Towards Heterogeneous Federated Learning: Analysis, Solutions, and Future Directions

Yongwei Lin, Yucheng Long, Zhili Zhou(✉), Yan Pang, and Chunsheng Yang

Artificial Intelligence Research Institute, Guangzhou University, Guangzhou 510006, China
{YongweiLin,YuchengLong}@e.gzhu.edu.cn, zhou_zhili@163.com, chunsheng.yang@gzhu.edu.cn

Abstract. With the rapid growth of edge devices such as smartphones, wearables, and mobile networks, how to effectively utilize a large amount of private data stored on these devices has become a challenging issue. To address this issue, federated learning has emerged as a promising solution. Federated learning allows multiple devices to train machine learning models collaboratively while keeping the data decentralized and following local privacy policies. However, the heterogeneous differences in data distributions, model structures, network environments, and devices pose challenges in realizing collaboration. In this paper, we reviewed the heterogeneous federated learning (HFL) approaches and classified them into data heterogeneity, device heterogeneity, communication heterogeneity, and model heterogeneity. Also, we concluded their advantages and disadvantages and gave the solutions to the limitations in detail. Meanwhile, this paper introduces the commonly used methods for evaluating the performance of federated learning and suggests the future directions of the HFL framework.

Keywords: Heterogeneous Federated Learning · Trustworthy AI · Federated Learning

1 Introduction

In today's technologically advanced society, edge devices such as smartphones, wearable devices, and mobile networks are widely used in real-world applications. However, it is challenging to use the large amount of personal data stored on these devices without compromising privacy. To address this challenge, federated learning has emerged as a promising solution. Federated learning allows multiple devices to train machine learning models collaboratively while keeping

This work is supported in part by the National Natural Science Foundation of China under Grant 62372125, in part by the Guangdong Natural Science Funds for Distinguished Young Scholar under Grant 2023B1515020041.

J. Vaidya et al. (Eds.): AIS&P 2023, LNCS 14509, pp. 172–189, 2024.
https://doi.org/10.1007/978-981-99-9785-5_13

the data decentralized, adhering to the data locality principle. In this paradigm, the devices participating in the federated learning system are called clients. Federated learning is a secure and distributed machine learning framework based on encryption techniques, enabling organizations to engage in collaborative model training while safeguarding data privacy. Federated learning(FL) [1,2], a collaborative learning model paradigm, has attracted increasing attention from industry and academia. Extensive research on this approach has been conducted in various real scenarios, including healthcare [3], recommendation systems [4], anti-money laundering [5], and data security [6].

Federated learning has achieved significant success. However, since most existing research in federated learning is based on the assumption of homogeneous data that can be easily aggregated. there are numerous challenges [7], including variations in data distributions, model structures, network environments, and edge devices, making federated collaboration hard to implement. The heterogeneity issues have existed in various aspects of the learning process, including data heterogeneity, device heterogeneity, communication heterogeneity, and model heterogeneity. As shown in Fig. 1, the specific challenges are summarized as follows.

(1) Data heterogeneity: Due to the Non-Independent Identical Distribution (Non-IID) problem of the client's local data, the results obtained from model training on one client may not be able to be generalized to other client's data, resulting in a decline in the overall performance of the model.
(2) Device heterogeneity: Due to the differences in client's storage, computation, and communication capabilities, the computational power of some clients is weak, and they cannot perform complex model training or gradient calculation, resulting in an imbalance between the devices involved in federated learning, affecting the overall training effect.
(3) Communication heterogeneity: Due to the differences in the network environment in which the client is located, there may be communication delays and bandwidth constraints, resulting in a blockage of the aggregation process of the model parameters, which affects the model's updating and convergence speed.
(4) Model heterogeneity: In various application scenarios, different tasks require different models, so customers need to effectively integrate different types of models. However, this is a challenging task that requires solving the problem of model fusion and integration.

To address the above heterogeneity problems, we provide a comprehensive survey of research work on HFL in this paper. We conduct a comprehensive investigation into the fundamental causes of heterogeneity in federated learning and subsequently classify HFL approaches into four categories: data heterogeneity, device heterogeneity, communication heterogeneity, and model heterogeneity. This study sufficiently analyzes the solutions to address these challenges. Additionally, it employs commonly employed performance evaluation methodologies to evaluate the performances of existing HFL approaches and also gives four potential research directions of the HFL framework.

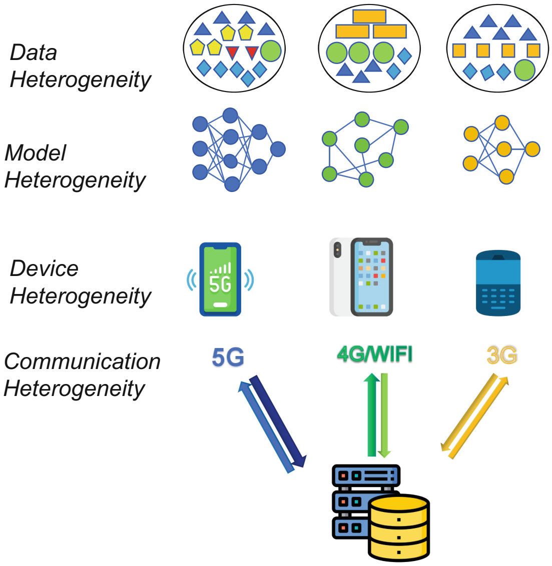

Fig. 1. Schematic of heterogeneous federated learning

1.1 Related Surveys

A survey conducted by Yang et al. [2] was quite influential in establishing the fundamental principles and concepts of Federated Learning. It also proposed an extensive and robust FL framework. Kairouz et al. [8] expanded the applications of FL to various scenarios. Wahab et al. [9] conducted a thorough investigation and synthesis, presenting a multi-level classification methodology and evaluation criteria, and exploring the prospects of federated learning within communication and network systems. In a more recent survey [10], the domain of Personalized Federated Learning (PFL) was introduced, accompanied by an exploration of the fundamental challenges of privacy-preserved machine learning on heterogeneous data. The survey described PFL techniques, pivotal concepts, and future research directions.

However, it is worth noting that several surveys focus on HFL. The study [11] offered an all-encompassing assessment of the profound impact of heterogeneity on quality and fairness in federated learning, highlighting significant effects on model performance and fairness in mixed heterogeneity scenarios. The concept of HFL was initially introduced by Gao et al. [12], who endeavored to tackle the intricate challenges posed by heterogeneity in federated learning through the comprehensive investigation of various aspects, including data space, statistics, systems, and model heterogeneity. In a recent survey, Ye et al. [13] provided a systematic examination and comprehensive review of the practical challenges and innovative solutions of HFL. The survey research challenges in HFL, a thorough review of recent advancements, analysis of existing approaches, and an insightful on future research directions.

2 A Taxonomy of Heterogeneous Federated Learning

2.1 Definition

The concept of HFL aims to address the inherent heterogeneity among participants in terms of data, devices, communication, and models. The primary objective of HFL is to facilitate the integration of knowledge across diverse participants, thereby enhancing model performance and generalization capabilities.

2.2 Analysis

Data Heterogeneity. Data Heterogeneity is often regarded as statistical heterogeneity, where the data deviates from complete independence and identical distribution, commonly known as non-independent and identically distributed (non-i.i.d.).

Model Heterogeneity. Model heterogeneity presents numerous technological and algorithmic challenges in the field of federated learning. Primarily, models with different architectures may have different quantities of parameters and follow distinct update rules. Consequently, this makes the aggregation of model parameters notably complex during the federated learning process. Moreover, model heterogeneity gives rise to disparities in model performance, as divergent model types may exhibit variances in data processing and learning tasks.

Device Heterogeneity. In a pristine federation environment, clients demonstrate a diverse range of device configurations, including variations in GPU, CPU, software, and network conditions. This heterogeneity leads to significant discrepancies in device overhead, such as compute time and resource utilization, when striving to accomplish the same task. Consequently, this exacerbates the performance degradation of a global model.

Communication Heterogeneity. In real-world implementations of the Internet of Things (IoT), devices are commonly deployed in diverse network environments, each characterized by distinct network connectivity settings. Consequently, that leads to variations in communication attributes such as bandwidth, latency, and reliability, resulting in what is widely known as communication heterogeneity.

3 Heterogeneous Federated Learning Taxonomy

3.1 Federal Learning Strategies with Heterogeneous Data

The presence of non-i.i.d. data among clients poses a challenge known as data heterogeneity. Addressing the detrimental effects of non-i.i.d. data remains an ongoing research endeavor. Non-i.i.d. data exhibits distributional skews, often observed as label distribution skew, feature distribution skew, and quantity skew.

Label Distribution Skew: In the context of label distribution skew [14], the heterogeneous distribution of target labels (or classes) can lead to significant disparities among diverse clients.

Feature Distribution Skew: Feature distribution skew refers to discrepancies in the distributions of input features among clients [15]. The variation in feature distributions could increase due to divergent data collection processes employed by different clients.

Quantity Skew: Quantity skew refers to disparities in the available data volume among clients. Certain clients may possess a substantial amount of data [16], while others may have a limited number of data samples. Consequently, clients with abundant data can exert undue influence on the training process. Effectively managing quantity skew requires techniques to mitigate the impact of data imbalance and prevent clients with limited data from being overshadowed.

Based on the classification of data heterogeneity problems, there are several potential solutions to consider. One such solution is data augmentation, which involves enriching or amplifying data by incorporating supplementary information or features. Fedmix [17] aims to integrate Mixup techniques into federated learning to enhance the mean-based federated learning paradigm. This innovative approach introduces mean-enhancement techniques within the federated learning framework, approximating the benefits of Mixup. As a result, Fedmix effectively addresses challenges such as overfitting, enhances model generalization, and mitigates issues associated with imbalanced data distribution. However, it is important to note that Fedmix has its limitations. The collection of local data distributions may introduce potential information leakages, raising concerns about privacy and security. Astraea [18] tackles the challenges of data imbalance and model bias through adaptive sample selection and uncertainty-driven model updating strategies. The efficacy of these approaches is rigorously demonstrated across various datasets encompassing mobile deep-learning applications. However, it is worth noting that some researchers have expressed concerns about Astraea's method, suggesting that the disclosure of local data distributions during upload could inadvertently expose vulnerabilities and make it susceptible to malicious intrusion.

This section delves into the data heterogeneity issues encountered in federated learning, encompassing label distribution skewness, feature distribution skewness, and quantity skewness. To address these challenges, the section examines potential solutions, including data augmentation and the Fedmix method. In future research, it would be valuable to explore approaches that effectively mitigate the limitations of these methods and further enhance their efficacy and security.

3.2 Federal Learning Strategies with Heterogeneous Model

Model heterogeneity presents a challenge when attempting to transfer knowledge between the clients that employ different models with a model-independent approach. To tackle this issue, we classify model heterogeneity into two categories: partial heterogeneity and full heterogeneity.

Partial Heterogeneity. Partial heterogeneity refers to the differences in the model architectures adopted by various clients, while certain components or layers of the models remain consistent across clients. In other words, there exists a partial overlap in the model architectures. This variability may arise due to hardware limitations, individual requirements, or variations in the task properties that clients aim to address.

Complete Heterogeneity. Complete heterogeneity in federated learning occurs when different clients use model architectures that have significant differences, leading to a wide variety of models. To effectively tackle this challenge, sophisticated strategies like meta-learning [19] or model-agnostic mechanisms are necessary. These strategies facilitate the generalization and transfer of knowledge while accommodating the diverse model structures.

Several solutions are proposed based on the above categorization of model heterogeneity into partial and complete heterogeneity. One of these solutions is knowledge distillation, which relaxes the stringent requirements for homogeneous local models by using logarithms as a representation of knowledge transfer. This approach allows for the creation of federated learning systems that can accommodate different model architectures [19]. Wang introduced the VFed-Trans framework for facilitating privacy-preserving data sharing and knowledge transfer among healthcare organizations [20]. This framework utilizes a joint modeling approach to extract a federated representation of shared samples by combining their features. However, researchers have expressed concerns regarding the effectiveness, scalability, and applicability of this approach in different scenarios of vertical federated learning. In the field of federated learning, Le et al. proposed FedLKD, an approach that utilizes layer-wise knowledge distillation [21]. The goal of FedLKD is to enhance the local training process by applying knowledge distillation between global and local models, using a small proxy dataset. However, it is important to carefully consider the potential impact of this method on privacy preservation, as emphasized by several researchers. Yu et al. proposed an innovative approach to address the inherent heterogeneity in joint learning through local adaptation [22]. This technique aims to enhance model efficiency and convergence by incorporating local model adaptation and parameter tuning. It enables each client to personalize model training based on its local data characteristics and device capabilities. However, it is important to consider the potential limitations of this method when utilizing logits, as it may result in insufficient integration of local information.

In summary, dealing with model heterogeneity presents a significant challenge in the field of federated learning, which can be categorized as partial heterogeneity and complete heterogeneity. To tackle this challenge, researchers have proposed several effective solutions, such as knowledge distillation, federated inter-layer distillation, local model adaptation, and parameter tuning. However, it is crucial to conduct further research and practical validation to improve the efficiency and feasibility of these approaches. Moreover, it is important to carefully consider essential aspects like privacy protection and potential limitations when implementing these methods.

3.3 Federal Learning Strategies with Heterogeneous Communication

Within the intricate landscape of the Internet of Things, the prevalence of communication heterogeneity poses significant challenges, characterized by high communication costs and suboptimal efficiency [23], thereby diminishing the efficacy of federated learning. Several methodologies have emerged as joint strategies to address the pervasive challenge of communication heterogeneity. These encompass the optimization of compression parameters and gradients, the reduction of communication rounds, and the implementation of asynchronous training techniques.

Compression Parameters and Gradients. Model parameter compression is an effective strategy for dealing with variations in communication during federated learning. It reduces the amount of data transmitted by compressing model parameters and can be personalized based on the characteristics and limitations of individual devices. By selectively transmitting gradient updates according to device characteristics and communication conditions, we can minimize communication overhead and improve efficiency.

From the perspective of Compression Parameters and Gradients, there are several methods to address the communication heterogeneity in federated learning. For example, Communication-Mitigated Federated Learning [24] addresses the transmission of inconsequential updates to the central server by evaluating the compliance of local updates with global updates. This method is effective in reducing the workload of communication transmission. However, it is crucial to take into account the limitations of this approach when dealing with networks that are highly diverse or unreliable. The Federated Deep Neural Networks Framework [25]introduces a transformative approach by substituting every fully connected (FC) layer with a pair of low-rank projection matrices, thereby achieving model compression within the DNNs architecture. The framework establishes a comprehensive global error function to reconstruct the output of the compressed DNNs model, ensuring fidelity in the compression process. In addition, FedSkel [26] enhances federated learning by improving computational efficiency and optimizing communication on edge devices. This is achieved through selective model updates that solely target the essential components, thereby reducing resource requirements. However, it is important to note that the scalability of the system, particularly concerning privacy and security concerns, has not been extensively analyzed.

Reducing Communication Rounds. Reducing the number of communication rounds is an effective strategy to handle communication heterogeneity. It helps to minimize the overall communication overhead between participants. In the context of federated learning, FedMMD [27] improves the optimization process by introducing the Maximum Mean Discrepancy constraint into the loss function. This integration leads to a reduction in the required communication rounds. Another approach, FedSeq [28] enhances the algorithm's performance and convergence rate by setting a predefined communication round budget. This approach effectively manages resource allocation and streamlines the learning process.

Reducing the number of communication rounds is an effective approach to address communication differences in federated learning. FedMMD integrates the Maximum Mean Discrepancy constraint into the loss function, aiming to minimize communication rounds. Meanwhile, FedSeq enhances performance and convergence by setting a predetermined limit on the communication rounds allowed.

Asynchronous Training. In asynchronous training, participants have the freedom to update model parameters independently without waiting for others to finish their updates. This concurrent process has the potential to improve communication efficiency, especially in situations with high communication latency [7]. In the context of asynchronous training, several methods have been proposed to handle communication differences in federated learning. For example, FedSeC [29] introduces a framework for differential privacy that incorporates an optimization technique based on updates. On the other hand, FedSA [30] uses a semi-asynchronous mechanism that relies on the sequential order of model updates. Additionally, FedHe [31] applies a knowledge distillation-like approach to reduce communication overhead.

Communication heterogeneity in federated learning poses significant challenges, including high communication costs and suboptimal efficiency. To address this issue, researchers have developed several methodologies, including compression parameters and gradients, reducing communication rounds, and asynchronous training techniques. Methods such as Communication-Mitigated Federated Learning, Federated Deep Neural Networks Framework, and FedSkel optimize compression and computational efficiency, while FedMMD and FedSeq reduce communication rounds. Asynchronous training methods such as FedSeC, FedSA, and FedHe also address communication heterogeneity.

3.4 Federal Learning Strategies with Heterogeneous Devices

Device heterogeneity in federated learning arises due to disparities in device configurations, including hardware, software, and network conditions [7]. To address this challenge, methodologies such as fine-tuning training tasks and client selection are utilized to allocate suitable tasks to edge devices, aiming to optimize overall efficiency.

Training Tasks Adjustment. To optimize global efficiency in federated learning, it is essential to allocate appropriate tasks based on the computational capabilities of each device, while also considering factors like fairness, privacy, and data diversity. Intelligent algorithms play a crucial role in navigating device heterogeneity and enabling effective collaboration. For instance, Abdellatif et al. [32] propose an efficient user and resource allocation scheme for horizontal federated learning. This system leverages the vast volumes of data generated by Internet of Things (IoT) devices to train deep learning models, addressing the challenges and requirements posed by data privacy and resource-constrained environments.FedSAE [33] tackles the issue of performance degradation in federated learning through two key mechanisms: automatic adjustment of device training task capabilities and participant selection. This approach utilizes comprehensive information about a device's history of training tasks to predict its training load capacity, enabling adaptive participant selection. However, it is important to note that refining workload allocation based on client training history may introduce temporal delays.

Client Selection. Client selection is a critical aspect of federated learning, aiming to identify suitable clients for each iteration based on their constraints, such as network bandwidth, computation capability, and local resources. Selection strategy plays a crucial role in accelerating convergence and improving model accuracy. Wang et al. [34] have made significant contributions in this field. Their research tackles the challenges posed by non-IID data in federated learning. They propose a reinforcement learning framework specifically designed for this scenario, including an effective data representation method, an optimized task allocation strategy, and a model aggregation mechanism. It is important to note that reinforcement learning models require a substantial amount of data for effective training. Furthermore, in addressing the challenges arising from device heterogeneity, client selection is often combined with task adjustment. Researchers have developed methodologies like CFL-HC [35]and HeteroSAg [36]to handle the varying computational capabilities of edge devices. These approaches effectively mitigate the impact of device heterogeneity, ensuring optimal performance and efficiency in federated learning settings.

In federated learning, addressing device heterogeneity requires the allocation of suitable tasks to edge devices and the selection of appropriate clients based on their constraints. Fine-tuning training tasks and client selection strategies are crucial for maximizing overall efficiency while considering factors like fairness, privacy, and data diversity. To tackle the challenges posed by device heterogeneity and ensure optimal performance and efficiency in federated learning settings, researchers have developed methods such as FedSAE, CFL-HC, and HeteroSAg. These methodologies effectively handle the varying computational capabilities of edge devices. Additionally, reinforcement learning models show promise in addressing the issue of non-IID data.

More details of the contributions and limitations of the existing Heterogeneous FL method can be found in the supplemental material. Given the evolving nature of this field, it is essential to establish widely ac-knowledged benchmarking and evaluation frameworks for heterogeneous scenarios with different complexities.

4 Heterogeneous Federal Learning Evaluation Methods

The concept of collaborative learning was first introduced by McMahan et al. [1]. In this rapidly evolving field, it is crucial to establish widely recognized benchmarking and evaluation frameworks for different scenarios with varying complexities. Empirical evaluation plays a vital role in examining a verifiable federated learning approach in simulated or real-world error-prone environments. It allows for a comprehensive exploration of its effectiveness in complex computational landscapes while ensuring the reliability of the findings.

Model Performance and Communication Overhead. The evaluation of federated learning methods takes into account precision, convergence velocity, and factors such as client heterogeneity and disparate data distributions [37]. It is crucial to strike a balance among precision, communication overhead, and model performance when assessing these methods. Researchers commonly use metrics such as accuracy, precision, recall, F1 score, and convergence velocity to measure the effectiveness of federated learning approaches. By analyzing both model performance and communication overhead, potential areas for optimization can be identified.

Robustness. Robustness is an essential metric for assessing the resilience of federated learning methods against adversarial scenarios [38,39]. It ensures that the model maintains its performance and accuracy in the federated learning environment. Evaluation techniques commonly include adversarial attacks such as model inversion, membership inference, and data contamination. Metrics such as accuracy degradation, model divergence, and anomaly detection are used to quantify robustness.

Privacy Protection. Privacy protection is an important consideration when evaluating the effectiveness of HFL methods. In federated learning, participants often have sensitive data, so ensuring the security of this data is paramount. Researchers evaluate the effectiveness of methods in preserving individual privacy using metrics such as differential privacy [40], information entropy, and data aggregation. These metrics allow for quantifying the level of privacy protection provided. By enhancing privacy protection measures, researchers aim to ensure data security and privacy during the federated learning process.

Customer Contribution. Analyzing client contributions is a crucial aspect of federated learning. It involves quantifying individual contributions by considering factors such as the quality and quantity of data, computational capabilities, and reliability. Metrics like data quality, data quantity, and computational resources are used to assess client contributions. Understanding the varying degrees of contribution is important for optimizing the federated learning process, improving model performance, and addressing data heterogeneity. For instance, FedCav [41]introduces an algorithm for model aggregation that takes into account client contributions in the presence of heterogeneous data.

5 Future Directions

Our empirical investigation unequivocally demonstrates the burgeoning prominence of HFL research. Nonetheless, many difficulties persist, necessitating their resolution to empower this technology to confront the difficulties encountered in real-world applications. We will look over the next steps for future inquiry to enhance the efficacy of addressing heterogeneous predicaments within forthcoming Federated Learning systems.

5.1 Privacy Protection

In HFL, participants often have sensitive personal data, so privacy protection measures are necessary. Future research should prioritize the development of efficient and secure privacy-preserving mechanisms to ensure participants have control over their privacy while sharing data. Differential privacy offers a mathematical guarantee that statistical results can be publicly released while safeguarding individual privacy [46]. By combining differential privacy with federated learning, it becomes possible to prevent the disclosure of sensitive information during model training and aggregation. Future research should focus on improving differential privacy algorithms and mechanisms that can accommodate diverse data types and privacy requirements in HFL. Another important research direction is investigating the use of Secure Multi-Party Computation in the context of federated learning [47]. However, it is important to note that these solutions may need to be adapted to account for system heterogeneity.

5.2 Improving Communication Efficiency

Communication plays a pivotal role in coordinating the collaborative learning process among heterogeneous participants, but it often incurs substantial costs in terms of bandwidth, latency, and energy consumption [42,43,45]. To enhance communication efficiency in HFL, researchers can explore the following aspects. (1)Integration of differential privacy techniques: By incorporating differential privacy techniques, the amount of information exchanged during the federated learning process can be effectively reduced. (2)Gradient compression techniques: These techniques aim to minimize the size of gradients communicated during federated learning, thereby reducing the communication overhead. (3)Leveraging edge computing capabilities and enabling local model updates: By utilizing the computational capabilities of edge devices and facilitating local model updates, the dependency on frequent communication with the central server can be decreased, leading to improved communication efficiency. (4)Knowledge transfer techniques: Exploring techniques that enable knowledge transfer among clients can significantly reduce the extensive communication requirements. Methods such as transfer learning, model personalization, and parameter sharing facilitate the transfer of learned knowledge from high-resource clients to low-resource clients, thereby mitigating overall communication needs.

In conclusion, enhancing communication efficiency is a critical area for future research in HFL. By employing techniques such as differential privacy, gradient compression, edge computing, and federated learning with knowledge transfer, we can effectively reduce communication overhead and enhance the scalability and efficiency of federated learning in heterogeneous settings.

5.3 Federated Fairness

Federal equity is a crucial consideration in the design and implementation of federated learning systems.The presence of diverse and distributed data sources among heterogeneous participants introduces biases and inequalities, and thus effective mitigation is required urgently [6]. Future research should prioritize the development of strong frameworks and algorithms that actively promote fairness, equality, and nondiscrimination in federated learning. One way to enhance fairness in federated learning is to focus on privacy-preserving methods that safeguard sensitive data and prevent unauthorized access or misuse [44]. By exploring privacy-enhancing technologies, we can facilitate collaborative learning while respecting individual privacy rights. This not only mitigates the risk of biased model updates but also fosters fairness in the process of aggregating data. Additionally, it is crucial to design federated learning algorithms explicitly to tackle the challenges posed by data heterogeneity and fairness requirements [45]. Traditional federated learning methods may unintentionally favor participants with more extensive or representative data, resulting in biased models and persistent inequality. To address this, future research should concentrate on innovative techniques such as sample weighting, domain adaptation, and model regularization. These approaches effectively account for data heterogeneity and ensure fairness throughout the model training and aggregation processes.

In conclusion, it is crucial to prioritize addressing equity at the federal level in federated learning. Researchers can promote fairness, equality, and nondiscrimination in federated learning systems by developing methods that protect privacy, exploring technologies that enhance privacy, and designing algorithms that explicitly address the diversity of data and fairness.

5.4 Uniform Benchmarks

The growing fascination with HFL is evident based on the results of our recent survey. However, as we delve further into this domain, numerous challenges arise that require immediate attention to make this technology suitable for practical applications. A crucial aspect for future research directions in addressing heterogeneity in FL systems revolves around the establishment of standardized benchmarks.

Improved Datasets. To accurately represent the diverse nature of real-world federated learning scenarios, it is crucial to develop comprehensive and realistic datasets. Improving datasets is a key area for future advancements in the field of heterogeneous federated learning. Researchers should focus on different aspects of dataset construction, including being aware of heterogeneity, using representative data sampling techniques, assessing and enhancing data quality, generating privacy-preserving datasets, creating benchmark datasets, and incorporating real-world data. By addressing the challenges associated with data heterogeneity using these strategies, researchers can enhance the performance and effectiveness of federated learning models in diverse settings. The availability of these improved datasets will enable more realistic and impactful research in the field of HFL.

Enhanced Evaluation Metrics. Establishing clear and consistent evaluation metrics is crucial for effectively measuring the performance of Horizontal Federated Learning. It is essential to advance the field by developing enhanced evaluation metrics that can provide a comprehensive understanding of the strengths and limitations of federated learning systems. A key focus of future research should be on expanding existing models such as FedEval [48]. The objective should be to create metrics that consider heterogeneity awareness, privacy preservation, fairness orientation, robustness emphasis, resource efficiency, and real-world performance. These enhanced evaluation metrics will drive progress in the field and contribute to the development of more effective and equitable federated learning systems.

6 Conclusion

This paper aims to provide a comprehensive definition and analysis of HFL. It categorizes HFL into four types of heterogeneity: data, model, device, and communication, based on the underlying causes of heterogeneity in federated learning. The study offers a meticulous examination of potential solutions to address these challenges, ultimately enhancing the reader's comprehension of the impact of heterogeneity on federated learning. Furthermore, it succinctly summarizes commonly employed performance evaluation methods and proposes future directions for the development of the HFL framework. These insightful discussions hold significant value in contributing to the advancement of the HFL community. HFL presents itself as an engaging research avenue, necessitating collaborative efforts from the machine learning, systems, and data privacy communities (Tables 1, 2, 3 and 4).

Appendix

Table 1. Heterogeneous data methods

Methods	Key Contributions	Limitations
Zhang et al. [14]	FedlC solves the problem of skewed label distribution in federated learning by calibrating logits and introducing label boundaries	The effectiveness in dealing with extreme labeling distribution skewness still needs further research and improvement.
Luo et al. [15]	DFL solves the problem of uneven attribute distribution on the performance and convergence stability of federated learning	Challenges remain in dealing with complex relationships between domain-specific and cross-invariant attributes.
Yoon et al. [17]	FedMix for improving the performance of federated learning with non-independent Identically distributed Data and Addressing Privacy Preservation	Collecting local data distributions may bring potential information leakage.
Duan et al. [18]	Astraea for Improving Classification Accuracy in Mobile Deep Learning Applications	Disclosure of local data distribution during upload may inadvertently expose vulnerabilities and make it susceptible to malicious intrusion

Table 2. Heterogeneous model methods

Methods	Key Contributions	Limitations
Fallah et al. [19]	MAML uses a personalized version of joint averaging algorithm and evaluates its performance against gradient specification of the non-convexloss function	verlooking other potential approaches or techniques that could enhance personalization in federated learning.
Wang et al. [20]	VFKF proposes a vertical federated knowledge transfer mechanism for feature enrichment in cross-party machine learning systems	The scalability and applicability of vertical federated learning in different scenarlos are not apparent.
Le et al. [21]	FedLKD effectively addresses the statistical heterogeneity challenge by leveraging knowledge istillation between global and local models	Its effectiveness and privacy preservation may vary depending on the specific characteristics of the dataset and the selection of proxy data.
Yu et al. [22]	They alleviate the issue of overfitting in personalized updates by augmenting the coherence of logits between the global and local models	The exploitation of logits may engender inadequate assimilation of local information

Table 3. Heterogeneous communication methods

Methods	Key Contributions	Limitations
Hou et al. [23]	FedChain combines the advantages of local and global update methods infederated learning,achieving fast convergence while leveraging data similarity	Devices may connect slowly, rendering them expensive and unreliable communicate.
Lu et al. [24]	CMFL avoids transmitting irrelevant updates to the server by measuring the consistency of local updates with global updates	Difficult to handle highly heterogeneous or unreliable network environments.
Li et al. [25]	They presents a concise and efficient federated learning framework fortraining deep neural networks on resource-constrained mobile device	Lack of in-depth analysis of potential privacy or security implications of proposed frameworks.
Luo et al. [26]	Fedskel enables federated learning for efficient computation and efficient communication on edge devices by updating the essential parts of the mode	Limited scalability analysis of the system with privacy or security concerns

Table 4. Heterogeneous device methods

Methods	Key Contributions	Limitations
Abdellatif et al. [32]	Allow massive amounts of data generated by IoT devices to train deep learning models	Failure to minimize communication overhead in hierarchical joint learning.
Li et al. [33]	FedSAE effectively addresses systems heterogeneity by adjusting the training tasks of devices and actively selecting participants	Refining the allocation of workloads in accordance with the client straining history may encounter temporal delays.
Wang et al. [34]	Favor dynamically curates the optimal cohort of clients to engage in iterations of federated learning	Raining reinforcement learning models necessitates a substantial volume of data

The four tables above summarize the solutions to federated learning data heterogeneity, model heterogeneity, communication heterogeneity, and device heterogeneity, and analyze the main contributions and limitations of each approach. These valuable discussions can contribute to the high-quality development of the heterogeneous federated learning community.

References

1. McMahan, B., et al.: Communication-efficient learning of deep networks from decentralized data. In: Artificial Intelligence and Statistics. PMLR (2017)
2. Yang, Q., et al.: Federated machine learning: Concept and applications. ACM Trans. Intell. Syst. Technol. (TIST) **10**(2), 1–19 (2019)
3. Dayan, I., et al.: Federated learning for predicting clinical outcomes in patients with COVID-19. Nat. Med. **27**(10), 1735–1743 (2021)
4. Wu, C., et al.: FedGNN: federated graph neural network for a privacy-preserving recommendation. arXiv preprint arXiv:2102.04925 (2021)
5. Suzumura, T., et al.: Towards federated graph learning for collaborative financial crimes detection. arXiv preprint arXiv:1909.12946 (2019)

6. Usynin, D., et al.: Adversarial interference and its mitigations in privacy-preserving collaborative machine learning. Nat. Mach. Intell. **3**(9), 749–758 (2021)
7. Li, T., et al.: Federated learning: challenges, methods, and future directions. IEEE Signal Process. Mag. **37**(3), 50–60 (2020)
8. Kairouz, P., et al.: Advances and open problems in federated learning. Found. Trends® Mach. Learn. **14**(1-2), 1–210 (2021)
9. Wahab, O.A., et al.: Federated machine learning: survey, multi-level classification, desirable criteria and future directions in communication and networking systems. IEEE Commun. Surv. Tutor. **23**(2), 1342–1397 (2021)
10. Tan, A.Z., et al.: Towards personalized federated learning. IEEE Trans. Neural Netw. Learn. Syst. **34**, 9587–9603 (2022)
11. Abdelmoniem, A.M., et al.: A comprehensive empirical study of heterogeneity in federated learning. IEEE Internet Things J. **10**, 14071–14083 (2023)
12. Gao, D., Yao, X., Yang, Q.: A survey on heterogeneous federated learning. arXiv preprint arXiv:2210.04505 (2022)
13. Ye, M., et al.: Heterogeneous Federated Learning: State-of-the-art and Research Challenges. arXiv preprint arXiv:2307.10616 (2023)
14. Zhang, J., et al.: Federated learning with label distribution skew via logits calibration. In: International Conference on Machine Learning. PMLR (2022)
15. Luo, Z., et al.: Disentangled federated learning for tackling attributes skew via invariant aggregation and diversity transferring. arXiv preprint arXiv:2206.06818 (2022)
16. Zhu, H., et al.: Federated learning on non-IID data: a survey. Neurocomputing **465**, 371–390 (2021)
17. Yoon, T., et al.: Fedmix: Approximation of mixup under mean augmented federated learning. arXiv preprint arXiv:2107.00233 (2021)
18. Duan, M., et al.: Astraea: self-balancing federated learning for improving classification accuracy of mobile deep learning applications. In: 2019 IEEE 37th International Conference on Computer Design (ICCD). IEEE (2019)
19. Fallah, A., Mokhtari, A., Ozdaglar, A.: Personalized federated learning with theoretical guarantees: a model-agnostic meta-learning approach. Adv. Neural. Inf. Process. Syst. **33**, 3557–3568 (2020)
20. Wang, L., Huang, C., Han, X.: Vertical federated knowledge transfer via representation distillation. In: FL-IJCAI Workshop (2022)
21. Le, H.Q., et al.: Layer-wise Knowledge Distillation for Cross-Device Federated Learning. In: 2023 International Conference on Information Networking (ICOIN). IEEE (2023)
22. Yu, T., Bagdasaryan, E., Shmatikov, V.: Salvaging federated learning by local adaptation. arXiv preprint arXiv:2002.04758 (2020)
23. Hou, C., et al.: FeDChain: Chained algorithms for near-optimal communication cost in federated learning. arXiv preprint arXiv:2108.06869 (2021)
24. Luping, W., Wei, W., Bo, L.: CMFL: mitigating communication overhead for federated learning. In: 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS). IEEE (2019)
25. Li, X., et al.: A unified federated DNNs framework for heterogeneous mobile devices. IEEE Internet Things J. **9**(3), 1737–1748 (2021)
26. Luo, J., et al.: Fedskel: efficient federated learning on heterogeneous systems with skeleton gradients update. In: Proceedings of the 30th ACM International Conference on Information & Knowledge Management (2021)
27. Yao, X., et al.: Federated learning with additional mechanisms on clients to reduce communication costs. arXiv preprint arXiv:1908.05891 (2019)

28. Zaccone, R., et al.: Speeding up heterogeneous federated learning with sequentially trained superclients. In: 2022 26th International Conference on Pattern Recognition (ICPR). IEEE (2022)
29. Gao, Z., et al.: FedSeC: a robust differential private federated learning framework in heterogeneous networks. In: 2022 IEEE Wireless Communications and Networking Conference (WCNC). IEEE (2022)
30. Ma, Q., et al.: FedSA: a semi-asynchronous federated learning mechanism in heterogeneous edge computing. IEEE J. Sel. Areas Commun. **39**(12), 3654–3672 (2021)
31. Chan, Y.H., Edith, C.H.N.: Fedhe: heterogeneous models and communication-efficient federated learning. In: 2021 17th International Conference on Mobility, Sensing and Networking (MSN). IEEE (2021)
32. Abdellatif, A.A., et al.: Communication-efficient hierarchical federated learning for IoT heterogeneous systems with imbalanced data. Future Gener. Comput. Syst. **128**, 406–419 (2022)
33. Li, L., et al.: FedSAE: a novel self-adaptive federated learning framework in heterogeneous systems. In: 2021 International Joint Conference on Neural Networks (IJCNN). IEEE (2021)
34. Wang, H., et al.: Optimizing federated learning on non-IID data with reinforcement learning. In: IEEE INFOCOM 2020-IEEE Conference on Computer Communications. IEEE (2020)
35. Wang, D., et al.: CFL-HC: a coded federated learning framework for heterogeneous computing scenarios. In: 2021 IEEE Global Communications Conference (GLOBECOM). IEEE (2021)
36. Elkordy, A.R., Salman Avestimehr, A.: Heterosag: secure aggregation with heterogeneous quantization in federated learning. IEEE Trans. Commun. **70**(4), 2372–2386 (2022)
37. Li, Y., et al.: FedH2L: Federated learning with model and statistical heterogeneity. arXiv preprint arXiv:2101.11296 (2021)
38. Takahashi, H., Liu, J., Liu, Y.: Breaching FedMD: image recovery via paired-logits inversion attack. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2023)
39. Liu, Y., et al.: A secure federated learning framework for 5G networks. IEEE Wirel. Commun. **27**(4), 24–31 (2020)
40. Ding, J., et al.: Differentially private and communication efficient collaborative learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35. No. 8. (2021)
41. Zeng, H., et al.: FedCAV: contribution-aware model aggregation on distributed heterogeneous data in federated learning. In: Proceedings of the 50th International Conference on Parallel Processing (2021)
42. Bibikar, S., et al.: Federated dynamic sparse training: computing less, communicating less, yet learning better. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36. No. 6 (2022)
43. Hardt, M., Price, E., Srebro, N.: Equality of opportunity in supervised learning. In: Advances in Neural Information Processing Systems, vol. 29 (2016)
44. Lyu, L., et al.: Towards fair and privacy-preserving federated deep models. IEEE Trans. Parallel Distrib. Syst. **31**(11), 2524–2541 (2020)
45. Gálvez, B.R., et al.: Enforcing fairness in private federated learning via the modified method of differential multipliers. In: NeurIPS 2021 Workshop Privacy in Machine Learning (2021)
46. Sun, L., Lyu, L.: Federated model distillation with noise-free differential privacy. arXiv preprint arXiv:2009.05537 (2020)

47. Bonawitz, K., et al.: Practical secure aggregation for privacy-preserving machine learning. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (2017)
48. Chai, D., et al.: FedEval: A Holistic Evaluation Framework for Federated Learning. arXiv preprint arXiv:2011.09655 (2020)