Article

Advancing Privacy-Preserving AI: A Survey on Federated Learning and Its Applications

Eustace Nowell and Sameera Gallus *

Posted Date: 9 January 2025

doi: 10.20944/preprints202501.0685.v1

Keywords: Federated Learning; Distributed Machine Learning; Privacy Preserving AI; Data Heterogeneity; Scalability; Communication Efficiency; Robust Aggregation; Personalization; Adversarial Robustness; Real-World Applications; Ethical AI; Decentralized Learning

 doi:10.20944/preprints202501.0685.v1

Article

Advancing Privacy-Preserving AI: A Survey on Federated Learning and its Applications

Eustace Nowell and Sameera Gallus *

Department of Computer Science, The University of Bath, United Kingdom; eustace.nowell@bath.ac.uk
* Correspondence: sameera.gallus@bath.ac.uk

Abstract: Federated Learning (FL) has emerged as a transformative approach to distributed machine learning, enabling the collaborative training of models across decentralized and private datasets. Unlike traditional centralized learning paradigms, FL ensures data privacy by keeping raw data localized on client devices while leveraging aggregated updates to build global models. This survey explores the critical aspects of efficient federated learning, including communication reduction, robustness to system and data heterogeneity, and scalability in real-world applications. We discuss key techniques such as model compression, asynchronous updates, personalized learning, and robust aggregation to address challenges posed by resource-constrained devices, non-IID data distributions, and adversarial environments. Applications of FL across diverse domains, including healthcare, finance, smart cities, and autonomous systems, highlight its potential to transform industries while preserving privacy and compliance with regulatory frameworks. The survey also identifies open challenges in scalability, privacy guarantees, fairness, and ethical considerations, providing future research directions to address these gaps. As FL continues to evolve, it holds the promise of enabling privacy-preserving, collaborative intelligence on a global scale, fostering innovation while addressing critical societal and technical challenges.

Keywords: federated learning; distributed machine learning; privacy preserving AI; data heterogeneity, scalability; communication efficiency; robust aggregation; personalization; adversarial robustness; real-world applications; ethical AI; decentralized learning

1. Introduction

The exponential growth of connected devices, fueled by advancements in Internet-of-Things (IoT), edge computing, and mobile technology, has led to a surge in data generation. While this data holds immense potential for driving machine learning (ML) innovations, the conventional approach of aggregating data on centralized servers raises significant concerns regarding privacy, security, and scalability. As a response to these challenges, federated learning (FL) has emerged as a promising decentralized machine learning paradigm, where multiple clients collaborate to train a shared global model while keeping their data localized on their devices. By ensuring that raw data never leaves the client's environment, FL inherently addresses privacy concerns and adheres to increasingly stringent data protection regulations, such as the General Data Protection Regulation (GDPR) [1]. Federated learning's decentralized nature offers several advantages over traditional centralized learning. These include enhanced data privacy, reduced latency for model updates, and the ability to leverage distributed computational resources. However, this paradigm shift also introduces unique challenges that must be addressed to unlock its full potential. Three primary challenges in FL include communication overhead, system heterogeneity, and data heterogeneity [2]. Communication efficiency is a critical bottleneck in FL systems, as frequent exchanges of model updates between the server and clients can strain bandwidth and delay convergence. This is particularly pronounced when dealing with large-scale models or operating in environments with limited or unreliable network connectivity. To mitigate this, researchers have explored a variety of techniques, including gradient compression,

 doi:10.20944/preprints202501.0685.v1

quantization, and sparsification, to reduce the amount of data transmitted during training. System heterogeneity further complicates federated learning [3]. The participating clients in an FL system often exhibit significant variability in their computational capabilities, network conditions, and availability [4]. For instance, some devices may be equipped with powerful hardware and stable network connections, while others may be constrained by limited resources and intermittent connectivity. These disparities can lead to inefficiencies in resource utilization, stragglers delaying model updates, and reduced system performance [5]. Addressing these issues requires designing adaptive algorithms that balance workload distribution and accommodate diverse client environments. Data heterogeneity, also referred to as non-IID (non-independent and identically distributed) data, is another critical challenge in federated learning [6]. Unlike traditional ML, where training data is often assumed to follow a unified distribution, FL must contend with data generated independently by individual clients, often reflecting their unique usage patterns and preferences [7]. This lack of uniformity can lead to biased local models that, when aggregated, may fail to generalize effectively [8]. Techniques such as personalization, model aggregation strategies, and client clustering have been proposed to tackle these issues. In recent years, the research community has made significant strides in addressing these challenges, proposing innovative solutions to enhance the efficiency and scalability of federated learning systems [9]. These advancements span a wide range of areas, including communication-efficient algorithms, robust aggregation techniques, adaptive scheduling mechanisms, and novel optimization frameworks [10]. Furthermore, researchers have explored practical deployment strategies for FL in resource-constrained environments, such as edge devices and IoT networks [11]. This survey aims to provide a comprehensive overview of the state-of-the-art techniques and strategies for improving the efficiency of federated learning [12]. Specifically, it categorizes existing approaches into three broad domains: communication efficiency, system optimization, and heterogeneity management. For each domain, we review key methodologies, highlight recent innovations, and discuss their strengths and limitations. Additionally, we explore the interplay between these domains and their impact on the overall performance of FL systems. The remainder of this paper is organized as follows [13]. Section 2 introduces the fundamentals of federated learning, its core principles, and the associated challenges [14]. Section 3 delves into techniques aimed at improving communication efficiency, including model compression, update sparsification, and asynchronous communication [15]. Section 4 focuses on strategies for managing system and data heterogeneity, such as adaptive scheduling, personalized models, and robust aggregation methods [16]. Section 5 highlights real-world applications of federated learning across various domains, including healthcare, finance, and smart cities [17]. Section **??** outlines open research problems and promising directions for future exploration [18]. Finally, Section 7 summarizes the key takeaways and implications of this survey.

2. Background and Fundamentals of Federated Learning

Federated learning (FL) represents a paradigm shift in machine learning, offering a decentralized framework for collaboratively training models across distributed data sources [19]. Unlike traditional centralized learning approaches, where data is aggregated on a single server for processing, FL emphasizes data privacy and minimizes risks associated with data breaches and unauthorized access [20]. In this section, we provide a detailed overview of FL, its core principles, and the fundamental challenges it faces.

2.1. Federated Learning Overview

Federated learning operates on a client-server architecture, where multiple clients, such as mobile devices, IoT nodes, or edge devices, contribute to the training of a global model under the orchestration of a central server [21]. The training process typically involves the following steps:

1. **Model Initialization:** The central server initializes a global model and shares it with participating clients.

 doi:10.20944/preprints202501.0685.v1

2. **Local Training:** Each client trains the model locally on its private data, generating updated model parameters [22].
3. **Model Aggregation:** Clients send their locally trained model updates to the server, which aggregates them to update the global model.
4. **Model Dissemination:** The updated global model is redistributed to clients for the next training round.

This iterative process continues until the global model converges to a satisfactory level of performance [23].

2.2. Challenges in Federated Learning

While FL offers distinct advantages, it also introduces challenges stemming from its decentralized and distributed nature [24]. The key challenges are as follows:

2.2.1. Communication Overhead

The frequent exchange of model updates between clients and the central server imposes significant communication costs [25]. This overhead is exacerbated in bandwidth-constrained environments or when the global model size is large [26]. Moreover, synchronous communication can lead to delays caused by stragglers, i.e., slow or unavailable clients [27].

2.2.2. System Heterogeneity

Clients in FL systems are often diverse in terms of hardware capabilities, network conditions, and power availability [28]. These disparities result in non-uniform resource utilization and may hinder the overall performance of the system. Designing mechanisms to accommodate such heterogeneity is critical for scalable FL deployments.

2.2.3. Data Heterogeneity

In most FL applications, client data is non-IID and unbalanced [29]. This data heterogeneity arises due to differences in user behavior, preferences, and geographical factors. Aggregating models trained on such data can lead to suboptimal global models with reduced generalizability [30].

2.2.4. Privacy and Security

Although FL reduces the risks associated with sharing raw data, it is not immune to privacy and security threats [31]. Adversaries can infer sensitive information from shared model updates (e.g., via model inversion attacks) or disrupt the training process through malicious behavior [32].

2.3. Types of Federated Learning

Federated learning can be broadly categorized based on the distribution of data and clients:

- **Horizontal Federated Learning (HFL):** Clients have data with similar feature spaces but different samples [33]. HFL is common in scenarios where clients operate in the same domain, such as hospitals sharing patient records.
- **Vertical Federated Learning (VFL):** Clients have data with different feature spaces but overlapping samples. VFL is relevant when organizations collaborate on common users with complementary datasets [34].
- **Federated Transfer Learning (FTL):** Combines the principles of FL and transfer learning to enable collaboration between clients with limited overlap in both features and samples [35].

2.4. Metrics for Evaluating Federated Learning

The success of FL systems is evaluated using several metrics:

- **Model Performance:** Accuracy, precision, recall, or other metrics used to assess the quality of the global model.

 doi:10.20944/preprints202501.0685.v1

- **Communication Efficiency:** Measured in terms of the number of communication rounds or the amount of data exchanged [36].
- **System Efficiency:** Computational cost, resource utilization, and scalability across heterogeneous clients [37].
- **Fairness:** The extent to which the global model benefits all clients equitably, particularly in the presence of data heterogeneity.
- **Privacy and Security:** Robustness against attacks and the level of privacy guarantees provided.

This background establishes the foundational concepts and challenges of federated learning, setting the stage for a deeper exploration of techniques to enhance its efficiency in the subsequent sections [38].

3. Techniques for Improving Communication Efficiency

Communication efficiency is one of the most critical aspects of federated learning (FL) due to the iterative exchange of model updates between the clients and the central server. Inefficient communication can significantly slow down convergence, particularly in bandwidth-constrained environments or when dealing with large-scale models [39]. This section delves into the strategies and techniques proposed to reduce communication overhead while maintaining or improving the performance of FL systems.

3.1. Model Compression Techniques

Model compression techniques aim to reduce the size of updates transmitted during communication. This reduction can be achieved through various approaches:

- **Gradient Quantization:** Gradients are quantized to a lower number of bits, reducing the size of the transmitted data [40]. For instance, techniques like QSGD (Quantized Stochastic Gradient Descent) apply fixed-point representations to gradients.
- **Gradient Sparsification:** Instead of transmitting all gradient updates, only a subset of significant updates is communicated [41]. The top-k sparsification method, for example, transmits only the k largest gradients, while the others are approximated or stored locally for future rounds [42].
- **Model Pruning:** Unimportant parameters of the model are removed, resulting in a smaller model that requires less communication overhead [43]. This technique can be applied both statically (before training) and dynamically (during training) [44,45].

3.2. Federated Averaging and Update Frequency

Federated Averaging (FedAvg) is one of the most widely used algorithms for FL, primarily because of its communication efficiency [46]. Instead of transmitting model updates after every batch, FedAvg performs multiple local updates before transmitting the averaged model updates to the server. By increasing the local update frequency, the number of communication rounds required for convergence can be reduced, although this introduces a trade-off between local computation and global synchronization [47].

3.3. Asynchronous Communication

Traditional FL frameworks often rely on synchronous communication, where all clients must complete their updates before the global model is aggregated [48]. However, this approach can be inefficient due to stragglers. Asynchronous communication allows clients to send updates independently, enabling faster aggregation at the server. Techniques such as asynchronous stochastic gradient descent (ASGD) and server buffering have been developed to support asynchronous operations while minimizing potential staleness in model updates.

 doi:10.20944/preprints202501.0685.v1

3.4. Communication-Efficient Aggregation

Aggregation strategies play a crucial role in reducing communication overhead [49]. Some of the prominent techniques include:

- **Weighted Aggregation:** Aggregating updates based on client contributions, such as the size of local datasets or the quality of local models, ensures efficient use of communication resources.
- **Hierarchical Aggregation:** Clients are organized into clusters, and intermediate aggregations are performed at the cluster level before updates are sent to the central server [50]. This reduces the number of direct client-to-server communications [51].
- **Error Feedback Mechanisms:** Techniques like federated dropout and residual feedback allow clients to focus on transmitting updates that contribute most to model improvement, while others are approximated or compressed.

3.5. Advanced Techniques and Hybrid Methods

Recent advancements have combined multiple techniques to achieve even greater communication efficiency. For example:

- **Split Learning with FL:** Splitting the model across clients and the server reduces the size of updates, as only specific layers or features are communicated [52].
- **Distillation-Based FL:** Knowledge distillation is used to share distilled model representations instead of full model updates, significantly reducing communication costs.
- **Adaptive Communication Schedules:** Dynamically adjusting the frequency of communication based on training progress, model convergence, or client conditions reduces redundant communications.

3.6. Challenges and Trade-offs

While communication-efficient techniques enhance the scalability of FL, they introduce trade-offs that must be carefully managed:

- **Accuracy vs. Communication Trade-off:** Over-aggressive compression or sparsification may degrade model performance, requiring careful tuning of compression parameters.
- **Computation vs. Communication Trade-off:** Increasing local computations to reduce communication rounds may overwhelm resource-constrained devices.
- **Staleness and Convergence:** Asynchronous and sparse updates may lead to stale gradients or slower convergence, requiring robust aggregation mechanisms to mitigate these effects.

This section has outlined key strategies for improving communication efficiency in federated learning [53]. These techniques, while promising, often require fine-tuning and a careful balance of trade-offs to achieve optimal results [54]. In the next section, we explore methods for managing system and data heterogeneity, which are equally crucial for the scalability and effectiveness of FL.

4. Managing System and Data Heterogeneity

Federated learning (FL) systems must operate effectively in environments characterized by significant heterogeneity [55]. System heterogeneity arises from variations in device capabilities, network conditions, and resource availability among clients. Data heterogeneity, on the other hand, stems from differences in data distributions across clients, often reflecting non-IID (non-independent and identically distributed) patterns. Both forms of heterogeneity pose significant challenges to the design, performance, and fairness of FL systems. This section discusses key approaches to address these challenges and ensure robust and scalable federated learning.

4.1. Techniques for Managing System Heterogeneity

System heterogeneity in FL arises from the diverse nature of participating devices, which may range from high-performance servers to resource-constrained IoT devices. The following techniques address this challenge:

 doi:10.20944/preprints202501.0685.v1

4.1.1. Adaptive Client Participation

In many FL scenarios, not all clients need to participate in every communication round. Adaptive participation mechanisms select a subset of clients based on criteria such as resource availability, device performance, or data quality [56]. This reduces the impact of stragglers and ensures efficient utilization of system resources [57].

4.1.2. Dynamic Model Architectures

To accommodate varying computational capacities, FL systems can use dynamic or lightweight model architectures. Techniques like model partitioning or distillation allow clients with limited resources to work on smaller sub-models while contributing to the global training process.

4.1.3. Federated Dropout

Federated dropout is a technique where clients only train and update a subset of the model parameters. This reduces computational and communication costs while allowing resource-constrained devices to participate in training.

4.1.4. Robust Aggregation in Heterogeneous Systems

Robust aggregation algorithms, such as adaptive weighting schemes, ensure that contributions from slower or less capable clients are appropriately balanced. These algorithms can assign weights based on the quality or relevance of the client updates, mitigating the effects of system disparities.

4.2. Techniques for Addressing Data Heterogeneity

Data heterogeneity, where clients generate data distributions that are non-IID or imbalanced, is one of the most critical challenges in FL. The following techniques are commonly employed to address this issue:

4.2.1. Personalized Federated Learning

Personalized FL aims to tailor global models to individual clients by incorporating client-specific adaptations. Approaches include:

- Fine-tuning the global model on local data.
- Training personalized layers while keeping shared layers common across clients [58].
- Using meta-learning techniques, such as Model-Agnostic Meta-Learning (MAML), to optimize for client-specific learning tasks.

4.2.2. Clustered Federated Learning

In scenarios with significant data heterogeneity, clients can be grouped into clusters based on similarities in their data distributions or model updates [59]. Separate models are trained for each cluster, ensuring that the global models are more aligned with the underlying data distributions.

4.2.3. Regularization-Based Techniques

Regularization methods can mitigate the effects of data heterogeneity by constraining local updates [60]. For instance, FedProx introduces a proximal term to the loss function, ensuring that local updates do not deviate significantly from the global model.

4.2.4. Data Sharing and Synthetic Data Generation

In some cases, clients may share a small subset of anonymized or synthetic data with the server to improve model consistency. Alternatively, the server can generate synthetic data to bridge the gap between heterogeneous client distributions.

 doi:10.20944/preprints202501.0685.v1

4.3. Fairness and Equity in Federated Learning

Fairness is a critical consideration in FL, particularly in heterogeneous environments where disparities in data quality or resource capabilities can lead to biased models. Key approaches to ensure fairness include:

- **Fair Aggregation:** Adjusting the aggregation process to prevent dominant clients from overshadowing underrepresented clients.
- **Equitable Resource Allocation:** Ensuring that clients with limited resources are not excluded from participation or penalized during training.
- **Performance Parity:** Designing global models that provide equitable benefits across diverse clients, even in the presence of significant heterogeneity [61].

4.4. Challenges and Open Problems

While significant progress has been made in addressing heterogeneity in FL, several challenges remain:

- **Scalability:** Managing heterogeneity becomes increasingly complex as the number of clients grows [62].
- **Trade-offs:** Balancing efficiency, accuracy, and fairness often involves trade-offs that depend on the specific FL application [63].
- **Dynamic Environments:** Clients' data distributions and system conditions may evolve over time, requiring adaptive mechanisms [64].
- **Evaluation Metrics:** Establishing robust and universally applicable metrics for evaluating the effectiveness of heterogeneity management techniques [65].

By addressing these challenges, federated learning systems can achieve improved scalability, fairness, and robustness in heterogeneous environments [66]. In the next section, we explore real-world applications of FL, showcasing how these techniques are applied in practice.

5. Applications of Federated Learning

Federated learning (FL) has demonstrated immense potential across a wide range of real-world applications where privacy, scalability, and distributed data are critical concerns. Its ability to enable collaborative machine learning while preserving data locality has made FL a promising solution in domains such as healthcare, finance, smart cities, and beyond [67]. This section provides an overview of key application areas, highlighting how FL is transforming these fields.

5.1. Healthcare and Medical Research

Healthcare is one of the most prominent domains benefiting from federated learning due to the sensitivity of patient data and the distributed nature of healthcare systems. FL enables hospitals, research institutions, and healthcare providers to collaboratively train models on sensitive medical data without violating patient privacy or regulatory compliance [68].

5.1.1. Disease Diagnosis and Prognosis

FL is used to develop models for disease diagnosis, prognosis, and treatment recommendation by aggregating knowledge from diverse patient datasets [69]. For example, federated learning has been employed in training models for detecting diseases like cancer, diabetes, and COVID-19 using medical imaging data such as X-rays, CT scans, and MRIs.

5.1.2. Drug Discovery and Genomics

In drug discovery and genomics, FL facilitates collaboration among pharmaceutical companies and research labs by enabling the training of models on proprietary datasets. This approach accelerates the development of personalized medicine while maintaining data security [70].

 doi:10.20944/preprints202501.0685.v1

5.2. Finance and Banking

In the financial sector, privacy and security are paramount due to the sensitive nature of customer data. FL enables financial institutions to leverage distributed data for building advanced predictive models without compromising confidentiality.

5.2.1. Fraud Detection

FL is applied to train fraud detection models by aggregating transaction data across multiple banks or financial platforms [71]. This collaboration enhances the detection of fraudulent activities while preserving the privacy of customer transactions [72].

5.2.2. Credit Scoring and Risk Assessment

Credit scoring models benefit from FL by learning from distributed customer profiles across different institutions. This approach ensures more accurate credit assessments while maintaining compliance with data protection regulations.

5.3. Smart Cities and IoT Networks

The proliferation of IoT devices and smart city initiatives has resulted in vast amounts of distributed data generated by sensors, cameras, and edge devices [73]. FL offers a scalable solution for training models on this data while addressing privacy and connectivity concerns [74].

5.3.1. Traffic Management

FL is employed in smart transportation systems to optimize traffic flow, reduce congestion, and predict accidents [75]. Models trained on distributed traffic data from sensors and vehicles help improve urban mobility.

5.3.2. Energy Management

In smart grids, FL facilitates the development of energy optimization models by aggregating data from distributed energy meters [76]. These models support load balancing, demand prediction, and energy efficiency [77].

5.4. Natural Language Processing and Recommendation Systems

FL has significant applications in natural language processing (NLP) and personalized recommendation systems, where user privacy is a primary concern.

5.4.1. Predictive Text and Language Models

Techniques like next-word prediction and personalized language models, such as those used in virtual keyboards, leverage FL to improve accuracy while keeping user data private on devices [78].

5.4.2. Personalized Recommendations

FL enables the development of recommendation systems for e-commerce, media streaming, and social platforms by training models on user behavior data without centralizing it [79]. This approach ensures personalized experiences without compromising user privacy [80].

5.5. Autonomous Systems and Robotics

Federated learning is also transforming the fields of autonomous systems and robotics, where distributed data from multiple agents or sensors is used to enhance system performance [81].

5.5.1. Autonomous Vehicles

FL allows autonomous vehicles to collaboratively learn from distributed driving data, improving models for tasks like object detection, lane navigation, and collision avoidance without sharing raw data [82].

 doi:10.20944/preprints202501.0685.v1

5.5.2. Robotic Swarms

In robotics, FL supports the training of decentralized models for collaborative tasks performed by robot swarms, such as search-and-rescue missions or warehouse automation [83].

5.6. Challenges in Real-World Applications

Despite its success in various domains, FL faces several challenges in real-world applications:

- **Scalability:** Managing large-scale FL deployments with thousands or millions of clients.
- **Privacy and Security:** Addressing threats such as model inversion attacks or poisoning attacks in sensitive applications.
- **Heterogeneous Data:** Adapting to diverse data distributions and quality across clients in practical scenarios.
- **Regulatory Compliance:** Navigating complex regulatory frameworks and ensuring FL implementations align with legal requirements.

The application of federated learning continues to grow, driven by its ability to balance privacy, efficiency, and scalability. The next section explores open problems and future directions in FL, paving the way for its broader adoption and advancement.

6. Open Challenges and Future Directions

Federated learning (FL) has emerged as a transformative approach to distributed machine learning, yet it faces numerous unresolved challenges that hinder its broader adoption and optimal performance [84]. These challenges span technical, practical, and ethical dimensions, requiring interdisciplinary research and innovation. In this section, we discuss key open challenges and propose future directions to advance the field of FL [85].

6.1. Scalability and Communication Efficiency

As FL systems grow to include millions of clients, ensuring scalability becomes increasingly challenging [86]. Communication remains a significant bottleneck, particularly in resource-constrained environments [87]. While techniques like model compression and adaptive communication schedules help, they often introduce trade-offs in accuracy or convergence speed [88,89]. **Future Directions:**

- Development of ultra-efficient communication protocols that minimize data exchange without sacrificing model performance [90].
- Exploration of hierarchical FL architectures to manage large-scale deployments by leveraging intermediate aggregators or edge servers.
- Dynamic client participation schemes that prioritize high-value clients while maintaining fairness and diversity [91].

6.2. Robustness to Adversarial Attacks

The decentralized nature of FL exposes it to various security threats, including model poisoning, data poisoning, and adversarial attacks [92]. Malicious clients can disrupt global model performance or compromise its integrity [93]. **Future Directions:**

- Designing robust aggregation algorithms that detect and mitigate the influence of malicious clients [94].
- Incorporating blockchain-based frameworks to enhance trust and accountability in FL systems [95].
- Developing adversarial training techniques tailored to distributed and heterogeneous environments.

6.3. Privacy Enhancements

While FL reduces the need to share raw data, privacy concerns persist due to potential leakage of sensitive information from shared model updates [96]. Techniques like differential privacy and

 doi:10.20944/preprints202501.0685.v1

secure multiparty computation provide solutions but often come at the cost of increased computational overhead and reduced model accuracy. **Future Directions:**

- Advancing lightweight privacy-preserving methods that ensure strong privacy guarantees with minimal impact on system efficiency [97].
- Integrating federated learning with privacy-enhancing technologies, such as homomorphic encryption and trusted execution environments.
- Establishing formal frameworks to quantify privacy risks and trade-offs in FL systems.

6.4. Handling System and Data Heterogeneity

The heterogeneity of client devices and data distributions remains one of the most fundamental challenges in FL. Non-IID data, in particular, can lead to suboptimal global models that fail to generalize well across clients. **Future Directions:**

- Developing adaptive aggregation methods that account for client diversity and data heterogeneity [98].
- Designing personalized FL frameworks that balance global model accuracy with individual client performance [99].
- Leveraging federated transfer learning to handle scenarios with minimal overlap in data distributions across clients [100].

6.5. Evaluation and Benchmarking

The lack of standardized evaluation metrics and benchmarks for FL systems hampers the ability to compare techniques effectively [101]. Existing benchmarks often fail to capture the complexity of real-world FL deployments. **Future Directions:**

- Creating comprehensive FL benchmarks that account for diverse application scenarios, client heterogeneity, and privacy constraints [102].
- Establishing metrics that measure trade-offs among accuracy, efficiency, fairness, and privacy [103].
- Conducting large-scale, real-world FL experiments to validate theoretical advancements and identify practical bottlenecks.

6.6. Cross-Disciplinary Collaboration and Ethical Considerations

The deployment of FL raises ethical questions about fairness, accountability, and transparency. Additionally, advancing FL requires collaboration across disciplines, including machine learning, systems engineering, cryptography, and policy-making. **Future Directions:**

- Developing frameworks for fairness-aware FL that ensure equitable model performance across diverse populations [104].
- Investigating explainable federated learning to enhance the interpretability of models in sensitive domains such as healthcare and finance.
- Establishing guidelines and standards for the ethical deployment of FL systems, considering societal impacts and regulatory compliance.

6.7. Emerging Applications and Novel Paradigms

As FL evolves, new application domains and paradigms are emerging, such as federated reinforcement learning, multi-task federated learning, and cross-device FL at unprecedented scales. **Future Directions:**

- Exploring FL in emerging fields like autonomous systems, metaverse applications, and decentralized AI ecosystems.
- Developing hybrid paradigms that integrate FL with other machine learning approaches, such as transfer learning and self-supervised learning [105].

 doi:10.20944/preprints202501.0685.v1

- Investigating cross-silo FL deployments that involve collaboration among organizations with varying trust levels and regulatory constraints [106].

The future of federated learning lies in addressing these open challenges through innovative research, interdisciplinary collaboration, and ethical foresight [107]. As FL continues to mature, it holds the promise of enabling privacy-preserving, scalable, and robust AI solutions across diverse domains and applications [108].

7. Conclusion

Federated learning (FL) has emerged as a groundbreaking approach to distributed machine learning, enabling collaborative model training across diverse and decentralized datasets while preserving data privacy [109]. This survey has provided a comprehensive exploration of FL, focusing on techniques to enhance efficiency, address heterogeneity, and tackle the challenges associated with real-world deployments. We began by examining the fundamental concepts of FL, highlighting its potential to revolutionize fields such as healthcare, finance, smart cities, and autonomous systems. Techniques for improving communication efficiency, such as model compression, asynchronous communication, and adaptive client participation, were discussed in detail, demonstrating their critical role in enabling scalable FL systems. We also explored strategies for managing system and data heterogeneity, emphasizing the importance of fairness, robustness, and personalization in FL deployments. The discussion extended to applications of FL across various domains, showcasing its ability to balance privacy, performance, and scalability in practical scenarios. Finally, we outlined open challenges and future directions, identifying areas where innovation is needed to address limitations in scalability, robustness, privacy, and fairness [110]. Despite its remarkable progress, FL remains a rapidly evolving field with significant opportunities for advancement. Addressing the identified challenges will require interdisciplinary research and collaboration, integrating insights from machine learning, cryptography, systems design, and ethics. Moreover, establishing standardized benchmarks, metrics, and guidelines will play a pivotal role in accelerating the development and deployment of FL systems [111]. In conclusion, federated learning represents a paradigm shift in how we approach machine learning in a decentralized and privacy-conscious world. As research and technology continue to advance, FL is poised to unlock new possibilities, fostering collaboration across industries and transforming the way we leverage data for innovation. By addressing its current limitations and embracing its future potential, FL can pave the way for equitable, secure, and intelligent systems that serve the collective good [112].

References

1. Ryabinin, M.; Gorbunov, E.; Plokhotnyuk, V.; Pekhimenko, G. Moshpit sgd: Communication-efficient decentralized training on heterogeneous unreliable devices. *Advances in Neural Information Processing Systems* **2021**, *34*, 18195–18211.
2. Ozkara, K.; Singh, N.; Data, D.; Diggavi, S. QuPeD: Quantized Personalization via Distillation with Applications to Federated Learning. *Advances in Neural Information Processing Systems* **2021**, *34*, 3622–3634.
3. Yu, P.; Liu, Y. Federated object detection: Optimizing object detection model with federated learning. In Proceedings of the Proceedings of the 3rd International Conference on Vision, Image and Signal Processing, 2019, pp. 1–6.
4. Lu, Y.; Huang, X.; Zhang, K.; Maharjan, S.; Zhang, Y. Low-latency federated learning and blockchain for edge association in digital twin empowered 6G networks. *IEEE Transactions on Industrial Informatics* **2020**, *17*, 5098–5107.
5. Mashhadi, M.B.; Shlezinger, N.; Eldar, Y.C.; Gündüz, D. Fedrec: Federated learning of universal receivers over fading channels. In Proceedings of the 2021 IEEE Statistical Signal Processing Workshop (SSP). IEEE, 2021, pp. 576–580.
6. Li, T.; Sanjabi, M.; Beirami, A.; Smith, V. Fair resource allocation in federated learning. *arXiv preprint arXiv:1905.10497* **2019**.

 doi:10.20944/preprints202501.0685.v1

7. Shi, S.; Wang, Q.; Zhao, K.; Tang, Z.; Wang, Y.; Huang, X.; Chu, X. A Distributed Synchronous SGD Algorithm with Global Top-k Sparsification for Low Bandwidth Networks. In Proceedings of the 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), 2019, pp. 2238–2247.
8. Sui, D.; Chen, Y.; Zhao, J.; Jia, Y.; Xie, Y.; Sun, W. Feded: Federated learning via ensemble distillation for medical relation extraction. In Proceedings of the Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP), 2020, pp. 2118–2128.
9. Cha, H.; Park, J.; Kim, H.; Bennis, M.; Kim, S.L. Proxy experience replay: Federated distillation for distributed reinforcement learning. *IEEE Intelligent Systems* **2020**, *35*, 94–101.
10. He, Y.; Zenk, M.; Fritz, M. CosSGD: Nonlinear Quantization for Communication-efficient Federated Learning. *CoRR* **2020**, *abs/2012.08241*, [2012.08241].
11. Shi, S.; Wang, Q.; Chu, X.; Li, B.; Qin, Y.; Liu, R.; Zhao, X. Communication-Efficient Distributed Deep Learning with Merged Gradient Sparsification on GPUs. In Proceedings of the IEEE INFOCOM 2020 - IEEE Conference on Computer Communications, 2020, pp. 406–415.
12. Salehi, M.; Hossain, E. Federated learning in unreliable and resource-constrained cellular wireless networks. *IEEE Transactions on Communications* **2021**, *69*, 5136–5151.
13. Xu, G.; Li, H.; Liu, S.; Yang, K.; Lin, X. Verifynet: Secure and verifiable federated learning. *IEEE Transactions on Information Forensics and Security* **2019**, *15*, 911–926.
14. Xu, H.; Ho, C.Y.; Abdelmoniem, A.M.; Dutta, A.; Bergou, E.H.; Karatsenidis, K.; Canini, M.; Kalnis, P. Compressed communication for distributed deep learning: Survey and quantitative evaluation. Technical report, 2020.
15. Sun, J.; Chen, T.; Giannakis, G.B.; Yang, Q.; Yang, Z. Lazily aggregated quantized gradient innovation for communication-efficient federated learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2020**.
16. Haddadpour, F.; Kamani, M.M.; Mokhtari, A.; Mahdavi, M. Federated learning with compression: Unified analysis and sharp guarantees. In Proceedings of the International Conference on Artificial Intelligence and Statistics. PMLR, 2021, pp. 2350–2358.
17. Li, S.; Qi, Q.; Wang, J.; Sun, H.; Li, Y.; Yu, F.R. GGS: General Gradient Sparsification for Federated Learning in Edge Computing. In Proceedings of the ICC 2020 - 2020 IEEE International Conference on Communications (ICC), 2020, pp. 1–7.
18. Alistarh, D.; Grubic, D.; Li, J.; Tomioka, R.; Vojnovic, M. QSGD: Communication-efficient SGD via gradient quantization and encoding. *Advances in neural information processing systems* **2017**, *30*.
19. Amiri, M.M.; Kulkarni, S.R.; Poor, H.V. Federated learning with downlink device selection. In Proceedings of the 2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2021, pp. 306–310.
20. *White paper for federated learning in mobile communication networks*; China Mobile Communications Research Institute, 2021.
21. Sattler, F.; Korjakow, T.; Rischke, R.; Samek, W. Fedaux: Leveraging unlabeled auxiliary data in federated learning. *IEEE Transactions on Neural Networks and Learning Systems* **2021**.
22. Zhao, Z.; Xia, J.; Fan, L.; Lei, X.; Karagiannidis, G.K.; Nallanathan, A. System optimization of federated learning networks with a constrained latency. *IEEE Transactions on Vehicular Technology* **2021**, *71*, 1095–1100.
23. Sun, L.; Lyu, L. Federated model distillation with noise-free differential privacy. *arXiv preprint arXiv:2009.05537* **2020**.
24. 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, 2017, pp. 1273–1282.
25. Das, R.; Acharya, A.; Hashemi, A.; Sanghavi, S.; Dhillon, I.S.; Topcu, U. Faster non-convex federated learning via global and local momentum. *arXiv preprint arXiv:2012.04061* **2020**.
26. Zhang, L.; Shen, L.; Ding, L.; Tao, D.; Duan, L.Y. Fine-tuning global model via data-free knowledge distillation for non-iid federated learning. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 10174–10183.
27. Chen, S.; Shen, C.; Zhang, L.; Tang, Y. Dynamic aggregation for heterogeneous quantization in federated learning. *IEEE Transactions on Wireless Communications* **2021**, *20*, 6804–6819.
28. Liu, Y.; James, J.; Kang, J.; Niyato, D.; Zhang, S. Privacy-preserving traffic flow prediction: A federated learning approach. *IEEE Internet of Things Journal* **2020**, *7*, 7751–7763.

 doi:10.20944/preprints202501.0685.v1

29. Yang, H.; Qiu, P.; Liu, J.; Yener, A. Over-the-Air Federated Learning with Joint Adaptive Computation and Power Control. *arXiv preprint arXiv:2205.05867* **2022**.
30. Xia, S.; Zhu, J.; Yang, Y.; Zhou, Y.; Shi, Y.; Chen, W. Fast convergence algorithm for analog federated learning. In Proceedings of the ICC 2021-IEEE International Conference on Communications. IEEE, 2021, pp. 1–6.
31. 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. *Foundations and Trends® in Machine Learning* **2021**, *14*, 1–210.
32. Sturluson, S.P.; Trew, S.; Muñoz-González, L.; Grama, M.; Passerat-Palmbach, J.; Rueckert, D.; Alansary, A. FedRAD: Federated Robust Adaptive Distillation. *arXiv preprint arXiv:2112.01405* **2021**.
33. Chen, M.; Shlezinger, N.; Poor, H.V.; Eldar, Y.C.; Cui, S. Communication-efficient federated learning. *Proceedings of the National Academy of Sciences* **2021**, *118*, e2024789118.
34. Wall, M.E.; Rechtsteiner, A.; Rocha, L.M. Singular value decomposition and principal component analysis. In *A practical approach to microarray data analysis*; Springer, 2003; pp. 91–109.
35. Zhu, Z.; Hong, J.; Zhou, J. Data-free knowledge distillation for heterogeneous federated learning. In Proceedings of the International Conference on Machine Learning. PMLR, 2021, pp. 12878–12889.
36. Lin, Y.; Han, S.; Mao, H.; Wang, Y.; Dally, W.J. Deep gradient compression: Reducing the communication bandwidth for distributed training. *arXiv preprint arXiv:1712.01887* **2017**.
37. Mao, Y.; Zhao, Z.; Yang, M.; Liang, L.; Liu, Y.; Ding, W.; Lan, T.; Zhang, X.P. SAFARI: Sparsity enabled Federated Learning with Limited and Unreliable Communications. *arXiv preprint arXiv:2204.02321* **2022**.
38. Sattler, F.; Marban, A.; Rischke, R.; Samek, W. Cfd: Communication-efficient federated distillation via soft-label quantization and delta coding. *IEEE Transactions on Network Science and Engineering* **2021**.
39. Aji, A.F.; Heafield, K. Sparse communication for distributed gradient descent. *arXiv preprint arXiv:1704.05021* **2017**.
40. Cho, Y.J.; Manoel, A.; Joshi, G.; Sim, R.; Dimitriadis, D. Heterogeneous Ensemble Knowledge Transfer for Training Large Models in Federated Learning. *arXiv preprint arXiv:2204.12703* **2022**.
41. Azam, S.S.; Hosseinalipour, S.; Qiu, Q.; Brinton, C. Recycling Model Updates in Federated Learning: Are Gradient Subspaces Low-Rank? In Proceedings of the International Conference on Learning Representations, 2021.
42. Wang, H.; Sievert, S.; Liu, S.; Charles, Z.; Papailiopoulos, D.; Wright, S. Atomo: Communication-efficient learning via atomic sparsification. *Advances in Neural Information Processing Systems* **2018**, *31*.
43. Wei, X.; Shen, C. Federated learning over noisy channels. In Proceedings of the ICC 2021-IEEE International Conference on Communications. IEEE, 2021, pp. 1–6.
44. Shi, S.; Tang, Z.; Wang, Q.; Zhao, K.; Chu, X. Layer-wise adaptive gradient sparsification for distributed deep learning with convergence guarantees. *arXiv preprint arXiv:1911.08727* **2019**.
45. Zniyed, Y.; Nguyen, T.P.; et al. Efficient tensor decomposition-based filter pruning. *Neural Networks* **2024**, *178*, 106393.
46. Philippenko, C.; Dieuleveut, A. Bidirectional compression in heterogeneous settings for distributed or federated learning with partial participation: tight convergence guarantees. *arXiv preprint arXiv:2006.14591* **2020**.
47. Wu, W.; He, L.; Lin, W.; Mao, R.; Maple, C.; Jarvis, S. SAFA: A semi-asynchronous protocol for fast federated learning with low overhead. *IEEE Transactions on Computers* **2020**, *70*, 655–668.
48. Muhammad, K.; Ullah, A.; Lloret, J.; Del Ser, J.; de Albuquerque, V.H.C. Deep learning for safe autonomous driving: Current challenges and future directions. *IEEE Transactions on Intelligent Transportation Systems* **2020**, *22*, 4316–4336.
49. Zamir, R.; Feder, M. On universal quantization by randomized uniform/lattice quantizers. *IEEE Transactions on Information Theory* **1992**, *38*, 428–436.
50. Fan, X.; Wang, Y.; Huo, Y.; Tian, Z. Communication-efficient federated learning through 1-bit compressive sensing and analog aggregation. In Proceedings of the 2021 IEEE International Conference on Communications Workshops (ICC Workshops). IEEE, 2021, pp. 1–6.
51. Mothukuri, V.; Parizi, R.M.; Pouriyeh, S.; Huang, Y.; Dehghantanha, A.; Srivastava, G. A survey on security and privacy of federated learning. *Future Generation Computer Systems* **2021**, *115*, 619–640.
52. Jia, X.; Song, S.; He, W.; Wang, Y.; Rong, H.; Zhou, F.; Xie, L.; Guo, Z.; Yang, Y.; Yu, L.; et al. Highly scalable deep learning training system with mixed-precision: Training imagenet in four minutes. *arXiv preprint arXiv:1807.11205* **2018**.

 doi:10.20944/preprints202501.0685.v1

53. Yu, C.; Tang, H.; Renggli, C.; Kassing, S.; Singla, A.; Alistarh, D.; Zhang, C.; Liu, J. Distributed learning over unreliable networks. In Proceedings of the International Conference on Machine Learning. PMLR, 2019, pp. 7202–7212.
54. Jhunjhunwala, D.; Gadhikar, A.; Joshi, G.; Eldar, Y.C. Adaptive quantization of model updates for communication-efficient federated learning. In Proceedings of the ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021, pp. 3110–3114.
55. Cha, H.; Park, J.; Kim, H.; Kim, S.L.; Bennis, M. Federated reinforcement distillation with proxy experience memory. *arXiv preprint arXiv:1907.06536* **2019**.
56. Eghlidi, N.F.; Jaggi, M. Sparse communication for training deep networks. *arXiv preprint arXiv:2009.09271* **2020**.
57. Bottou, L. Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010*; Springer, 2010; pp. 177–186.
58. Li, C.; Li, G.; Varshney, P.K. Communication-efficient federated learning based on compressed sensing. *IEEE Internet of Things Journal* **2021**, *8*, 15531–15541.
59. Kairouz, P.; Oh, S.; Viswanath, P. Extremal mechanisms for local differential privacy. *Advances in neural information processing systems* **2014**, *27*.
60. Al-Qizwini, M.; Barjasteh, I.; Al-Qassab, H.; Radha, H. Deep learning algorithm for autonomous driving using googlenet. In Proceedings of the 2017 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2017, pp. 89–96.
61. Rieke, N.; Hancox, J.; Li, W.; Milletari, F.; Roth, H.R.; Albarqouni, S.; Bakas, S.; Galtier, M.N.; Landman, B.A.; Maier-Hein, K.; et al. The future of digital health with federated learning. *NPJ digital medicine* **2020**, *3*, 1–7.
62. Xiao, Y.; Shi, G.; Krunz, M. Towards ubiquitous AI in 6G with federated learning. *arXiv preprint arXiv:2004.13563* **2020**.
63. Jeong, E.; Oh, S.; Kim, H.; Park, J.; Bennis, M.; Kim, S.L. Communication-efficient on-device machine learning: Federated distillation and augmentation under non-iid private data. *arXiv preprint arXiv:1811.11479* **2018**.
64. Konečnỳ, J.; McMahan, H.B.; Yu, F.X.; Richtárik, P.; Suresh, A.T.; Bacon, D. Federated learning: Strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492* **2016**.
65. Sahoo, A.K.; Pradhan, C.; Barik, R.K.; Dubey, H. DeepReco: deep learning based health recommender system using collaborative filtering. *Computation* **2019**, *7*, 25.
66. Arandjelovic, O.; Shakhnarovich, G.; Fisher, J.; Cipolla, R.; Darrell, T. Face recognition with image sets using manifold density divergence. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). IEEE, 2005, Vol. 1, pp. 581–588.
67. Lin, Z.; Liu, H.; Zhang, Y.J.A. Relay-assisted cooperative federated learning. *IEEE Transactions on Wireless Communications* **2022**.
68. Ahn, J.H.; Simeone, O.; Kang, J. Wireless federated distillation for distributed edge learning with heterogeneous data. In Proceedings of the 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC). IEEE, 2019, pp. 1–6.
69. Zhang, X.; Hong, M.; Dhople, S.; Yin, W.; Liu, Y. Fedpd: A federated learning framework with optimal rates and adaptivity to non-iid data. *arXiv preprint arXiv:2005.11418* **2020**.
70. Jiang, Z.; Wang, W.; Li, B.; Yang, Q. Towards Efficient Synchronous Federated Training: A Survey on System Optimization Strategies. *IEEE Transactions on Big Data* **2022**.
71. Feynman, R.; Vernon Jr., F. The theory of a general quantum system interacting with a linear dissipative system. *Annals of Physics* **1963**, *24*, 118–173. https://doi.org/10.1016/0003-4916(63)90068-X.
72. Mahmoudi, A.; Ghadikolaei, H.S.; Júnior, J.M.B.D.S.; Fischione, C. FedCau: A Proactive Stop Policy for Communication and Computation Efficient Federated Learning. *arXiv preprint arXiv:2204.07773* **2022**.
73. Zeng, Q.; Du, Y.; Huang, K.; Leung, K.K. Energy-efficient radio resource allocation for federated edge learning. In Proceedings of the 2020 IEEE International Conference on Communications Workshops (ICC Workshops). IEEE, 2020, pp. 1–6.
74. Ahn, J.H.; Simeone, O.; Kang, J. Cooperative learning via federated distillation over fading channels. In Proceedings of the ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020, pp. 8856–8860.
75. Oh, S.; Park, J.; Jeong, E.; Kim, H.; Bennis, M.; Kim, S.L. Mix2FLD: Downlink federated learning after uplink federated distillation with two-way mixup. *IEEE Communications Letters* **2020**, *24*, 2211–2215.
76. Jiang, D.; Shan, C.; Zhang, Z. Federated learning algorithm based on knowledge distillation. In Proceedings of the 2020 International Conference on Artificial Intelligence and Computer Engineering (ICAICE). IEEE, 2020, pp. 163–167.

 doi:10.20944/preprints202501.0685.v1

77. Zhang, L.; Wu, D.; Yuan, X. FedZKT: Zero-Shot Knowledge Transfer towards Resource-Constrained Federated Learning with Heterogeneous On-Device Models. In Proceedings of the 2022 IEEE 42nd International Conference on Distributed Computing Systems (ICDCS). IEEE, 2022, pp. 928–938.
78. Vargaftik, S.; Basat, R.B.; Portnoy, A.; Mendelson, G.; Ben-Itzhak, Y.; Mitzenmacher, M. DRIVE: one-bit distributed mean estimation. *Advances in Neural Information Processing Systems* **2021**, *34*, 362–377.
79. Nishio, T.; Yonetani, R. Client selection for federated learning with heterogeneous resources in mobile edge. In Proceedings of the ICC 2019-2019 IEEE international conference on communications (ICC). IEEE, 2019, pp. 1–7.
80. Lian, X.; Zhang, C.; Zhang, H.; Hsieh, C.J.; Zhang, W.; Liu, J. Can decentralized algorithms outperform centralized algorithms? a case study for decentralized parallel stochastic gradient descent. *Advances in Neural Information Processing Systems* **2017**, *30*.
81. Chen, M.; Poor, H.V.; Saad, W.; Cui, S. Convergence time optimization for federated learning over wireless networks. *IEEE Transactions on Wireless Communications* **2020**, *20*, 2457–2471.
82. Wu, C.; Zhu, S.; Mitra, P. Federated Unlearning with Knowledge Distillation. *arXiv preprint arXiv:2201.09441* **2022**.
83. Grigorescu, S.; Trasnea, B.; Cocias, T.; Macesanu, G. A survey of deep learning techniques for autonomous driving. *Journal of Field Robotics* **2020**, *37*, 362–386.
84. Chen, M.; Mao, B.; Ma, T. FedSA: A staleness-aware asynchronous Federated Learning algorithm with non-IID data. *Future Generation Computer Systems* **2021**, *120*, 1–12.
85. Chang, H.; Shejwalkar, V.; Shokri, R.; Houmansadr, A. Cronus: Robust and heterogeneous collaborative learning with black-box knowledge transfer. *arXiv preprint arXiv:1912.11279* **2019**.
86. Nori, M.K.; Yun, S.; Kim, I.M. Fast federated learning by balancing communication trade-offs. *IEEE Transactions on Communications* **2021**, *69*, 5168–5182.
87. Itahara, S.; Nishio, T.; Koda, Y.; Morikura, M.; Yamamoto, K. Distillation-Based Semi-Supervised Federated Learning for Communication-Efficient Collaborative Training with Non-IID Private Data. *IEEE Transactions on Mobile Computing* **2021**.
88. Sun, C.; Jiang, T.; Zonouz, S.; Pompili, D. Fed2KD: Heterogeneous Federated Learning for Pandemic Risk Assessment via Two-Way Knowledge Distillation. In Proceedings of the 2022 17th Wireless On-Demand Network Systems and Services Conference (WONS). IEEE, 2022, pp. 1–8.
89. Zniyed, Y.; Nguyen, T.P.; et al. Enhanced network compression through tensor decompositions and pruning. *IEEE Transactions on Neural Networks and Learning Systems* **2024**.
90. Gray, R.M.; Neuhoff, D.L. Quantization. *IEEE transactions on information theory* **1998**, *44*, 2325–2383.
91. Han, P.; Wang, S.; Leung, K.K. Adaptive Gradient Sparsification for Efficient Federated Learning: An Online Learning Approach. In Proceedings of the 2020 IEEE 40th International Conference on Distributed Computing Systems (ICDCS), 2020, pp. 300–310.
92. Zaw, C.W.; Pandey, S.R.; Kim, K.; Hong, C.S. Energy-aware resource management for federated learning in multi-access edge computing systems. *IEEE Access* **2021**, *9*, 34938–34950.
93. Tramèr, F.; Kurakin, A.; Papernot, N.; Goodfellow, I.; Boneh, D.; McDaniel, P. Ensemble adversarial training: Attacks and defenses. *arXiv preprint arXiv:1705.07204* **2017**.
94. Malekijoo, A.; Fadaeieslam, M.J.; Malekijou, H.; Homayounfar, M.; Alizadeh-Shabdiz, F.; Rawassizadeh, R. Fedzip: A compression framework for communication-efficient federated learning. *arXiv preprint arXiv:2102.01593* **2021**.
95. Suresh, A.T.; Felix, X.Y.; Kumar, S.; McMahan, H.B. Distributed mean estimation with limited communication. In Proceedings of the International conference on machine learning. PMLR, 2017, pp. 3329–3337.
96. Xu, H.; Kostopoulou, K.; Dutta, A.; Li, X.; Ntoulas, A.; Kalnis, P. DeepReduce: A Sparse-tensor Communication Framework for Federated Deep Learning. In Proceedings of the Advances in Neural Information Processing Systems; Ranzato, M.; Beygelzimer, A.; Dauphin, Y.; Liang, P.; Vaughan, J.W., Eds. Curran Associates, Inc., 2021, Vol. 34, pp. 21150–21163.
97. Imteaj, A.; Amini, M.H. Fedar: Activity and resource-aware federated learning model for distributed mobile robots. In Proceedings of the 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, 2020, pp. 1153–1160.
98. Reisizadeh, A.; Mokhtari, A.; Hassani, H.; Jadbabaie, A.; Pedarsani, R. Fedpaq: A communication-efficient federated learning method with periodic averaging and quantization. In Proceedings of the International Conference on Artificial Intelligence and Statistics. PMLR, 2020, pp. 2021–2031.

 doi:10.20944/preprints202501.0685.v1

99. Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J.; et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning* **2011**, *3*, 1–122.
100. Hard, A.; Rao, K.; Mathews, R.; Ramaswamy, S.; Beaufays, F.; Augenstein, S.; Eichner, H.; Kiddon, C.; Ramage, D. Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604* **2018**.
101. Donoho, D.L. Compressed sensing. *IEEE Transactions on information theory* **2006**, *52*, 1289–1306.
102. *Study on enablers for network automation for the 5G System (5GS)*; 3GPP TR 23.700-91, 2020.
103. Zhu, Z.; Hong, J.; Drew, S.; Zhou, J. Resilient and Communication Efficient Learning for Heterogeneous Federated Systems. In Proceedings of the International Conference on Machine Learning. PMLR, 2022, pp. 27504–27526.
104. Zhang, N.; Tao, M. Gradient statistics aware power control for over-the-air federated learning in fading channels. In Proceedings of the 2020 IEEE International Conference on Communications Workshops (ICC Workshops). IEEE, 2020, pp. 1–6.
105. Yang, Z.; Chen, M.; Saad, W.; Hong, C.S.; Shikh-Bahaei, M.; Poor, H.V.; Cui, S. Delay minimization for federated learning over wireless communication networks. *arXiv preprint arXiv:2007.03462* **2020**.
106. Wu, W.; He, L.; Lin, W.; Mao, R. Accelerating federated learning over reliability-agnostic clients in mobile edge computing systems. *IEEE Transactions on Parallel and Distributed Systems* **2020**, *32*, 1539–1551.
107. Wu, C.; Wu, F.; Liu, R.; Lyu, L.; Huang, Y.; Xie, X. Fedkd: Communication efficient federated learning via knowledge distillation. *arXiv preprint arXiv:2108.13323* **2021**.
108. Sergeev, A.; Del Balso, M. Horovod: fast and easy distributed deep learning in TensorFlow. *arXiv preprint arXiv:1802.05799* **2018**.
109. Bernstein, J.; Wang, Y.X.; Azizzadenesheli, K.; Anandkumar, A. signSGD: Compressed optimisation for non-convex problems. In Proceedings of the International Conference on Machine Learning. PMLR, 2018, pp. 560–569.
110. Lin, T.; Kong, L.; Stich, S.U.; Jaggi, M. Ensemble distillation for robust model fusion in federated learning. *Advances in Neural Information Processing Systems* **2020**, *33*, 2351–2363.
111. Yang, Z.; Chen, M.; Saad, W.; Hong, C.S.; Shikh-Bahaei, M. Energy efficient federated learning over wireless communication networks. *IEEE Transactions on Wireless Communications* **2020**, *20*, 1935–1949.
112. Stich, S.U.; Cordonnier, J.B.; Jaggi, M. Sparsified SGD with memory. *Advances in Neural Information Processing Systems* **2018**, *31*.