in situations with high granularity requirements. Additionally, the deep monitor framework improved the IDS application's attack detection performance by 22.83% compared to the FlowStat solution.

In reference [86], the authors suggest a method called FRL for automatically controlling software-defined networking (SDN)-based IoT systems while prioritizing data security. This approach allows IoT devices to independently learn and adjust to the network's changing conditions without needing a central controller. The experiments show that the proposed approach can effectively optimize network performance while reducing the computational burden on individual devices. They utilized the actor–critic PPO, which is a kind of reinforcement learning algorithm. They also introduced two federation policies—transfer learning and gradient sharing—to enhance learning speed and overall performance.

Wireless networks are vulnerable to security breaches and intrusions due to the lack of clear boundaries. As cyber intruders continue to grow, the risk of compromising critical applications monitored by networked systems has also increased. A federated reinforcement learning-based intrusion detection system (FRL-IDS) has been proposed in [91], for healthcare infrastructures in the IoT networks to address these security concerns. The system is designed to manage the challenge of detecting attacks in IoT-enabled healthcare systems, which are increasingly vulnerable to security threats due to many related devices and the sensitive character of the data they control. FRL allows multiple IDSs to collaborate and learn from each other without sharing sensitive data, thereby addressing the privacy issue in healthcare systems. The suggested system comprises a central server that manages the training of intrusion detection system (IDS) models using federated reinforcement learning (FRL), along with several edge devices that operate local IDSs. The edge devices train their IDSs on local data and periodically upload the trained models to the central server, aggregating them to enhance the system's overall performance.

Current research on IIoT routing primarily concentrates on latency and routing reliability but often overlooks the importance of privacy and security in the routing process. Wang et al. [102] present a solution for quality of service (QoS) and privacy-aware routing in the context of the 5G-enabled IIoT. The approach considers the constraints of the IIoT environment and the communication requirements to create an efficient routing mechanism that ensures QoS while protecting user privacy. This approach aims to optimize routing decisions while ensuring that the privacy of the data transmitted is preserved. The proposed approach utilizes an FRL method to optimize the QoS while ensuring the confidentiality of the user. The experiment results show that the quality of service and privacy-aware routing (QoSPR) protocol can serve as a routing method that considers data privacy concerns, effectively reducing the average and maximum latency and ensuring optimal load balancing in 5G-enabled IIoT networks.

The authors of [22] propose a new framework called MARL-FRL to improve the security of ICPS. It encourages agents to work in the system's best interest, thereby reducing the risks of security threats and malicious attacks. They introduce the MA-FRL algorithm, which aims to address the problem of nonstationarity that arises due to frequent interaction between devices in FL without compromising privacy by sharing sensitive information. The suggested methodology requires several agents to interact with each other and acquire knowledge from their shared experiences, enabling the system to adjust to evolving

circumstances and enhance its security measures. The authors illustrate the efficacy of their suggested method through simulations, and the results indicate that the suggested MARL-based mechanism can improve the security of intelligent cyber-physical systems by providing a more robust and adaptive approach to managing incentives.

Virtual network functions (VNFs) are a collection of VNFs linked in a particular sequence to make different network services compatible and flexible. The application of network function virtualization (NFV) enables these VNF groups to work together in a federated manner. Nevertheless, meeting the rising demands of network services through NFV execution can be challenging, mainly because of the fixed orchestration of service function chains (SFCs). NFV is necessary to obtain compatibility and scalability for different network services in service function chains (SFCs). The NFV execution is difficult to achieve the ever-increasing conditions of network services, especially due to the static orchestrations of SFCs. To address this challenge, the authors of [103] present a scheme based on scalable SFC orchestration (SSCO) for NFV-enabled networks via FRL. SSCO has unique features that differentiate it from the prior work, such as (1) it allows for training an international education prototype that utilizes time-variant regional sample investigations. This framework makes it possible to orchestrate scalable benefit function chains (SFC) on a large scale while also ensuring that stakeholder data is kept private. (2) The SSCO approach permits parameter updates between local clients and the cloud server solely at the start and finish of every episode. This guarantees that distributed clients can enhance the model while reducing communication expenses. This approach is designed to improve the security of NFV-enabled networks by optimizing SFC orchestration. This can enhance network performance while minimizing the risk of security threats and attacks.

Similar to [88,95], Yu et al. in [104] discussed the 5G network. They proposed a framework combining DRL and FL to allow intelligent resource management in 5G ultra-dense networks. The framework utilizes a hierarchical architecture with global and local controllers to handle resources at various timescales. It seeks to enhance resource allocation and task offloading in a multi-access edge computing environment. They created a unique and authentic direction to DRL called the "two-timescale deep reinforcement learning (2Ts-DRL) approach". This method contains two learning processes: one on a fast-timescale and another on a slow-timescale. In addition, they use FL to train the 2Ts-DRL model in a decentralized manner to safeguard data privacy on edge devices.Although the paper does not mention any particular security concerns, it does address the overall security challenges facing MEC networks. Such networks are susceptible to several security threats, such as attacks on edge devices, data breaches, and denial-of-service attacks. The proposed approach uses DRL and FL to optimize resource allocation and management to mitigate these threats, enhancing network performance and lowering the risk of attacks and other security threats.

While FRL has many potential applications, current research must address two crucial problems. Firstly, there needs to be more theoretical analysis of the convergence of FRL algorithms. Secondly, the recent works do not consider the effect of random system failures or adversarial attacks on the execution of FRL. Xiaofeng et al. [105] propose the first FRL framework that ensures convergence and is resilient to the failure or malicious behavior of up to half of the participating agents.The proposed approach called FT-FRL with theoretical guarantees has been presented to improve the security of FRL. This can effectively address the fault tolerance issue and enhance FRL's overall security. This framework uses a Byzantine fault-tolerant algorithm that allows the agents to exchange information securely while ensuring that malicious agents cannot disrupt the learning process. The FT-FRL algorithm is compared with existing reinforcement learning algorithms, demonstrating that it performs better in system performance and accuracy in a decentralized environment.

In reference [86], the authors present an FRL architecture in which each agent operates independently on their respective IoT device and shares their learning experience with other agents in a decentralized way. This scheme addresses the security concerns related to training control policies for IoT devices that arise due to scalability. Leveraging FRL to

optimize the training process can significantly improve the overall security of these devices. This technique lets agents learn from each other's experiences without sharing raw data, enhancing the system's general performance. Using this FRL architecture, the data privacy on the IoT devices is maintained, and the agents can cooperate effectively to acquire optimal control policies for the given task. They combine the actor–critic proximal policy optimization (actor–critic PPO) algorithm into each agent in the suggested collaborative framework. They also present an effective method for exchanging gradients and transferring model parameters to the agent.

The main focus of the authors of [106], is to tackle the security and privacy concerns in FL, where clients provide their data to a central server for model training. The proposed solution is to enhance the security and privacy of FL by carefully choosing reliable clients to participate in the learning process. They offer a novel method for determining trustworthy and reliable clients in federated learning systems. They implement their proposed method in the healthcare sector, specifically for detecting COVID-19 using IoT devices. They use a dataset of COVID-19 chest X-rays and apply the federated learning framework with the proposed client selection mechanism to train a model for COVID-19 detection.

Finally, Table 1 provides a summary of all the works that were reviewed, taking into account their areas of study, the technologies used, and the objectives pursued.

Table 1. Overview of FRL-based solutions for secure IoT operations.

Number	Work	Domain	Technology	Research Purpose	Research Problem	Dataset	Year
1	Li et al. [94]	Wireless communication networks, dynamic spectrum access	FDRL	Optimizing spectrum allocation efficiency	Spectrum sharing optimization problem	Simulated data	2022
2	Lu et al. [95]	5G networks	FL, blockchain, and 5G	Improving 5G networks, integration of blockchain and FL	Security, privacy, and scalability challenges of 5G networks	Simulated data	2020
3	Ali et al. [88]	Dynamic spectrum	FRL	Improved performance	Incumbent interference	Simulated environment	2021
4	Tiwari [96]	Fog computing and IoT	Federated computing	Service provisioning	Criticality management	Simulated data	2021
5	Xu et al. [97]	Cloud-enabled IoT	FDRL	Privacy-preserving offloading	Privacy preservation	Own dataset	2022
6	Miao et al. [98]	IoT and data sharing	FDRL	Secure data sharing	Secure and efficient data sharing	Simulated data	2021
7	Zheng et al. [99]	Edge computing and IoT	DRL, FL, and edge computing	Resource optimization	Privacy-preserving	Real-world dataset	2022
8	Anwar et al. [100]	Federated learning	FL, RL, and federated RL framework	Multi-task RL efficiency	Privacy-preserving federated learning	N/A	2021
9	Xue et al. [20]	Healthcare decision system	FL, edge computing, and RL	Privacy-preserving, clinical decision	Resource constraints, privacy, and clinical decision-making	MIMIC III dataset [107]	2021
10	Nguyen et al. [101]	Networking and IoT	FL, DRL, and SDN	Traffic monitoring optimization	Traffic monitoring scalability, and privacy	N/A	2021
11	Lim et al. [86]	SDN-based IoT networks	FRL, IoT, and SDN	IoT network optimization, automation	IoT control complexity, IoT network performance, and network management efficiency	Simulated environment	2020
12	Wang et al. [102]	Wireless IoT networks	FRL, IIoT, and 5G	Improve IIoT routing, improve IIoT communication	Routing, privacy, and QoS	Simulated data	2021
13	Xu et al. [22]	AI, multi-agent systems	MARL, FL, and cybersecurity	Secure multi-agent cyber-physical systems	Secure multi-agent collaboration and privacy-preserving incentives	FMNIST dataset	2021
14	Huang et al. [103]	Network function virtualization (NFV) and service function chain (SFC)	NFV, SFC, and FRL	NFV service orchestration and SFC orchestration enhancement	SFC inflexibility and scalability problem	Simulated data	2021
15	Yu et al. [104]	Networking and resource management	FL, DL, and multi-access edge computing	Optimizing network resources and edge computing,	Resource allocation optimization	Simulated environment	2020

Table 1. *Cont.*

Number	Work	Domain	Technology	Research Purpose	Research Problem	Dataset	Year
16	Xiaofeng et al. [105]	Distributed systems and fault-tolerance	FRL, FL, and multi-agent systems	Secure FRL convergence, develop fault-tolerant FRL	Faulty agent resilience and faulty multi-agent learning	N/A	2021
17	Lim et al. [86]	Privacy preservation and IoT	FL, RL	Federated IoT control	Federated IoT control	Simulated environments	2020
18	rjoub et al. [106]	Federated ML trust	FL, IoT	Federated client selection, enhancing FL security	Client trustworthiness evaluation	COVID-19 radiography database [108]	2022

3.2. FRL Applications in the IoT in Terms of Sustainability and Efficiency

Qiu et al. [109] models an energy and carbon allowance trading mechanism as a multiagent reinforcement learning (MARL) problem for a building community. The multi-energy system (MES) is a pivotal pillar towards the future low-carbon energy system, but some challenges harden its current exploration. For example, factors that make this exploration complex include the complex operation of combining multi-energy sectors, the privacy aspect of decentralizing the energy system, and the integration of energy and carbon emission schemes. Thereupon, the authors adopt an abstract critic network using a deep deterministic policy gradient method. This pipeline is, then, integrated into a federated learning (FL) framework to preserve the information private of each building in the community. The experiment showed that the proposed strategy achieves 5.87% lower total energy cost and 8.02% total environment costs compared to baselines.

The authors of [110] introduce an edge-based backhaul selection method to enhance traffic delivery based on multi-objective feedback. IoT adoption has substantially increased backhaul traffic congestion. This new scenario demands effective traffic management optimization at the network edge. In fact, edge devices can forward IoT traffic through the backhaul network by choosing appropriate links for collected data flows. This selection challenge requires efficient strategies to learn how to handle partially observable components of the network. The authors employ different advantage-actor–critic deep reinforcement learning (DRL) and federated learning (FL) to train a shared backhaul selection policy. Finally, the proposed solution can solve the backhaul selection problem effectively.

The contributions of [111–114] focus on offloading optimization. The authors of [111] introduce FedAdapt as an adaptive offloading FL framework to tackle the efficiency aspects of FL, e.g., consideration of limited computational capabilities, computational heterogeneity, and variable network bandwidths. To accomplish this, this approach adopts the proximal policy optimization (PPO) to identify which deep neural network (DNN) layers can be offloaded for IoT devices onto a server to handle computational heterogeneity and changing network bandwidth. Zang et al. [112] propose a federated DRL-based online task offloading and resource allocation (FDOR) technique. The provision of cloud-like services to IoT offered by mobile edge computing (MEC) is the target of this contribution. The authors also consider wireless powered communication (WPC) technology, given that a base station (BS) can transmit energy to edge users (EUs) and execute tasks via task off-loading. Thereupon, DRL is executed in EUs with aggregated parameters and an adaptive learning rate.

The authors use multiple DRL agents situated on edge nodes to specify the offloading choices of IoT devices in [113]. By transferring computationally demanding tasks to edge nodes, the idea of offloading allows IoT devices to conserve energy and sustain the quality of service. Conversely, federation and intricate resource management are dynamically determined in real-time, considering varying workloads and radio conditions. To tackle this, FL trains DRL agents in a spread approach. Furthermore, Chen et al.'s [114] objective was to reduce the amount of energy used by IoT devices while satisfying the threshold for delaying and resource constraints. They created a collaborative optimization challenge that involved both task offloading and resource allocation in accomplishing this. The authors emphasize that privacy disclosure is a present issue in MEC data exchange and that FL can support these transactions. In this sense, a two-timescale federated DRL technique based on the deep deterministic policy gradient (DDPG) is proposed.

Similarly, Zarandi et al. [115] introduce a federated DRL framework focused on multi-objective optimization problems to decrease the delay and energy usage of IoT devices' long-term task completion. The main goal is to handle, in a distributed fashion, the offloading decisions, resources, and transmit power allocation. The experiments demonstrated that the proposed framework is effective in the cases considered.

Moreover, resource allocation is another aspect considered in many efforts. The authors of [116] focus on resource allocation in device-to-device (D2D)-enabled 6G using FL. The suggested approach takes into account a D2D-enabled wireless network in the underlay mode and decentralized resource allocation to maximize the capacity while minimizing

the overall energy usage. In fact, resource allocation and further improving spectrum utilization are challenges faced nowadays. Thereupon, the authors considered the quality of service (QoS) requirements of both cellular users and D2D users and demonstrated that this method achieves significant network performance.

Tainqing et al. [117] propose a resource allocation method named concurrent federated reinforcement learning. The challenge considered relies on the limited information faced when resource allocation is planned at edge hosts, whereas lack of privacy is the result of moving this process to central servers. The primary concept is to control the privacy-preserving aspect of FL combined with the RL efficiency. Then, the authors adopt concurrency as joint decision-making to achieve global solutions. Similarly, the authors of [118] adopt a DQN to optimize decisions regarding energy and WiFi channels without any pre-existing network information. This method outperforms baseline approaches and maximizes successful transmissions while reducing energy and channel expenses to a minimum.

Cui et al. [119] present an FL protocol to enhance the efficiency of FL systems powered by renewable energy sources. This effort is motivated by the challenges of limited device resources faced by industrial FL deployments. The authors focus on using RL for scheduling devices to adjust to inconsistent renewable energy supply. Moreover, the authors introduce an efficient bandwidth management scheme focused on communication efficiency. The experiments showed that this proposal outperforms state-of-the-art methods. The authors of [120] present a residential energy management system (EMS) using a personalized federated DRL (PFDRL) system to address the issue of reducing standby energy consumption. This privacy-preserving and cloud-free proposal is motivated by (i) the challenges of handling various in standby mode, consuming energy while waiting for wireless communication, and (ii) the potential personal data leakage of existing solutions. Moreover, global collaborative models produce unsatisfactory energy management performance since capturing individual residential characteristics is complex.

The authors of [121] present a federated DRL-based cooperative edge caching (FADE) framework to solve the lack of self-adaptivity in dynamic environments of most existing methods. Current proposals focus on centralized solutions, so the authors enable base stations (BSs) to develop a predictive model that can be shared. This strategy provides fast training and separates learning from storage, relying on a distributed-centralized procedure. Similarly, Majidi et al. [122] introduce a hierarchical federated DRL (HFDRL) method to predict future users' requests to identify appropriate content replacement strategies. The efficiency of edge caching reduces access time and optimizes content transfer, and novel smart caching solutions have been suggested during recent years. In addition to existing methods, the authors categorize edge devices hierarchically and improve local and global performance. The results obtained in the experiments showed that the proposed strategy presents improvements in, e.g., hit rate and delay have improved over traditional methods.

Finally, the authors of [123] propose an RL-based request service provisioning system as a component of smart edge orchestration. Although IoT is heavily supported by edge computing and its short response times, solutions are required to maximize profitability while minimizing response time. One possible approach is to integrate edge nodes forming a federation. In this context, the authors implement the DRL dispatcher and compare it with baseline methods, showing that their proposal is efficient in the cases considered. Table 2 summarizes all of the works reviewed, considering their domains, technologies, and objectives.

Table 2. Overview of FRL-based solutions for sustainable and efficient IoT operations.

Number	Work	Domain	Technology	Goal	Year
1	Qiu et al. [109]	Multi-energy systems (MES)	Deep deterministic policy gradient and abstract critic network	Method to address the joint P2P energy and carbon trading (JPC) problem in a local community	2023
2	Jarwan et al. [110]	IoT traffic management	Advantage-actor–critic	Propose an edge-based backhaul selection to improve traffic delivery based on multi-objective feedback	2022
3	Wu et al. [111]	Offloading optimization	Proximal policy optimization (PPO)	Adaptive offloading FL framework to tackle efficiency challenges	2022
4	Zang et al. [112]	Offloading optimization	Federated DRL-based online task offloading and resource allocation (FDOR)	Online task offloading and resource allocation in WPC-MEC Networks	2022
5	Ren et al. [113]	Offloading optimization	Double deep Q-learning (DDQN)	FL and DRL to optmize IoT computation offloading	2019
6	Chen et al. [114]	Offloading optimization	Deep deterministic policy gradient (DDPG)	Task offloading and resource allocation	2022
7	Guo et al. [116]	Resource allocation	Deep Q-network (DQN)	FRL-based resource allocation in D2D-enabled 6G	2022
8	Tianqing et al. [117]	Resource allocation	Deep Q-network (DQN)	Resource allocation in IoT edge computing	2022
9	Nguyen et al. [118]	Resource allocation	Deep Q-network (DQN), Deep Q-learning (DQL), and double deep Q-network (DDQN)	Resource allocation in mobility-aware networks	2020
10	Zarandi et al. [115]	Offloading optimization	Double deep Q-network (DDQN)	Delay and energy minimization in IoT networks	2021
11	Cui et al. [119]	Energy management	Adapter multi-armed bandit (MAB) algorithm	Device scheduling for renewable energy-powered federated learning	2022
12	Gao et al. [120]	Energy management	Personalized deep federated reinforcement (PDRL)	Residential energy management system to tackle the standby energy reduction in residential buildings	2022
13	Wang et al. [121]	Caching	Double deep Q-network (DDQN)	Federated DRL-based cooperative edge caching (FADE) framework to enable collaborative learning of shared predictive models	2020
14	Majidi et al. [122]	Caching	Hierarchical federated deep reinforcement learning (HFDRL)	Prediction of user's future requests to determine appropriate content replacement strategies	2021
15	Baghban et al. [123]	Service provision	Actor–critic (AC) reinforcement learning (RL)	Propose an RL-based request service provisioning system	2022

3.3. FRL Applications of IoT in the Vehicular Industry

The authors of [124] present an online federated deep Q-learning-based offloading technique for Vehicular fog computing (FedDOVe). In an urban environment, connected autonomous vehicles (CAVs) offload processing jobs to RSUs with restricted power, computational abilities, and storage powered by renewable energy. In this context, although vehicular fog computing can perform computation-intensive tasks, defining which RSUs can be associated with fog servers is challenging since offloading demands robust power consumption and varying offloading rates across uneven computation loads. To tackle these issues, the authors optimize RSUs' energy consumption and perform load balancing across fog servers using a model-free RL approach based on global information to identify appropriate connections among fog servers and RSUs. The experiments demonstrate that this strategy reduces energy consumption and enhances load balancing compared to existing offloading methods.

In reference [125], the authors face the ultra reliable low-latency communications (URLLC) resource slicing and scheduling challenge focusing on trustworthy 6G vehicular services. New technologies aim to connect vehicles to roadside units (RSUs), which can lead to security problems. The authors target mitigating malicious attacks by unauthorized edge access points using a reputation score based on a personal logic sample. In this scenario, offloading is conducted based on such a reputation, which is supported by a federated asynchronous RL algorithm. In addition to that, the authors of [126] present a federated multi-agent DRL method to optimize task-offloading decisions at local and global scales. Focusing on vehicular fog computing, this effort focuses on fast convergence by fostering a local learning approach with limited information sharing, reducing the communication overhead and improving overall privacy.

Lee et al. [127] focus on developing an unmanned aerial vehicles (UAVs) swarm system for aerial remote sensing. By merging FL with RL, this strategy establishes better, trustworthy, and more robust swarm intelligence (SI) in the UAV system. Similarly, Salameh et al. [128] present a cooperative FRL approach to support search missions using UAVs. The central idea is to enable cooperation through experience exchange while maintaining privacy.

Zhang et al. [129] propose a vehicle–road–base position partnership architecture, along-side task offloading and aid distribution algorithm in CAVs. The main goal is to decrease performance delays with various restrictions. In fact, despite the importance of joint optimization of multiple resources to ensure the implementation of automatic driving protection, existing efforts do not target low-latency requirements in exceptional cases (e.g., raw perception data sharing with specific constraints). This aspect can intimidate suitable automatic driving safeness in CAVs networks. Then, the authors introduce a DRL method to perform optimal assignment offloading and resource allowance and an FRL-enabled algorithm to minimize the implementation delay.

Ye et al. [130] introduce a novel FRL strategy to optimize signal control policy generation for multi-intersection traffic scenarios. Given the current problems of elevated mean vehicle travel duration and delayed optimization faced by existing solutions, the authors focus on fostering knowledge sharing in a decentralized procedure. Indeed, despite the current difficulty concerning optimization targets at a global level for complicated traffic situations, the authors demonstrate that it is possible to enhance both the general convergence rate and the quality of control.

Kwon et al. [131] present an FL-baed multiagent DRL in the context of internet-of-underwater-things (IoUT) devices that operate in the ocean environment. The main goal is to design a joint cell association and resource allocation (JCARA) technique in an application that faces challenges in setting up reliable links. Each device in the IoUT conducts local training, and the accumulated knowledge is merged at a centralized system located in the smart ocean base stations (BSs). The experiments performed proved that the suggested approach results in better performance than alternative methods in terms of downlink throughput.

Finally, Table 3 summarizes all of the works reviewed in this context, considering their domains, technologies, and objectives.

Table 3. Overview of FRL-based solutions for vehicular IoT operations.

Number	Work	Domain	Technology	Goal	Year
1	Sethi et al. [124]	Offloading Optimization	Deep Q-learning	Optimize energy consumption across Roadside Units (RSUs) and load balancing across fog servers	2022
2	Hao et al. [125]	Offloading Optimization	Asynchronous Advantage actor–critic (A3C)	Resource slicing and scheduling for trustworthy 6G vehicular services	2021
3	Salameh et al. [128]	Search Cooperation	SARSA and Q-learning	Uncertain Deceptive Target Detection	2023
4	Lee et al. [127]	Aerial Remote Sensing	Proximal Policy Optimization (PPO)	Propose a UAV Swarm System for Aerial Remote Sensing	2022
5	Zhang et al. [129]	Offloading Optimization	Deep Reinforcement Learning (DRL)	Optimization of Resources Allocation in Connected Automated Vehicles Networks	2022
6	Ye et al. [130]	Traffic Control	Advantage actor–critic (A2C)	Autonomous Multi-Intersection Traffic Signal Control	2021
7	Shabir et al. [126]	Offloading Optimization	Asynchronous Advantage actor–critic (A3C)	Optimization of task-offloading decisions at multiple tiers in vehicular fog computing	2022
8	Kwon et al. [131]	Resource Allocation	Multiagent Deep Deterministic Policy Gradient (MADDPG)	Propose a method for joint cell association and resource allocation	2020

3.4. FRL Applications in the Industrial Internet of Things

Guo et al. [132] developed a federated learning-based approach that allows for efficient and adaptable management of mobile edge computing (MEC) in the context of the industrial internet of things (IIoT). More specifically, to handle the problem of network optimization and resource allocation in IIoT networks, the authors deployed a DRL algorithm with FL settings. The proposed system can optimize three essential attributes, i.e., the proportion of task offloading, the allocation of bandwidth, and transmission power. Comprehensive

experimental results showed that their optimization approach can decrease the cost of the system while also decreasing the cost of communication, expressed in normalized terms. Lim et al. [133] proposed a DRL-based framework to determine the most efficient control approach and learning effectiveness in practical use cases, such as autonomous driving and robotics. In particular, the multi-agent environment is considered to perform training and share learning parameters (e.g., gradient) to enhance the training quality and performance. Similar to [134], the actor–critic PPO algorithm is deployed within different RL experimental simulations, namely, OpenAI Gym's CartPole, MoutainCar, Acrobot, and Pendulum. The designed weighted federation policy management enables parameter sharing to solve the multi-agent control problem more effectively and efficiently. At last, the authors demonstrated that the proposed policy control system can be used in more complex and realistic application scenarios. Zhang et al. [135] focused on the research problem of device assignment and resource allocation in distributed IIoT platforms. They designed a three-layer collaborative architecture in FL settings for optimizing device selection and computational resource allocation. By minimizing the training loss with a stochastic optimization approach, the proposed architecture can reduce the delay in data transmission and reduce long-term energy consumption. Moreover, a reinforcement on federated (RoF) scheme is designed and executed in a decentralized manner at edge servers based on deep multi-agent RL. By utilizing a device refinement subroutine that is integrated into the RoF-based method, the suggested architecture can speed up the rate of convergence while still keeping the on-device energy consumption low. The comprehensive simulated experimental evaluation demonstrated that the proposed scheme outperforms the benchmark techniques with regard to both performance and efficiency.

With the explosive availability of smart devices in IIoT, Industry 4.0, and digital twin, many smart devices should be deployed and executed at the same time. As a result, device task assignment and scheduling have become another fundamental problem in industrial applications. Zhang et al. [136] introduced a three-layer collaborative FL-based architecture to find a solution for the resource management and the problem of managing the schedule for devices in IIoT. In particular, the DNN model is trained locally at each industrial device side, and then the model parameters (e.g., the gradients) are aggregated, as described in FedAvg algorithm. In order to enhance the efficiency in FL training among resource-limited devices, an optimization algorithm is created to provide a solution for the resource allocation issue while still strictly following the requirements of the FL epoch and device resource consumption. Consequently, the resource management problem is transferred to Markov's decision process based on a DRL model, which can significantly facilitate the FL training process with satisfied performance. Ho et al. [134] studied the issue of organizing the sequence of tasks in automated warehouses by considering the heterogeneous nature of autonomous robotics. Specifically, the authors defined the long-perspective non-convex queuing control system as an optimization problem by reducing the number of tasks waiting to be processed in the system. Unlike the traditional task-scheduling solutions, the proposed techniques utilized the DRL approach, which can employ the proximal policy optimization method (PPO) to handle the stochastic nature of the task flow and the significant quantity of robots in the system. After that, a proximal weighted FL-based algorithm is implemented to enhance the performance of the PPO agents in geographically distributed warehouses. The experimental results showed that the proposed technique outperforms the current approaches based on simulated data. Yang et al. [137] defined a novel digital twin architecture empowered by IIoT, where the real-time status of the industrial devices can be captured and processed for better intelligent decision making. In the Industry 4.0 environment, regarding the prevalence of the availability of smart industrial devices, the efficiency of data transmission and potential privacy leakage poses significant challenges for both researchers and industrial participants. The proposed architecture can cope with the heterogeneous IIoT devices by optimizing the digital twin platform with FL and DRL. Systematic experiments are conducted to present the advantages of the suggested method over the existing solutions. It is shown that an

asynchronous FL architecture is better for solving the discrete effects issues due to the heterogeneous nature of the IIoT devices. Additionally, the designed digital twin system can make the convergence faster while still keeping high performance in the training phase.

In addition, numerous studies are concentrating on how to enhance the performance of FL-based IIoT applications and the quality of online services as well. Sun et al. [138] deployed a unique structure of digital twin (DT) to authorize IIoT for Industry 4.0 applications. Specifically, in their designed framework, the DT sensors are capable of detecting the features of industrial devices in order to aid and enhance FL performance. In order to alleviate the effects of estimation deviations from the actual value of the DT device state, the authors adaptively adjust and design the aggregation frequency of FL models based on the Lyapunov dynamic deficit queue and DRL under resource constraints. Furthermore, the heterogeneity of IIoT devices is adapted with a clustering-based asynchronous FL framework. Their experimental results confirmed the excellence of their suggested framework over the current benchmark approaches. Messaoud et al. [139] presented a new DRL scheme for federated and dynamic network management in IIoT applications. More specifically, under the circumstance that IIoT devices have a stronger computational capacity, the designed architecture is developed to solve the quality-of-service (QoS) satisfaction and secure data sharing issues. By taking into account transmission power and spreading factors across IIoT slices, the authors can deploy effective resource allocation solutions for differentiated QoS services. There are major two steps in the proposed architecture. First, a multi-agent deep Q-learning-based technique is designed for maximizing self-QoS requirements. Second, a DRL-based framework is implemented for optimal decision making on transmission power and spreading factors based on the shared information among agents. The simulated results showed that the proposed architecture is more sufficient than the latest techniques.

Even though FL is popular in real-world industrial applications, there are still some limitations that may hinder its further adaptation. One of the major challenges is that FL suffers from the heterogeneous nature of the participating workers because the FL system treats all the workers equally without fully considering their computational resources and capacity. To cope with the above issues, several studies focus on optimizing the worker-selection problem to increase training efficiency in FL. Zeng [140] focused on the task-assignment problem in which the participant with weak computational capacity will significantly drag the model training process in the synchronous FL architecture. The key idea of the proposed technique is that worker selection in FL should be based on computational resources . Workers with high/low training resources and computational capacities should be assigned with more/less training intensity. After formulating this research problem to a novel heterogeneous training intensity task, the authors deployed an optimal deterministic algorithm and a DRL approach to evaluate each worker's current capability and network condition. Finally, workers can be assigned adaptively based on their training intensities in a real-time manner. The simulation results demonstrated that the proposed scheme is effective in terms of reducing the waiting time, accelerating the convergence rate and improving the overall training speed in FL applications. In a very similar study, Pang et al. [141] proposed an RL-based algorithm that can be used to select more qualified workers with higher probabilities. In particular, the FL center platform will first evaluate the situation of different collaborations based on each worker's rating feedback in real-time. After that, each worker's weight will be updated iteratively until an optimal group of workers is selected. The proposed solution was found to be superior to other existing approaches based on the experimental results obtained from a real-world dataset.

Except for the major research problems of quality assurance, task assignment, and task scheduling, cybersecurity has become an emerging and extremely important topic for every IIoT application. As malicious attacking models become more and more sophisticated, a large number of novel studies are proposed to cope with the potential privacy and security issues. Wang et al. [142] aimed at detecting anomaly smart devices in IIoT environment.

The motivation behind this work is that even though IIoT is emerging and has significant potential for improving production efficiency and performing better industrial decision making, the anomalies of the smart devices in IIoT are still the major issues and concerns, which may cause serious privacy leakage consequences and cyber threats to real-world IIoT applications. Particularly, the authors deployed the FL technique for establishing a universal anomaly detection framework, where each local model was trained by the DRL algorithm. Since the consideration of the DRL approach is on the local side, the privacy of the client's data can be further protected. The experimental results are evaluated by two novel proposed metrics, namely, the privacy leakage degree and the action relation. Their results validated that the proposed DRL-based architecture can achieve better performance in terms of high accuracy and low latency. Zhang et al. [143] studied the problem of managing time series data efficiently and securely in IIoT applications in a wireless network environment. In particular, DRL was applied to IIoT equipment nodes with accurate models to manage industrial smart equipment data in a wireless network environment. The proposed technology considers both privacy and utility in model training and presents valid corroboration on the effectiveness and security of real-world datasets, such as MNIST, fashion MNIST, and CIFAR-10. Zhang et al. [144] proposed an AI-based collaboration architecture for secure data computation and offloading in cloud-edge power internet of things (PIoT). With blockchain-based techniques, the proposed architecture enables secure and flexible data sharing, resource allocation, energy scheduling, access authentication, and differentiated services. After that, the authors designed and developed a blockchain-empowered FL algorithm for addressing the secure and low-latency issue in computation offloading with the consideration of long-term security constraints and short-term queuing delay. Experimental results verified the efficacy of the suggested techniques.

In summary, both DRL and FL have been widely used in the real-world industrial environments for facilitating intelligent business decision-making, improving the efficiency and quality of the production line, and collaborating with multi-agents simultaneously. In this section, we summarize related works in terms of the domain of interest, proposed technologies, the research problem that is being addressed, and the major research motivations. The detailed comparison between different methods is presented in Table 4. Even though the major issues are well studied and addressed in the existing solutions, cybersecurity-related problems are not defined and evaluated systematically. For instance, even though data privacy can be protected in FL due to the fact that the local data is not transferred to the cloud center. Many studies [145–147] have proved that there is still a big privacy leakage potential in FL settings since the model parameters (e.g., gradients) can still reveal important information about the original data and can be used to infer sensitive knowledge by the attackers. Therefore, knowing how to protect the privacy of the model parameters during their transmission between local clients to the cloud server can be considered an interesting and essential research topic for future work.

Table 4. Overview of FRL-based solutions for industrial IoT (IIoT) operations.

Number	Work	Domain	Technology	Research Purpose	Research Problem	Dataset	Year
1	Wang et al. [142]	IIoT	DRL, FL	Privacy preservation	Anomaly detection	Simulated data	2022
2	Yang et al. [137]	IIoT, digital twin, and industrial 4.0	DRL, FL	Mitigation for data transmission Burden	Real-time device evaluation	MNIST data	2022
3	Zeng [140]	IoT	DRL, FL	Efficient improvement	Task assignment in FL	Simulated Data, MNIST, and CIFAR-10	2022
4	Ho et al. [134]	Autonomous robotic system	DRL, proximal policy optimization	Efficient queue control	Task scheduling	Simulated data	2022
5	Zhang et al. [144]	Power internet of things (PioT)	DRL, FL, and blockchain	Data security, intelligent computation	Cloud-edge collaboration	Simulated data	2022
6	Guo et al. [132]	IIoT, mobile-edge computing	DRL, FL	Efficient resource allocation	Management optimization	Simulated data	2021
7	Sun et al. [138]	IIoT, digital twin	DRL, FL	Performance enhancement with resource constraints	Deviation reduction	MNIST data	2021
8	Messaoud et al. [139]	IIoT	DRL, deep federated Q-learning	Resource allocation, data sharing	Quality-of-service	Simulated data	2021
9	Zhang et al. [143]	IIoT, industrial 4.0	DRL, FL	Efficient and secure data training	Training data management	MNIST, Fashion MNIST, and CIFAR-10	2021
10	Lim et al. [133]	Robotics, autonomous driving	DRL	Control policy optimization	Multi-agent control and management	Simulated data and QUBE-servo system	2021
11	Zhang et al. [135]	IIoT	DRL, FL	Device assignment and resource allocation	On-device resource-consumption management	Simulated data	2021
12	Zhang et al. [136]	IIoT	DRL, FL	Efficient resource allocation	Management optimization	Simulated data	2021
13	Pang et al. [141]	IoT	DRL, FL	Efficient improvement	Task assignment in FL	MNIST, Fashion MNIST, and CIFAR-10	2020

4. Open Challenges

The insights described in the previous section show that there are multiple research gaps in specific areas. Figure 4 summarizes important aspects of the efforts reviewed in this research and highlights their main focuses.

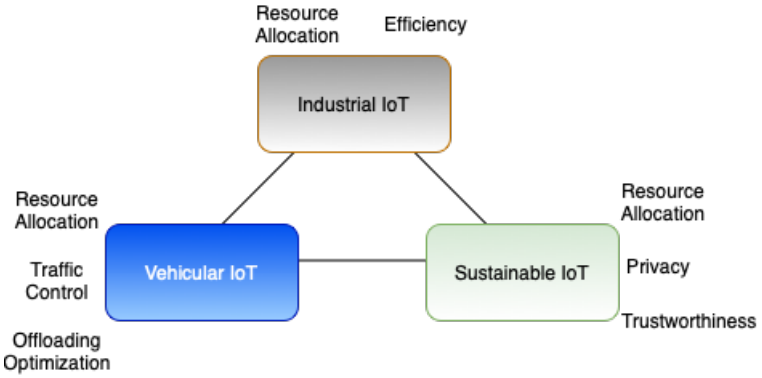

Figure 4. Summary of the main focus of state-of-the-art FRL solutions regarding IoT subdomains.

Based on that, several future directions can be identified. Considering all of the applications of FRL in IoT presented in the previous sections, the immediate future directions are as follows:

- **Integration of adaptable offloading methods, resource allocation, and energy management:** Different solutions focus on individual challenges alongside specific constraints. Future efforts are expected to introduce an integrated approach to cover multiple tasks while offering flexibility regarding important systems' constraints. In fact, the combination of energy management constraints with offloading and allocation solutions can further optimize existing solutions;
- **Mitigation of security threats again distributed learning:** In the past few years, there has been an increase in attacks again federated learning (FL) methods. These attacks threaten the integrity of the knowledge shared and can impact the performance of the overall system. Thereupon, new solutions are needed to mitigate attacks against FRL initiatives to ensure the learning process is not severally affected in different applications;
- **Caching across multi-tenant applications:** As presented in the reviewed papers, caching is a pivotal solution to improve the system performance in different IoT appli-

cations. In this sense, one future direction relies on caching solutions for multi-tenant and multi-service applications, in which separate logical infrastructures operated on the same physical topology with traffic isolation;

- **Prioritized training:** Due to the increasing number of IoT devices across different applications, it becomes challenging to identify trustworthy data feeds throughout the IoT network topology. In this context, another future direction focuses on proving a trust score for different training agents in order to define priorities in the FL aggregation procedure. This can mitigate attacks and ensure legitimate feeds are prioritized;
- **Layered knowledge sharing based on service license agreements (SLA):** Assuming that all knowledge can be shared across different agents, an application can diverge from real-world requirements, in which different organizations can establish special agreements regarding resource sharing. In this sense, the definition of logical channels for knowledge sharing in different FRL applications in IoT is an important direction for future works.

Moreover, there are areas that require solutions for long-term challenges. These areas refer to fundamental contributions in how FRL is used in this context. Examples of long-term open challenges are as follows:

- **Application-specific solutions:** The reviewed efforts adopt different RL methods to solve multiple problems. However, future efforts are expected to include problem-specific mechanisms in the agents' internal training process. This also extends to the aggregation procedure, e.g., methods focused on offloading optimization can have tailored training mechanisms not necessarily present in caching solutions;
- **Continuous adaptability:** The training process of FRL applications in IoT considers several components and constraints. A future direction in this regard relies on adopting dynamic constraints in which initial assumptions evolve throughout the system's operation. These changes can comprise states, prioritized, temporary goals, and special conditions;
- **Large-scale solutions:** Unfolding the solutions proposed by the reviewed works, it is possible to generalize different applications to operate on a global scale. However, scalability can bring multiple obstacles to efficient operations (e.g., global knowledge sharing, the balance of local and global influence, and multi-regional collaborations). In this sense, future endeavors can focus on the aspects of scaling FRL applications in different IoT domains;
- **New FL aggregation methods:** In recent years, there has been an increase in the number of FL aggregation algorithms. In fact, a future direction relies on designing and evaluating new aggregation methods that can consider specific aspects related to the IoT operation. These new methods can simplify global convergence as well as enable more secure training procedures;
- **Deployment and evaluation in a real environment:** Although the solutions reviewed present high-performance solutions for the cases investigated, there is a need for evaluating such strategies in realistic testbeds. Indeed, these efforts involve both the replication of realistic testbeds (e.g., topologies, devices, and connections) and the use of the proposed methods in real operations. Future endeavors can focus on establishing a safe and secure environment for testing such methods.

Finally, several future directions for FRL architectures and approaches can also be identified. Such directions include the design of IoT-specific vertical and horizontal FL, the intersection between security and efficiency, and the use of shared models in IoT and non-IoT scenarios.

5. Conclusions

In the past few years, new FRL solutions have been proposed. Some of these new efforts focus on providing efficient solutions to different IoT problems. In this context, this research presented a literature review of FRL applications in IoT, focusing on analyzing

applications in multiple areas. Throughout the paper, we highlighted existing solutions, their characteristics, and research gaps. Finally, critical short- and long-term challenges were identified to foster the development of new solutions. The application of this promising paradigm in IoT can be beneficial in several ways, and this paper highlighted new research opportunities given the review presented.

Author Contributions: Conceptualization, E.C.P.N., S.S., X.Z. and S.D.; methodology, E.C.P.N., S.S., X.Z. and S.D.; validation, E.C.P.N., S.S., X.Z. and S.D.; formal analysis, E.C.P.N., S.S., X.Z. and S.D.; investigation, E.C.P.N., S.S. and X.Z.; resources, E.C.P.N., S.S., X.Z. and S.D.; data curation, E.C.P.N., S.S., X.Z. and S.D.; writing—original draft preparation, E.C.P.N., S.S., X.Z. and S.D.; writing—review and editing, E.C.P.N., S.S., X.Z. and S.D.; visualization, E.C.P.N., S.S., X.Z. and S.D.; supervision, X.Z. and S.D.; project administration, X.Z. and S.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rose , K.; Eldridge, S.; Chapin, L. The internet of things: An overview. *Internet Soc. ISOC* **2015**, *80*, 1–50.
2. Tan, L.; Wang, N. Future internet: The internet of things. In Proceedings of the 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), Chengdu, China, 20–22 August 2010; Volume 5.
3. Yang, F.; Wang, S.; Li, J.; Liu, Z.; Sun, Q. An overview of internet of vehicles. *China Commun.* **2014**, *11*, 1–15. [CrossRef]
4. Ding, Y.; Jin, M.; Li, S.; Feng, D. Smart logistics based on the internet of things technology: An overview. *Int. J. Logist. Res. Appl.* **2021**, *24*, 323–345. [CrossRef]
5. Ramlowat, D.D.; Pattanayak, B.K. Exploring the internet of things (IoT) in education: A review. In *Proceedings of the Information Systems Design and Intelligent Applications: Proceedings of 5th International Conference INDIA 2018*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 245–255.
6. Verdouw, C.; Wolfert, S.; Tekinerdogan, B. Internet of things in agriculture. *CABI Rev.* **2016**, 1–12. [CrossRef]
7. Pan, J.; Yang, Z. Cybersecurity challenges and opportunities in the new "edge computing + IoT" world. In Proceedings of the ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization, Tempe, AZ, USA, 19–21 March 2018; pp. 29–32.
8. Kaur, B.; Dadkhah, S.; Shoeleh, F.; Neto, E.C.P.; Xiong, P.; Iqbal, S.; Lamontagne, P.; Ray, S.; Ghorbani, A.A. Internet of things (IoT) security dataset evolution: Challenges and future directions. *Internet Things* **2023**, *22*, 100780. [CrossRef]
9. Danso, P.K.; Neto, E.C.P.; Dadkhah, S.; Zohourian, A.; Molyneaux, H.; Ghorbani, A.A. Ensemble-based Intrusion Detection for internet of things Devices. In Proceedings of the IEEE 19th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET), Marietta, GE, USA, 10–12 October 2022; pp. 034–039.
10. Leminen, S.; Rajahonka, M.; Wendelin, R.; Westerlund, M. Industrial internet of things business models in the machine-to-machine context. *Ind. Mark. Manag.* **2020**, *84*, 298–311. [CrossRef]
11. Roy, S.; Rawat, U.; Karjee, J. A lightweight cellular automata based encryption technique for IoT applications. *IEEE Access* **2019**, *7*, 39782–39793. [CrossRef]
12. Cecere, G.; Le Guel, F.; Soulié, N. Perceived Internet privacy concerns on social networks in Europe. *Technol. Forecast. Soc. Chang.* **2015**, *96*, 277–287. [CrossRef]
13. Singh, R.; Dwivedi, A.D.; Srivastava, G.; Chatterjee, P.; Lin, J.C.W. A Privacy Preserving internet of things Smart Healthcare Financial System. *IEEE Internet Things J.* 2023, *Early Access*. [CrossRef]
14. Sfar, A.R.; Challal, Y.; Moyal, P.; Natalizio, E. A game theoretic approach for privacy preserving model in IoT-based transportation. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 4405–4414. [CrossRef]
15. Sutton, R.S. Reinforcement learning: Past, present and future. In *Proceedings of the Simulated Evolution and Learning: Second Asia-Pacific Conference on Simulated Evolution and Learning, SEAL'98, Canberra, Australia, 24–27 November 1998*; Springer: Berlin/Heidelberg, Germany, 1999; pp. 195–197.
16. Sutton, R.S. Open theoretical questions in reinforcement learning. In *Proceedings of the Computational Learning Theory: 4th European Conference, EuroCOLT'99, Nordkirchen, Germany, 29–31 March 1999*; Springer: Berlin/Heidelberg, Germany, 1999; pp. 11–17.
17. Li, T.; Sahu, A.K.; Talwalkar, A.; Smith, V. Federated learning: Challenges, methods, and future directions. *IEEE Signal Process. Mag.* **2020**, *37*, 50–60. [CrossRef]
18. Jin, H.; Peng, Y.; Yang, W.; Wang, S.; Zhang, Z. Federated reinforcement learning with environment heterogeneity. In Proceedings of the International Conference on Artificial Intelligence and Statistics, Virtual Conference, 28–30 March 2022; pp. 18–37.
19. Fu, Y.; Li, C.; Yu, F.R.; Luan, T.H.; Zhang, Y. A Selective federated reinforcement learning Strategy for Autonomous Driving. *IEEE Trans. Intell. Transp. Syst.* **2022**, *24*, 1655–1668. [CrossRef]

20. Xue, Z.; Zhou, P.; Xu, Z.; Wang, X.; Xie, Y.; Ding, X.; Wen, S. A resource-constrained and privacy-preserving edge-computing-enabled clinical decision system: A federated reinforcement learning approach. *IEEE Internet Things J.* **2021**, *8*, 9122–9138. [CrossRef]
21. Wang, H.; Kaplan, Z.; Niu, D.; Li, B. Optimizing federated learning on non-iid data with reinforcement learning. In Proceedings of the IEEE INFOCOM 2020—IEEE Conference on Computer Communications, Online, 6–9 July 2020; pp. 1698–1707.
22. Xu, M.; Peng, J.; Gupta, B.; Kang, J.; Xiong, Z.; Li, Z.; Abd El-Latif, A.A. Multiagent federated reinforcement learning for Secure Incentive Mechanism in Intelligent Cyber–Physical Systems. *IEEE Internet Things J.* **2021**, *9*, 22095–22108. [CrossRef]
23. Qi, J.; Zhou, Q.; Lei, L.; Zheng, K. Federated reinforcement learning: Techniques, applications, and open challenges. *arXiv* **2021**, arXiv:2108.11887.
24. Sutton, R.S.; Barto, A.G. Reinforcement learning. *J. Cogn. Neurosci.* **1999**, *11*, 126–134.
25. Sutton, R.S.; Precup, D.; Singh, S. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artif. Intell.* **1999**, *112*, 181–211. [CrossRef]
26. Sutton, R.S.; McAllester, D.; Singh, S.; Mansour, Y. Policy gradient methods for reinforcement learning with function approximation. *Adv. Neural Inf. Process. Syst.* **1999**, *12*, 1058–1063.
27. Dayan, P.; Niv, Y. Reinforcement learning: The good, the bad and the ugly. *Curr. Opin. Neurobiol.* **2008**, *18*, 185–196. [CrossRef]
28. Li, Y. Deep reinforcement learning: An overview. *arXiv* **2017**, arXiv:1701.07274.
29. Stone, P.; Sutton, R.S. Scaling reinforcement learning toward RoboCup soccer. *ICML* **2001**, *1*, 537–544.
30. Kaelbling, L.P.; Littman, M.L.; Moore, A.W. Reinforcement learning: A survey. *J. Artif. Intell. Res.* **1996**, *4*, 237–285. [CrossRef]
31. Kurach, K.; Raichuk, A.; Stańczyk, P.; Zając, M.; Bachem, O.; Espeholt, L.; Riquelme, C.; Vincent, D.; Michalski, M.; Bousquet, O.; et al. Google research football: A novel reinforcement learning environment. In Proceedings of the AAAI Conference on artificial intelligence, Hilton, NY, USA, 7–12 February 2020; pp. 4501–4510.
32. Zhang, H.; Feng, S.; Liu, C.; Ding, Y.; Zhu, Y.; Zhou, Z.; Zhang, W.; Yu, Y.; Jin, H.; Li, Z. Cityflow: A multi-agent reinforcement learning environment for large scale city traffic scenario. In Proceedings of the World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; pp. 3620–3624.
33. Tizhoosh, H.R. Reinforcement learning based on actions and opposite actions. In Proceedings of the International Conference on Artificial Intelligence and Machine Learning, Hong Kong, China, 14–16 November 2005.
34. Branavan, S.R.; Chen, H.; Zettlemoyer, L.; Barzilay, R. Reinforcement learning for mapping instructions to actions. In Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP, Singapore, 2–7 August 2009; pp. 82–90.
35. Matignon, L.; Laurent, G.J.; Le Fort-Piat, N. Reward function and initial values: Better choices for accelerated goal-directed reinforcement learning. In *Proceedings of the Artificial Neural Networks–ICANN 2006: 16th International Conference, Athens, Greece, 10–14 September 2006*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 840–849.
36. Singh, S.; Jaakkola, T.; Littman, M.L.; Szepesvári, C. Convergence results for single-step on-policy reinforcement-learning algorithms. *Mach. Learn.* **2000**, *38*, 287–308. [CrossRef]
37. Barreto, A.; Hou, S.; Borsa, D.; Silver, D.; Precup, D. Fast reinforcement learning with generalized policy updates. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 30079–30087. [CrossRef]
38. Galatzer-Levy, I.R.; Ruggles, K.V.; Chen, Z. Data science in the Research Domain Criteria era: Relevance of machine learning to the study of stress pathology, recovery, and resilience. *Chronic Stress* **2018**, *2*, 2470547017747553. [CrossRef] [PubMed]
39. Yu, J.; Su, Y.; Liao, Y. The path planning of mobile robot by neural networks and hierarchical reinforcement learning. *Front. Neurorobot.* **2020**, *14*, 63. [CrossRef]
40. Watkins, C.J.; Dayan, P. Q-learning. *Mach. Learn.* **1992**, *8*, 279–292. [CrossRef]
41. Bianchi, R.A.; Ros, R.; Lopez de Mantaras, R. Improving reinforcement learning by using case based heuristics. In *Proceedings of the Case-Based Reasoning Research and Development: 8th International Conference on Case-Based Reasoning, ICCBR 2009, Seattle, WA, USA, 20–23 July 2009*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 75–89.
42. Rahmani, A.M.; Ali, S.; Malik, M.H.; Yousefpoor, E.; Yousefpoor, M.S.; Mousavi, A.; Khan, F.; Hosseinzadeh, M. An energy-aware and Q-learning-based area coverage for oil pipeline monitoring systems using sensors and internet of things. *Sci. Rep.* **2022**, *12*, 9638. [CrossRef]
43. Aihara, N.; Adachi, K.; Takyu, O.; Ohta, M.; Fujii, T. Q-learning aided resource allocation and environment recognition in LoRaWAN with CSMA/CA. *IEEE Access* **2019**, *7*, 152126–152137. [CrossRef]
44. Fan, J.; Wang, Z.; Xie, Y.; Yang, Z. A theoretical analysis of deep Q-learning. In Proceedings of the Learning for Dynamics and Control PMLR, Palo Alto, CA, USA, 10–11 June 2020; pp. 486–489.
45. Brim, A. Deep reinforcement learning pairs trading with a double deep Q-network. In Proceedings of the 10th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 6–8 January 2020; pp. 0222–0227.
46. Zhu, J.; Song, Y.; Jiang, D.; Song, H. A new deep-Q-learning-based transmission scheduling mechanism for the cognitive internet of things. *IEEE Internet Things J.* **2017**, *5*, 2375–2385. [CrossRef]
47. Salh, A.; Audah, L.; Alhartomi, M.A.; Kim, K.S.; Alsamhi, S.H.; Almalki, F.A.; Abdullah, Q.; Saif, A.; Algethami, H. Smart packet transmission scheduling in cognitive IoT systems: DDQN based approach. *IEEE Access* **2022**, *10*, 50023–50036. [CrossRef]
48. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal policy optimization algorithms. *arXiv* **2017**, arXiv:1707.06347.

49. Li, S.; Bing, S.; Yang, S. Distributional advantage actor–critic. *arXiv* **2018**, arXiv:1806.06914.
50. Peng, B.; Li, X.; Gao, J.; Liu, J.; Chen, Y.N.; Wong, K.F. Adversarial advantage actor–critic model for task-completion dialogue policy learning. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 6149–6153.
51. Chen, G.; Xu, X.; Zeng, Q.; Zhang, Y.D. A Vehicle-Assisted Computation Offloading Algorithm Based on Proximal Policy Optimization in Vehicle Edge Networks. *Mob. Netw. Appl.* **2022**, 1–15. [CrossRef]
52. Li, K.; Ni, W.; Yuan, X.; Noor, A.; Jamalipour, A. Deep-Graph-Based reinforcement learning for Joint Cruise Control and Task Offloading for Aerial Edge internet of things (EdgeIoT). *IEEE Internet Things J.* **2022**, *9*, 21676–21686. [CrossRef]
53. Qiu, C.; Hu, Y.; Chen, Y.; Zeng, B. Deep deterministic policy gradient (DDPG)-based energy harvesting wireless communications. *IEEE Internet Things J.* **2019**, *6*, 8577–8588. [CrossRef]
54. Nie, L.; Sun, W.; Wang, S.; Ning, Z.; Rodrigues, J.J.; Wu, Y.; Li, S. Intrusion detection in green internet of things: A deep deterministic policy gradient-based algorithm. *IEEE Trans. Green Commun. Netw.* **2021**, *5*, 778–788. [CrossRef]
55. Babaeizadeh, M.; Frosio, I.; Tyree, S.; Clemons, J.; Kautz, J. Reinforcement learning through asynchronous advantage actor–critic on a gpu. *arXiv* **2016**, arXiv:1611.06256.
56. Zare, M.; Sola, Y.E.; Hasanpour, H. Towards distributed and autonomous IoT service placement in fog computing using asynchronous advantage actor–critic algorithm. *J. King Saud Univ.-Comput. Inf. Sci.* **2023**, *35*, 368–381. [CrossRef]
57. Kairouz, P.; McMahan, H.B.; Avent, B.; Bellet, A.; Bennis, M.; Bhagoji, A.N.; Bonawitz, K.; Charles, Z.; Cormode, G.; Cummings, R.; et al. Advances and open problems in federated learning. *Found. Trends Mach. Learn.* **2021**, *14*, 1–210. [CrossRef]
58. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; y Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the artificial intelligence and Statistics, PMLR, Fort Lauderdale, FL, USA, 20–22 April 2017; pp. 1273–1282.
59. Zhang, X.; Lu, R.; Shao, J.; Wang, F.; Zhu, H.; Ghorbani, A.A. FedSky: An efficient and privacy-preserving scheme for federated mobile crowdsensing. *IEEE Internet Things J.* **2021**, *9*, 5344–5356. [CrossRef]
60. Hao, M.; Li, H.; Luo, X.; Xu, G.; Yang, H.; Liu, S. Efficient and privacy-enhanced federated learning for industrial artificial intelligence. *IEEE Trans. Ind. Inf.* **2019**, *16*, 6532–6542. [CrossRef]
61. Lu, Y.; Huang, X.; Dai, Y.; Maharjan, S.; Zhang, Y. Federated learning for data privacy preservation in vehicular cyber-physical systems. *IEEE Netw.* **2020**, *34*, 50–56. [CrossRef]
62. Kim, H.; Park, J.; Bennis, M.; Kim, S.L. Blockchained on-device federated learning. *IEEE Commun. Lett.* **2019**, *24*, 1279–1283. [CrossRef]
63. Kumar, P.; Gupta, G.P.; Tripathi, R. PEFL: Deep Privacy-Encoding based federated learning Framework for Smart Agriculture. *IEEE Micro* **2021**, *42*, 33–40. [CrossRef]
64. Yang, Q.; Liu, Y.; Chen, T.; Tong, Y. Federated Machine Learning: Concept and Applications. *ACM Trans. Intell. Syst. Technol.* **2019**, *10*, 1–19. [CrossRef]
65. Lo, S.K.; Lu, Q.; Zhu, L.; Paik, H.Y.; Xu, X.; Wang, C. Architectural patterns for the design of federated learning systems. *J. Syst. Softw.* **2022**, *191*, 111357. [CrossRef]
66. Liu, Y.; Kang, Y.; Zou, T.; Pu, Y.; He, Y.; Ye, X.; Ouyang, Y.; Zhang, Y.Q.; Yang, Q. Vertical federated learning. *arXiv* **2022**, arXiv:2211.12814.
67. Chen, T.; Jin, X.; Sun, Y.; Yin, W. Vafl: A method of vertical asynchronous federated learning. *arXiv* **2020**, arXiv:2007.06081.
68. Liu, Y.; Zhang, X.; Wang, L. Asymmetrical vertical federated learning. *arXiv* **2020**, arXiv:2004.07427.
69. Gao, D.; Ju, C.; Wei, X.; Liu, Y.; Chen, T.; Yang, Q. Hhhfl: Hierarchical heterogeneous horizontal federated learning for electroencephalography. *arXiv* **2019**, arXiv:1909.05784.
70. Karimireddy, S.P.; Jaggi, M.; Kale, S.; Mohri, M.; Reddi, S.; Stich, S.U.; Suresh, A.T. Breaking the centralized barrier for cross-device federated learning. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 28663–28676.
71. ur Rehman, M.H.; Dirir, A.M.; Salah, K.; Damiani, E.; Svetinovic, D. TrustFed: A framework for fair and trustworthy cross-device federated learning in IIoT. *IEEE Trans. Ind. Inf.* **2021**, *17*, 8485–8494. [CrossRef]
72. Yang, W.; Wang, N.; Guan, Z.; Wu, L.; Du, X.; Guizani, M. A practical cross-device federated learning framework over 5g networks. *IEEE Wirel. Commun.* **2022**, *29*, 128–134. [CrossRef]
73. Huang, C.; Huang, J.; Liu, X. Cross-Silo federated learning: Challenges and Opportunities. *arXiv* **2022**, arXiv:2206.12949.
74. Huang, Y.; Chu, L.; Zhou, Z.; Wang, L.; Liu, J.; Pei, J.; Zhang, Y. Personalized cross-silo federated learning on non-iid data. In Proceedings of the AAAI Conference on artificial intelligence, Online, 2–9 February 2021; pp. 7865–7873.
75. Jiang, Z.; Wang, W.; Liu, Y. Flashe: Additively symmetric homomorphic encryption for cross-silo federated learning. *arXiv* **2021**, arXiv:2109.00675.
76. Zhang, Y.; Zeng, D.; Luo, J.; Xu, Z.; King, I. A Survey of Trustworthy federated learning with Perspectives on Security, Robustness, and Privacy. *arXiv* **2023**, arXiv:2302.10637.
77. Yang, Z.; Shi, Y.; Zhou, Y.; Wang, Z.; Yang, K. Trustworthy federated learning via blockchain. *IEEE Internet Things J.* **2022**, *10*, 92–109. [CrossRef]
78. Nguyen, T.D.; Marchal, S.; Miettinen, M.; Fereidooni, H.; Asokan, N.; Sadeghi, A.R. DÏoT: A federated self-learning anomaly detection system for IoT. In Proceedings of the IEEE 39th International conference on distributed computing systems (ICDCS), Dallas, TX, USA, 7–9 July 2019; pp. 756–767.

79. Mothukuri, V.; Khare, P.; Parizi, R.M.; Pouriyeh, S.; Dehghantanha, A.; Srivastava, G. Federated-learning-based anomaly detection for iot security attacks. *IEEE Internet Things J.* **2021**, *9*, 2545–2554. [CrossRef]
80. Zhang, C.; Li, M.; Wu, D. Federated Multidomain Learning With Graph Ensemble Autoencoder GMM for Emotion Recognition. *IEEE Trans. Intell. Transp. Syst.* 2022, *Early Access*. [CrossRef]
81. Hamid, O.H. Data-Centric and Model-Centric AI: Twin Drivers of Compact and Robust Industry 4.0 Solutions. *Appl. Sci.* **2023**, *13*, 2753. [CrossRef]
82. Hamid, O.H.; Braun, J. Reinforcement learning and attractor neural network models of associative learning. In *Proceedings of the Computational Intelligence: 9th International Joint Conference, IJCCI 2017, Funchal, Portugal, 1–3 November 2017*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 327–349.
83. Espeholt, L.; Soyer, H.; Munos, R.; Simonyan, K.; Mnih, V.; Ward, T.; Doron, Y.; Firoiu, V.; Harley, T.; Dunning, I.; et al. Impala: Scalable distributed deep-rl with importance weighted actor-learner architectures. In Proceedings of the International Conference on Machine Learning, PMLR, Vienna, Austria, 25–31 July 2018; pp. 1407–1416.
84. Hoffman, M.W.; Shahriari, B.; Aslanides, J.; Barth-Maron, G.; Momchev, N.; Sinopalnikov, D.; Stańczyk, P.; Ramos, S.; Raichuk, A.; Vincent, D.; et al. Acme: A research framework for distributed reinforcement learning. *arXiv* **2020**, arXiv:2006.00979.
85. Tolpegin, V.; Truex, S.; Gursoy, M.E.; Liu, L. Data poisoning attacks against federated learning systems. In Proceedings of the Computer Security–ESORICS 2020: 25th European Symposium on Research in Computer Security, ESORICS 2020, Guildford, UK, 14–18 September 2020; Part I 25; Springer: Berlin/Heidelberg, Germany, 2020; pp. 480–501.
86. Lim, H.K.; Kim, J.B.; Heo, J.S.; Han, Y.H. Federated reinforcement learning for training control policies on multiple IoT devices. *Sensors* **2020**, *20*, 1359. [CrossRef] [PubMed]
87. Liang, X.; Liu, Y.; Chen, T.; Liu, M.; Yang, Q. Federated transfer reinforcement learning for autonomous driving. In *Federated and Transfer Learning*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 357–371.
88. Ali, R.; Zikria, Y.B.; Garg, S.; Bashir, A.K.; Obaidat, M.S.; Kim, H.S. A federated reinforcement learning framework for incumbent technologies in beyond 5G networks. *IEEE Netw.* **2021**, *35*, 152–159. [CrossRef]
89. Rjoub, G.; Bentahar, J.; Wahab, O.A. Explainable AI-based federated deep reinforcement learning for Trusted Autonomous Driving. In Proceedings of the 2022 International Wireless Communications and Mobile Computing (IWCMC), Marrakesh, Morocco, 19–23 June 2022; pp. 318–323.
90. Na, S.; Krajník, T.; Lennox, B.; Arvin, F. Federated reinforcement learning for Collective Navigation of Robotic Swarms. *arXiv* **2022**, arXiv:2202.01141.
91. Otoum, S.; Guizani, N.; Mouftah, H. Federated reinforcement learning-supported IDS for IoT-steered healthcare systems. In Proceedings of the ICC 2021-IEEE International Conference on Communications, Virtual, 4–23 June 2021; pp. 1–6.
92. Zhu, R.; Li, M.; Liu, H.; Liu, L.; Ma, M. Federated deep reinforcement learning-Based Spectrum Access Algorithm With Warranty Contract in Intelligent Transportation Systems. *IEEE Trans. Intell. Transp. Syst.* **2022**, *24*, 1178–1190. [CrossRef]
93. Tiwari, P.; Lakhan, A.; Jhaveri, R.H.; Gronli, T.M. Consumer-Centric Internet of Medical Things for Cyborg Applications based on federated reinforcement learning. *IEEE Trans. Consum. Electron.* 2023, *Early Access*. [CrossRef]
94. Li, F.; Shen, B.; Guo, J.; Lam, K.Y.; Wei, G.; Wang, L. Dynamic spectrum access for internet-of-things based on federated deep reinforcement learning. *IEEE Trans. Veh. Technol.* **2022**, *71*, 7952–7956. [CrossRef]
95. Lu, Y.; Huang, X.; Zhang, K.; Maharjan, S.; Zhang, Y. Blockchain and federated learning for 5G beyond. *IEEE Netw.* **2020**, *35*, 219–225. [CrossRef]
96. Tiwari, M.; Misra, S.; Bishoyi, P.K.; Yang, L.T. Devote: Criticality-aware federated service provisioning in fog-based iot environments. *IEEE Internet Things J.* **2021**, *8*, 10631–10638. [CrossRef]
97. Xu, Y.; Bhuiyan, M.Z.A.; Wang, T.; Zhou, X.; Singh, A.K. C-fdrl: Context-aware privacy-preserving offloading through federated deep reinforcement learning in cloud-enabled IoT. *IEEE Trans. Ind. Inf.* **2022**, *19*, 1155–1164. [CrossRef]
98. Miao, Q.; Lin, H.; Wang, X.; Hassan, M.M. Federated deep reinforcement learning based secure data sharing for internet of things. *Comput. Netw.* **2021**, *197*, 108327. [CrossRef]
99. Zheng, J.; Li, K.; Mhaisen, N.; Ni, W.; Tovar, E.; Guizani, M. Exploring Deep-Reinforcement-Learning-Assisted federated learning for Online Resource Allocation in Privacy-Preserving EdgeIoT. *IEEE Internet Things J.* **2022**, *9*, 21099–21110. [CrossRef]
100. Anwar, A.; Raychowdhury, A. Multi-task federated reinforcement learning with adversaries. *arXiv* **2021**, arXiv:2103.06473.
101. Nguyen, T.G.; Phan, T.V.; Hoang, D.T.; Nguyen, T.N.; So-In, C. Federated deep reinforcement learning for traffic monitoring in SDN-based IoT networks. *IEEE Trans. Cogn. Commun. Netw.* **2021**, *7*, 1048–1065. [CrossRef]
102. Wang, X.; Hu, J.; Lin, H.; Garg, S.; Kaddoum, G.; Piran, M.J.; Hossain, M.S. QoS and privacy-aware routing for 5G-enabled industrial internet of things: A federated reinforcement learning approach. *IEEE Trans. Ind. Inf.* **2021**, *18*, 4189–4197. [CrossRef]
103. Huang, H.; Zeng, C.; Zhao, Y.; Min, G.; Zhu, Y.; Miao, W.; Hu, J. Scalable orchestration of service function chains in NFV-enabled networks: A federated reinforcement learning approach. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 2558–2571. [CrossRef]
104. Yu, S.; Chen, X.; Zhou, Z.; Gong, X.; Wu, D. When deep reinforcement learning meets federated learning: Intelligent multitimescale resource management for multiaccess edge computing in 5G ultradense network. *IEEE Internet Things J.* **2020**, *8*, 2238–2251. [CrossRef]
105. Xiaofeng Fan, F.; Ma, Y.; Dai, Z.; Jing, W.; Tan, C.; Low, B.K.H. Fault-Tolerant federated reinforcement learning with Theoretical Guarantee. *arXiv* **2021**, arXiv:2110.14074v2.

106. Rjoub, G.; Wahab, O.A.; Bentahar, J.; Cohen, R.; Bataineh, A.S. Trust-augmented deep reinforcement learning for federated learning client selection. *Inf. Syst. Front.* **2022**, 1–18. [CrossRef]
107. Raghu, A.; Komorowski, M.; Ahmed, I.; Celi, L.; Szolovits, P.; Ghassemi, M. Deep reinforcement learning for sepsis treatment. *arXiv* **2017**, arXiv:1711.09602.
108. Tahir, A.M.; Chowdhury, M.E.; Khandakar, A.; Rahman, T.; Qiblawey, Y.; Khurshid, U.; Kiranyaz, S.; Ibtehaz, N.; Rahman, M.S.; Al-Maadeed, S.; et al. COVID-19 infection localization and severity grading from chest X-ray images. *Comput. Biol. Med.* **2021**, *139*, 105002. [CrossRef] [PubMed]
109. Qiu, D.; Xue, J.; Zhang, T.; Wang, J.; Sun, M. Federated reinforcement learning for smart building joint peer-to-peer energy and carbon allowance trading. *Appl. Energy* **2023**, *333*, 120526. [CrossRef]
110. Jarwan, A.; Ibnkahla, M. Edge-Based federated deep reinforcement learning for IoT Traffic Management. *IEEE Internet Things J.* **2022**, *10*, 3799–3813. [CrossRef]
111. Wu, D.; Ullah, R.; Harvey, P.; Kilpatrick, P.; Spence, I.; Varghese, B. Fedadapt: Adaptive offloading for iot devices in federated learning. *IEEE Internet Things J.* **2022**, *9*, 20889–20901. [CrossRef]
112. Zang, L.; Zhang, X.; Guo, B. Federated deep reinforcement learning for online task offloading and resource allocation in WPC-MEC networks. *IEEE Access* **2022**, *10*, 9856–9867. [CrossRef]
113. Ren, J.; Wang, H.; Hou, T.; Zheng, S.; Tang, C. Federated learning-Based Computation Offloading Optimization in Edge Computing-Supported internet of things. *IEEE Access* **2019**, *7*, 69194–69201. [CrossRef]
114. Chen, X.; Liu, G. Federated deep reinforcement learning-based task offloading and resource allocation for smart cities in a mobile edge network. *Sensors* **2022**, *22*, 4738. [CrossRef]
115. Zarandi, S.; Tabassum, H. Federated double deep Q-learning for joint delay and energy minimization in IoT networks. In Proceedings of the IEEE International Conference on Communications Workshops (ICC Workshops), Virtual, 14–15 June 2021; pp. 1–6.
116. Guo, Q.; Tang, F.; Kato, N. Federated reinforcement learning-Based Resource Allocation in D2D-Enabled 6G. *IEEE Netw.* 2022, *Early Access*. [CrossRef]
117. Tianqing, Z.; Zhou, W.; Ye, D.; Cheng, Z.; Li, J. Resource allocation in IoT edge computing via concurrent federated reinforcement learning. *IEEE Internet Things J.* **2021**, *9*, 1414–1426. [CrossRef]
118. Nguyen, H.T.; Luong, N.C.; Zhao, J.; Yuen, C.; Niyato, D. Resource allocation in mobility-aware federated learning networks: A deep reinforcement learning approach. In Proceedings of the IEEE 6th World Forum on internet of things (WF-IoT), New Orleans, LO, USA, 2–16 June 2020; pp. 1–6.
119. Cui, Y.; Cao, K.; Wei, T. Reinforcement learning-Based Device Scheduling for Renewable Energy-Powered federated learning. *IEEE Trans. Ind. Inf.* **2022**, *19*, 6264–6274. [CrossRef]
120. Gao, J.; Wang, W.; Campbell, B. Residential Energy Management System Using Personalized Federated deep reinforcement learning. In Proceedings of the 2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), Milano, Italy, 4–6 May 2022; pp. 541–542.
121. Wang, X.; Wang, C.; Li, X.; Leung, V.C.; Taleb, T. Federated deep reinforcement learning for internet of things with decentralized cooperative edge caching. *IEEE Internet Things J.* **2020**, *7*, 9441–9455. [CrossRef]
122. Majidi, F.; Khayyambashi, M.R.; Barekatain, B. Hfdrl: An intelligent dynamic cooperate cashing method based on hierarchical federated deep reinforcement learning in edge-enabled iot. *IEEE Internet Things J.* **2021**, *9*, 1402–1413. [CrossRef]
123. Baghban, H.; Rezapour, A.; Hsu, C.H.; Nuannimnoi, S.; Huang, C.Y. Edge-AI: IoT Request Service Provisioning in Federated Edge Computing Using actor–critic reinforcement learning. *IEEE Trans. Eng. Manag.* 2022, *Early Access*. [CrossRef]
124. Sethi, V.; Pal, S. FedDOVe: A Federated Deep Q-learning-based Offloading for Vehicular fog computing. *Future Gener. Comput. Syst.* **2023**, *141*, 96–105. [CrossRef]
125. Hao, M.; Ye, D.; Wang, S.; Tan, B.; Yu, R. URLLC resource slicing and scheduling for trustworthy 6G vehicular services: A federated reinforcement learning approach. *Phys. Commun.* **2021**, *49*, 101470. [CrossRef]
126. Shabir, B.; Rahman, A.U.; Malik, A.W.; Buyya, R.; Khan, M.A. A federated multi-agent deep reinforcement learning for vehicular fog computing. *J. Supercomput.* **2022**, *79*, 6141–6167. [CrossRef]
127. Lee, W. Federated reinforcement learning-Based UAV Swarm System for Aerial Remote Sensing. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 4327380. [CrossRef]
128. Salameh, H.B.; Alhafnawi, M.; Masadeh, A.; Jararweh, Y. Federated reinforcement learning approach for detecting uncertain deceptive target using autonomous dual UAV system. *Inf. Process. Manag.* **2023**, *60*, 103149. [CrossRef]
129. Zhang, Q.; Wen, H.; Liu, Y.; Chang, S.; Han, Z. Federated-Reinforcement-Learning-Enabled Joint Communication, Sensing, and Computing Resources Allocation in Connected Automated Vehicles Networks. *IEEE Internet Things J.* **2022**, *9*, 23224–23240. [CrossRef]
130. Ye, Y.; Zhao, W.; Wei, T.; Hu, S.; Chen, M. Fedlight: Federated reinforcement learning for autonomous multi-intersection traffic signal control. In Proceedings of the 2021 58th ACM/IEEE Design Automation Conference (DAC), San Francisco, CA, USA, 5–9 December 2021; pp. 847–852.
131. Kwon, D.; Jeon, J.; Park, S.; Kim, J.; Cho, S. Multiagent DDPG-based deep learning for smart ocean federated learning IoT networks. *IEEE Internet Things J.* **2020**, *7*, 9895–9903. [CrossRef]

132. Guo, Y.; Zhao, Z.; He, K.; Lai, S.; Xia, J.; Fan, L. Efficient and flexible management for industrial internet of things: A federated learning approach. *Comput. Netw.* **2021**, *192*, 108122. [CrossRef]
133. Lim, H.K.; Kim, J.B.; Ullah, I.; Heo, J.S.; Han, Y.H. Federated reinforcement learning acceleration method for precise control of multiple devices. *IEEE Access* **2021**, *9*, 76296–76306. [CrossRef]
134. Ho, T.M.; Nguyen, K.K.; Cheriet, M. Federated deep reinforcement learning for task scheduling in heterogeneous autonomous robotic system. *IEEE Trans. Autom. Sci. Eng.* 2022, *Early Access*. [CrossRef]
135. Zhang, W.; Yang, D.; Wu, W.; Peng, H.; Zhang, N.; Zhang, H.; Shen, X. Optimizing federated learning in distributed industrial IoT: A multi-agent approach. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 3688–3703. [CrossRef]
136. Zhang, W.; Yang, D.; Wu, W.; Peng, H.; Zhang, H.; Shen, X.S. Spectrum and computing resource management for federated learning in distributed industrial IoT. In Proceedings of the 2021 IEEE International Conference on Communications Workshops (ICC Workshops), Montreal, QC, Canada, 14–23 June 2021; pp. 1–6.
137. Yang, W.; Xiang, W.; Yang, Y.; Cheng, P. Optimizing federated learning with deep reinforcement learning for digital twin empowered industrial IoT. *IEEE Trans. Ind. Inf.* **2022**, *19*, 1884–1893. [CrossRef]
138. Sun, W.; Lei, S.; Wang, L.; Liu, Z.; Zhang, Y. Adaptive federated learning and digital twin for industrial internet of things. *IEEE Trans. Ind. Inf.* **2020**, *17*, 5605–5614. [CrossRef]
139. Messaoud, S.; Bradai, A.; Ahmed, O.B.; Quang, P.T.A.; Atri, M.; Hossain, M.S. Deep federated Q-learning-based network slicing for industrial IoT. *IEEE Trans. Ind. Inf.* **2020**, *17*, 5572–5582. [CrossRef]
140. Zeng, M.; Wang, X.; Pan, W.; Zhou, P. Heterogeneous Training Intensity for federated learning: A Deep reinforcement learning Approach. *IEEE Trans. Netw. Sci. Eng.* **2022**, *10*, 990–1002. [CrossRef]
141. Pang, J.; Huang, Y.; Xie, Z.; Han, Q.; Cai, Z. Realizing the heterogeneity: A self-organized federated learning framework for IoT. *IEEE Internet Things J.* **2020**, *8*, 3088–3098. [CrossRef]
142. Wang, X.; Garg, S.; Lin, H.; Hu, J.; Kaddoum, G.; Piran, M.J.; Hossain, M.S. Toward accurate anomaly detection in Industrial internet of things using hierarchical federated learning. *IEEE Internet Things J.* **2021**, *9*, 7110–7119. [CrossRef]
143. Zhang, P.; Wang, C.; Jiang, C.; Han, Z. Deep reinforcement learning assisted federated learning algorithm for data management of IIoT. *IEEE Trans. Ind. Inf.* **2021**, *17*, 8475–8484. [CrossRef]
144. Zhang, S.; Wang, Z.; Zhou, Z.; Wang, Y.; Zhang, H.; Zhang, G.; Ding, H.; Mumtaz, S.; Guizani, M. Blockchain and federated deep reinforcement learning Based Secure Cloud-Edge-End Collaboration in Power IoT. *IEEE Wirel. Commun.* **2022**, *29*, 84–91. [CrossRef]
145. Melis, L.; Song, C.; De Cristofaro, E.; Shmatikov, V. Exploiting unintended feature leakage in collaborative learning. In Proceedings of the IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 19–23 May 2019; pp. 691–706.
146. Fredrikson, M.; Jha, S.; Ristenpart, T. Model inversion attacks that exploit confidence information and basic countermeasures. In Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, Denver, CO, USA, 12–16 October 2015; pp. 1322–1333.
147. Hitaj, B.; Ateniese, G.; Perez-Cruz, F. Deep models under the GAN: Information leakage from collaborative deep learning. In Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, Dallas, TX, USA, 30 October –3 November 2017; pp. 603–618.