TCSA
https://dergipark.org.tr/tr/pub/tcsa
Vol.1 No.2
May 2024: 1-2
Accepted 7 Nov 2024
Published 20 Dec 2024

Transactions on Computer Science and Applications

Galatasaray University TCSA Transactions on Computer Science and Applications

ISSN(E): 3023-8129

Federated Learning and Resource-Constrained Embedded Systems: A Comprehensive Survey

Eda Bahar [a*], Ozgun Pinarer [b],
[a] Galatasaray University, Computer Engineering Department, Istanbul, Turkey, 34349, ORCID: 0009-0000-2243-9266.
[b] Galatasaray University, Computer Engineering Department, Istanbul, Turkey, 34349, ORCID: 0000-0002-0280-3689.

*** Corresponding author email address**: ebahar@gsu.edu.tr, **phone:** +902122244480

Abstract

Federated Learning (FL) has become a transformative approach in machine learning, allowing decentralized training of models across multiple devices while preserving data privacy. This paradigm addresses critical concerns related to data privacy, security, and communication overhead, making it particularly relevant for applications in domains such as healthcare, finance, and the Internet of Things (IoT). Resource-constrained FL extends this concept to environments where computational, communication, and energy resources are limited, such as edge networks and IoT devices. This extension focuses on optimizing various aspects of the learning process to enable effective model training even in resource-limited settings. The primary aim of this survey is to provide a comprehensive and structured overview of the current state of research in FL and resource-constrained FL. By examining 62 key publications, this survey synthesizes insights and developments across these domains, highlighting advancements, challenges, and gaps that exist. This survey aims to provide a holistic view of the advancements and ongoing challenges in FL and resource-constrained FL. It identifies research gaps and proposes future directions, such as improving communication efficiency, developing adaptive learning algorithms, and enhancing resource management strategies. This survey serves as a valuable resource for researchers, practitioners, and stakeholders in the field, offering practical insights and guiding future exploration and innovation in FL and its applications in resource-constrained environments.

Keywords: Decentralized Training, Edge Computing, Privacy-Preserving Algorithms, IoT Applications, Computational Efficiency

1. Introduction

Federated Learning (FL) is an innovative paradigm in machine learning that enables decentralized training of models across multiple devices while ensuring that data remains localized. This approach addresses critical concerns related to data privacy, security, and communication overhead, making it particularly suitable for applications in healthcare, finance, the Internet of Things (IoT), and other domains where sensitive data is involved. Unlike traditional centralized learning, FL leverages the computational power of edge devices to collaboratively learn a shared model without the need to centralize the data, thus mitigating privacy risks and reducing the dependency on robust centralized infrastructure.

While several surveys on FL already exist, these works primarily focus on the general principles and applications of FL, often overlooking the nuanced challenges that arise in resource-constrained environments. This survey fills this gap by providing a detailed analysis of FL in contexts where computational, communication, and energy resources are severely limited. Resource-constrained FL extends the basic FL paradigm to environments such as edge networks, IoT devices, and embedded systems, where efficient resource utilization is critical for practical deployment. Existing surveys either broadly address FL or provide only a brief overview of resource-constrained environments, without delving into the specific technical challenges involved.

Our work is distinct from previous surveys in several key ways. First, we narrow the focus to resource-constrained FL, addressing specific issues like energy efficiency, limited bandwidth, and the impact of non-independent and identically distributed (non-IID) data on model performance. Second, this survey synthesizes recent advancements, particularly from 2023 and 2024, in addressing these challenges—areas that are underrepresented in earlier surveys. Third, we categorize the literature based on research themes, including security, privacy, scalability, and optimization in constrained settings, making it easier to identify research gaps and future directions.

The primary aim of this survey is to provide a comprehensive and structured overview of the current state of research in FL and resource-constrained FL. By examining 62 key publications, this survey synthesizes insights and developments across these domains, highlighting the advancements, challenges, and gaps that exist. By categorizing the literature into targeted sections, this survey offers a unique contribution by focusing on the intersection of FL and resource-limited environments, offering fresh perspectives on the challenges that remain unsolved in this emerging field.

The remainder of the study is organized as: Section 2 presents the enabled technologies on FL. Section 3 introduces the existing survey publications on FL. Resource-constrained FL applications are presented in Section 4 and the survey publications are in Section 5. Section 6 gives a global evaluation of these examined studies. Section 7 presents the impact of stragglers on FL performance. Finally, the conclusion and discussion are given in Section 7.

2. Enabled Technologies on Federated Learning

FL has emerged as a revolutionary approach that enables decentralized model training by using data distributed across multiple edge devices, thus preserving data privacy and reducing latency. This section explores various existing studies that propose concrete methodologies and applications of FL across diverse domains. These studies address challenges such as model accuracy, communication efficiency, data heterogeneity, and privacy preservation. By examining these contributions, we gain insight into the advancements and practical implementations of FL, showcasing its potential to transform fields ranging from healthcare and finance to IoT and personalized services. Table 1 summarizes the existing studies on this research area.

Nguyen et al. (2020a) propose a fast-convergent FL algorithm that optimizes device sampling to accelerate model convergence. It adapts to communication and computation heterogeneity, achieving near-optimal loss reduction per communication round. The study demonstrates superior accuracy, convergence speed, and model stability across various tasks and datasets compared to existing algorithms.

Ziller et al. (2021) introduce PySyft, an open-source library designed to facilitate secure and private machine learning algorithms by integrating with popular frameworks like PyTorch. PySyft aims to popularize privacy-preserving techniques and provide a flexible platform for implementing new FL, Multi-Party Computation, and Differential Privacy methods. The chapter details PySyft's functionalities, demonstrates a FL workflow, reviews its academic applications, and introduces Duet, a tool for simplifying FL for researchers and data owners.

Chen et al. (2020) propose a novel framework called collaborative federated learning (CFL) to enhance the deployment of FL in resource-constrained systems. Unlike traditional FL, which requires direct connections to a central controller, CFL reduces this dependency by enabling edge devices to collaborate more autonomously. The paper discusses various communication techniques and performance metrics to optimize CFL and highlights future research opportunities. The CFL framework aims to improve FL's applicability in large-scale wireless systems and IoT environments, potentially transforming applications such as mobile keyboard prediction and autonomous vehicle monitoring.

Li et al. (2021a) introduce Ditto, a framework designed to enhance fairness and robustness in FL. Ditto addresses the trade-off between these two goals—fairness in performance across devices and robustness against data and model poisoning attacks—by offering a personalized FL approach. The paper provides both theoretical analysis and empirical evidence demonstrating that Ditto effectively balances fairness and robustness, outperforming existing methods in terms of accuracy and resilience across various federated datasets.

Li, He, and Song (2021b) introduce MOON, a novel FL framework that addresses the challenge of heterogeneous local data distributions by using model-level contrastive learning. MOON improves FL performance by aligning model representations across parties, enhancing local training accuracy. Their experiments demonstrate that MOON outperforms existing FL methods in image classification tasks, offering a more effective solution for handling data distribution variability.

Wang et al. (2020) introduce Favor, a framework that optimizes FL by addressing challenges posed by non-IID data. Favor uses reinforcement learning to intelligently select client devices for each training round, reducing communication rounds and speeding up convergence. By applying Deep Q-learning, the framework effectively profiles data distribution based on model weights and reduces the number of communication rounds required by up to 49% on MNIST, 23% on FashionMNIST, and 42% on CIFAR-10, compared to traditional methods like Federated Averaging.

Ghosh et al. (2020) introduce the Iterative Federated Clustering Algorithm (IFCA) to address FL with users grouped into clusters, each with distinct learning tasks. IFCA alternates between estimating cluster identities and optimizing model parameters, demonstrating exponential convergence rates in both linear and non-convex settings. The framework effectively handles ambiguous clustering and improves efficiency in clustered FL scenarios, outperforming existing methods on various benchmarks.

3. Surveys on FL

The rapid evolution of FL has prompted extensive research and the publication of numerous surveys that synthesize and evaluate existing knowledge in the field. This section provides an overview of these comprehensive surveys, highlighting key themes, methodologies, and emerging trends in FL research. These surveys serve as invaluable resources for understanding the current state of the art, identifying research gaps, and guiding future investigations. By systematically reviewing these surveys, we aim to present a cohesive picture of the FL landscape, encompassing its theoretical foundations, technological advancements, and practical applications. Table 2 presents the classification of these surveys.

Li, Fan, Tse, and Lin (2020a) review FL and its applications in industrial engineering, highlighting its role in addressing data silos and privacy issues. The study identifies key research areas and challenges in FL development and provides insights into optimization strategies and future applications. It aims to guide future research and application of FL in both industrial engineering and computer science.

Mammen (2021) explores FL, highlighting its potential for privacy-preserving machine learning across critical domains like healthcare and finance. The study discusses FL's opportunities and challenges, noting its vulnerability to attacks despite its privacy advantages. It emphasizes the need for research on developing Byzantine-tolerant FL models that ensure privacy with low computational costs.

Table 1: Existing Studies on FL.

Study	Objective	Methodology	Key Findings	Applications
Nguyen et al. (2020a)	Improve efficiency in FL for resource-constrained devices	Utilizes adaptive model aggregation and client selection	Achieves better model performance with fewer devices, reducing computational and communication costs	IoT networks with limited resources
Ziller et al. (2021)	Enhance privacy in FL systems	Proposes a privacy-preserving aggregation scheme with differential privacy	Demonstrates reduced privacy risks while maintaining model accuracy	Healthcare and sensitive data applications
Chen et al. (2020)	Optimize resource usage in FL	Introduces a resource-aware FL framework with dynamic resource allocation	Improves resource utilization and reduces energy consumption	Edge computing environments
Li et al. (2021a)	Scale FL to large numbers of devices	Proposes a scalable model aggregation method with hierarchical clustering	Enhances scalability and convergence speed with large device networks	Large-scale IoT networks
Li, He, and Song (2021b)	Adapt FL to dynamic network conditions	Uses adaptive learning rates and dynamic aggregation frequency	Improves model performance and network efficiency under varying conditions	Dynamic IoT and edge computing environments
Wang et al. (2020)	Facilitate collaboration among heterogeneous devices	Implements a collaborative learning framework with heterogeneous device support	Achieves high model accuracy across diverse device types and data	Diverse and heterogeneous IoT devices
Ghosh et al. (2020)	Enhance learning efficiency through local updates	Proposes a method for local model updates before global aggregation	Reduces communication overhead and improves convergence speed	IoT networks with limited bandwidth

Nguyen et al. (2021) provide a comprehensive survey of FL applications on the IoT, emphasizing FL's role in enabling privacy-preserving and scalable AI for IoT devices. The study covers various IoT services such as data sharing, attack detection, and localization, and explores key applications in smart healthcare, transportation, UAVs, smart cities, and industries. It highlights current challenges and suggests directions for future research in the FL-IoT domain.

Zhang et al. (2021a) provide a comprehensive survey of FL, detailing its mechanisms of local computing and model transmission to ensure data privacy. The study reviews FL's application in data partitioning, privacy mechanisms, machine learning models, communication architecture, and system heterogeneity. It discusses current challenges, future research directions, and practical applications, emphasizing FL's potential in privacy-preserving AI and cross-platform data security.

Rieke et al. (2020) explore FL as a solution to privacy challenges in digital health, enabling collaborative ML model training without data centralization. The study highlights FL's potential to improve medical image analysis, precision medicine, and drug discovery while addressing privacy concerns. FL is poised to significantly enhance patient care and drive innovations in digital healthcare.

Gafni et al. (2022) explore FL from a signal processing perspective, highlighting its role in enabling model training across multiple edge devices without data exchange. The study examines the unique challenges and requirements of FL, drawing connections to classical signal processing and communication problems. It emphasizes the importance of dedicated schemes from these areas for the successful implementation of FL on mobile edge devices.

Zhang et al. (2022a) examine FL for the IoT, highlighting its potential to address the high communication, storage costs, and privacy concerns of centralized data processing. The study discusses opportunities and challenges of implementing FL in IoT networks, identifying seven critical challenges and recent approaches to overcome them. It emphasizes FL's role in enabling diverse IoT applications amid the rapid growth of IoT devices.

Kairouz et al. (2021) provide an extensive overview of FL, where multiple clients collaboratively train a model while keeping data decentralized, mitigating privacy risks associated with centralized ML. The study discusses recent advances in FL and identifies numerous open problems, highlighting interdisciplinary challenges and opportunities that span fields like distributed optimization, cryptography, and differential privacy. This comprehensive work serves as a crucial resource for understanding FL's current state and future research directions.

Khan et al. (2021) explore the application of FL in IoT, emphasizing its potential to enable on-device machine learning while maintaining data privacy. They present recent advances, a detailed taxonomy, and key evaluation metrics for FL in IoT networks. The study also identifies critical challenges, such as privacy concerns and high communication resource consumption, offering solutions and discussing the future role of FL in the context of 5G and 6G networks.

Pfitzner et al. (2021) provide a systematic review of FL in the medical field, emphasizing its role in preserving data privacy while enabling valuable research. They discuss how FL, which involves sharing machine learning models rather than raw data, addresses privacy concerns and allows for the integration of data from multiple sites without compromising patient anonymity. The review highlights the growing research and applications of FL in handling confidential healthcare data and explores its potential for advancing medical research and personalized treatments.

Xu et al. (2021) review the potential of FL in healthcare informatics, focusing on its ability to integrate fragmented and sensitive health data from various sources while preserving privacy. They discuss how FL can address challenges such as data fragmentation and privacy concerns by training global models without transferring raw data. The survey highlights key issues including data quality, the incorporation of expert knowledge, incentive mechanisms for participation, personalization of models, and improving model precision. It provides insights into current solutions and identifies open questions for future research in applying FL to healthcare.

Wen et al. (2023) provide a comprehensive survey on FL, highlighting its effectiveness in secure, distributed machine learning while addressing data privacy. The study reviews key aspects of FL, including its foundational concepts, privacy and security mechanisms, communication overhead, and challenges related to heterogeneity. It summarizes current research achievements and practical applications, and explores future research directions, emphasizing the need for advancements in privacy protection, client cooperation, fairness, and personalization to enhance the deployment and efficiency of FL systems.

Aledhari et al. (2020) offer a detailed survey on FL, emphasizing its enabling technologies, protocols, and real-world applications. FL, which trains models across multiple devices without sharing raw data, enhances privacy and security. The paper provides a thorough review of technical aspects, including protocols and platforms, and examines various use cases and applications. It also identifies key challenges such as system heterogeneity and fault tolerance and discusses the need for ongoing research to address these issues and integrate FL effectively into diverse industries.

Lyu et al. (2020) survey the vulnerabilities in FL, focusing on threats to data privacy from both internal and external adversaries. The paper introduces FL and provides a taxonomy of threat models, emphasizing two major attack types: poisoning attacks and inference attacks. It reviews key techniques and assumptions underlying these attacks and discusses future research directions for enhancing privacy preservation in FL systems. This study highlights the critical need for designing more robust FL algorithms to address emerging privacy threats.

Banabilah et al. (2022) provide a comprehensive review of FL, covering its fundamentals, enabling technologies, and future applications. The study highlights the current use of FL in applications such as predictive text and virtual assistants and examines its integration with technologies like AI, IoT, blockchain, and autonomous vehicles. The paper also addresses challenges and open questions in FL, including privacy, fairness, and personalization, while exploring its potential in various sectors such as healthcare, education, and industry. This survey serves as a broad reference for researchers and practitioners interested in the evolving applications and prospects of FL.

Niknam et al. (2020) discuss the relevance of FL in wireless communications, particularly within 5G networks. They highlight the limitations of traditional, centralized machine learning approaches due to data privacy concerns and high communication overhead. FL, which allows for decentralized data processing, addresses these issues by keeping data local. The paper provides an overview of FL, explores its potential applications in 5G networks, and outlines key challenges and open research areas for advancing FL in wireless communications.

Liu et al. (2024) provide a comprehensive review of Vertical Federated Learning (VFL), where different parties with unique feature sets about the same users collaboratively train machine learning models while preserving data privacy. The study covers VFL concepts, algorithms, and advancements, and addresses challenges related to effectiveness, efficiency, and privacy. They introduce VFLow, a unified framework that integrates communication, computation, privacy, and fairness constraints, and discuss recent industrial applications and future research directions in VFL.

Tan et al. (2022) explore Personalized Federated Learning (PFL), focusing on addressing the challenges of heterogeneous data in real-world datasets. The survey provides a comprehensive overview of PFL, including motivations, key techniques, and challenges. It introduces a taxonomy of PFL methods based on challenges and personalization strategies and discusses future research directions, including new architectural designs and realistic benchmarking. The study aims to guide researchers and practitioners in advancing PFL and addressing open problems in the field.

Bharati et al. (2022) provide a detailed review of FL with a focus on healthcare applications. The paper examines FL frameworks, architectures, and various privacy methods, such as secure multiparty computation and homomorphic encryption. It also covers challenges like privacy protection, communication costs, and system heterogeneity, and discusses different types of FL, including horizontal, vertical, and federated transfer learning. The study highlights recent advancements and outlines future research directions to address unresolved issues in the field.

Li et al. (2020b) discuss FL as a privacy-preserving machine learning technique that allows decentralized data training while keeping personal data on local devices. The paper highlights significant challenges in FL, including attacks on privacy and solutions such as hiding updates, optimizing communication, and defending against inference and model poisoning attacks. It also explores potential future applications in areas like IoT, genomics, smart cities, and finance, providing a comprehensive overview of ongoing research and promising directions for enhancing data privacy in FL.

Ma et al. (2022) provide a comprehensive review of the challenges and solutions related to non-IID (non-independent and identically distributed) data in FL. They analyze how non-IID data affects the performance of FL models and user participation and categorize existing methods for addressing these issues into data-side and model-side approaches. The paper highlights the need for advancements in FL algorithms, models, and frameworks to handle non-IID data effectively. It serves as an up-to-date, authoritative review aimed at guiding future research and implementation in this area.

Zhan et al. (2021) provide a survey on incentive mechanisms for FL, emphasizing the need to motivate clients to participate in the training process. They identify challenges unique to FL, such as evaluating the value of client data and modeling performance across different algorithms. The paper presents a taxonomy of existing incentive mechanisms and discusses various approaches in depth. It also highlights future research directions for developing effective incentives to ensure active and reliable client participation in FL.

Antunes et al. (2022) provide a comprehensive review of FL applications in healthcare, focusing on electronic health records (EHR). The study highlights FL's potential to train machine learning models while preserving data privacy, aligning with healthcare regulations. The paper reviews current research, case studies, and proposed architectures, emphasizing the importance of privacy and confidentiality. It identifies ongoing challenges such as improving aggregation mechanisms and dataset normalization. The review also notes that while there are successful applications, further research is needed to address privacy concerns and enhance FL's effectiveness in healthcare settings.

Liu et al. (2022) provide a comprehensive survey on FL, outlining its evolution from distributed machine learning. The paper describes a functional architecture for FL systems, including parallelism types, aggregation algorithms, and data manipulation techniques. It discusses the operational aspects of FL, such as diverse parallelism strategies, communication methods, and security measures. The review also highlights four prominent FL systems and identifies key research directions, including benchmarks, interpretability, decentralized aggregation, and applications to distributed intelligent systems.

Mothukuri et al. (2021) survey the security and privacy aspects of FL, emphasizing its potential in handling sensitive data through decentralized AI. The paper highlights current security threats, including communication bottlenecks, poisoning, and backdoor attacks, as well as privacy concerns like inference-based attacks. It provides a detailed review of these issues and suggests future research directions to enhance the security and adaptability of FL for broader adoption.

AbdulRahman et al. (2020) provide a comprehensive survey on FL, highlighting its evolution from centralized to decentralized on-device machine learning. FL preserves privacy by training models locally on user devices and aggregating the learned models centrally, minimizing data communication. The paper introduces a new classification of FL research topics and challenges, covering system models, applications, privacy, and resource management. It also outlines future research directions to enhance FL systems, reflecting the paradigm's potential to address privacy concerns and computational advancements in various domains.

Nguyen et al. (2022) present a comprehensive survey on FL for smart healthcare, emphasizing its role in addressing data privacy and scalability issues inherent in centralized AI systems. The paper reviews recent FL advancements and designs for healthcare, such as resource-aware, secure, and personalized FL. It explores FL applications in health data management, remote monitoring, medical imaging, and COVID-19 detection, and highlights key lessons from real-world projects. The survey concludes with an outline of current challenges and future research directions, underscoring FL's potential to enhance privacy and scalability in smart healthcare systems.

Bouacida and Mohapatra (2021) survey the security vulnerabilities inherent in FL, a decentralized approach that maintains user privacy by avoiding data sharing. The study identifies unique security threats faced by FL systems, including potential attacks that could compromise model integrity. It reviews existing defenses and highlights the need for further research to address these challenges, emphasizing that without resolving these vulnerabilities, the widespread adoption of FL could be impeded. The paper aims to bring attention to these issues and outline promising research directions for enhancing the security of FL environments.

Li et al. (2021c) review FL systems, highlighting their importance in enabling collaborative machine learning while preserving data privacy. The paper categorizes FLSs based on data distribution, model type, privacy mechanisms, communication architecture, federation scale, and motivation, providing insights into system design and effectiveness. It emphasizes the need for continued research and development in FLSs, identifying key areas for improvement such as benchmarking and integration with emerging technologies like blockchain. The study aims to guide future research and development in FL systems by summarizing current efforts and highlighting opportunities for advancement.

Xia et al. (2021) survey FL in edge computing, highlighting its potential for benefiting local data and computation power while enhancing privacy. The study examines key challenges and solutions in edge FL, including applications, development tools, communication efficiency, and security. It underscores the nascent stage of this field and emphasizes the need for further research to address open problems and improve collaborative training methods. The paper provides insights into future directions for advancing edge FL.

Long et al. (2020) explore the application of FL in open banking, a sector poised for decentralized data ownership and new financial services. FL's decentralized approach protects user privacy by avoiding the collection of sensitive data. The study discusses the potential challenges and solutions for implementing FL in open banking ecosystems, highlighting its ability to support privacy-preserving intelligent model training in a distributed framework.

Posner et al. (2021) explore Federated Vehicular Networks (FVN) as a novel approach to using FL within vehicular environments. They describe FVN as a robust distributed network that combines DSRC and mmWave communication technologies for scalable and stable performance. The study discusses the architecture of FVN, including its Federated Vehicular Cloud (FVC) component for data and model sharing, and highlights the potential of FVN to support data- and computation-intensive applications. The authors identify key challenges and future research directions, such as optimizing routing algorithms, adapting machine learning techniques for FVN, and addressing data transmission issues.

Ghimire and Rawat (2022) review recent advances in applying FL to enhance cybersecurity for the IoT. They highlight federated cybersecurity (FC) as a revolutionary concept for detecting and mitigating security threats in IoT networks while maintaining data privacy. The survey compares centralized, distributed, and FL approaches and explores the performance and security challenges of FL in IoT. The authors also discuss ongoing research efforts, emerging trends, and potential solutions to improve the security and efficiency of IoT environments using FL.

Zhang et al. (2022b) survey the challenges and future directions of secure FL, an algorithm that protects user privacy by aggregating model parameters instead of collecting raw data. They highlight the vulnerabilities due to its distributed nature, such as potential attacks from malicious data uploads and information recovery from parameters. The survey reviews state-of-the-art techniques, discusses improvements, and identifies open issues and solutions, pointing out future research directions to enhance FL security.

4. Resource-constrained FL Studies

Resource-constrained FL focuses on optimizing the performance of FL systems in environments with limited computational, communication, and energy resources. This section delves into the existing studies that address these critical challenges by proposing innovative solutions to resource-constrained settings. We explored various strategies such as adaptive device selection, efficient communication protocols, and energy-aware training algorithms. These studies are pivotal in extending the applicability of FL to edge networks, IoT devices, and other resource-limited platforms, ensuring efficient and sustainable model training without compromising performance. Table 3 presents the existing studies on this domain.

Saha et al. (2020) propose FogFL, a fog-enabled FL framework designed to address communication overheads and high computational requirements in resource-constrained IoT environments. By introducing geospatially placed fog nodes as local aggregators, FogFL reduces reliance on a centralized server, minimizing communication latency and energy consumption by 85% and 92%, respectively. The framework enhances system reliability and efficiency, demonstrating significant improvements over traditional FL methods.

Wang et al. (2019) present an adaptive FL approach for resource-constrained edge computing systems, addressing bandwidth, storage, and privacy concerns by avoiding the transfer of raw data to a central location. They propose a control algorithm that optimizes the trade-off between local updates and global parameter aggregation to minimize the loss function under resource constraints. Extensive experiments with real datasets demonstrate the effectiveness of their algorithm, achieving near-optimal performance across various machine learning models and data distributions.

Imteaj and Amini (2021b) introduce FedPARL, a lightweight FL model for resource-constrained and heterogeneous IoT environments. FedPARL incorporates sample-based pruning to reduce model size, assesses clients' resource availability, and adjusts task assignments accordingly. It also monitors client activity and updates trust scores to mitigate the impact of misbehaved clients. This tri-layer approach enhances the robustness and convergence of FL by adapting the Federated Averaging algorithm to accommodate variable workloads and resource constraints, demonstrating improved performance in diverse IoT settings.

Li et al. (2022) propose a fairness-aware FL method to improve performance and fairness in resource-constrained IoT networks with unreliable links. They introduce a statistically reweighted aggregation (SRA) scheme and its extension, reliable SRA (RSRA), which accounts for node fairness and unreliable parameter transmissions.

The RSRA scheme is analytically proven to enhance stability and achieve unbiased aggregation. An adaptive local training scheme, guided by the derived convergence bound, further optimizes performance. Extensive experiments validate that the proposed FL approach outperforms existing methods in both global performance and node fairness.

Nguyen et al. (2020b) propose an efficient FL algorithm for resource allocation in wireless IoT networks, addressing challenges such as non-i.i.d. data and user equipment (UE) heterogeneity. They extend FedAvg with a weight-based proximal term, allowing a small number of UEs to participate per round while ensuring convergence. The proposed algorithm minimizes energy consumption or completion time using a path-following approach. Empirical results demonstrate the algorithm's robustness, efficiency in training time, and energy consumption compared to traditional FL methods, advocating its use in bandwidth-constrained IoT networks.

Table 2: Survey publications on FL.

Focus Area	Studies	Description
General Purpose	Li, Fan, Tse, and Lin (2020a) Mammen (2021) Gafni et al. (2022) Kairouz et al. (2021) Khan et al. (2021) Xu et al. (2021) Aledhari et al. (2020) Niknam et al. (2020) AbdulRahman et al. (2020) Bouacida and Mohapatra (2021) Xia et al. (2021) Ghimire and Rawat (2022)	Broad reviews covering various FL techniques, applications, and advancements.
IoT	Nguyen et al. (2021) Banabilah et al. (2022) Nguyen et al. (2022)	Focused on FL applications and challenges in the IoT environments.
Industrial Engineering	Zhang et al. (2021a) Zhang et al. (2022a) Liu et al. (2024)	Examines FL techniques and their applications within industrial contexts.
Health	Rieke et al. (2020)	Survey on the use of FL techniques in healthcare and medical applications.
Data Privacy	Pfitzner et al. (2021) Zhang et al. (2022b)	Focus on methods and strategies for ensuring privacy in FL systems.
Signal Processing	Lyu et al. (2020) Liu et al. (2022)	Reviews FL applications and challenges in signal processing contexts.
Wireless Communication	Mothukuri et al. (2021)	Surveys FL methods specific to wireless communication systems.
Personalized FL	Wen et al. (2023)	Focus on FL approaches to individual user needs and personalization.
VFL (Vertical Federated Learning)	Ma et al. (2022)	Surveys vertical FL techniques and challenges.
Non-IID Data	Zhan et al. (2021)	Focus on FL techniques addressing non-IID data scenarios.
Decentralized System	Antunes et al. (2022)	Overview of FL approaches within decentralized system environments.
Security	Li et al. (2021c)	Focus on security aspects and methods within FL frameworks.
Finance	Tan et al. (2022)	Examines FL applications and challenges in financial contexts.
Distributed System	Posner et al. (2021)	Reviews FL in the context of distributed systems.

Nguyen et al. (2020b) propose an efficient FL algorithm for resource allocation in wireless IoT networks, addressing challenges such as non-i.i.d. data and user equipment (UE) heterogeneity. They extend FedAvg with a weight-based proximal term, allowing a small number of UEs to participate per round while ensuring convergence. The proposed algorithm minimizes energy consumption or completion time using a path-following approach. Empirical results demonstrate the algorithm's robustness, efficiency in training time, and energy consumption compared to traditional FL methods, advocating its use in bandwidth-constrained IoT networks.

Chen et al. (2022) propose a Communication-Efficient Federated Learning (CEFL) framework for wireless IoT networks, addressing the challenges of lossy communication channels and limited resources. They divide the optimization problem into client scheduling and resource allocation subproblems, introducing a new scheduling policy that reuses stale local model parameters. Using a Lagrange multiplier method and linear-search-based allocation, the CEFL framework maximizes communication efficiency and resource utilization. Extensive experiments show that CEFL significantly improves communication efficiency and learning performance compared to basic FL approaches with uniform resource allocation.

Salehi and Hossain (2021) present a FL algorithm for unreliable and resource-constrained cellular wireless networks. They address issues like communication failures and limited resources by incorporating success probability into the averaging step. The paper provides theoretical proof of convergence and a sub-optimal scheduling policy to enhance convergence rates. Experiments on real and synthetic datasets confirm the algorithm's effectiveness, highlighting that traditional FL methods may not perform well without accounting for wireless network unreliability.

Imteaj et al. (2021c) introduce the FedResilience algorithm, designed to enhance the resilience of resource-constrained critical infrastructures such as energy and transportation systems. Their approach uses FL to enable distributed training without sharing local data, addressing issues of data privacy and computational limitations. The study demonstrates that selecting proficient agents and allowing partial task execution can mitigate the effects of resource constraints and stragglers, thereby improving model performance and convergence. Simulations show that FedResilience outperforms traditional methods like FedAvg in terms of learning efficiency and accuracy in heterogeneous environments.

Gupta et al. (2023) introduce FedCare, a FL framework designed to address the challenges of resource-constrained healthcare devices in the Internet of Medical Things (IoMT) systems. Unlike traditional methods, which discard resource-constrained stragglers, FedCare incorporates them through a split learning approach, enabling collaborative training with edge nodes. This method reduces training time and maintains high accuracy, achieving a global accuracy of 90.32% while accommodating heterogeneous devices. FedCare enhances the efficiency of video-based data processing and provides a robust solution for monitoring vital signs in social IoMT systems, with potential applications in disease prediction and personal health analysis.

Ma et al. (2021) address the synchronization barrier in FL for heterogeneous IoT devices by proposing an adaptive batch size algorithm. This approach adjusts batch sizes and learning rates to minimize waiting times and conserve battery life while maintaining model convergence. The study includes theoretical analysis and extensive experiments, showing that the adaptive method effectively reduces waiting time and improves efficiency. Future work will explore the impact of parameter estimation and retransmission on system performance.

Zhang et al. (2021b) propose a deep reinforcement learning approach for adaptive client selection in FL systems within a Mobile Edge Computing (MEC) framework. The method models client selection as a Markov Decision Process (MDP) and uses Double Deep Q-Learning (DDQN) to optimize client participation, reducing energy consumption by up to 50% and training delay by up to 20.70%. The approach improves scalability and performance compared to static algorithms and baseline methods, demonstrating the system's effectiveness across different tasks and non-IID settings. Future work will address the challenge of action space expansion for growing client numbers.

Du et al. (2023) introduce AdapCom-DFL, a decentralized FL algorithm designed for resource-constrained IoT networks. This method adaptively adjusts the compression ratio of transmitted data to manage communication latency and includes network topology pruning to eliminate inefficient links and a power allocation strategy to optimize communication performance. Extensive simulations show that AdapCom-DFL, combined with these strategies, significantly improves performance and reduces bandwidth usage compared to existing methods while maintaining accuracy and meeting resource constraints.

Ficco et al. (2024) present a technique that enhances TinyML for IoT devices by integrating FL with transfer learning (TL) to enable on-board training of machine learning models. This approach improves accuracy in classification and regression tasks compared to traditional FL and TensorFlow Lite methods. The study shows that FL with TL achieves high accuracy while maintaining low power consumption and efficient training on devices with limited resources. Additionally, the method demonstrates robustness in handling unbalanced datasets and continuous training, offering a viable solution for privacy-preserving, local ML model training on IoT devices.

Gao et al. (2021) introduce RaFed, a resource allocation scheme designed to reduce training latency in FL for Industrial Internet of Things (IIoT) systems. They address the challenge of balancing the number of active devices to minimize interference and optimize convergence time, which is crucial for timely processing in automated manufacturing. RaFed performs a heuristic algorithm to select the most effective devices, reducing latency by 29.9% compared to existing methods. This approach improves the tradeoff between device interference and training efficiency in dense IIoT environments.

Sun et al. (2020) present an adaptive FL framework integrated with digital twin (DT) technology for Industrial IoT. Their approach uses DTs to capture device characteristics and address estimation deviations through trusted-based aggregation. They apply a combination of Lyapunov dynamic deficit queues and deep reinforcement learning (DRL) to adjust aggregation frequency and enhance learning performance under resource constraints. The proposed asynchronous FL framework, which incorporates clustering to manage heterogeneity and eliminate stragglers, demonstrates improved accuracy, convergence, and energy efficiency compared to existing methods.

Salh et al. (2023) present an energy-efficient FL framework for edge intelligence in the IoT, targeting green IoT networks. The paper proposes an optimization approach integrating bandwidth allocation, CPU frequency, and transmission power to minimize energy consumption while meeting FL time requirements. They introduce the Alternative Direction Algorithm (ADA), which adjusts computation and transmission resources to reduce energy use with minimal impact on FL time. Simulations show that ADA outperforms other methods in energy efficiency, balancing learning accuracy and resource usage effectively. Future work will focus on enhancing packet scheduling using double deep Q-learning algorithms.

Kushwaha et al. (2023) propose an optimal device selection method for FL in resource-constrained edge networks. Their approach minimizes redundant data training and improves resource utilization without compromising performance. Using the EMNIST dataset, they demonstrate that their method reduces device usage by up to 50% and resource consumption by up to 30%, while achieving near-baseline accuracy of 99%. This method enhances efficiency in edge networks by selecting a subset of devices for training, addressing issues of device heterogeneity and data variability.

Gad et al. (2024) present a communication-efficient FL framework for resource-constrained UAV-IoT systems, especially in rural areas. The authors propose a hybrid solution combining self-organizing maps (SOM) for UAV path optimization and knowledge distillation (KD)-based FL to minimize communication overhead. By formulating the drone's path as a traveling salesman problem (TSP) and using SOM for efficient route planning, the UAV coordinates FL among geographically dispersed nodes. The proposed CFedAKD algorithm compresses knowledge into small-sized synthetic labels (SLs), reducing data transmission to just 3 KB per node, compared to 250 KB in standard FL methods. Experimental results show this approach achieves competitive performance with significantly reduced communication costs.

Zhu et al. (2024) introduce ESFL (Efficient Split Federated Learning), a novel approach to optimize resource allocation in FL for heterogeneous wireless devices. The model is split between the central server and edge devices, maximizing the server's computational power while balancing the workload on resource-constrained devices. The optimization problem, framed as a mixed-integer nonlinear program, is solved using an iterative method to address device heterogeneity and improve efficiency. Simulation results demonstrate that ESFL significantly outperforms standard FL, split learning, and splitfed learning in terms of training efficiency across different scenarios.

Siddique et al. (2024) present a FL integrated with IoT for forest fire detection and classification, promoting environmentally conscious solutions. Using strategically placed cameras in fire-prone areas, the system enables early detection of forest fires. The framework employs Federated Stochastic Gradient Descent (FedSGD) to aggregate local models into a global model in the cloud. Trained on fire and non-fire image datasets distributed across five nodes, the system achieved 99.27% accuracy, highlighting its efficiency in collaborative learning. The approach reduces server load while ensuring high precision (98.92%) and recall (98.67%), making it a robust solution for wildfire detection.

5. Survey Publications on Resource-constrained FL

With the increasing deployment of FL in resource-constrained environments, several surveys have been conducted to consolidate research efforts and provide a holistic understanding of the field. This section reviews these surveys, which focus specifically on the unique challenges and solutions associated with resource-constrained FL. By analyzing these surveys, we aim to identify prevailing research directions, technological innovations, and best practices for implementing FL in constrained settings. This synthesis not only highlights the progress made but also underscores the need for continued exploration to address the evolving demands of resource-efficient FL. Table 4 introduces the key points of these surveys.

Imteaj et al. (2021a) provide a comprehensive survey on FL for resource-constrained IoT devices. They examine how FL, which relies on decentralized model training while preserving data privacy, can be adapted for devices with limited computational resources, bandwidth, and storage. The study highlights the challenges and limitations of applying FL in such environments and outlines potential solutions and future research directions to improve FL efficiency and scalability in heterogeneous IoT settings.

Imteaj et al. (2022) explore the use of FL for resource-constrained IoT devices, addressing challenges such as communication overhead, processing delay, privacy leakage, and security issues inherent in traditional centralized ML approaches. They highlight real-life FL applications suitable for IoT environments and emphasize

Table 3: Existing studies on resource-constrained FL applications.

Study	Focus Area	Key Points
Saha et al. (2020)	Energy Efficiency, IoT	Proposes methods for energy-efficient FL in IoT environments with resource constraints.
Wang et al. (2019)	Communication Efficiency, IoT	Focuses on reducing communication overhead in FL, specifically for IoT networks.
Imteaj and Amini (2021b)	Resource Allocation, IoT	Discusses resource allocation strategies in FL for IoT networks, addressing computation and communication constraints.
Li et al. (2022)	Edge Computing, Computational Constraints	Explores FL in edge computing environments, focusing on computational limitations and efficient algorithms.
Nguyen et al. (2020b)	Bandwidth Optimization, IoT	Proposes approaches to optimize bandwidth usage in FL systems with limited resources.
Chen et al. (2022)	Energy Efficiency, Mobile Devices	Investigates energy-efficient FL techniques for resource-constrained mobile devices.
Salehi and Hossain (2021)	Latency and Communication	Addresses latency issues and communication efficiency in FL for resource-constrained environments.
Imteaj et al. (2021c)	Resource Management, IoT	Reviews resource management techniques in FL, focusing on IoT devices and constraints.
Gupta et al. (2023)	Computation and Communication	Proposes solutions for balancing computation and communication resources in FL systems.
Ma et al. (2021)	Low-Power Devices, Federated Learning	Examines FL strategies for low-power and resource-constrained devices.
Zhang et al. (2021b)	Efficiency and Scalability	Focuses on enhancing efficiency and scalability of FL in constrained environments.
Du et al. (2023)	Resource-Constrained Edge Computing	Investigates FL techniques for resource-constrained edge computing environments.
Ficco et al. (2024)	On-Board Training, IoT	Explores FL with on-board training for IoT devices, addressing computational constraints.
Gao et al. (2021)	Latency Reduction, IIoT	Proposes a resource allocation scheme to reduce training latency in industrial IoT systems.
Sun et al. (2020)	Digital Twin, Industrial IoT	Combines FL with digital twins to manage resource constraints in industrial IoT systems.
Salh et al. (2023)	Energy Efficiency, Green IoT	Focuses on energy-efficient FL in green IoT networks, optimizing CPU frequency and transmission power.
Kushwaha et al. (2023)	Device Selection, Edge Networks	Proposes optimal device selection methods to improve resource utilization and performance in FL for edge networks.

core challenges related to resource limitations like memory, bandwidth, and energy. The paper also discusses open issues and future directions for implementing FL in resource-constrained settings.

Pfeiffer et al. (2023) survey FL challenges and solutions for computationally constrained heterogeneous devices, crucial in environments with diverse smart devices. They highlight the limitations of traditional FL approaches due to varying computational capacities and offer a detailed review of recent methods addressing this issue, including adaptations in neural network architectures and system-level approaches like Federated Averaging and distillation. The survey categorizes constraints into hard and soft, examines the effectiveness of current solutions, and identifies key open problems such as comparability and trade-offs in heterogeneous settings.

Da Silva et al. (2023) provide a systematic review of FL in the context of IoT, focusing on resource optimization challenges. They explore the adaptation of FL for IoT edge devices, emphasizing the need to balance processing and communication resources. The review highlights various strategies for optimizing FL, including reinforcement learning, heuristics, and game theory, and examines metrics like energy consumption, cost, and delay. The study identifies key research directions and scenarios, such as edge computing and 6G networks, noting trends and solutions for improving FL performance in resource-constrained IoT environments.

Brecko et al. (2022) provide a comprehensive survey on the application of federated learning (FL) in edge computing, emphasizing the challenges and solutions for deploying FL on resource-constrained edge devices. The paper discusses how FL enables decentralized training across edge devices, allowing for scalable, privacy-preserving machine learning without transferring raw data to a central server. The paper reviews popular FL frameworks, focusing on their compatibility with heterogeneous devices that have limited computational resources, different operating systems, and varying levels of connectivity. We focus on resource-constrained embedded systems such as IoT sensor nodes used in environmental monitoring, where limited memory and computational power are critical concerns. While Brecko et al. briefly discuss resource constraints, we focus on optimizing FL models for energy efficiency, specifically for embedded systems with limited battery life. Our focus includes analyzing the trade-off between communication costs and local computation to prolong device operation.

Hazra et al. (2022) explore the role of federated learning (FL) in enabling intelligent services at the edge of next-generation networks. The paper highlights the use of FL in edge networks to train models locally on devices with heterogeneous capabilities, such as differences in processing power and energy constraints, while maintaining data privacy and minimizing training time. Challenges identified include heterogeneity in device capabilities, energy consumption, and computational limitations, all of which need to be addressed for successful FL deployment in next-gen IoT networks. While Hazra et al. examine FL for general edge networks, including intelligent transportation and healthcare systems, our research targets resource-constrained embedded systems specifically, such as IoT meteorological sensor nodes with limited memory and computational power. Hazra et al. discuss the importance of minimizing training time and energy usage on edge devices, but their focus is on high-level system architecture and intelligent services. In contrast, our paper dives deeper into optimizing FL models for energy efficiency in low-power, embedded systems, ensuring that local computation balances energy consumption with model accuracy.

Table 4: Survey publications on resource-constrained FL.

Study	**Focus Area**	**Key Points**
Imteaj et al. (2021a)	Resource-Constrained Federated Learning	Reviews various approaches and challenges in FL under resource constraints. Discusses techniques for managing limited computational and communication resources.
Imteaj et al. (2022)	Resource-Constrained Federated Learning	Expands on resource management strategies in FL, focusing on energy efficiency, communication overhead, and computation constraints. Highlights advancements and gaps.
Pfeiffer et al. (2023)	Resource-Constrained Federated Learning	Provides a comprehensive survey on FL with a focus on computationally constrained heterogeneous devices. Reviews solutions for optimizing resource usage and performance.
Silva et al. (2023)	Resource-Constrained Federated Learning	Examines FL techniques specifically for resource-limited environments, addressing energy efficiency, bandwidth usage, and computational limitations. Presents recent developments and future directions.

Abreha et al. (2022) provide a systematic survey on federated learning (FL) in the context of edge computing (EC). The paper aims to address the intersection between these two emerging technologies by systematically reviewing the literature and identifying key challenges and solutions for FL in EC environments. Abreha et al. provide a comprehensive survey of existing studies and challenges in FL for edge computing, offering a high-level overview of the field. Our research, on the other hand, is empirical and focuses on improving FL performance in specific environments where resource limitations are a major constraint.

6. Global Evaluation of Federated Learning and Resource-Constrained Federated Learning

Surveys in FL offer a consolidated view of advancements and challenges, with notable works by Li et al. (2020a), Mammen (2021), and Gafni et al. (2022) providing comprehensive overviews of FL methodologies, trends, and research gaps. These surveys highlight the need for innovation in areas such as communication protocols, privacy-preserving techniques, and model aggregation methods. The complexities of deploying FL in resource-constrained environments—characterized by limited computational, communication, and energy resources—are addressed in studies by Imteaj et al. (2021a, 2022), Pfeiffer et al. (2023), and da Silva et al. (2023). These works introduce adaptive device selection, efficient communication protocols, and energy-aware training algorithms, underscoring FL's potential for edge networks and IoT applications.

To tackle these specific constraints, specialized surveys like those by Nguyen et al. (2021), Zhang et al. (2022a), and Bouacida and Mohapatra (2021) provide insights into technological innovations in resource-constrained FL, such as adaptive communication, energy-efficient algorithms, and edge computing integration. Studies by Saha et al. (2020), Wang et al. (2019), and Sun et al. (2020) further explore optimization methods that enhance computational and energy efficiency without compromising performance.

A synthesis of 62 key publications underscores the progress in addressing core FL challenges, particularly data privacy, communication efficiency, and model accuracy. In resource-constrained environments, research focuses on strategies for optimizing computational, communication, and energy resources. However, achieving universally applicable solutions, robust security, and efficient resource management remains a priority for future work.

6.1 Federated Learning: Achievements and Challenges

FL has been successfully applied across various domains such as healthcare, finance, and IoT, demonstrating its versatility and effectiveness. Studies such as those by Nguyen et al. (2020a), Ziller et al. (2021), and Chen et al. (2020) highlight the application of FL in enhancing data privacy and reducing communication overhead while maintaining high model accuracy. Key challenges addressed in these studies include data heterogeneity, communication efficiency, and model accuracy, each of which is critical for the practical deployment of FL systems.

Despite these achievements, FL faces significant challenges. Device heterogeneity, non-IID data distributions, and the need for robust security measures are recurring themes in the literature. Researchers like Li et al. (2021a) and Wang et al. (2020) have proposed various methodologies to tackle these issues, yet there remains a gap in creating universally applicable solutions that can be seamlessly integrated across different applications and environments.

6.2 Impact of Stragglers on Federated Learning Performance

One significant challenge in FL, particularly in resource-constrained environments, is the issue of stragglers—devices that are slower or less capable of processing data than others. Stragglers can significantly delay the global model's convergence since traditional FL approaches like FedAvg rely on synchronous updates. In this case, all devices must finish their local training before the server aggregates the model updates, leading to bottlenecks. Addressing the straggler problem is critical to improving both the efficiency and the overall quality of FL, especially in scenarios with heterogeneous devices such as IoT networks or edge computing environments. Solutions like partial aggregation, selective update mechanisms, or asynchronous aggregation have been proposed to mitigate the impact of stragglers. However, these approaches come with trade-offs, including the potential degradation of the global model's quality if updates from slower devices are frequently excluded. Further research is needed to develop more robust strategies that can balance efficiency with model performance, particularly when considering non-IID data and limited computational resources.

6.3 Asynchronicity and Buffered Asynchronous Federated Learning

Asynchronous Federated Learning (AFL) has emerged as a promising solution to mitigate the delays caused by stragglers, allowing devices to upload updates at different times without waiting for all participants to finish their local training. While AFL improves training efficiency, it introduces new challenges, particularly in security and privacy. The recent introduction of Buffered Asynchronous Federated Learning (BAFL) has highlighted these issues further, where devices maintain a buffer of model updates to account for varying device speeds. This approach enhances the flexibility of the system but creates additional vulnerabilities, such as exposing partial updates to adversarial attacks or model poisoning. As the field progresses, these security and privacy challenges must be addressed through enhanced encryption protocols, differential privacy, or secure aggregation techniques. Moreover, managing the trade-offs between communication efficiency and model accuracy in asynchronous settings remains an open research area, warranting further exploration. Incorporating these elements into future FL frameworks will be essential to ensure both scalable and secure implementations of FL in real-world resource-constrained scenarios.

7. Conclusion and Discussion

The field of FL represents a transformative approach to machine learning, particularly in the context of privacy-preserving computation and decentralized data processing. FL enables the training of machine learning models across multiple decentralized devices while keeping data localized, thus enhancing data privacy and reducing the communication overhead associated with traditional centralized training methods. This approach is particularly relevant in domains such as healthcare, finance, and IoT, where data privacy and security are paramount. Additionally, resource-constrained FL focuses on optimizing the performance of FL in environments with limited computational, communication, and energy resources, such as edge networks and IoT devices.

Despite the significant advancements in FL, several gaps persist in the literature. Key challenges include handling data heterogeneity, improving communication efficiency, and ensuring robust security measures across diverse applications. In resource-constrained environments, the challenges are further compounded by limited computational power, bandwidth, and energy resources. The current literature lacks universally applicable solutions that can seamlessly integrate FL across various domains and environments while addressing these challenges. Furthermore, there is a need for comprehensive surveys that consolidate the latest advancements and provide a holistic view of the field, particularly in resource-constrained FL.

The aim of this survey is to provide a comprehensive overview of the current state of FL and resource-constrained FL, highlighting the key challenges, advancements, and research gaps. By synthesizing insights from 62 key publications, this survey seeks to offer a structured understanding of the progress made in FL and the specific innovations for resource-constrained environments. The survey categorizes the literature into four sections: existing studies on FL, existing surveys on FL, existing studies on resource-constrained FL, and existing surveys on resource-constrained FL. This structured approach aims to provide clarity and direction for future research in these areas.

To address the gaps identified in the literature, several research directions can be pursued:

Data Heterogeneity: Developing advanced algorithms that can effectively handle non-iid data distributions across devices. This includes methods for robust model aggregation and adaptive learning that can accommodate diverse data sources.

Communication Efficiency: Innovating new communication protocols that minimize the data transfer requirements while maintaining model accuracy. Techniques such as gradient compression, quantization, and efficient update mechanisms can be explored.

Security and Privacy: Enhancing security measures to protect against adversarial attacks and ensuring the privacy of model updates. This includes the development of secure multiparty computation, differential privacy, and blockchain-based FL frameworks.

Resource Optimization: In resource-constrained environments, focusing on optimizing computational and energy resources through adaptive device selection, efficient training algorithms, and energy-aware communication protocols.

Scalability: Designing scalable FL frameworks that can efficiently manage a large number of devices and vast amounts of data, ensuring robustness and reliability in real-world applications.

This survey provides a detailed and structured examination of the advancements and ongoing challenges in the field of FL and resource-constrained FL. By categorizing the literature into targeted sections, it offers a clear roadmap for researchers and practitioners to understand the current state of the art and identify areas for future exploration. The insights gained from this survey highlight the collaborative efforts within the research community to tackle the multifaceted challenges of FL and extend its applicability to diverse and resource-constrained environments.

While the field of FL has seen numerous surveys, including recent ones on FL in edge computing and resource-constrained environments, this survey stands out by focusing specifically on addressing the unique challenges of resource-constrained devices in a federated setting. Unlike previous works, which either focus on general FL frameworks or specific edge applications, our survey identifies critical gaps in the current literature, such as the under-explored impact of heterogeneous hardware and limited power constraints on federated learning model performance. We also provide an updated classification of FL-enabling technologies, going beyond the traditional tools to include more recent and advanced frameworks. In addition to PySyft, we highlight other key libraries such as TensorFlow Federated and Flower. TensorFlow Federated offers a flexible framework for building FL models using the well-established TensorFlow ecosystem, making it a popular choice for research and real-world applications. Flower, on the other hand, focuses on scalability and ease of use, providing a platform for orchestrating FL across diverse devices and environments. By including these libraries, we emphasize the growing diversity of tools available for implementing FL in resource-constrained environments, each with distinct features catering to different use cases such as model customization, communication optimization, and cross-device learning efficiency. This broader overview helps to illuminate the technological landscape and guide practitioners in selecting the most appropriate tools for their specific FL needs.

Additionally, our survey extends the discussion by proposing future research directions that have been overlooked in prior studies, such as the optimization of communication efficiency and robustness to non-IID data for low-resource IoT devices. This unique focus provides a much-needed lens for understanding how FL can be better adapted to constrained environments.

In conclusion, while significant progress has been made in FL, there remain substantial challenges that require ongoing innovation and research. By addressing the identified gaps and pursuing the outlined research directions, the potential of FL can be fully realized, driving advancements in privacy-preserving, efficient, and scalable machine learning across various domains and environments. This survey serves as a valuable resource for guiding future research efforts and contributing to the evolution of FL.

References

AbdulRahman, S., Tout, H., Ould-Slimane, H., Mourad, A., Talhi, C., & Guizani, M. (2020). A survey on federated learning: The journey from centralized to distributed on-site learning and beyond. IEEE Internet of Things Journal, 8(7), 5476-5497.

Abreha, H. G., Hayajneh, M., & Serhani, M. A. (2022). Federated learning in edge computing: a systematic survey. Sensors, 22(2), 450.

Aledhari, M., Razzak, R., Parizi, R. M., & Saeed, F. (2020). Federated learning: A survey on enabling technologies, protocols, and applications. IEEE Access, 8, 140699-140725.

Antunes, R. S., André da Costa, C., Küderle, A., Yari, I. A., & Eskofier, B. (2022). Federated learning for healthcare: Systematic review and architecture proposal. ACM Transactions on Intelligent Systems and Technology (TIST), 13(4), 1-23.

Banabilah, S., Aloqaily, M., Alsayed, E., Malik, N., & Jararweh, Y. (2022). Federated learning review: Fundamentals, enabling technologies, and future applications. Information processing & management, 59(6), 103061.

Bharati, S., Mondal, M., Podder, P., & Prasath, V. B. (2022). Federated learning: Applications, challenges and future directions. International Journal of Hybrid Intelligent Systems, 18(1-2), 19-35.

Bouacida, N., & Mohapatra, P. (2021). Vulnerabilities in federated learning. IEEE Access, 9, 63229-63249.

Brecko, A., Kajati, E., Koziorek, J., & Zolotova, I. (2022). Federated learning for edge computing: A survey. Applied Sciences, 12(18), 9124.

Chen, M., Poor, H. V., Saad, W., & Cui, S. (2020). Wireless communications for collaborative federated learning. IEEE Communications Magazine, 58(12), 48-54.

Chen, H., Huang, S., Zhang, D., Xiao, M., Skoglund, M., & Poor, H. V. (2022). Federated learning over wireless IoT networks with optimized communication and resources. IEEE Internet of Things Journal, 9(17), 16592-16605.

da Silva, L. G. F., Sadok, D. F., & Endo, P. T. (2023). Resource optimizing federated learning for use with IoT: A systematic review. Journal of Parallel and Distributed Computing, 175, 92-108.

Du, M., Zheng, H., Gao, M., & Feng, X. (2023). Adaptive Decentralized Federated Learning in Resource-Constrained IoT Networks. IEEE Internet of Things Journal.

Ficco, M., Guerriero, A., Milite, E., Palmieri, F., Pietrantuono, R., & Russo, S. (2024). Federated learning for IoT devices: Enhancing TinyML with on-board training. Information Fusion, 104, 102189.

Gafni, T., Shlezinger, N., Cohen, K., Eldar, Y. C., & Poor, H. V. (2022). Federated learning: A signal processing perspective. IEEE Signal Processing Magazine, 39(3), 14-41.

Gao, W., Zhao, Z., Min, G., Ni, Q., & Jiang, Y. (2021). Resource allocation for latency-aware federated learning in industrial internet of things. IEEE Transactions on Industrial Informatics, 17(12), 8505-8513.

Ghimire, B., & Rawat, D. B. (2022). Recent advances on federated learning for cybersecurity and cybersecurity for federated learning for internet of things. IEEE Internet of Things Journal, 9(11), 8229-8249.

Ghosh, A., Chung, J., Yin, D., & Ramchandran, K. (2020). An efficient framework for clustered federated learning. Advances in Neural Information Processing Systems, 33, 19586-19597.

Gupta, A., Misra, S., Pathak, N., & Das, D. (2023). Fedcare: Federated learning for resource-constrained healthcare devices in iomt system. IEEE Transactions on Computational Social Systems, 10(4), 1587-1596.

Hazra, A., Adhikari, M., Nandy, S., Doulani, K., & Menon, V. G. (2022). Federated-learning-aided next-generation edge networks for intelligent services. IEEE Network, 36(3), 56-64.

Imteaj, A., Thakker, U., Wang, S., Li, J., & Amini, M. H. (2021a). A survey on federated learning for resource-constrained IoT devices. IEEE Internet of Things Journal, 9(1), 1-24.

Imteaj, A., & Amini, M. H. (2021b). FedPARL: Client activity and resource-oriented lightweight federated learning model for resource-constrained heterogeneous IoT environment. Frontiers in Communications and Networks, 2, 657653.

Imteaj, A., Khan, I., Khazaei, J., & Amini, M. H. (2021c). Fedresilience: A federated learning application to improve resilience of resource-constrained critical infrastructures. Electronics, 10(16), 1917.

Imteaj, A., Mamun Ahmed, K., Thakker, U., Wang, S., Li, J., & Amini, M. H. (2022). Federated learning for resource-constrained iot devices: Panoramas and state of the art. Federated and Transfer Learning, 7-27.

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhao, S. (2021). Advances and open problems in federated learning. Foundations and trends® in machine learning, 14(1–2), 1-210.

Khan, Latif U., et al. "Federated learning for internet of things: Recent advances, taxonomy, and open challenges." IEEE Communications Surveys & Tutorials 23.3 (2021): 1759-1799.

Kushwaha, D., Redhu, S., Brinton, C. G., & Hegde, R. M. (2023). Optimal device selection in federated learning for resource-constrained edge networks. IEEE Internet of Things Journal, 10(12), 10845-10856.

Li, L., Fan, Y., Tse, M., & Lin, K. Y. (2020a). A review of applications in federated learning. Computers & Industrial Engineering, 149, 106854.

Li, Z., Sharma, V., & Mohanty, S. P. (2020b). Preserving data privacy via federated learning: Challenges and solutions. IEEE Consumer Electronics Magazine, 9(3), 8-16.

Li, T., Hu, S., Beirami, A., & Smith, V. (2021a). Ditto: Fair and robust federated learning through personalization. In International conference on machine learning (pp. 6357-6368). PMLR.

Li, Q., He, B., & Song, D. (2021b). Model-contrastive federated learning. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10713-10722).

Li, Q., Wen, Z., Wu, Z., Hu, S., Wang, N., Li, Y., ... & He, B. (2021c). A survey on federated learning systems: Vision, hype and reality for data privacy and protection. IEEE Transactions on Knowledge and Data Engineering, 35(4), 3347-3366.

Li, Z., Zhou, Y., Wu, D., Tang, T., & Wang, R. (2022). Fairness-aware federated learning with unreliable links in resource-constrained Internet of things. IEEE Internet of Things Journal, 9(18), 17359-17371.

Liu, J., Huang, J., Zhou, Y., Li, X., Ji, S., Xiong, H., & Dou, D. (2022). From distributed machine learning to federated learning: A survey. Knowledge and Information Systems, 64(4), 885-917.

Liu, Y., Kang, Y., Zou, T., Pu, Y., He, Y., Ye, X., ... & Yang, Q. (2024). Vertical federated learning: Concepts, advances, and challenges. IEEE Transactions on Knowledge and Data Engineering.

Long, G., Tan, Y., Jiang, J., & Zhang, C. (2020). Federated learning for open banking. In Federated learning: privacy and incentive (pp. 240-254). Cham: Springer International Publishing.

Lyu, L., Yu, H., & Yang, Q. (2020). Threats to federated learning: A survey. arXiv preprint arXiv:2003.02133.

Ma, Z., Xu, Y., Xu, H., Meng, Z., Huang, L., & Xue, Y. (2021). Adaptive batch size for federated learning in resource-constrained edge computing. IEEE Transactions on Mobile Computing, 22(1), 37-53.

Ma, X., Zhu, J., Lin, Z., Chen, S., & Qin, Y. (2022). A state-of-the-art survey on solving non-iid data in federated learning. Future Generation Computer Systems, 135, 244-258.

Mammen, P. M. (2021). Federated learning: Opportunities and challenges. arXiv preprint arXiv:2101.05428.

Mothukuri, V., Parizi, R. M., Pouriyeh, S., Huang, Y., Dehghantanha, A., & Srivastava, G. (2021). A survey on security and privacy of federated learning. Future Generation Computer Systems, 115, 619-640.

Nguyen, H. T., Sehwag, V., Hosseinalipour, S., Brinton, C. G., Chiang, M., & Poor, H. V. (2020a). Fast-convergent federated learning. IEEE Journal on Selected Areas in Communications, 39(1), 201-218.

Nguyen, V. D., Sharma, S. K., Vu, T. X., Chatzinotas, S., & Ottersten, B. (2020b). Efficient federated learning algorithm for resource allocation in wireless IoT networks. IEEE Internet of Things Journal, 8(5), 3394-3409.

Nguyen, D. C., Ding, M., Pathirana, P. N., Seneviratne, A., Li, J., & Poor, H. V. (2021). Federated learning for internet of things: A comprehensive survey. IEEE Communications Surveys & Tutorials, 23(3), 1622-1658.

Nguyen, D. C., Pham, Q. V., Pathirana, P. N., Ding, M., Seneviratne, A., Lin, Z., ... & Hwang, W. J. (2022). Federated learning for smart healthcare: A survey. ACM Computing Surveys (Csur), 55(3), 1-37.

Niknam, S., Dhillon, H. S., & Reed, J. H. (2020). Federated learning for wireless communications: Motivation, opportunities, and challenges. IEEE Communications Magazine, 58(6), 46-51.

Pfeiffer, K., Rapp, M., Khalili, R., & Henkel, J. (2023). Federated learning for computationally constrained heterogeneous devices: A survey. ACM Computing Surveys, 55(14s), 1-27.

Pfitzner, B., Steckhan, N., & Arnrich, B. (2021). Federated learning in a medical context: a systematic literature review. ACM Transactions on Internet Technology (TOIT), 21(2), 1-31.

Posner, J., Tseng, L., Aloqaily, M., & Jararweh, Y. (2021). Federated learning in vehicular networks: Opportunities and solutions. IEEE Network, 35(2), 152-159.

Saha, R., Misra, S., & Deb, P. K. (2020). FogFL: Fog-assisted federated learning for resource-constrained IoT devices. IEEE Internet of Things Journal, 8(10), 8456-8463.

Salehi, M., & Hossain, E. (2021). Federated learning in unreliable and resource-constrained cellular wireless networks. IEEE Transactions on Communications, 69(8), 5136-5151.

Salh, A., Ngah, R., Audah, L., Kim, K. S., Abdullah, Q., Al-Moliki, Y. M., ... & Talib, H. N. (2023). Energy-efficient federated learning with resource allocation for green IoT edge intelligence in B5G. IEEE Access, 11, 16353-16367.

Siddique, A. A., Alasbali, N., Driss, M., Boulila, W., Alshehri, M. S., & Ahmad, J. (2024). Sustainable collaboration: Federated learning for environmentally conscious forest fire classification in green internet of things (IoT). Internet of Things, 25, 101013.

Sun, W., Lei, S., Wang, L., Liu, Z., & Zhang, Y. (2020). Adaptive federated learning and digital twin for industrial internet of things. IEEE Transactions on Industrial Informatics, 17(8), 5605-5614.

Tan, A. Z., Yu, H., Cui, L., & Yang, Q. (2022). Towards personalized federated learning. IEEE transactions on neural networks and learning systems, 34(12), 9587-9603.

Wang, S., Tuor, T., Salonidis, T., Leung, K. K., Makaya, C., He, T., & Chan, K. (2019). Adaptive federated learning in resource constrained edge computing systems. IEEE journal on selected areas in communications, 37(6), 1205-1221.

Wang, H., Kaplan, Z., Niu, D., & Li, B. (2020, July). Optimizing federated learning on non-iid data with reinforcement learning. In IEEE INFOCOM 2020-IEEE conference on computer communications (pp. 1698-1707). IEEE.

Wen, J., Zhang, Z., Lan, Y., Cui, Z., Cai, J., & Zhang, W. (2023). A survey on federated learning: challenges and applications. International Journal of Machine Learning and Cybernetics, 14(2), 513-535.

Xia, Q., Ye, W., Tao, Z., Wu, J., & Li, Q. (2021). A survey of federated learning for edge computing: Research problems and solutions. High-Confidence Computing, 1(1), 100008.

Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., & Wang, F. (2021). Federated learning for healthcare informatics. Journal of healthcare informatics research, 5, 1-19.

Zhan, Y., Zhang, J., Hong, Z., Wu, L., Li, P., & Guo, S. (2021). A survey of incentive mechanism design for federated learning. IEEE Transactions on Emerging Topics in Computing, 10(2), 1035-1044.

Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. (2021a). A survey on federated learning. Knowledge-Based Systems, 216, 106775.

Zhang, H., Xie, Z., Zarei, R., Wu, T., & Chen, K. (2021b). Adaptive client selection in resource constrained federated learning systems: A deep reinforcement learning approach. IEEE Access, 9, 98423-98432.

Zhang, T., Gao, L., He, C., Zhang, M., Krishnamachari, B., & Avestimehr, A. S. (2022). Federated learning for the internet of things: Applications, challenges, and opportunities. IEEE Internet of Things Magazine, 5(1), 24-29.

Zhang, K., Song, X., Zhang, C., & Yu, S. (2022b). Challenges and future directions of secure federated learning: a survey. Frontiers of computer science, 16, 1-8.

Zhu, G., Deng, Y., Chen, X., Zhang, H., Fang, Y., & Wong, T. F. (2024). ESFL: Efficient Split Federated Learning over Resource-Constrained Heterogeneous Wireless Devices. IEEE Internet of Things Journal.

Ziller, A., Trask, A., Lopardo, A., Szymkow, B., Wagner, B., Bluemke, E., ... & Kaissis, G. (2021). Pysyft: A library for easy federated learning. Federated Learning Systems: Towards Next-Generation AI, 111-139.

Author Biographies

Eda Bahar earned her B.S. degree in Computer Engineering from Galatasaray University in 2021. She is currently pursuing a master's degree in Smart Systems Engineering at the same institution. Presently, she works as a Research Assistant in the Computer Engineering Department at Galatasaray University. Her research interests include TinyML, Embedded Systems, Embedded ML, and IoT.

Ozgun Pinarer received his B.S. and M.S. degrees in Computer Engineering from Galatasaray University in 2010 and 2012. In 2017 he received his Ph.D. degree in Computer Engineering from INSA Lyon University. Currently he works in the Computer Engineering Department of Galatasaray University as an assistant professor. His research interests are IoT, edge computing, pervasive environments, sensor based smart systems.