

Review

Communication Efficiency and Non-Independent and Identically Distributed Data Challenge in Federated Learning: A Systematic Mapping Study

Basmah Alotaibi ^{1,2} , Fakhri Alam Khan ^{1,3,4,*}  and Sajjad Mahmood ^{1,3} 

- ¹ Department of Information and Computer Science, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia; g202008360@kfupm.edu.sa (B.A.); smahmood@kfupm.edu.sa (S.M.)
- ² Department of Computer Science, College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 13318, Saudi Arabia
- ³ Interdisciplinary Research Centre for Intelligent Secure Systems, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia
- ⁴ SDAIA-KFUPM Joint Research Center for Artificial Intelligence, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia
- * Correspondence: fakhri.khan@kfupm.edu.sa

Abstract: Federated learning has emerged as a promising approach for collaborative model training across distributed devices. Federated learning faces challenges such as Non-Independent and Identically Distributed (non-IID) data and communication challenges. This study aims to provide in-depth knowledge in the federated learning environment by identifying the most used techniques for overcoming non-IID data challenges and techniques that provide communication-efficient solutions in federated learning. The study highlights the most used non-IID data types, learning models, and datasets in federated learning. A systematic mapping study was performed using six digital libraries, and 193 studies were identified and analyzed after the inclusion and exclusion criteria were applied. We identified that enhancing the aggregation method and clustering are the most widely used techniques for non-IID data problems (used in 18% and 16% of the selected studies), and a quantization technique was the most common technique in studies that provide communication-efficient solutions in federated learning (used in 27% and 15% of the selected studies). Additionally, our work shows that label distribution skew is the most used case to simulate a non-IID environment, specifically, the quantity label imbalance. The supervised learning model CNN model is the most commonly used learning model, and the image datasets MNIST and Cifar-10 are the most widely used datasets when evaluating the proposed approaches. Furthermore, we believe the research community needs to consider the client's limited resources and the importance of their updates when addressing non-IID and communication challenges to prevent the loss of valuable and unique information. The outcome of this systematic study will benefit federated learning users, researchers, and providers.

Keywords: communication-efficient; federated learning; non-IID data; systematic mapping study



Citation: Alotaibi, B.; Khan, F.A.; Mahmood, S. Communication Efficiency and Non-Independent and Identically Distributed Data Challenge in Federated Learning: A Systematic Mapping Study. *Appl. Sci.* **2024**, *14*, 2720. <https://doi.org/10.3390/app14072720>

Received: 2 March 2024

Revised: 21 March 2024

Accepted: 22 March 2024

Published: 24 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Technology development in this era leads to a growth in the amount of data produced by devices. When used as training data, these data can provide intelligence to the end devices. Traditional machine learning approaches are centralized, requiring all data to be collected and sent to the central server for processing. The central server will collect the data to train a machine learning model to enable intelligence at the end devices. This approach raises concerns due to the amount of data traveling through the Internet and the possibility of privacy leaks. Federated learning offers a solution to address these concerns by allowing for the sharing of knowledge through a trained local model instead of sharing data with a central server [1–4].

Federated learning (FL) is a distributed collaborative artificial intelligence approach that trains a local model and shares it with the central server. Instead of sharing data with the central server, the FL approach shares the locally trained models, which the central server aggregates to build a global model, as shown in Figure 1 [1].

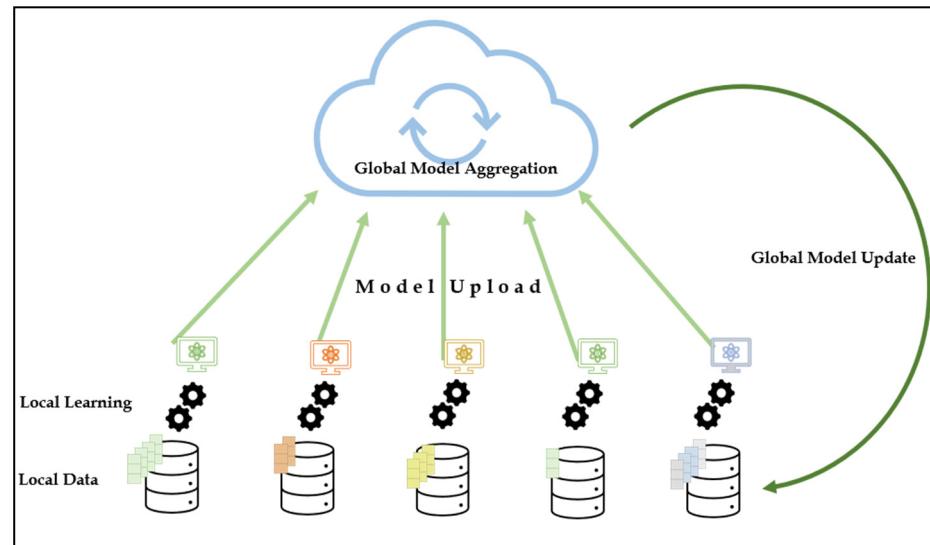


Figure 1. General architecture of federated learning.

FL technology provides many benefits compared to traditional learning approaches. It makes more efficient use of network bandwidth and preserves data privacy since there is no need to transfer raw data to the server. Additionally, FL can enhance the quality of the global model by leveraging computation resources and diverse datasets on clients' devices [1,5]. With these advantages that FL provides, it can be used in different applications such as healthcare, the Internet of Things (IoT), transportation, and mobile applications (for example, next-word prediction) [6,7].

However, FL faces some challenges due to its decentralized approach. The assumption of the Independent and Identically Distributed (IID) data used in machine learning algorithms is not applicable in FL. Usually, the data in FL are heterogeneous data caused by Non-Independent and Identically Distributed (non-IID) data. For that reason, FL encounters the challenge of data heterogeneity (non-IID data challenge) [5,8,9].

The nature of data in FL differs from the centralized approach since the training data depend on the device usage and location and can vary between clients. The data in each device can differ in quantity or class distribution. In some cases, clients may have more data than other clients, while in other cases, some clients may have data belonging to a specific class (label). These variations affect the FL process and model performance. In FL, the central server randomly selects a subset of clients to perform local training and receives the locally trained models for global aggregation. Clients attempt to minimize their loss function during the local model training based on their local data. However, if the selected clients' local data distribution varies, the obtained local models may differ significantly, resulting in a divergence of the global model from the optimal one. This inconsistency between the local and global models occurs because the local models fit the clients' data distribution and do not reflect the overall data distribution. Therefore, training local models using non-IID data affects global model performance and convergence speed. Figure 2 shows an example that illustrates the impact of non-IID data on the global model and how the model diverges from the optimal case; while the global model built using IID data is close to the optimal case, in the non-IID data case, the model will need more communication rounds to converge and reach the same accuracy as models trained using IID data. Therefore, the non-IID data problem must be solved to enhance the model's performance [5,10,11].

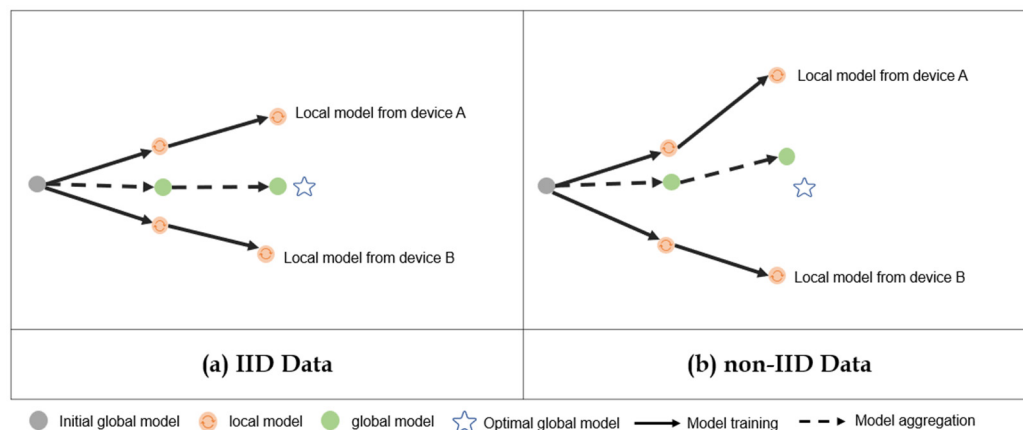


Figure 2. Example showing the model divergence for federated learning with IID and non-IID data.

Another challenge in FL is communication; the communication between clients and the central server is considered a bottleneck due to the issue of limited bandwidth in the network and the number of communication rounds as clients in FL train local models and share them with the central server in repetitive rounds [12,13]. Since communication rounds between clients and the central server can be costly [13], reducing the number of rounds or the number of parameters exchanged can lead to efficient communication in FL [14–16]. These challenges are important in the FL environment, and many studies in the literature have been proposed to address them. There are a few mapping studies on the federated learning environment; for example, the study in [17] covers the motivation for using FL, and the study in [18] focuses on applying federated learning on energy-constrained IoT devices. However, there are no comprehensive studies that aim to provide a systematic mapping study to cover the techniques utilized to provide communication efficiency and overcome the non-IID data challenge. Therefore, we present a systematic mapping of the literature to summarize and analyze the research carried out on these challenges.

Mapping studies are review studies considered an evidence-based technique that helps to analyze a research topic systematically. This method summarizes the research area and identifies the amount and kind of research and available results [19]. This work aims to conduct a systematic mapping study on the challenges mentioned earlier to provide in-depth knowledge in the FL environment. Specifically, we will examine articles that address non-IID data problems, improve communication efficiency in FL, or tackle both challenges simultaneously. This work also aims to identify techniques widely used to overcome these challenges in the FL environment. The non-IID data skewness utilized in the studies aims to overcome the non-IID data challenge. Additionally, the work covers the learning models and datasets most widely used to evaluate the proposed techniques. Furthermore, the work highlights the publication venues and years of these articles. The main contributions of this study are as follows:

- Providing in-depth knowledge about the techniques that have been proposed to overcome the non-IID data challenge in FL.
- Offering a deep understanding of the techniques that have been proposed to provide efficient communication in federated learning.
- Identifying the widely used learning models and datasets and associating the respective learning models with the utilized datasets.
- Highlighting promising research directions that can open up new opportunities for future studies.

The rest of this paper is organized as follows: Section 2 covers the preliminaries and related work. Section 3 provides the research methodology of this study. In Section 4, the results of this systematic study and future research directions are discussed. Section 5 provides the conclusion of the work.

2. Preliminaries and Related Work

2.1. Preliminaries

2.1.1. Federated Learning

In federated learning, the end devices known as clients hold the data and train a local model. The clients receive the global model from the central server and train their local model using the received global model and their data. Clients share their local model with the central server. Once the central server receives the local models from the participating clients, it aggregates them using an aggregation method to create a new global model [5,20]. The general process of how FL works is shown in the following steps:

1. The central server decides which devices are participating in training the model at this round.
2. The selected participating devices receive the global model from the central server.
3. The devices train a local model using their dataset and the received global model.
4. Each device uploads the trained local model to the central server for aggregation.
5. The received local models are aggregated to create the new global model.
6. The steps are repeated until the target performance is accomplished (the target can be specific accuracy) or the deadline is reached.

In 2016, McMahan et al. [21] introduced the concept of federated learning for the first time and proposed an aggregation algorithm called Federated Average (FedAvg). FedAvg is a weighted averaging scheme that weights the client's local model based on their dataset size.

$$w_{t+1} = \sum_{k \in S_t} \frac{n_k}{n} w_{t+1}^k \quad (1)$$

where w_{t+1} is the new global model, w_{t+1}^k is the local model received from client k , n_k is the client k data size, and n is the total data size for all clients participating in this round, and S_t is the set of clients participating in the training process at round t .

FL can be classified into three categories based on the data partitions, according to [22]. The three categories are (1) horizontal FL (HFL), (2) vertical FL (VFL), and (3) federated transfer learning (FTL). In HFL, the clients have local datasets with the same features, but the data samples differ. HFL is like having a large dataset divided horizontally between clients. The sample space will overlap in VFL, while the feature space overlap is minimal. This occurs when two datasets are created using the same samples but differ in extraction [23]. VFL is often related to an enterprise setting where the number of clients participating is much smaller than HFL, but privacy matters are paramount [24,25]. In FTL, devices have different samples, and the extracted features differ. This means the overlap between the sample and feature spaces between the devices is rare. Transfer learning is used to enhance the learning process. An initial model is trained using the overlapped samples and all features. The initial model is used as a baseline with the remaining samples (non-overlap) and the available features for the sample [22].

Furthermore, FL can be classified into cross-device FL and cross-silo FL based on the participating clients and training scale. The number of devices participating in the cross-device setting is more than in the cross-silo setting. In the cross-device setting, devices are typically small and have limited samples compared to the cross-silo setting, where clients are usually companies or organizations that hold large datasets [26,27].

2.1.2. Federated Learning Application

Federated learning has emerged as a promising approach for collaborative model training across distributed devices while preserving data privacy; it can be utilized in different applications such as healthcare, smart city, smart transport, and finance industry applications.

For example, in the smart healthcare industry, sharing patient records with a central server or cloud is necessary to develop intelligent approaches, e.g., intelligent imaging for disease detection. However, patient records are sensitive, and simply removing personal information is not enough to protect their privacy. This is especially true in complex

healthcare settings where multiple parties, such as hospitals and insurance companies, have access to healthcare databases for data analysis and processing. To address this issue, using FL can provide intelligence and knowledge from different patient records while protecting patient privacy, to enhance the healthcare system to provide a robust predictive model for disease diagnosis, healthcare management, etc. [1,28,29].

On the other hand, smart devices deployed in smart cities are assisting city officials in improving the efficiency of city operations while also enhancing the quality of life for residents by ensuring the seamless delivery of food, water, and energy to end-users. To facilitate smart cities, machine learning techniques have been widely adopted to process real-time big data from sensors, devices, and human activities. However, due to privacy concerns and the massive traffic generated, a centralized learning approach is not scalable. Therefore, FL can be leveraged in this field to enable decentralized smart city applications with high privacy and minimal communication delays. For example, it can be used in a smart grid system to learn power consumption patterns without revealing individual power traces; this can help create an interconnected and intelligent energy exchange network in the city [1,29].

Intelligent transport systems aim to ensure safe and efficient traffic flow by using various technologies to monitor and assess their performance. However, traditional intelligent transportation systems share data in untrusted environments, which may raise privacy concerns. Therefore, FL can be utilized in transport applications to enable intelligence without the need to share data; FL can be used in intelligent transport systems to enable traffic prediction and manage resource strategies for vehicle-to-vehicle networks [1,30].

Collaboration among various financial institutions has become a growing trend in the financial and insurance industry. Federated learning is a technology that financial institutions can utilize for risk management, fraud detection, marketing, and other purposes. The use of FL in this sector facilitates collaboration between different financial institutions without the need to share their clients' information, thus helping them to develop various financial task models, such as risk assessment machine learning models [31,32].

2.1.3. Non-IID Data in Federated Learning

The data used to train the local model on FL could be non-IID data. This is due to the heterogeneity of datasets among devices, since the local datasets depend on the device's performance and usage, resulting in non-IID data [33,34].

Each data sample has features and labels that identify them. Assuming that we can represent all features as F , containing several features (f_1, f_2, f_3, \dots) , each feature f_i has a domain of different values that identify the sample. The data label can be represented as L that contains a value from its different values (classes). So, we can say the device (i) has samples where each sample can be represented as (F, L) . The device samples follow a local distribution $P_i(F, L) = P_i(F|L) P_i(L) = P_i(L|F) P_i(F)$. For that, when we say the data are non-IID, we mean that P is different from device to device, with the difference being caused by differences in the features, labels, or both [33,35]. We can classify the types of non-IID data as shown in Figure 3.

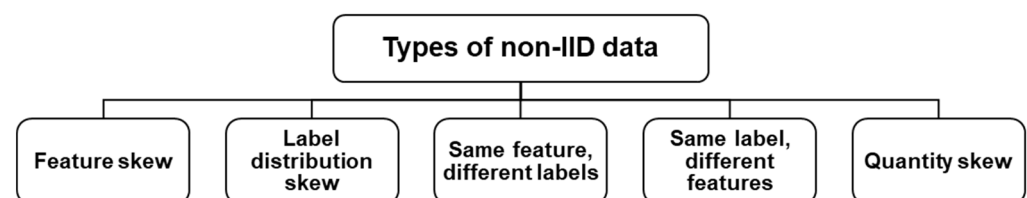


Figure 3. Classification of non-IID data.

- **Feature skew:** Feature skew indicates that the features differ among devices; this can be described as the $P_i(F)$ being different while $P_i(L|F)$ is the same. The features can be non-overlapped between devices, partially overlapped, or fully overlapped. In

non-overlapping feature skew, the different devices have different features; this case is similar to vertical federated learning; images with different angles are an example. While in partial overlapping feature skew, some features can be overlapped. Full overlapping feature skew is a case similar to horizontal federated learning; an example is the case of having two datasets for the same numbers (digits), one is written in a bold line while the other is written with a thin line [33–35].

- Label distribution skew: Label distribution skew indicates that the devices have different labels; this can be described as the $P_i(L)$ being different while $P_i(F|L)$ is the same. This skew can happen when the device tends to have local data with the same labels (for example, it can be caused by location variations between devices) or labels from some classes more than others. Label skew is defined in different ways in two studies. The study in [21] introduces the quantity label imbalance, and the study in [36] introduces the distribution label imbalance. Generally, the amount of data belonging to the same class is not equal and varies between devices.
 - Quantity label imbalance: This situation occurs when the devices have a predetermined number of labels they can own. For example, all devices have data from two class labels only. If we take device (i) and device (j), the labels in device (i) can be from class (c_1, c_2), while those in device (j) can be from class (c_3, c_4). This kind of distribution was first introduced in Federated Average (FedAvg) experiments. In this case, the smaller the label quantity, the stronger the label imbalance [21,33].
 - Distribution label imbalance: In this skew, each device has a proportion of the samples from each label class that follows Dirichlet distribution $Dir()$. The portion of the data that belongs to a specific class (c) is distributed on device (i) with a probability $p_c \sim Dir_i(\beta)$, where β is the concentration parameter that determines the imbalance level; a higher value indicates a high imbalance partition [33,35,36].
- Same feature, different labels: The case of the same feature with different labels implies that the distribution of $P_i(L|F)$ is different but $P_i(F)$ is the same. In this case, the same features indicate different classes (labels) on different devices; the data label for the same feature can be c_1 on the first device and c_2 on the other device. This could depend on the user preference; for example, in the same weather condition, some people may refer to a rainy day as good weather, while other people refer to the same rainy weather as a bad day [33,34].
- Same label, different features: The case of the same label with different features implies that the distribution of $P_i(F|L)$ is different but $P_i(L)$ is the same. Different features on different devices could belong to the same class. For example, the first device has images of a school building on a sunny day, and the second device has images of a school building on a rainy day; both can belong to the same class (school buildings), but they have different features [34]. For example, the first device has images of a residential building, and the second device has images of a factory building; both belong to the same class (buildings) but have different features.
- Quantity skew: The amount of local data varies between devices. For example, the number of training data for the first device equals 1000, while the number of training data for the second device equals 30. This skew can happen with any of the previously described categories [33–35].

2.2. Related Work

Several surveys have been conducted in the literature illustrating the federated learning concept and its challenges. For instance, the article in [1] provides a comprehensive survey of federated learning for the IoT. The authors cover state-of-the-art federated learning, the federated learning role for IoT services and applications, and federated learning challenges.

In [37], Lim et al. provide a comprehensive survey of the use of federated learning in mobile edge networks. Their work covers the concept and background of federated

learning, federated learning challenges, security, privacy issues, and federated learning applications in mobile edge networks.

Similarly, in [38], Li et al. explained the FL concept and system components, providing a taxonomy for different federated learning aspects. In [22], the work covers federated learning concepts and applications. Some works cover challenges along with the concept of federated learning and its applications [32,39]. In [33], Zhu et al. provided a survey focused on the non-IID data problem challenge in FL. The authors in [40] provide a review article on federated learning in smart cities; they illustrate the advantages and disadvantages of implementing federated learning in smart cities. The article focuses on privacy and security issues of federated learning in smart cities. The articles in [41–43] focus on reviewing blockchain in federated learning, while the article in [44] focuses on FL-enabled 6G technology, specifying its requirements, applications, and current challenges. In [45], Liu et al. focus on the concept of vertical federated learning and its challenges. The study in [17] provides a systematic mapping study highlighting why federated learning has been used and different machine learning pipelines used for federated learning, while a systematic mapping study focused on energy-constrained IoT devices [18].

To the best of our knowledge, our article is the first article that provides a systematic mapping study for federated learning challenges. This work aims to identify the most used techniques to overcome the non-IID data challenge in FL and the most used techniques to provide communication efficiency in federated learning. Furthermore, we studied the works that solve both problems in their work. Also, we provide information about the widely used local models and datasets as these are essential in any federated learning approach. Furthermore, the study also highlights the publication venues, types, and years of these articles.

3. Research Methodology

A systematic mapping study needs to follow a formal guideline when conducted; for that, we follow the guideline provided by Kitchenham and Charters [46] as it is a well-known guideline. All authors contributed during all phases of this study. The authors carefully discussed the paper selection to reduce personal bias; we used an Microsoft Excel 365 to carry out the process and examine the work.

3.1. Research Questions

To accomplish the objective of this research, we addressed the following research questions:

RQ1: Which non-IID type has been mainly addressed when overcoming the non-IID data challenge in federated learning?

RQ2: What are the techniques that are utilized to overcome the non-IID data challenge that federated learning faces?

RQ3: What are the techniques that are utilized to provide communication efficiency (to reduce the communication overhead) in federated learning?

RQ4: What are the learning models utilized in these studies to perform the learning process?

RQ5: What are the datasets utilized in these studies to evaluate the proposed work?

3.2. Search Strategy

The search strategy we followed in this work was as follows:

1. Search terms: We first started our work by identifying the search term and constructing the search string; our search scope was in the federated learning area; we focused on solutions for overcoming the non-IID data problem and on solutions for providing communication efficiency in federated learning. For that, we used the terms shown in Table 1.
2. Search string: The search string used in the search process within the digital library was created by identifying keywords from populations, interventions, and outcomes. The search terms were as follows: “Federated Learning” AND (“non-IID data” OR

- “non IID data” OR “non-I.I.D data” OR “not independent and identically distributed data”) OR (“Communication-efficiency” OR “Communication-efficient” OR “Communication efficiency” OR “Communication efficient”).
- Database: In this work, we used six popular digital databases to perform our search; the databases used are shown in Table 2. The search string was customized to suit each digital library search mechanism.

Table 1. Search terminology.

Term	Alternative Synonyms
Federated learning	---
non-IID data	non IID data, non-I.I.D data, not independent and identically distributed data
Communication-efficiency	Communication-efficient, Communication efficiency, Communication efficient

Table 2. Digital libraries used in our study.

Database	Link
ACM Digital library	https://dl.acm.org/ , accessed on 29 February 2024
IEEE Xplore	https://ieeexplore.ieee.org/ , accessed on 29 February 2024
Science Direct	https://www.sciencedirect.com/ , accessed on 29 February 2024
Springer Link	https://link.springer.com/ , accessed on 29 February 2024
John Wiley Online Library	https://onlinelibrary.wiley.com/ , accessed on 29 February 2024
Web of Science	https://www.webofscience.com/wos/woscc/basic-search , accessed on 29 February 2024

3.3. Study Inclusion Criteria

The search results obtained from the search string were filtered using the following initial selection criteria:

- Conference and journal publications.
- Publication published from 2016 until the end of 2022.
- Publications that include the search string in their title or abstract.
- Publication written in English language.

The search was initiated from 2016 since federated learning was introduced in that year [17]. Thus, we covered the studies that had been proposed after introduction of the FL concept. After the above-mentioned criteria were applied, 1078 publications were extracted from the selected digital libraries. The details for the number of publications extracted from each digital library are shown in Table 3.

Table 3. Primary results from each digital library.

Library	Number of Publications
ACM Digital library	124
IEEE Explore	355
Science Direct	34
Springer Link	165
John Wiley Online Library	3
Web of Science	397

We started filtering the obtained results by removing the duplicated versions of the publications. After that, we started scanning the article titles and abstracts to select relevant articles; when reviewing the articles, we excluded the publications that provided review studies and publications that did not provide any related solutions to our research questions. Furthermore, VFL and FTL were excluded since the federated learning process

differs from the HFL we focused on in this search. We included publications that followed the centralized approach with a central server responsible for aggregating the global model. Publications that aimed to solve the cross-silo architecture were excluded since the devices in this architecture are limited and have more data than the cross-device architecture. Publications that introduced a security solution were also excluded since the security solution may introduce overhead for communication and since we aimed for articles that proposed a solution using vanilla federated learning.

After scanning the title and abstract of the papers using the inclusion and exclusion criteria, we ended up with 362 papers and started the full scan of these papers. In the full paper scan, we excluded the inaccessible papers, i.e., those for which we could not access their entire content, so we ended up with 262 papers that we thoroughly reviewed, resulting in 193 papers passing this stage using the inclusion and exclusion criteria mentioned earlier and classified based on their category. Figure 4 summarizes the detailed numbers for each phase of the filtration process. The studies were then examined to extract useful information for conducting this work, and the studies were classified into three categories. The first category comprises 93 studies that aim to solve the non-IID data challenge. The second category includes 74 studies that aim to provide efficient communication in federated learning. The last category focuses on addressing both challenges and involves 26 studies. The references [47–139] are the non-IID studies, the studies [140–213] are the communication-efficient studies, and the remaining studies [214–239] are the studies that aimed to provide a solution for both categories.

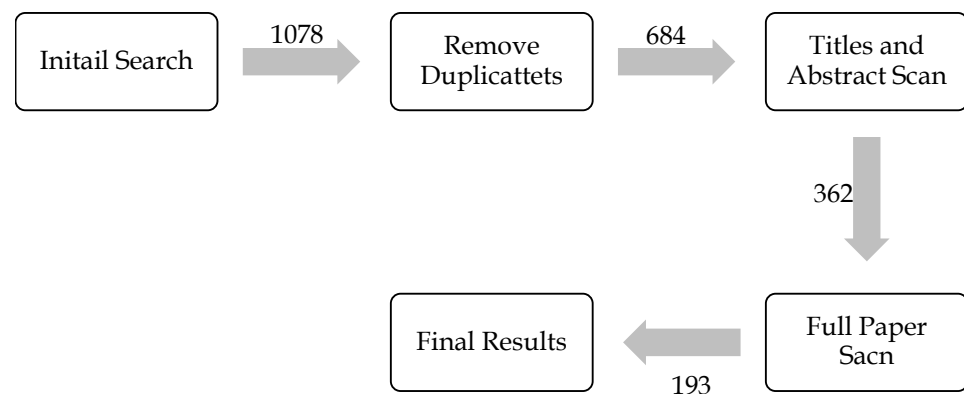


Figure 4. Flow diagram of the selection results.

4. Results and Discussion

This section illustrates the outcomes of the systematic mapping study we conducted. We divide the outcomes of our work into four subsections: the first subsection shows the publication years, source type, and publication venues of the selected study. The second subsection presents the results of studies related to solving the non-IID data problem, studies related to communication efficiency are shown in the third subsection, and the fourth subsection contains the results of studies related to solving both challenges. Furthermore, this section discusses the results and threats to validity.

4.1. Publication Years and Source Types

According to our results in Figure 4, we obtained 193 studies after applying the inclusion and exclusion criteria, indicating that FL gained more interest after its introduction in 2016. Numerous publications have been dedicated to addressing issues related to FL. Figure 5 shows the distribution of the studies over the years. As shown in the figure, the selected studies that offer solutions to the non-IID data problem or provide communication efficiency started in 2019 and increased over the years. The studies aim to solve both challenges that started in 2020 and increased in recent years.

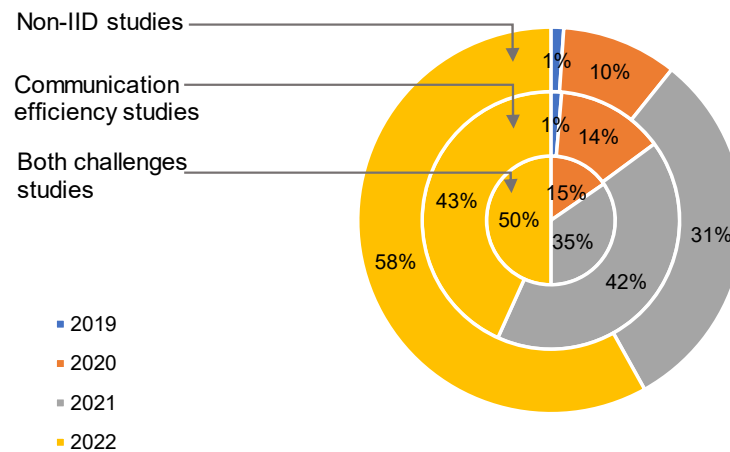


Figure 5. Percentage of selected studies published over the years.

Another research aspect of this work focuses on the selected studies' source types and publication venues. The selected studies have been published in four sources: conferences, journals, workshops, and symposiums, as shown in Figure 6. Most of the studies that addressed the non-IID data challenge were published at conferences (55% of the 93 selected studies). The studies that aim to provide efficient communication in federated learning and the studies that address both challenges were mainly published in journals (49% of the 74 selected studies and 54% of the 26 selected studies, respectively).

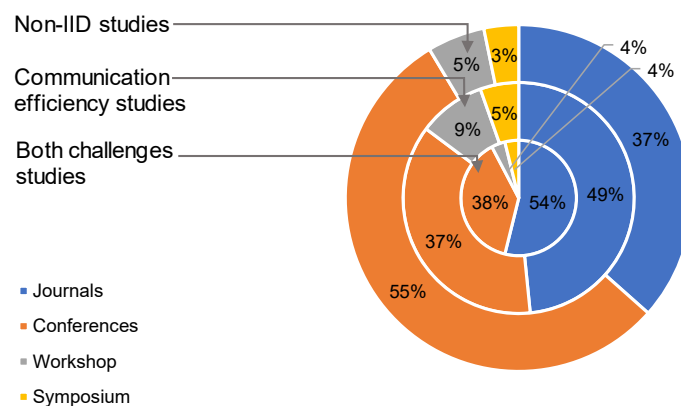


Figure 6. Percentage of selected studies based on source types.

Table 4 shows the most common publication venues that published more than one work of the selected studies that aimed to overcome the non-IID data problem. The table includes the types of publications, the number of studies, and the ratio of published studies to the total number of selected studies. In general, the selected studies were published in 77 publication venues. The leading conference is CVPR while the leading journal is IEEE Transactions on Parallel and Distributed Systems.

Table 5 shows the most common publication venues that published more than one work of the selected studies that aimed to provide communication efficiency in FL. The table includes the types of publications, the number of studies, and the ratio of published studies to the total number of selected studies. In general, the selected studies were published in 60 publication venues. The leading journal is IEEE Internet of Things Journal, with seven publications. We omitted the table for the publication venues for the studies that address both challenges since all venues had only one study published, except for one journal (Journal of Systems Architecture) that published two works.

Table 4. Publication venues with more than one selected study on solving the non-IID data problem.

Publication Venue	Type	No.	%
2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	Conference	3	3.23%
IEEE Transactions on Parallel and Distributed Systems	Journal	3	3.23%
2022 IEEE International Conference on Big Data (Big Data)	Conference	2	2.15%
2022 IEEE International Conference on Data Mining (ICDM)	Conference	2	2.15%
ICC 2020—2020 IEEE International Conference on Communications (ICC)	Conference	2	2.15%
2021 International Joint Conference on Neural Networks (IJCNN)	Conference	2	2.15%
2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys)	Conference	2	2.15%
KDD '21: proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining	Conference	2	2.15%
Machine learning and knowledge discovery in databases	Conference	2	2.15%
Computer Vision—ECCV 2022	Conference	2	2.15%
IEEE Transactions on Wireless Communications	Journal	2	2.15%
IEEE Transactions on Network Science and Engineering	Journal	2	2.15%
Future generation computer systems-the international journal of eScience	Journal	2	2.15%

Table 5. Publication venues with more than one selected study on providing communication efficiency in federated learning.

Publication Venue	Type	No.	%
IEEE Internet of Things Journal	Journal	7	9.46%
IEEE Transactions on Wireless Communications	Journal	3	4.05%
2021 17th International Conference on Mobility, Sensing and Networking (MSN)	Conference	2	2.70%
2021 IEEE 41st International Conference on Distributed Computing Systems (ICDCS)	Conference	2	2.70%
2021 IEEE International Conference on Communications Workshops (ICC Workshops)	Conference	2	2.70%
2022 IEEE Globecom Workshops (GC Wkshps)	Workshop	2	2.70%
GLOBECOM 2022—2022 IEEE Global Communications Conference	Conference	2	2.70%
IEEE Transactions on Network Science and Engineering	Journal	2	2.70%

4.2. Results for Non-IID Data Studies

This subsection illustrates the systematic mapping results of works that have proposed techniques for solving the non-IID data problem. A total of 93 papers passed the selection criteria. Figure 7 shows the non-IID types that have been simulated in the non-IID studies; as shown in the figure, label distribution skews such as quantity label imbalance and distribution label imbalance were widely addressed. The citation of these studies is presented in Appendix A Table A1.

Figure 8 shows the top ten techniques most utilized in these studies, representing the central focus of this study in guiding FL users, researchers, and providers when implementing an FL approach. The figure indicates that enhancing the aggregation method and clustering are the most widely used techniques to overcome the non-IID data problem, appearing in 18% (17 studies), and 16% (15 studies) of the selected works. Personalized federated learning was used in 13% (12 studies) of the works to design a personalized model incorporating client data. An adaptive approach was utilized in 11% (10 studies) of the works. Client selection and data sharing were used in 9% (eight studies) of the works. Regularization was used in 5% (five studies), while knowledge distillation was used in 4% (four studies) of the works. Hierarchical and hierarchical clustering techniques were used in 3% (three studies) of the works to overcome the non-IID data problem. The citation of these studies is presented in Appendix A Table A2.

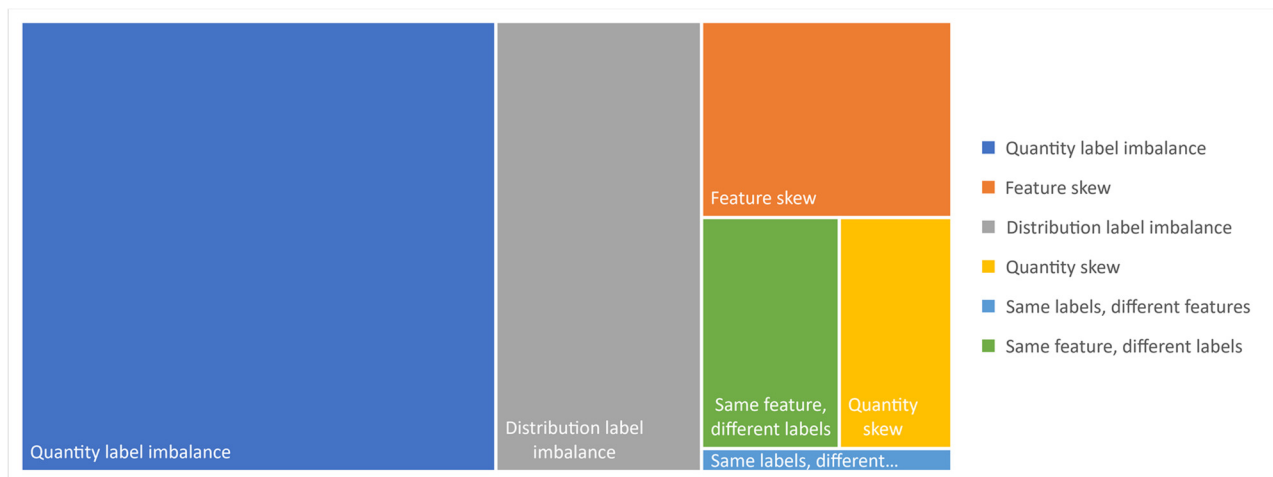


Figure 7. Non-IID types simulated in the non-IID selected studies.

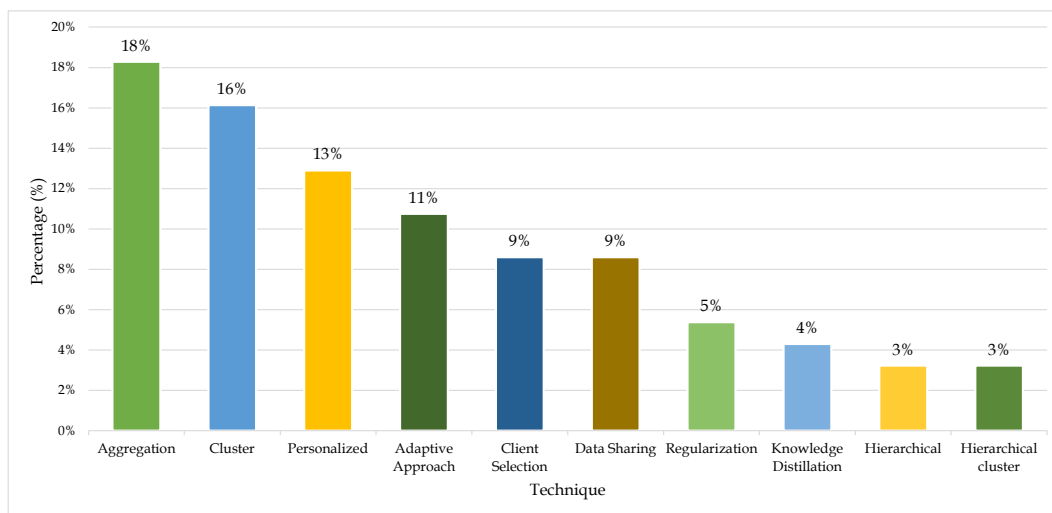


Figure 8. Percentage of the top ten techniques commonly used to solve the non-IID data problem.

In FL, clients train their local model by applying one of the learning techniques. The selected studies used different learning models to examine their proposed approach, with some studies using more than one learning model. Figure 9 shows the selected studies' top ten most used learning models. The results indicate that the convolutional neural network (CNN) is the most widely used model. However, the different studies utilize different numbers of layers when constructing the network; other studies utilize deeper networks such as ResNet and VGG when evaluating their work. Studies focusing on supervised learning methods such as FL often employ this approach.

The selected studies used different datasets to examine their proposed approach, with some studies utilizing more than one dataset. Figure 10 shows the percentage of the top ten most commonly used datasets in the 93 selected studies. The figure indicates that the Cifar-10 dataset is the most widely used; also, we can see that most of the studies utilize an image dataset.

In Appendix A Table A3, we map the topmost commonly used learning models with the respective datasets utilized when employing these models. We demonstrate only the datasets that are used in more than one study. As we can see, Cifar-10 is widely used among the different local models, as Cifar-10 was proven to be the most used dataset in non-IID studies, as shown in Figures 10 and 11. In Figure 11, the greener color indicates that more studies utilized the dataset with their respective learning models, while the red indicates that the dataset was not utilized with their respective learning models.

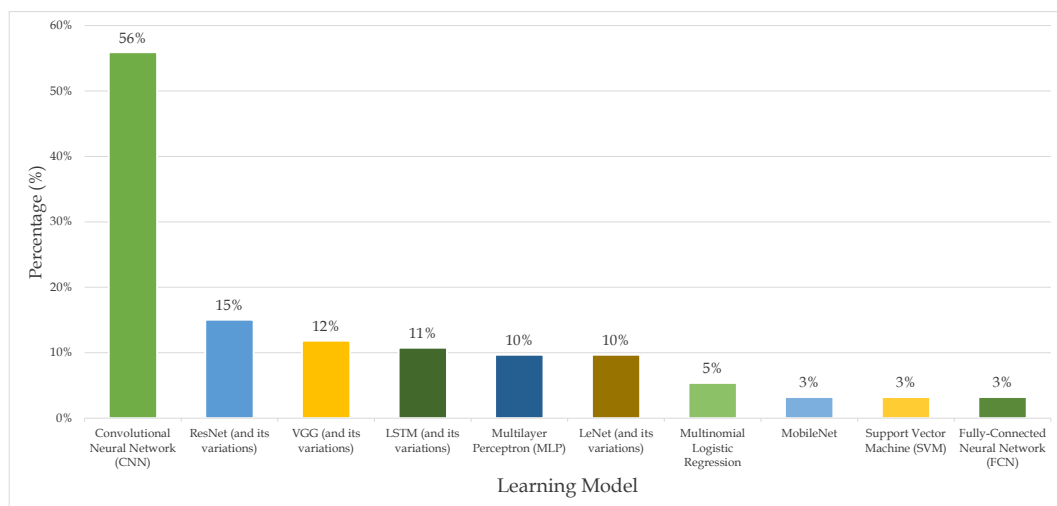


Figure 9. Percentage of the ten most commonly used learning models in non-IID data problem studies.

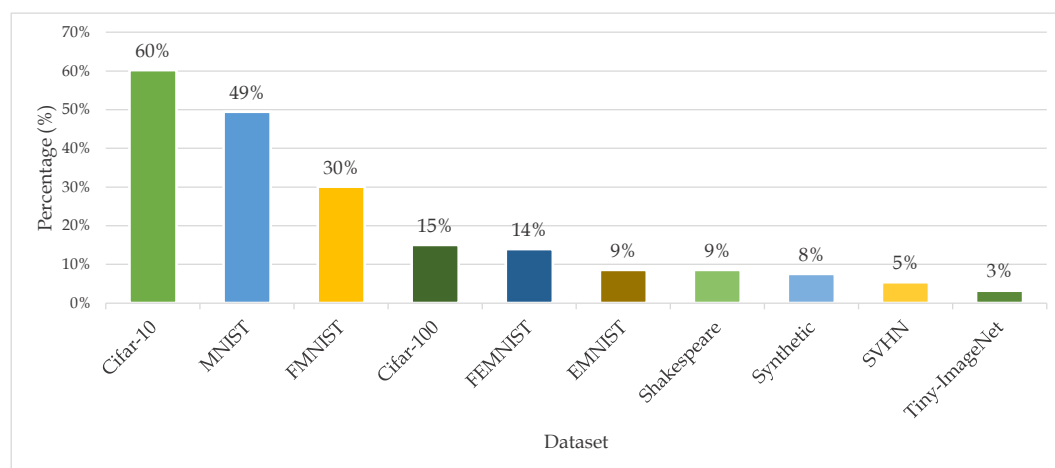


Figure 10. Percentage of the ten most commonly used datasets in non-IID data problem studies.

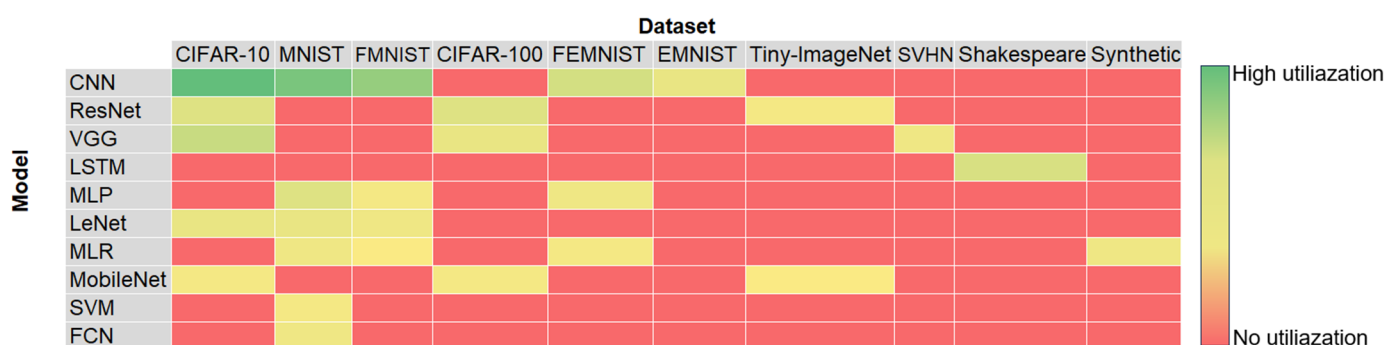


Figure 11. Heatmap between the most commonly used learning models and the respective datasets utilized in non-IID data problem studies.

4.3. Results for Communication-Efficient Studies

This subsection illustrates the systematic mapping results of works that present techniques for enhancing communication in FL. We focused on papers that provide efficient communication by reducing the number of updates or the number of bits shared. Out of the papers that we screened, a total of 74 met our selection criteria.

Figure 12 illustrates the top ten techniques most commonly used in the selected studies, representing a central focus of this study in guiding FL users, researchers, and

providers when implementing an FL approach. The results demonstrate that compression techniques such as quantization and sparsification are the most widely used techniques, used in 27% (20 studies) and 15% (11 studies) of the studies, respectively. Furthermore, other techniques are used, like the client selection technique where the server selects the participating clients, used in 9% (seven studies) of the studies. An asynchronous scheme was used in 7% (five studies) of the studies, and two-level aggregation where there is a middle node that aggregates the local models received from some clients before uploading them to the central server is used in 7% (five studies) of the studies. Select model updates where the irrelevant updates will not be uploaded are used in 7% (five studies) of the studies, and over-the-air computation is used in 5% (four studies). Cluster, periodic model averaging, and knowledge distillation are used in 4% (three studies) of the studies. All these techniques can enhance communication in federated learning. The citation of these studies is presented in Appendix A Table A4.

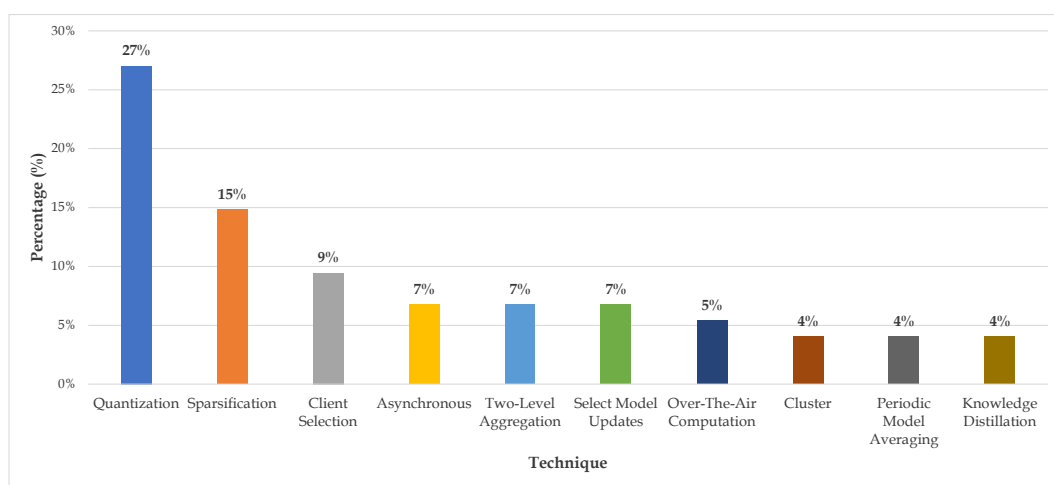


Figure 12. Percentage of the top ten techniques commonly used to provide communication efficiency in FL.

Figure 13 shows the selected studies' top ten most used learning models. The results indicate that even in the studies that aim to provide communication efficiency in federated learning, the most widely used learning model is the supervised learning model CNN.

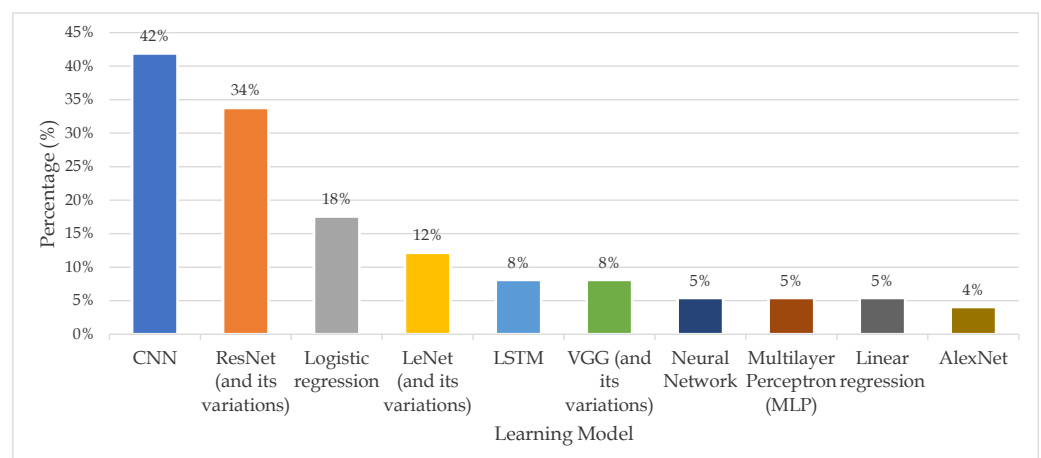


Figure 13. Percentage of the ten most commonly used learning models in communication-efficient studies.

The selected studies also utilized different datasets when examining their approach. Figure 14 shows the percentage of the datasets used in more than one study of the 74 selected studies. The figure indicates that the MNIST and Cifar-10 datasets are the most widely used.

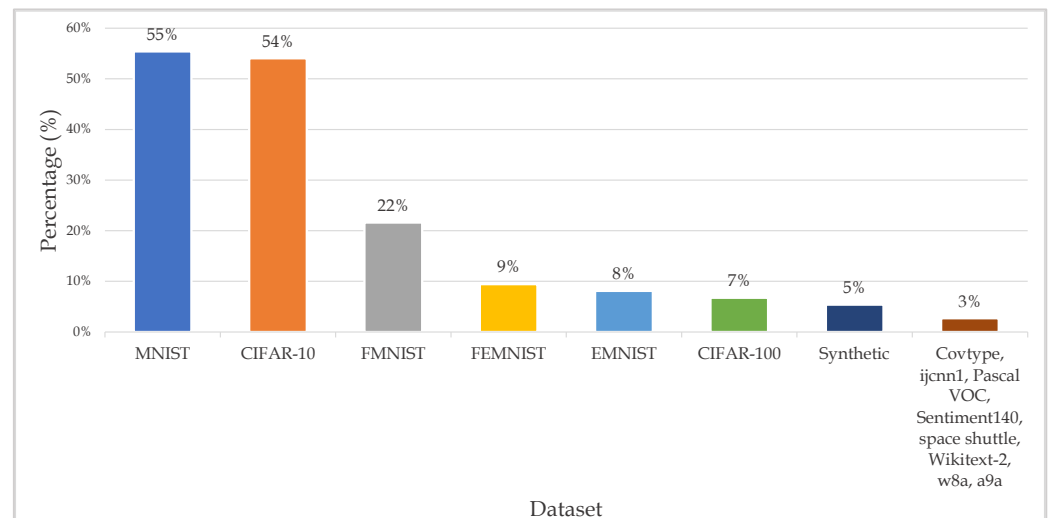


Figure 14. Percentage of the top commonly used datasets in communication-efficient studies.

Table A5 in Appendix A presents a mapping between the most used learning models and the respective datasets utilized when employing these models. We highlight only the datasets that are used in more than one study. As illustrated in Figure 15, MNIST and Cifar-10 are widely used among the different local models in the selected studies.

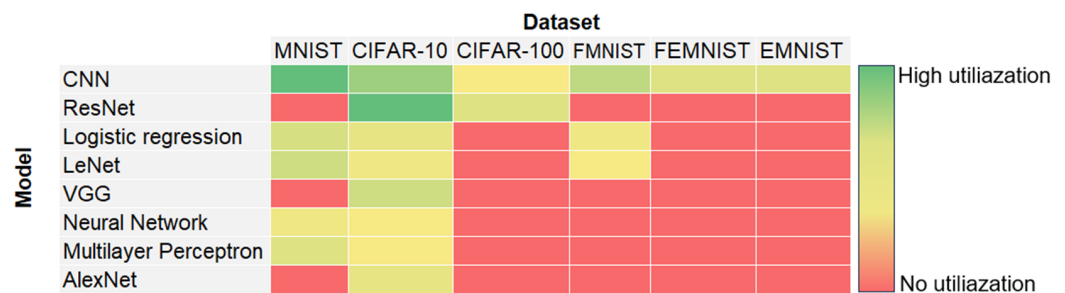


Figure 15. Heatmap between the most commonly used learning models and the respective datasets utilized in communication-efficient studies.

4.4. Results for Studies Providing Solutions for Both Challenges

This subsection illustrates the systematic mapping results of works that focus on addressing both problems. Out of all the papers that were reviewed, only 26 met our selection criteria.

Figure 16 illustrates the top ten techniques most used in the selected studies. Most of the works use a technique for each challenge; however, some techniques are used to solve both challenges, such as clustering, two-level aggregation, and client selections. The figure shows that knowledge distillation is the most widely used technique. Knowledge distillation can be used to solve both problems using the teacher–student approach to reduce communication overhead and overcome non-IID data problems by distilling knowledge. However, the selected studies used this technique along with another technique to solve the challenges. Knowledge distillation and personalized approach are used in 15% (four studies) of the studies, clustering is used in 12% (three studies) of the studies, and 8% (two studies) of the studies used one of the following techniques: adaptive approach, asynchronous method, client selection, lottery ticket, pruning method, quantization, or

two-level aggregation. The selected studies used these techniques to solve one or both challenges. The citation of these studies is presented in Appendix A Table A6.

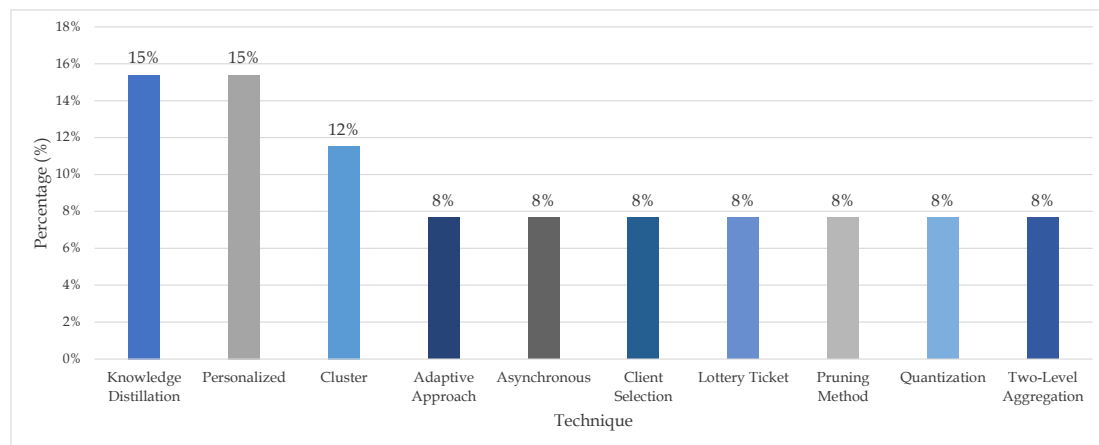


Figure 16. Percentage of the top ten techniques commonly used in the selected studies that provide solutions to both challenges.

In Figure 17, we show the most used learning models. The figure illustrates the percentage of learning models used in more than one study. The results indicate that CNN is the most widely used model, while in Figure 18, we show the percentage of the datasets used in more than one study. The figure indicates that the Cifar-10 and MNIST datasets are the most widely used.

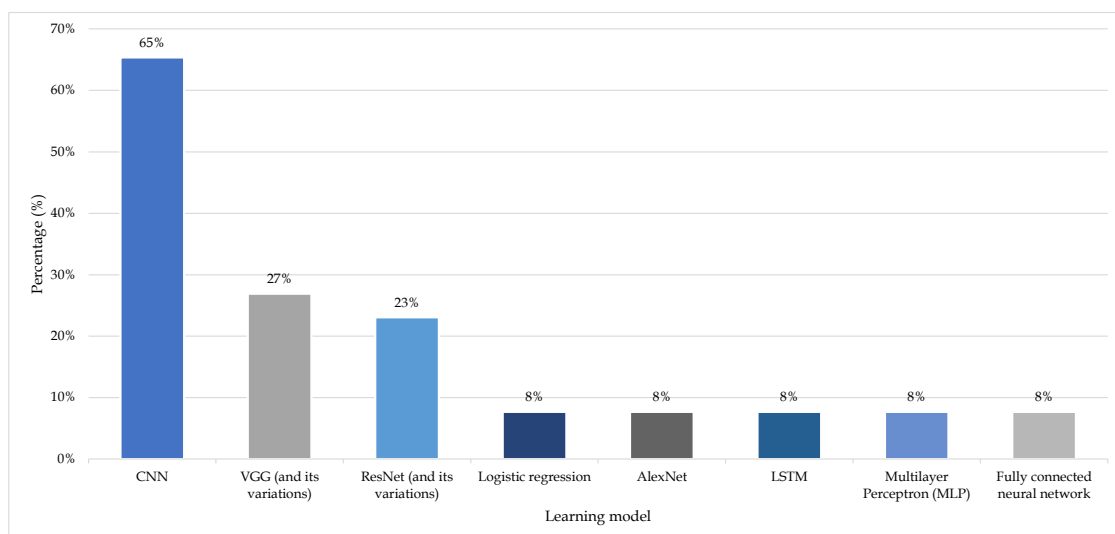


Figure 17. Percentage of the top commonly used learning models in the studies that provide solutions to both challenges.

Table A7 in Appendix A shows a mapping between the topmost commonly used learning models and the respective datasets utilized when employing these models. We show only three models, as the remaining model does not utilize a common dataset. They all use different datasets for each study, so we did not highlight them. As we can see, Cifar-10 is widely used between the different local models, as Cifar-10 was proven to be the most used dataset in these studies, as shown in Figure 19.

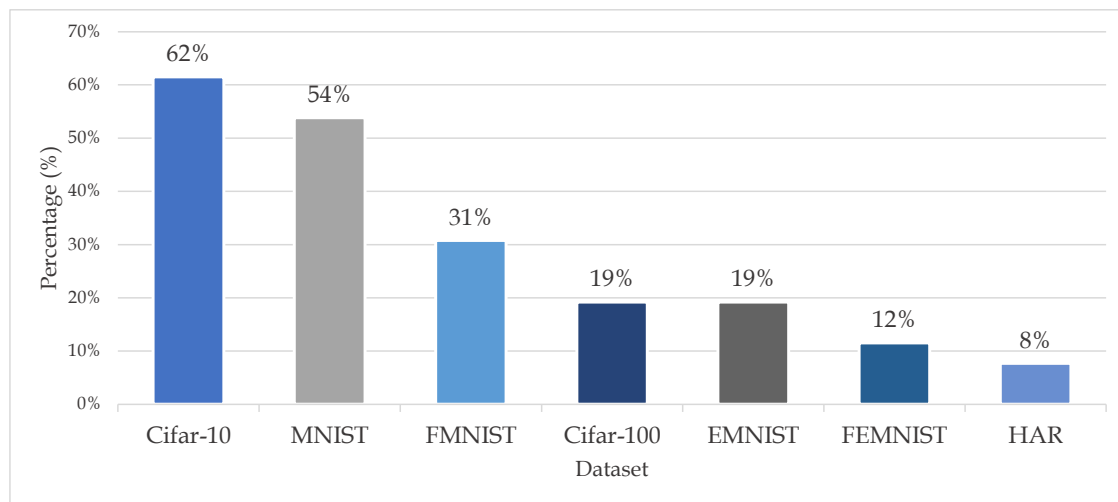


Figure 18. Percentage of the top commonly used datasets in the studies that provide solutions to both challenges.

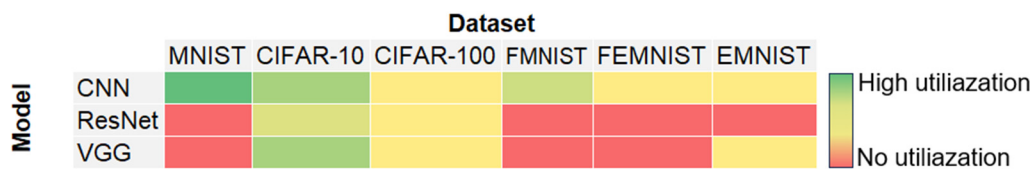


Figure 19. Heatmap between the most commonly used learning models and the respective datasets utilized in the studies that provide solutions to both challenges.

4.5. Discussion

This work aims to identify the techniques used to overcome the non-IID data problem and techniques that provide communication efficiency in an FL environment.

To address RQ1, we examined the selected studies that aim to overcome the non-IID data challenge in their work. When evaluating their work, these studies need to select a suitable distribution to introduce the non-IID data distribution between the clients. As shown in Figure 7, most of the studies introduce the label distribution skew, specifically quantity label imbalance, by dividing the datasets between the clients according to their class, such that the clients can access only a limited number of classes.

Also, the distribution label imbalance was widely addressed, where the clients had variations in their class distribution. To address the feature skew, the studies utilized the FEMNIST (Federated Extended MNIST) dataset that divides the datasets based on the writer, and in the selected studies, the writers were represented as the clients. However, we can see that most of the studies did not focus on the remaining skew types, and the label distribution skew was the widely addressed skew.

To address RQ2, we thoroughly examined the retrieved studies and analyzed the techniques utilized. Figure 8 shows the most widely used techniques to overcome the non-IID data challenge in FL. Based on the selected studies and findings outlined in this work, several critical research hotspots have emerged in FL. One of the prominent areas of interest is the development of novel aggregation schemes to address the non-IID data challenge used in 18% (17 studies) of the selected studies. These schemes aim to improve model accuracy and convergence by effectively aggregating local updates from heterogeneous client devices. Additionally, there is growing attention towards exploring clustering in FL, which is used in 16% (15 studies) of the studies. Clustering is a technique used to group clients into clusters to overcome the non-IID data challenge, using different clustering criteria such as model similarity and the client's dataset size.

Aggregation in FL: As the first study that proposed FL also highlighted the importance of the non-IID data problem and suggested an aggregation method that is a weighted averaging scheme (FedAvg) based on clients' dataset size to overcome this challenge, many studies were inspired by this work and tried to enhance it. For instance, the work in [50] delayed the aggregation process by sending the model back to some clients for further training to enhance the model, while the study in [105] enhanced the calculation of the weights to be based on indices of statistical heterogeneity instead of just the client dataset size. While the study in [131] adds to the FedAvg a regularization to lower the excess risk, the study in [133] also uses regularization in its aggregation scheme to penalize the diverging model.

Clustering in FL: Clustering is also used to overcome the non-IID data challenge, where the server clusters clients based on certain criteria. In [54,88], the clients are clustered based on the similarity of their models, whereas in [97], the central server clusters clients based on their dataset size. The work in [100] mandates every client to report specific statistics about their local dataset in order to perform clustering. Aggregation and clustering are the two widely used techniques that aim to overcome non-IID data on the server side by utilizing information extracted from the local model. However, these techniques raise security concerns as they involve analyzing the obtained local model and sharing information. Hence, it is necessary to develop a secure environment in FL where the clients can trust that their privacy will not be compromised in an analysis of their local model and the server will ensure the integrity of the received local model.

The techniques used to overcome the non-IID data challenge can be classified into server-side and client-side, according to where the non-IID data challenge is addressed. Aggregation, clustering, and client selection can be considered server-side techniques, where the server will be responsible for overcoming the non-IID data challenge. The work in [47,86] selects the clients that will participate in the next round based on the received local model in the current round. In contrast, in [53,84], the server selects the clients' models that will contribute to the new global model based on their model divergence; the models of the unselected clients will be abandoned.

Some techniques can be implemented either at the client side or the server side, like a personalized approach that, for instance, can be implemented at the server side by keeping a record for each client to provide a personalized model [83] or at the client side where each client can have their personalized trained model locally and share a general model [87]. An adaptive approach also can be performed on the server side, where the work in [48] proves that using a fixed batch size can degrade the model performance since the data distribution and size differ between clients, so it proposes a batch adaption technique to determine the suitable batch size for each client; [99] proposes an adaptive local epoch technique to avoid overfitting the model, by decreasing the local epoch value after a certain iteration based on the global model performance. The adaptive approach can be implemented on the client side, where in [49,95], the clients adapt their learning rate based on the global model received. The study [49] considers the deviations between the local model and global model and introduces a penalty term to force the local model to be inclined to learn the global model. In [81], the clients adapt their local model by having some local parameters used for local adaption. Knowledge distillation using a teacher–student model is also a technique that can be applied on the server side [71,124] or the client side [111,134]. The regularization technique is a technique used on the client side in [55,126,138] that can keep the local model closer to the global model by adding a regularization term to help the local model approach the global model. Each technique has its own benefits and limitations, as the server-side technique generally does not introduce extra computation at the client devices, but it may analyze the obtained local model, raising some privacy concerns. While client-side techniques preserve the client's privacy, extra computation steps are required to overcome the non-IID data challenge.

Collaboration between clients and servers is sometimes necessary for certain techniques, particularly when it comes to data sharing. Data sharing can be seen as a joint

effort between clients and the central server, where clients share some of their data with the server, and the server either uses the collected data to train a model at the server [79] or shares the collected data with clients [89]. Other techniques can change the architecture by introducing a middle layer between the client and server using a hierarchical architecture and hierarchical clustering architecture. Figure 20 shows a representation of the most widely used techniques that aim to overcome the non-IID data challenge based on their side. As we can see, most of the techniques are server-side, as the server has a more global look compared to the clients that can communicate only with the central server. However, there are techniques that can be implemented at the client-side that can help overcome the non-IID data challenge using the received global model or information obtained from the global server. Furthermore, introducing a middle layer between the client and server can help overcome the non-IID data challenge without extra computation at the client or the server.

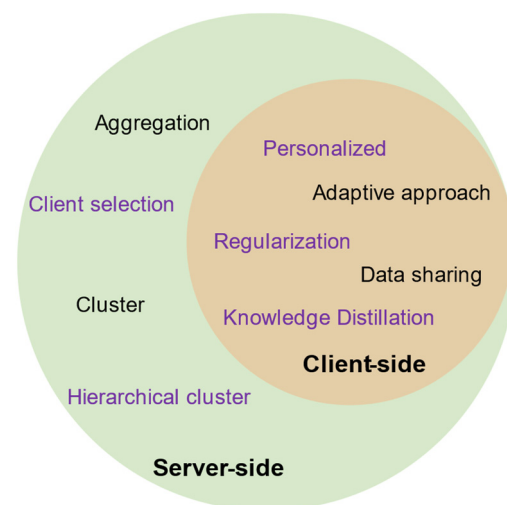


Figure 20. The classification of the techniques for solving the non-IID data challenge in federated learning.

To answer RQ3, the most widely used techniques to provide communication efficiency in FL are shown in Figure 12. Based on the selected studies and findings outlined in this work, compression techniques are a research hotspot in FL for providing efficient communication; compression techniques of particular importance are quantization and sparsification, which are used in 27% (20 studies) and 15% (11 studies) of the studies, respectively. These techniques reduce the number of bits transmitted by compressing the model.

Quantization in FL: Quantization is used to reduce the number of bits that represent a value; it is essential to set the quantization level carefully to avoid damaging the model and prevent the loss of useful information. For instance, the studies conducted in [143,145] used an adaptive level of quantization to minimize the error bound. In contrast, Refs. [152,156] used a 1-bit quantization scheme to quantize the local model to 1-bit data, while the work in [184] used ternary quantization that quantizes the value into one of three values (−1,0,1).

Sparsification in FL: Sparsification schemes are used to reduce the number of bits transmitted. Different studies used different methods to select these bits; for example, the work in [141] sets the non-important weight update elements to zero, while the works in [153,165] use a top-k selection-based gradient compression scheme. In addition, Ref. [155] uses block sparsification by dividing the local gradient vector into sub-vectors and then dropping some gradient entries with small magnitudes at each sub-vector. These compression techniques require clients to compress the model without losing any vital information and share the compressed model with the server, as losing the vital information may affect the global model performance and convergence speed. However, most of these techniques are applied

on the client side, which means that clients need to be capable of training the model and applying a compression technique, and this can be challenging due to the limited resources of clients.

We can classify the top techniques into techniques that aim to reduce the number of bits transmitted (reducing the model size) and techniques that aim to reduce the number of updates, as shown in Figure 21. Quantization, sparsification, and knowledge distillation techniques focus on reducing the number of bits shared to provide efficient communication. The client selection, select model updates, over-the-air computation, clustering, periodic model averaging, asynchronous, and two-level aggregation techniques aim to reduce the number of updates between the server and the clients.

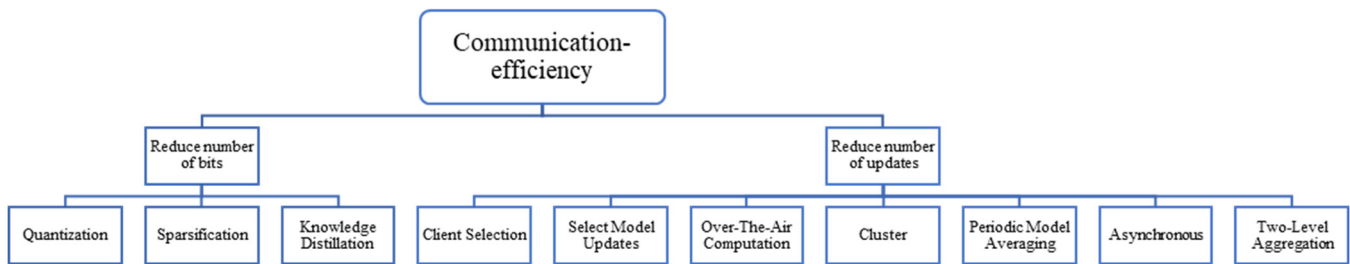


Figure 21. Categories of techniques to provide efficient communication in federated learning.

The client selection scheme is used to select the clients that can contribute more to enhance the global model, which results in a reduction in the communication rounds [147,185,191]. Using asynchronous communication can enhance the global model performance by allowing the aggregation of the received model without waiting for all clients [146,171,211]. Select model update is a technique that uploads the trained model that can help model coverage and ignore irrelevant updates [157,189]. Over-the-air computation [179,197] is a technique used by exploiting the waveform superposition property, which directly obtains the aggregated model. This technique can provide faster and more communication-efficient training in federated learning. Clustering is where similar clients are clustered together and a representative is selected to share the updated model [172,177,185]. Periodic model averaging [198,209] is where the local model is uploaded periodically to reduce the number of updates. Two-level aggregation [164,175] is where a middle layer near the clients aggregates the model to reduce the communication rounds with the central server.

Even though most of the techniques are based on reducing the number of updates, most of the selected studies employ techniques to reduce the communication bits such as quantization and sparsification, while a few studies aimed to reduce the number of bits and the number of rounds in their approaches by employing different techniques such as quantization with client selection [191] and with periodic model averaging [198]; as shown in Figure 22, 61% of the selected studies apply a technique that aims to reduce the communication bits.

Figure 16 shows the top techniques that are commonly used in studies that address both challenges (non-IID data challenge and providing efficient communication). Most studies used a technique for each challenge, while some studies used one technique to overcome both challenges. These techniques are clustering, two-level aggregation, and client selection. Clustering is the most widely used technique for overcoming both challenges, appearing in 16% of non-IID data studies, 4% of communication-efficient studies, and 11% of studies that provided solutions to both challenges. In the cluster technique, the central server mostly clusters the clients into different groups based on different criteria such that similar clients are grouped to enhance the learning process and reduce the communication rounds. Two-level aggregation also can help reduce the communication rounds between the central server and the clients by introducing a partial aggregation near the clients that can reduce the number of models transmitted and enhance the model training. Client selection is an essential step in the federated learning performed by the server; selecting

the clients based on their model or resources enhances the training process and reduces the communication rounds.

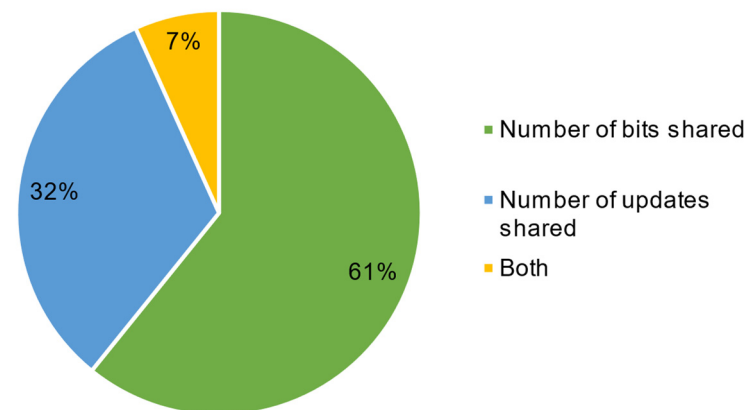


Figure 22. Percentage of the objective used to provide efficient communication in federated learning.

To answer RQ4, we extracted the most commonly used learning models in the selected studies, as shown in Figures 9, 13 and 17. The figures demonstrate that most of the studies utilized CNN as their learning model; it was used in 56% of non-IID data studies, 42% of communication-efficient studies, and 65% of studies that provided solutions to both challenges. Different learning models such as ResNet and VGG were also widely utilized in these studies. These learning models are supervised learning models and are often employed in FL.

To address RQ5, we examined the selected studies and retrieved the most commonly used datasets, as shown in Figures 10, 14 and 18. The figures demonstrate that most studies use image datasets; in particular, they utilize Cifar-10 and MNIST datasets to evaluate their work. The Cifar-10 dataset was used in 60% of non-IID data studies, 54% of communication-efficient studies, and 62% of studies that provided solutions to both challenges, while the MNIST dataset was used in 49% of non-IID data studies, 55% of communication-efficient studies, and 54% of studies that provided solutions to both challenges. These datasets are used as benchmarks in FL to simulate IID and non-IID data distribution. Many studies simulate this skew using MNIST and Cifar-10 datasets, as they can be easily understood and evaluated by others. Besides, FL is a relatively new approach, researchers focus on addressing the challenges in clear-cut scenarios, especially the non-IID data challenge, as the studies need to simulate a non-IID data distribution that can be easily understood and evaluated by others. Researchers can use commonly used datasets such as MNIST and Cifar-10 to gain valuable insights into the effectiveness of proposed methodologies and compare them with existing approaches. Furthermore, the first study that proposed FL utilized these datasets and the Shakespeare dataset in their work. However, many studies utilize the FEMNIST dataset, which is a federated version of the EMNIST dataset built by partitioning the data between the clients based on their handwriting; the FEMNIST dataset was utilized in 14% of non-IID data studies, 9% of communication-efficient studies, and 12% of studies that provided solutions to both challenges. Other than image datasets, text datasets such as the Shakespeare dataset were utilized in 9% of the non-IID data studies; Sentiment140 and Wikitext-2 are also text datasets used in 3% of communication-efficient studies. The MNIST and Cifar-10 datasets are widely used with the CNN model, as shown in Figures 11, 15 and 19. The deeper neural networks ResNet and VGG are commonly used with the colored image datasets Cifar-10 and Cifar-100 as they have more features than the greyscale MNIST dataset. Based on the datasets utilized, we can conclude that image classification tasks are the most used tasks in the selected studies.

4.6. Threats to Validity

The obtained results are affected by the choice of the database; to reduce this effect, we use different databases that we believe contain the most relevant studies published in the federated learning domain. Furthermore, the choice of the search terms affects the obtained results and can be considered a threat; we used different alternative synonyms to reduce the effect of this threat.

Another potential limitation is that this work focuses on studies that provide communication efficiency by decreasing the number of rounds or bits transmitted, while there might be other studies that offer communication efficiency from other aspects.

4.7. Future Research Directions

Non-IID data challenge: Many studies have tried to overcome the non-IID data challenge on the server side, often resulting in the abandonment of a trained model or a decrease in its effectiveness. This can result in losing crucial information, mainly if the model contains infrequent but essential data. To improve the training process, addressing the non-IID data challenge on the client end or using a hierarchical architecture could be more beneficial since the trained model will not be wasted, for instance, using edge computing or fog computing to help the IoT devices to overcome the skewness of the client data. Furthermore, resolving the non-IID data challenge before the training process could streamline training and reduce its duration.

Privacy concern in non-IID data: FL was proposed to preserve the privacy of the client data. However, some techniques require analyzing the received client model to extract information to overcome the non-IID data challenge. This analysis may unintentionally lead to the exposure of sensitive information. Therefore, it is crucial to incorporate security measures, such as anonymous sharing of models, when using these techniques to prevent privacy leakage.

Security in FL: FL training is a collaborative process between different parties that exposes it to various security threats, especially when clients have non-IID data; the existence of these data can facilitate backdoor attacks by malicious clients who may mislabel the samples to compromise the global model's performance and convergence speed. Additionally, a malicious server can analyze the obtained local models to expose clients' privacy. Therefore, when developing a scheme to overcome the non-IID data challenges, it is necessary to consider these attacks as some approaches can mitigate the impact of backdoor attacks and the Byzantine problems inherently, such as a personalized approach and aggregation. However, it is important to design them carefully so as not to introduce vulnerabilities that adversaries could exploit to compromise the security of the federated learning; hence, developing trusted federated learning is essential for ensuring a safe FL environment.

Encryption in FL: FL is a distributed approach that exchanges the model through a network. However, this approach exposes the model to various security threats, such as snooping and modification by attackers. If the attacker modifies the captured model, it can compromise the training process and increase the communication rounds. Attackers can also capture and analyze local models to expose clients' private information. Moreover, some techniques that overcome the non-IID data challenge, such as clustering and data sharing, can facilitate the attacks. For instance, clustering methods may group clients with similar data distributions, making them more vulnerable to targeted attacks. Data-sharing techniques may also expose sensitive data points across clients, creating opportunities for attackers to exploit. Therefore, an encryption scheme in federated learning can mitigate these attacks and enhance the overall training process; however, the limited resources of clients and networks need to be considered.

Generalizing of the global model: FL was proposed to facilitate the sharing of knowledge among different devices and learning from various environments. However, some techniques propose a personalized approach or cluster clients based on the similarity of their data. These approaches enable the central server to provide clients with a model that fits their own data. Although this approach can be beneficial in some specific scenarios

(such as next-word prediction), some other applications require more global knowledge to achieve accurate results with unseen data. Therefore, a generalized approach to FL is required to facilitate the sharing of knowledge among different clients while still providing accurate and personalized results.

Communication-efficiency challenge: It is crucial to take into account the storage and capability limitations of the clients' devices when developing an approach for efficient communication. Some techniques overlook these constraints and only focus on the network communication cost. However, some of these techniques can result in the loss of critical information that affects the model convergence. For that, to ensure the effectiveness of FL, it is necessary to adopt an approach that facilitates efficient communication without placing an extra burden on the clients.

Local model challenge: Researchers face the choice of selecting the appropriate learning model when evaluating their proposed work, which may impact the model's performance. It is worth noting that training the model in the FL environment takes place on client devices and necessitates sharing the trained model with a central server over multiple rounds, which can be challenging given the limited communication and computation resources of the client's devices. It would be beneficial to investigate the effects of various learning models on FL performance, particularly those that can be applied to the same tasks and have different model sizes, such as CNN, ResNet, and VGG.

Real-world deployment and evaluation: It is important to evaluate the FL approach in real-world scenarios with actual datasets. Many studies focus on proposing solutions to overcome challenges in FL, but they do not always consider the limitations of real-world scenarios. For example, using a deep learning neural network to evaluate FL performance could improve the model's performance, but it may not be feasible due to limited resources on client devices and networks. Therefore, there is a need to carry out research studies and evaluate the performance of FL using a real-world deployment which will help guide this field in the future.

5. Conclusions

This work presents a systematic mapping study to identify the most commonly used techniques for overcoming non-IID data problems and communication challenges in FL. A total of 193 articles that met our inclusion and exclusion criteria were selected using a systematic mapping study. We categorized these articles into three groups based on the problem they aimed to solve: articles that addressed the non-IID data problem, articles that aimed to provide communication-efficient solutions in FL, and articles that provided solutions to both challenges. To answer RQ1, we analyzed the selected articles that aimed to overcome the non-IID data challenge; we concluded that label distribution skew, specifically the quantity label imbalance, where clients had some missing labels, was most commonly used.

To answer RQ2, we analyzed the selected articles and identified the most commonly used techniques for overcoming non-IID data problems; we found that enhancing the aggregation method and clustering are the two most commonly used techniques. For RQ3, we analyzed the selected articles and identified the most commonly used techniques for providing communication-efficient solutions in FL; we found that quantization and sparsification are the two most commonly used techniques. Some techniques can be used to provide solutions to both challenges, such as clustering, two-level aggregation, and client selection.

For RQ4, we extracted the most commonly used learning models in the selected articles and found that the supervised learning model CNN is the most used. For RQ5, we extracted the most commonly used dataset in the selected articles, and we found that the image datasets Cifar-10 and MNIST datasets are the most commonly used datasets to evaluate the proposed work in the selected studies.

Author Contributions: Conceptualization, B.A., F.A.K. and S.M.; methodology, B.A., F.A.K. and S.M.; validation, B.A., F.A.K. and S.M.; formal analysis, B.A., F.A.K. and S.M.; investigation, B.A., F.A.K. and S.M.; data curation, B.A., F.A.K. and S.M.; writing—original draft preparation, B.A.; writing—review and editing, F.A.K. and S.M.; visualization, B.A., F.A.K. and S.M.; supervision, F.A.K. and S.M.; funding acquisition, F.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Saudi Data and AI Authority (SDAIA) and King Fahd University of Petroleum and Minerals (KFUPM) under SDAIA-KFUPM Joint Research Center for Artificial Intelligence Grant No. JRC-AI-RFP-12.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

Appendix A. Additional Tables

Table A1. The citation of the non-IID data types used in the studies aiming to solve non-IID data problems.

Non-IID Data Type	Studies Referenced
Quantity label imbalance	[48,50,51,54,55,57–60,65–67,70–72,75,77,78,81,82,84,87–94,98–100,104–115,117,119,121–126,128–131,134,138]
Quantity label imbalance	[214,217,218,221–228,230–233,235,237,238]
Distribution label imbalance	[48,49,54,62,71,72,74–76,78,80,86,87,92,94,96,97,99,100,102,103,112,115,116,118,122,136,137,215,224,234,236]
Feature skew	[50,56,67,74,76,77,88,93,99,109,111,113–115,135,237,238]
Quantity skew	[61,68,91,109,116,119,214,218,232]
Same features, different labels	[50,67,76,99,111,113–115,125,237]
Same labels, different features	[52,103]

Table A2. The citation of the top ten techniques commonly used to solve non-IID data problems.

Techniques	Studies Referenced
Aggregation	[50,61,68,74–76,91,99,102,105,107,110,114,118,130,131,133]
Cluster	[54,56,77,82,97,100,103,108,109,116,122,125,127,135,138]
Personalized	[52,65,83,87,88,96,111,113,120,125,129,139]
Adaptive Approach	[48,49,59,81,95,97,99,112,115,116]
Client Selection	[47,50,53,56,63,84,86,128]
Data Sharing	[57,58,68,78,79,89,90,121]
Regularization	[55,126,131,133,138]
Knowledge Distillation	[71,111,124,134]
Hierarchical	[56,85,117]
Hierarchical Clustering	[67,88,119]

Table A3. The non-IID data challenge studies for the top most commonly used learning models with the respective datasets utilized when employing these models.

Model	Dataset	Studies Referenced
CNN	Cifar-10	[48,53–55,60–62,66,70,73,76,79,84,86,97–103,105,108,111,115,119,122,128]
	MNIST	[48,54,59–61,66,67,70,72,75,82,84,86,93,97,98,100,104,105,108,114,119,122,128]
	FMNIST	[48,51,53,59–61,73,77,79,81,86,89,97,99,100,105,117,122,126]
	FEMNIST	[56,76,77,99,109,113–115]
	EMNIST	[72,93,95,104]
	Other	[50,62,73,80,81,88,90,93,98,102,112,114,134]
ResNet	Cifar-10	[71,74,78,91,107,118]
	Cifar-100	[52,69,71,76,83,136]
	Tiny ImageNet	[55,62]
	Other	[52,55,69,71,74,101,118]
VGG	Cifar-10	[69,72,74,75,106,113,114,117,124,130]
	Cifar-100	[72,106,113,114]
	SVHN	[69,74,124]
	Other	[69,124,136]
LSTM	Shakespeare	[50,76,99,111,114,115,125]
	Other	[51,95,111,123]
MLP	MNIST	[57,69,77,98,102,108]
	FEMNIST	[69,77,109]
	FMNIST	[77,129]
	Other	[77,108,123,129]
	Other	[77,108,123,129]
LeNet	MNIST	[68,101,107,117]
	Cifar-10	[52,92,93,96]
	FMNIST	[107,117,130]
	Other	[92,96]
MLR	Synthetic	[77,84,125]
	MNIST	[59,77,84]
	FEMNIST	[77,111]
	FMNIST	[59]
MobileNet	Cifar-10	[106,107]
	Cifar-100	[106,136]
	Tiny ImageNet	[136]
SVM	MNIST	[68,129]
	Other	[123]
FCN	MNIST	[50,62,131]
	Other	[131]

Table A4. The citation of the top ten techniques commonly used to provide communication efficiency in FL.

Techniques	Studies Referenced
Quantization	[142–145,148,152,156,158,168,169,174,176,183,184,188,191–193,198,199]
Sparsification	[140,141,151,153,155,165,174,186,200,202,204]
Client Selection	[147,166,172,185,191,198,207]
Asynchronous	[146,171,190,203,211]
Two-Level Aggregation	[164,175,180,182,185]
Select Model Updates	[149,157,170,189,206]
Over-The-Air Computation	[162,178,179,197]
Cluster	[172,177,185]
Periodic Model Averaging	[198,207,209]
Knowledge Distillation	[171,205,210]

Table A5. The communication-efficient studies for the top most commonly used learning models with the respective datasets utilized when employing these models.

Model	Dataset	Studies Referenced
CNN	MNIST	[146,149,152,154,165,169–171,180,182,185,189,192,202,203,208–210,212]
	Cifar-10	[140,146,167,171,176,184,188–191,204,210]
	FMNIST	[143,152,167–169,190,203,212]
	FEMNIST	[140,142,167,185]
	EMNIST	[150,169,203,210]
	Cifar-100	[190]
ResNet	Cifar-10	[143,144,148,159,161,168,174,179,180,182,184–186,193,200,201,208,212,213]
	Cifar-100	[144,161,199,213]
	Other	[141,160,166,200–202,206,208,213]
Logistic Regression	MNIST	[157,176,180,190,198]
	Cifar-10	[173,180,192]
	FMNIST	[178,207]
	Other	[147,157,162,167,178,193]
LeNet	MNIST	[148,174,181,183,195,211]
	Cifar-10	[151,159]
	FMNIST	[158]
LSTM	Other	[142,149,159,165,167,195]
VGG	Cifar-10	[141,144,148,158,161,195]
	Other	[141]
Neural Network	MNIST	[155,157]
	Other	[198,209]

Table A5. *Cont.*

Model	Dataset	Studies Referenced
MLP	MNIST	[146,156,161,184]
	Cifar-10	[146]
Linear Regression	Other	[157,163,173,190]
AlexNet	Cifar-10	[148,180,192]

Table A6. The citation of the top ten techniques commonly used in the selected studies that provide solutions to both challenges.

Techniques	Studies Referenced
Knowledge Distillation	[215,233,236,239]
Personalized	[228,234,237,238]
Cluster	[222,227,232]
Adaptive Approach	[217,236]
Asynchronous	[218,229]
Client Selection	[216,223]
Lottery Ticket	[224,234]
Pruning Method	[237,238]
Quantization	[219,233]
Two-Level Aggregation	[226,230]

Table A7. The top most commonly used learning models with the respective datasets utilized when employing these models in the studies that provide solutions for both challenges.

Model	Dataset	Studies Referenced
CNN	MNIST	[217,221,223,224,226,227,229,231–233,239]
	Cifar-10	[217,218,224,229,232–234]
	FMNIST	[215,218,231,232,239]
	FEMNIST	[232,235]
	Cifar-100	[233,235]
	EMNIST	[224,233]
	Other	[225,229,230,232,234]
ResNet	Cifar-10	[221,231,233,234]
	Cifar-100	[218,233]
	Other	[234,235]
VGG	Cifar-10	[215,221,223,225,228,237,238]
	Cifar-100	[221,228]
	EMNIST	[237,238]

References

1. Nguyen, D.C.; Ding, M.; Pathirana, P.N.; Seneviratne, A.; Li, J.; Poor, H.V. Federated learning for internet of things: A comprehensive survey. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 1622–1658. [\[CrossRef\]](#)
2. Xia, Q.; Ye, W.; Tao, Z.; Wu, J.; Li, Q. A survey of federated learning for edge computing: Research problems and solutions. *High-Confid. Comput.* **2021**, *1*, 100008. [\[CrossRef\]](#)
3. Song, S.; Liang, X. Federated Pseudo-Sample Clustering Algorithm: A Label-Personalized Federated Learning Scheme Based on Image Clustering. *Appl. Sci.* **2024**, *14*, 2345. [\[CrossRef\]](#)

4. Zhang, C.; Li, M.; Wu, D. Federated multidomain learning with graph ensemble autoencoder GMM for emotion recognition. *IEEE Trans. Intell. Transp. Syst.* **2022**, *24*, 7631–7641. [\[CrossRef\]](#)
5. Ting, D.; Hamdan, H.; Kasmiran, K.A.; Yaakob, R. Federated learning optimization techniques for non-IID data: A review. *Int. J. Adv. Res. Eng. Technol.* **2020**, *11*, 1315–1329.
6. Aledhari, M.; Razzak, R.; Parizi, R.M.; Saeed, F. Federated learning: A survey on enabling technologies, protocols, and applications. *IEEE Access* **2020**, *8*, 140699–140725. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Li, T.; Sahu, A.K.; Talwalkar, A.; Smith, V. Federated learning: Challenges, methods, and future directions. *IEEE Signal Process. Mag.* **2020**, *37*, 50–60. [\[CrossRef\]](#)
8. Briggs, C.; Fan, Z.; Andras, P. A review of privacy-preserving federated learning for the Internet-of-Things. In *Federated Learning Systems*; Springer: Cham, Switzerland, 2021; pp. 21–50. [\[CrossRef\]](#)
9. Wang, C.; Xia, H.; Xu, S.; Chi, H.; Zhang, R.; Hu, C. FedBnR: Mitigating federated learning Non-IID problem by breaking the skewed task and reconstructing representation. *Future Gener. Comput. Syst.* **2024**, *153*, 1–11. [\[CrossRef\]](#)
10. Chen, H.; Frikha, A.; Krompass, D.; Gu, J.; Tresp, V. FRAug: Tackling federated learning with Non-IID features via representation augmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 4–6 October 2023.
11. Mengistu, T.M.; Kim, T.; Lin, J.-W. A Survey on Heterogeneity Taxonomy, Security and Privacy Preservation in the Integration of IoT, Wireless Sensor Networks and Federated Learning. *Sensors* **2024**, *24*, 968. [\[CrossRef\]](#)
12. Hamer, J.; Mohri, M.; Suresh, A.T. Fedboost: A communication-efficient algorithm for federated learning. In Proceedings of the 37th International Conference on Machine Learning, Virtual, 13–18 July 2020.
13. Khan, A.; Thij, M.T.; Wilbik, A. Communication-Efficient Vertical Federated Learning. *Algorithms* **2022**, *15*, 273. [\[CrossRef\]](#)
14. Li, K.; Wang, H.; Zhang, Q. FedTCR: Communication-efficient federated learning via taming computing resources. *Complex Intell. Syst.* **2023**, *9*, 5199–5219. [\[CrossRef\]](#)
15. Liu, Y.; Yuan, X.; Xiong, Z.; Kang, J.; Wang, X.; Niyato, D. Federated learning for 6G communications: Challenges, methods, and future directions. *China Commun.* **2020**, *17*, 105–118. [\[CrossRef\]](#)
16. Tian, T.; Shi, H.; Ma, R.; Liu, Y. FedACQ: Adaptive clustering quantization of model parameters in federated learning. *Int. J. Web Inf. Syst.* **2024**, *20*, 88–110. [\[CrossRef\]](#)
17. Lo, S.K.; Lu, Q.; Wang, C.; Paik, H.-Y.; Zhu, L. A systematic literature review on federated machine learning: From a software engineering perspective. *ACM Comput. Surv.* **2021**, *54*, 1–39. [\[CrossRef\]](#)
18. El Mokadem, R.; Maissa, Y.B.; El Akkaoui, Z. Federated learning for energy constrained devices: A systematic mapping study. *Clust. Comput.* **2023**, *26*, 1685–1708. [\[CrossRef\]](#)
19. Petersen, K.; Feldt, R.; Mujtaba, S.; Mattsson, M. Systematic mapping studies in software engineering. In Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering (EASE), Bari, Italy, 26–27 June 2008. [\[CrossRef\]](#)
20. Li, L.; Fan, Y.; Tse, M.; Lin, K.-Y. A review of applications in federated learning. *Comput. Ind. Eng.* **2020**, *149*, 106854. [\[CrossRef\]](#)
21. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.A.; y Arcas, "Communication-efficient learning of deep networks from decentralized data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, Lauderdale, FL, USA, 20–22 April 2017.
22. Yang, Q.; Liu, Y.; Chen, T.; Tong, Y. Federated machine learning: Concept and applications. *ACM Trans. Intell. Syst. Technol.* **2019**, *10*, 1–19. [\[CrossRef\]](#)
23. Zhang, C.; Xie, Y.; Bai, H.; Yu, B.; Li, W.; Gao, Y. A survey on federated learning. *Knowl. Based Syst.* **2021**, *216*, 106775. [\[CrossRef\]](#)
24. Li, Q.; Thapa, C.; Ong, L.; Zheng, Y.; Ma, H.; Camtepe, S.A.; Fu, A.; Gao, Y. Vertical Federated Learning: Taxonomies, Threats, and Prospects. *arXiv* **2023**, arXiv:2302.01550. [\[CrossRef\]](#)
25. Kang, Y.; Luo, J.; He, Y.; Zhang, X.; Fan, L.; Yang, Q. A Framework for Evaluating Privacy-Utility Trade-off in Vertical Federated Learning. *arXiv* **2022**, arXiv:2209.03885. [\[CrossRef\]](#)
26. Gao, D.; Yao, X.; Yang, Q. A Survey on Heterogeneous Federated Learning. *arXiv* **2022**, arXiv:2210.04505. [\[CrossRef\]](#)
27. Huang, C.; Huang, J.; Liu, X. Cross-Silo Federated Learning: Challenges and Opportunities. *arXiv* **2022**, arXiv:2206.12949. [\[CrossRef\]](#)
28. AbdulRahman, S.; Tout, H.; Ould-Slimane, H.; Mourad, A.; Talhi, C.; Guizani, M. A survey on federated learning: The journey from centralized to distributed on-site learning and beyond. *IEEE Internet Things J.* **2020**, *8*, 5476–5497. [\[CrossRef\]](#)
29. Zhang, T.; Gao, L.; He, C.; Zhang, M.; Krishnamachari, B.; Avestimehr, A.S. Federated Learning for Internet of Things: Applications, Challenges, and Opportunities. *IEEE Internet Things Mag.* **2022**, *5*, 24–29. [\[CrossRef\]](#)
30. Pandya, S.; Srivastava, G.; Jhaveri, R.; Babu, M.R.; Bhattacharya, S.; Maddikunta, P.K.R.; Mastorakis, S.; Piran, M.J.; Gadekallu, T.R. Federated learning for smart cities: A comprehensive survey. *Sustain. Energy Technol. Assess.* **2023**, *55*, 102987. [\[CrossRef\]](#)
31. Mammen, P.M. Federated Learning: Opportunities and Challenges. *arXiv* **2021**, arXiv:2101.05428. [\[CrossRef\]](#)
32. Wen, J.; Zhang, Z.; Lan, Y.; Cui, Z.; Cai, J.; Zhang, W. A survey on federated learning: Challenges and applications. *Int. J. Mach. Learn. Cybern.* **2023**, *14*, 513–535. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Zhu, H.; Xu, J.; Liu, S.; Jin, Y. Federated Learning on Non-IID Data: A Survey. *Neurocomputing* **2021**, *465*, 371–390. [\[CrossRef\]](#)
34. Ma, X.; Zhu, J.; Lin, Z.; Chen, S.; Qin, Y. A state-of-the-art survey on solving non-IID data in Federated Learning. *Future Gener. Comput. Syst.* **2022**, *135*, 244–258. [\[CrossRef\]](#)
35. Li, Q.; Diao, Y.; Chen, Q.; He, B. Federated learning on non-iid data silos: An experimental study. In Proceedings of the 2022 IEEE 38th International Conference on Data Engineering (ICDE), Kuala Lumpur, Malaysia, 9–12 May 2022. [\[CrossRef\]](#)

36. Yurochkin, M.; Agarwal, M.; Ghosh, S.; Greenewald, K.; Hoang, N.; Khazaeni, Y. Bayesian nonparametric federated learning of neural networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 10–15 June 2019.
37. Lim, W.Y.B.; Luong, N.C.; Hoang, D.T.; Jiao, Y.; Liang, Y.-C.; Yang, Q.; Niyato, D.; Miao, C. Federated learning in mobile edge networks: A comprehensive survey. *IEEE Commun. Surv. Tutor.* **2020**, *22*, 2031–2063. [\[CrossRef\]](#)
38. Li, Q.; Wen, Z.; Wu, Z.; Hu, S.; Wang, N.; Li, Y.; Liu, X.; He, B. A survey on federated learning systems: Vision, hype and reality for data privacy and protection. *IEEE Trans. Knowl. Data Eng.* **2021**, *35*, 3347–3366. [\[CrossRef\]](#)
39. Rahman, K.J.; Ahmed, F.; Akhter, N.; Hasan, M.; Amin, R.; Aziz, K.E.; Islam, A.M.; Mukta, M.S.H.; Islam, A.N. Challenges, applications and design aspects of federated learning: A survey. *IEEE Access* **2021**, *9*, 124682–124700. [\[CrossRef\]](#)
40. Rasha, A.-H.; Li, T.; Huang, W.; Gu, J.; Li, C. Federated learning in smart cities: Privacy and security survey. *Inf. Sci.* **2023**, *632*, 833–857. [\[CrossRef\]](#)
41. Li, D.; Han, D.; Weng, T.-H.; Zheng, Z.; Li, H.; Liu, H.; Castiglione, A.; Li, K.-C. Blockchain for federated learning toward secure distributed machine learning systems: A systemic survey. *Soft Comput.* **2022**, *26*, 4423–4440. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Qu, Y.; Uddin, M.P.; Gan, C.; Xiang, Y.; Gao, L.; Yearwood, J. Blockchain-enabled federated learning: A survey. *ACM Comput. Surv.* **2022**, *55*, 1–35. [\[CrossRef\]](#)
43. Hou, D.; Zhang, J.; Man, K.L.; Ma, J.; Peng, Z. A systematic literature review of blockchain-based federated learning: Architectures, applications and issues. In Proceedings of the 2021 2nd Information Communication Technologies Conference (ICTC), Nanjing, China, 7–9 May 2021. [\[CrossRef\]](#)
44. Hasan, M.K.; Habib, A.A.; Islam, S.; Safie, N.; Ghazal, T.M.; Khan, M.A.; Alzahrani, A.I.; Alalwan, N.; Kadry, S.; Masood, A. Federated learning enables 6G communication technology: Requirements, applications, and integrated with intelligence framework. *Alex. Eng. J.* **2024**, *91*, 658–668. [\[CrossRef\]](#)
45. Liu, Y.; Kang, Y.; Zou, T.; Pu, Y.; He, Y.; Ye, X.; Ouyang, Y.; Zhang, Y.-Q.; Yang, Q. Vertical Federated Learning: Concepts, Advances, and Challenges. *IEEE Trans. Knowl. Data Eng.* **2024**, 1–20. [\[CrossRef\]](#)
46. Kitchenham, B.; Charters, S.M. Guidelines for Performing Systematic Literature Reviews in Software Engineering. 2007. Available online: https://www.researchgate.net/profile/Barbara-Kitchenham/publication/302924724_Guidelines_for_performing_Systematic_Literature_Reviews_in_Software_Engineering/links/61712932766c4a211c03a6f7/Guidelines-for-performing-Systematic-Literature-Reviews-in-Software-Engineering.pdf (accessed on 22 March 2024).
47. Qiao, D.; Guo, S.; Liu, D.; Long, S.; Zhou, P.; Li, Z. Adaptive federated deep reinforcement learning for proactive content caching in edge computing. *IEEE Trans. Parallel Distrib. Syst.* **2022**, *33*, 4767–4782. [\[CrossRef\]](#)
48. Zhang, J.; Guo, S.; Qu, Z.; Zeng, D.; Zhan, Y.; Liu, Q.; Akerkar, R. Adaptive federated learning on non-iid data with resource constraint. *IEEE Trans. Comput.* **2021**, *71*, 1655–1667. [\[CrossRef\]](#)
49. Tu, K.; Zheng, S.; Wang, X.; Hu, X. Adaptive federated learning via mean field approach. In Proceedings of the 2022 IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCoM) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics), Espoo, Finland, 22–25 August 2022; pp. 168–175.
50. Xue, Y.; Klabjan, D.; Luo, Y. Aggregation delayed federated learning. In Proceedings of the 2022 IEEE International Conference on Big Data (Big Data), Osaka, Japan, 17–20 December 2022; pp. 85–94.
51. Chen, Y.; Ning, Y.; Slawski, M.; Rangwala, H. Asynchronous online federated learning for edge devices with non-iid data. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 10–13 December 2020; pp. 15–24.
52. Shen, Y.; Zhou, Y.; Yu, L. Cd2-pfed: Cyclic distillation-guided channel decoupling for model personalization in federated learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 10041–10050.
53. Zhang, W.; Wang, X.; Zhou, P.; Wu, W.; Zhang, X. Client selection for federated learning with non-iid data in mobile edge computing. *IEEE Access* **2021**, *9*, 24462–24474. [\[CrossRef\]](#)
54. Xiao, Y.; Shu, J.; Jia, X.; Huang, H. Clustered federated multi-task learning with non-iid data. In Proceedings of the 2021 IEEE 27th International Conference on Parallel and Distributed Systems (ICPADS), Beijing, China, 14–16 December 2021; pp. 50–57.
55. Chen, Z.; Wu, Z.; Wu, X.; Zhang, L.; Zhao, J.; Yan, Y.; Zheng, Y. Contractible regularization for federated learning on non-iid data. In Proceedings of the 2022 IEEE International Conference on Data Mining (ICDM), Orlando, FL, USA, 28 November–1 December 2022; pp. 61–70.
56. Li, Z.; He, Y.; Yu, H.; Kang, J.; Li, X.; Xu, Z.; Niyato, D. Data heterogeneity-robust federated learning via group client selection in industrial iot. *IEEE Internet Things J.* **2022**, *9*, 17844–17857. [\[CrossRef\]](#)
57. Sun, Y.; Zhou, S.; Gündüz, D. Energy-aware analog aggregation for federated learning with redundant data. In Proceedings of the ICC 2020–2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 7–11 June 2020; pp. 1–7.
58. Shullary, M.H.; Abdellatif, A.A.; Massoudn, Y. Energy-efficient active federated learning on non-iid data. In Proceedings of the 2022 IEEE 65th International Midwest Symposium on Circuits and Systems (MWSCAS), Fukuoka, Japan, 7–10 August 2022; pp. 1–4.
59. Wu, H.; Wang, P. Fast-convergent federated learning with adaptive weighting. *IEEE Trans. Cogn. Commun. Netw.* **2021**, *7*, 1078–1088. [\[CrossRef\]](#)

60. Gong, Y.; Li, Y.; Freris, N.M. Fedadmm: A robust federated deep learning framework with adaptivity to system heterogeneity. In Proceedings of the 2022 IEEE 38th International Conference on Data Engineering (ICDE), Kuala Lumpur, Malaysia, 9–12 May 2022; pp. 2575–2587.
61. Idrissi, M.J.; Berrada, I.; Noubir, G. Fedbs: Learning on non-iid data in federated learning using batch normalization. In Proceedings of the 2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI), Washington, DC, USA, 1–3 November 2021; pp. 861–867.
62. Gao, L.; Fu, H.; Li, L.; Chen, Y.; Xu, M.; Xu, C.-Z. Feddc: Federated learning with non-iid data via local drift decoupling and correction. In Proceedings of the IEEE/CVF Conference on Computer Vision And Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 10112–10121.
63. Zou, S.; Xiao, M.; Xu, Y.; An, B.; Zheng, J. Feddcs: Federated learning framework based on dynamic client selection. In Proceedings of the 2021 IEEE 18th International Conference on Mobile Ad Hoc and Smart Systems (MASS), Denver, CO, USA, 4–7 October 2021; pp. 627–632.
64. Kesanapalli, S.A.; Bharath, B. Federated algorithm with bayesian approach: Omni-fedge. In Proceedings of the ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021; pp. 3075–3079.
65. Gkillas, A.; Ampeliotis, D.; Berberidis, K. Federated dictionary learning from non-iid data. In Proceedings of the 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP), Nafplio, Greece, 26–29 June 2022; pp. 1–5.
66. Zhang, L.; Luo, Y.; Bai, Y.; Du, B.; Duan, L.-Y. Federated learning for noniid data via unified feature learning and optimization objective alignment. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 4420–4428.
67. Briggs, C.; Fan, Z.; Andras, P. Federated learning with hierarchical clustering of local updates to improve training on non-iid data. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; pp. 1–9.
68. Zhao, Z.; Feng, C.; Hong, W.; Jiang, J.; Jia, C.; Quek, T.Q.S.; Peng, M. Federated learning with non-iid data in wireless networks. *IEEE Trans. Wirel. Commun.* **2021**, *21*, 1927–1942. [[CrossRef](#)]
69. Li, X.-C.; Xu, Y.-C.; Song, S.; Li, B.; Li, Y.; Shao, Y.; Zhan, D.-C. Federated learning with position-aware neurons. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 10082–10091.
70. Mao, Z.; Dai, W.; Li, C.; Xu, Y. Fedexg: Federated learning with model exchange. In Proceedings of the 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Seville, Spain, 12–14 October 2020; pp. 1–5.
71. Shang, X.; Lu, Y.; Cheung, Y.-M.; Wang, H. Fedic: Federated learning on non-iid and long-tailed data via calibrated distillation. In Proceedings of the 2022 IEEE International Conference on Multimedia and Expo (ICME), Taipei, Taiwan, 18–22 July 2022; pp. 1–6.
72. Lian, Z.; Liu, W.; Cao, J.; Zhu, Z.; Zhou, X. Fednorm: An efficient federated learning framework with dual heterogeneity coexistence on edge intelligence systems. In Proceedings of the 2022 IEEE 40th International Conference on Computer Design (ICCD), Olympic Valley, CA, USA, 23–26 October 2022; pp. 619–626.
73. Zhu, Y.; Markos, C.; Zhao, R.; Zheng, Y.; James, J. Fedova: One-vs-all training method for federated learning with non-iid data. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 18–22 July 2021; pp. 1–7.
74. Nguyen, D.-V.; Tran, A.-K.; Zettsu, K. Fedprob: An aggregation method based on feature probability distribution for federated learning on non-iid data. In Proceedings of the 2022 IEEE International Conference on Big Data (Big Data), Osaka, Japan, 17–20 December 2022; pp. 2875–2881.
75. Kang, Y.; Li, B.; Zeyl, T. Fedrl: Improving the performance of federated learning with non-iid data. In Proceedings of the GLOBECOM 2022–2022 IEEE Global Communications Conference, Rio de Janeiro, Brazil, 4–8 December 2022; pp. 3023–3028.
76. Fan, Z.; Wang, Y.; Yao, J.; Lyu, L.; Zhang, Y.; Tian, Q. Fedskip: Combatting statistical heterogeneity with federated skip aggregation. In Proceedings of the 2022 IEEE International Conference on Data Mining (ICDM), Orlando, FL, USA, 28 November–1 December 2022; pp. 131–140.
77. Duan, M.; Liu, D.; Ji, X.; Wu, Y.; Liang, L.; Chen, X.; Tan, Y.; Ren, A. Flexible clustered federated learning for clientlevel data distribution shift. *IEEE Trans. Parallel Distrib. Syst.* **2021**, *33*, 2661–2674.
78. Zhao, J.; Li, R.; Wang, H.; Xu, Z. Hotfed: Hot start through self-supervised learning in federated learning. In Proceedings of the 2021 IEEE 23rd International Conference on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys), Haikou, China, 20–22 December 2021; pp. 149–156.
79. Yoshida, N.; Nishio, T.; Morikura, M.; Yamamoto, K.; Yonetani, R. Hybridfl for wireless networks: Cooperative learning mechanism using non-iid data. In Proceedings of the ICC 2020–2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 7–11 June 2020; pp. 1–7.
80. Cetinkaya, E.; Akin, M.; Sagioglu, S. Improving performance of federated learning based medical image analysis in non-iid settings using image augmentation. In Proceedings of the 2021 International Conference on Information Security and Cryptology (ISCTURKEY), Ankara, Turkey, 2–3 December 2021; pp. 69–74.
81. Lin, X.; Pan, J.; Xu, J.; Chen, Y.; Zhuo, C. Lithography hotspot detection via heterogeneous federated learning with local adaptation. In Proceedings of the 2022 27th Asia and South Pacific Design Automation Conference (ASP-DAC), Taipei, Taiwan, 17–20 January 2022; pp. 166–171.

82. Feng, C.; Yang, H.H.; Hu, D.; Zhao, Z.; Quek, T.Q.; Min, G. Mobility-aware cluster federated learning in hierarchical wireless networks. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 8441–8458. [\[CrossRef\]](#)
83. Cai, S.; Zhao, Y.; Liu, Z.; Qiu, C.; Wang, X.; Hu, Q. Multi-granularity weighted federated learning in heterogeneous mobile edge computing systems. In Proceedings of the 2022 IEEE 42nd International Conference on Distributed Computing Systems (ICDCS), Bologna, Italy, 10–13 July 2022; pp. 436–446.
84. Wu, H.; Wang, P. Node selection toward faster convergence for federated learning on non-iid data. *IEEE Trans. Netw. Sci. Eng.* **2022**, *9*, 3099–3111. [\[CrossRef\]](#)
85. Mhaisen, N.; Abdellatif, A.A.; Mohamed, A.; Erbad, A.; Guizani, M. Optimal user-edge assignment in hierarchical federated learning based on statistical properties and network topology constraints. *IEEE Trans. Netw. Sci. Eng.* **2021**, *9*, 55–66. [\[CrossRef\]](#)
86. Wang, H.; Kaplan, Z.; Niu, D.; Li, B. Optimizing federated learning on non-iid data with reinforcement learning. In Proceedings of the IEEE INFOCOM 2020-IEEE Conference on Computer Communications, Toronto, ON, Canada, 6–9 July 2020; pp. 1698–1707.
87. Wu, P.; Imbiriba, T.; Park, J.; Kim, S.; Closas, P. Personalized federated learning over non-iid data for indoor localization. In Proceedings of the 2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Lucca, Italy, 27–30 September 2021; pp. 421–425.
88. Yoo, J.H.; Son, H.M.; Jeong, H.; Jang, E.-H.; Kim, A.Y.; Yu, H.Y.; Jeon, H.J.; Chung, T.-M. Personalized federated learning with clustering: Non-iid heart rate variability data application. In Proceedings of the 2021 International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Republic of Korea, 20–22 October 2021; pp. 1046–1051.
89. Lian, Z.; Zeng, Q.; Su, C. Privacy-preserving blockchain-based global data sharing for federated learning with non-iid data. In Proceedings of the 2022 IEEE 42nd International Conference on Distributed Computing Systems Workshops (ICDCSW), Bologna, Italy, 10 July 2022; pp. 193–198.
90. Zhu, Y.; Zhang, S.; Liu, Y.; Niyato, D.; James, J. Robust federated learning approach for travel mode identification from non-iid gps trajectories. In Proceedings of the 2020 IEEE 26th International Conference on Parallel and Distributed Systems (ICPADS), Hong Kong, China, 2–4 December 2020; pp. 585–592.
91. Zhang, Z.; Ma, S.; Nie, J.; Wu, Y.; Yan, Q.; Xu, X.; Niyato, D. Semi-supervised federated learning with noniid data: Algorithm and system design. In Proceedings of the 2021 IEEE 23rd International Conference on High Performance Computing & Communications; 7th International Conference on Data Science & Systems; 19th International Conference on Smart City; 7th International Conference on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys), Haikou, China, 20–22 December 2021; pp. 157–164.
92. Zacccone, R.; Rizzardi, A.; Caldarella, D.; Ciccone, M.; Caputo, B. Speeding up heterogeneous federated learning with sequentially trained superclients. In Proceedings of the 2022 26th International Conference on Pattern Recognition (ICPR), Montreal, QC, Canada, 21–25 August 2022; pp. 3376–3382.
93. Zhou, Z.; Li, Y.; Ren, X.; Yang, S. Towards efficient and stable k-asynchronous federated learning with unbounded stale gradients on non-iid data. *IEEE Trans. Parallel Distrib. Syst.* **2022**, *33*, 3291–3305. [\[CrossRef\]](#)
94. Chen, S.; Li, B. Towards optimal multi-modal federated learning on noniid data with hierarchical gradient blending. In Proceedings of the IEEE INFOCOM 2022-IEEE Conference on Computer Communications, London, UK, 2–5 May 2022; pp. 1469–1478.
95. Mo, K.; Chen, C.; Li, J.; Xu, H.; Xue, C.J. Two-dimensional learning rate decay: Towards accurate federated learning with non-iid data. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 18–22 July 2021; pp. 1–7.
96. Mestoukirdi, M.; Zecchin, M.; Gesbert, D.; Li, Q.; Gresset, N. User-centric federated learning. In Proceedings of the 2021 IEEE Globecom Workshops (GC Wkshps), Madrid, Spain, 7–11 December 2021; pp. 1–6.
97. Jeong, Y.; Kim, T. A cluster-driven adaptive training approach for federated learning. *Sensors* **2022**, *22*, 7061. [\[CrossRef\]](#)
98. Hu, K.; Wu, J.; Weng, L.; Zhang, Y.; Zheng, F.; Pang, Z.; Xia, M. A novel federated learning approach based on the confidence of federated kalman filters. *Int. J. Mach. Learn. Cybern.* **2021**, *12*, 3607–3627. [\[CrossRef\]](#)
99. Ma, T.; Mao, B.; Chen, M. A two-phase half-async method for heterogeneityaware federated learning. *Neurocomputing* **2022**, *485*, 134–154. [\[CrossRef\]](#)
100. Gong, B.; Xing, T.; Liu, Z.; Wang, J.; Liu, X. Adaptive clustered federated learning for heterogeneous data in edge computing. *Mob. Netw. Appl.* **2022**, *27*, 1520–1530. [\[CrossRef\]](#)
101. Wang, L.; Xu, S.; Wang, X.; Zhu, Q. Addressing class imbalance in federated learning. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 2–9 February 2021; Volume 35, pp. 10165–10173.
102. Ou, J.; Shen, Y.; Wang, F.; Liu, Q.; Zhang, X.; Lv, H. Aggenhance: Aggregation enhancement by class interior points in federated learning with noniid data. *ACM Trans. Intell. Syst. Technol.* **2022**, *13*, 1–25. [\[CrossRef\]](#)
103. Fu, Y.; Liu, X.; Tang, S.; Niu, J.; Huang, Z. Cic-fl: Enabling class imbalanceaware clustered federated learning over shifted distributions. In Proceedings of the Database Systems for Advanced Applications: 26th International Conference, DASFAA 2021, Taipei, Taiwan, 11–14 April 2021; Springer: Berlin/Heidelberg, Germany, 2021. Part I 26. pp. 37–52.
104. Hu, F.; Zhou, W.; Liao, K.; Li, H. Contribution-and participation-based federated learning on non-iid data. *IEEE Intell. Syst.* **2022**, *37*, 35–43. [\[CrossRef\]](#)
105. Chen, A.; Fu, Y.; Wang, L.; Duan, G. Dwfed: A statistical-heterogeneitybased dynamic weighted model aggregation algorithm for federated learning. *Front. Neurobotics* **2022**, *16*, 1041553. [\[CrossRef\]](#) [\[PubMed\]](#)

106. Yu, F.; Zhang, W.; Qin, Z.; Xu, Z.; Wang, D.; Liu, C.; Tian, Z.; Chen, X. Fed2: Feature-aligned federated learning. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, Virtual Event, Singapore, 14–18 August 2021; pp. 2066–2074.
107. Duan, J.-H.; Li, W.; Lu, S. Feddna: Federated learning with decoupled normalization-layer aggregation for non-iid data. In Proceedings of the Machine Learning and Knowledge Discovery in Databases. Research Track: European Conference, ECML PKDD 2021, Bilbao, Spain, 13–17 September 2021; Springer: Berlin/Heidelberg, Germany, 2021. Part I 21. pp. 722–737.
108. Lu, C.; Deng, S.; Wu, Y.; Zhou, H.; Ma, W. Federated learning based on optics clustering optimization. *Discret. Dyn. Nat. Soc.* **2022**, *2022*, 7151373. [\[CrossRef\]](#)
109. Jamali-Rad, H.; Abdizadeh, M.; Singh, A. Federated learning with taskonomy for non-iid data. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *34*, 8719–8730. [\[CrossRef\]](#)
110. Yu, P.; Liu, Y. Federated object detection: Optimizing object detection model with federated learning. In Proceedings of the 3rd International Conference on Vision, Image and Signal Processing, Vancouver, BC, Canada, 26–28 August 2019; pp. 1–6.
111. Ni, X.; Shen, X.; Zhao, H. Federated optimization via knowledge codistillation. *Expert Syst. Appl.* **2022**, *191*, 116310. [\[CrossRef\]](#)
112. Jiang, C.; Yin, K.; Xia, C.; Huang, W. Fedhgcdroid: An adaptive multidimensional federated learning for privacy-preserving android malware classification. *Entropy* **2022**, *24*, 919. [\[CrossRef\]](#)
113. Li, X.-C.; Zhan, D.-C.; Shao, Y.; Li, B.; Song, S. Fedphp: Federated personalization with inherited private models. In Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Bilbao, Spain, 13–17 September 2021; pp. 587–602.
114. Li, X.-C.; Zhan, D.-C. Fedrs: Federated learning with restricted softmax for label distribution non-iid data. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, Virtual Event, Singapore, 14–18 August 2021; pp. 995–1005.
115. Chen, M.; Mao, B.; Ma, T. Fedrsa: A staleness-aware asynchronous federated learning algorithm with non-iid data. *Future Gener. Comput. Syst.* **2021**, *120*, 1–12. [\[CrossRef\]](#)
116. Agrawal, S.; Sarkar, S.; Alazab, M.; Maddikunta, P.K.R.; Gadekallu, T.R.; Pham, Q.-V. Genetic cfl: Hyperparameter optimization in clustered federated learning. *Comput. Intell. Neurosci.* **2021**, *2021*, 7156420. [\[CrossRef\]](#)
117. Cai, Y.; Xi, W.; Shen, Y.; Peng, Y.; Song, S.; Zhao, J. High-efficient hierarchical federated learning on non-iid data with progressive collaboration. *Future Gener. Comput. Syst.* **2022**, *137*, 111–128. [\[CrossRef\]](#)
118. Mou, Y.; Geng, J.; Welten, S.; Rong, C.; Decker, S.; Beyan, O. Optimized federated learning on class-biased distributed data sources. In Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Bilbao, Spain, 13–17 September 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 146–158.
119. Zhong, J.; Wu, Y.; Ma, W.; Deng, S.; Zhou, H. Optimizing multi-objective federated learning on non-iid data with improved nsga-iii and hierarchical clustering. *Symmetry* **2022**, *14*, 1070. [\[CrossRef\]](#)
120. Anaissi, A.; Suleiman, B.; Alyassine, W. Personalised federated learning framework for damage detection in structural health monitoring. *J. Civ. Struct. Health Monit.* **2023**, *13*, 295–308. [\[CrossRef\]](#)
121. Tian, P.; Chen, Z.; Yu, W.; Liao, W. Towards asynchronous federated learning based threat detection: A dc-adam approach. *Comput. Secur.* **2021**, *108*, 102344. [\[CrossRef\]](#)
122. Liu, T.; Ding, J.; Wang, T.; Pan, M.; Chen, M. Towards fast and accurate federated learning with non-iid data for cloud-based iot applications. *J. Circuits Syst. Comput.* **2022**, *31*, 2250235. [\[CrossRef\]](#)
123. Gong, Q.; Ruan, H.; Chen, Y.; Su, X. Cloudyfl: A cloudlet-based federated learning framework for sensing user behavior using wearable devices. In Proceedings of the 6th International Workshop on Embedded and Mobile Deep Learning, Portland, Oregon, 1 July 2022; pp. 13–18.
124. Zhu, S.; Qi, Q.; Zhuang, Z.; Wang, J.; Sun, H.; Liao, J. Fednkd: A dependable federated learning using fine-tuned random noise and knowledge distillation. In Proceedings of the 2022 International Conference on Multimedia Retrieval, Newark, NJ, USA, 27–30 June 2022; pp. 185–193.
125. Yang, L.; Huang, J.; Lin, W.; Cao, J. Personalized federated learning on non-iid data via group-based meta-learning. *ACM Trans. Knowl. Discov. Data* **2023**, *17*, 1–20. [\[CrossRef\]](#)
126. Zhou, C.; Tian, H.; Zhang, H.; Zhang, J.; Dong, M.; Jia, J. Tea-fed: Timeefficient asynchronous federated learning for edge computing. In Proceedings of the 18th ACM International Conference on Computing Frontiers, Virtual Event, Italy, 11–13 May 2021; pp. 30–37.
127. Huang, X.; Chen, Z.; Chen, Q.; Zhang, J. Federated learning based qosaware caching decisions in fog-enabled internet of things networks. *Digit. Commun. Netw.* **2023**, *9*, 580–589. [\[CrossRef\]](#)
128. Cao, M.; Zhang, Y.; Ma, Z.; Zhao, M. C2s: Class-aware client selection for effective aggregation in federated learning. *High-Confid. Comput.* **2022**, *2*, 100068. [\[CrossRef\]](#)
129. Baccarelli, E.; Scarpiniti, M.; Momenzadeh, A.; Ahrabi, S.S. Afafed—Asynchronous fair adaptive federated learning for iot stream applications. *Comput. Commun.* **2022**, *195*, 376–402. [\[CrossRef\]](#)
130. Yeganeh, Y.; Farshad, A.; Navab, N.; Albarqouni, S. Inverse distance aggregation for federated learning with non-iid data. In Proceedings of the Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning: Second MICCAI Workshop, DART 2020, and First MICCAI Workshop, DCL 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, 4–8 October 2020; Springer: Berlin/Heidelberg, Germany; pp. 150–159.

131. Wei, B.; Li, J.; Liu, Y.; Wang, W. Federated learning for non-iid data: From theory to algorithm. In Proceedings of the Pacific Rim International Conference on Artificial Intelligence, Hanoi, Vietnam, 8–12 November 2021; Springer: Berlin/Heidelberg, Germany; pp. 33–48.
132. Wang, J.; Huang, Z.; Xiao, J. Fedsmart: An auto updating federated learning optimization mechanism. In Proceedings of the Asia-Pacific Web (APWeb) and WebAge Information Management (WAIM) Joint International Conference on Web and Big Data, Tianjin, China, 18–20 September 2020; Springer: Berlin/Heidelberg, Germany; pp. 716–724.
133. Khan, M.I.; Jafaritadi, M.; Alhoniemi, E.; Kontio, E.; Khan, S.A. Adaptive weight aggregation in federated learning for brain tumor segmentation. In Proceedings of the International MICCAI Brainlesion Workshop, Virtual Event, 27 September 2021; Springer: Berlin/Heidelberg, Germany; pp. 455–469.
134. Gudur, G.K.; Perepu, S.K. Resource-constrained federated learning with heterogeneous labels and models for human activity recognition. In Proceedings of the International Workshop on Deep Learning for Human Activity Recognition, Kyoto, Japan, 8 January 2021; Springer: Berlin/Heidelberg, Germany; pp. 57–69.
135. Zeng, S.; Li, Z.; Yu, H.; He, Y.; Xu, Z.; Niyato, D.; Yu, H. Heterogeneous federated learning via grouped sequential-to-parallel training. In Proceedings of the International Conference on Database Systems for Advanced Applications, Virtual Event, 11–14 April 2022; Springer: Berlin/Heidelberg, Germany; pp. 455–471.
136. Dong, X.; Zhang, S.Q.; Li, A.; Kung, H. Sphered: Hyperspherical federated learning. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 165–184.
137. Caldarola, D.; Caputo, B.; Ciccone, M. Improving generalization in federated learning by seeking flat minima. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer: Berlin/Heidelberg, Germany; pp. 654–672.
138. Lu, C.; Ma, W.; Wang, R.; Deng, S.; Wu, Y. Federated learning based on stratified sampling and regularization. *Complex Intell. Syst.* **2023**, *9*, 2081–2099. [[CrossRef](#)]
139. Anaissi, A.; Suleiman, B.; Alyassine, W. A personalized federated learning algorithm for one-class support vector machine: An application in anomaly detection. In Proceedings of the International Conference on Computational Science, London, UK, 21–23 June 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 373–379.
140. Li, X.; Li, Y.; Li, S.; Zhou, Y.; Chen, C.; Zheng, Z. A unified federated dnns framework for heterogeneous mobile devices. *IEEE Internet Things J.* **2021**, *9*, 1737–1748. [[CrossRef](#)]
141. Becking, D.; Kirchhoffer, H.; Tech, G.; Haase, P.; Muller, K.; Schwarz, H.; Samek, W. Adaptive differential filters for fast and communication-efficient federated learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 3367–3376.
142. Nader, B.; Jiahui, H.; Hui, Z.; Xin, L. Adaptive federated dropout: Improving communication efficiency and generalization for federated learning. In Proceedings of the IEEE Conference on Computer Communications Workshops, Vancouver, BC, Canada, 10–13 May 2021; pp. 1–6.
143. 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), Toronto, ON, Canada, 6–11 June 2021; pp. 3110–3114.
144. Lian, Z.; Cao, J.; Zuo, Y.; Liu, W.; Zhu, Z. Agqfl: Communication-efficient federated learning via automatic gradient quantization in edge heterogeneous systems. In Proceedings of the 2021 IEEE 39th International Conference on Computer Design (ICCD), Storrs, CT, USA, 24–27 October 2021; pp. 551–558.
145. Mahmoudi, A.; Júnior, J.M.B.D.S.; Ghadikolaei, H.S.; Fischione, C. Alaq: Adaptive lazily aggregated quantized gradient. In Proceedings of the 2022 IEEE Globecom Workshops (GC Wkshps), Rio de Janeiro, Brazil, 4–8 December 2022; pp. 1828–1833.
146. Elmahallawy, M.; Luo, T. Asyncfleo: Asynchronous federated learning for leo satellite constellations with high-altitude platforms. In Proceedings of the 2022 IEEE International Conference on Big Data (Big Data), Osaka, Japan, 17–20 December 2022; pp. 5478–5487.
147. Cho, Y.J.; Gupta, S.; Joshi, G.; Ya, O. Bandit-based communicationefficient client selection strategies for federated learning. In Proceedings of the 2020 54th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 1–4 November 2020; pp. 1066–1069.
148. Sattler, F.; Marban, A.; Rischke, R.; Samek, W. Cfd: Communicationefficient federated distillation via soft-label quantization and delta coding. *IEEE Trans. Netw. Sci. Eng.* **2021**, *9*, 2025–2038. [[CrossRef](#)]
149. Luping, W.; Wei, W.; Bo, L. Cmfl: Mitigating communication overhead for federated learning. In Proceedings of the 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS), Dallas, TX, USA, 7–10 July 2019; pp. 954–964.
150. Xie, R.; Zhou, X. Communication efficient federated learning framework with local momentum. In Proceedings of the 2022 15th International Conference on Human System Interaction (HSI), Melbourne, Australia, 28–31 July 2022; pp. 1–6.
151. Seo, S.; Ko, S.-W.; Park, J.; Kim, S.-L.; Bennis, M. Communication-efficient and personalized federated lottery ticket learning. In Proceedings of the 2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Lucca, Italy, 27–30 September 2021; pp. 581–585.
152. Li, C.; Li, G.; Varshney, P.K. Communication-efficient federated learning based on compressed sensing. *IEEE Internet Things J.* **2021**, *8*, 15531–15541. [[CrossRef](#)]

153. Liu, Y.; Kumar, N.; Xiong, Z.; Lim, W.Y.B.; Kang, J.; Niyato, D. Communication efficient federated learning for anomaly detection in industrial internet of things. In Proceedings of the GLOBECOM 2020–2020 IEEE Global Communications Conference, Taipei, Taiwan, 7–11 December 2020; pp. 1–6.
154. Lu, Y.; Huang, X.; Zhang, K.; Maharjan, S.; Zhang, Y. Communication-efficient federated learning for digital twin edge networks in industrial iot. *IEEE Trans. Ind. Inform.* **2020**, *17*, 5709–5718. [[CrossRef](#)]
155. Jeon, Y.-S.; Amiri, M.M.; Lee, N. Communication-efficient federated learning over mimo multiple access channels. *IEEE Trans. Commun.* **2022**, *70*, 6547–6562. [[CrossRef](#)]
156. 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), Montreal, QC, Canada, 14–23 June 2021; pp. 1–6.
157. Chen, Y.; Blum, R.S.; Sadler, B.M. Communication-efficient federated learning using censored heavy ball descent. *IEEE Trans. Signal Inf. Process. Over Netw.* **2022**, *8*, 983–996. [[CrossRef](#)]
158. Yue, K.; Jin, R.; Wong, C.-W.; Dai, H. Communication-efficient federated learning via predictive coding. *IEEE J. Sel. Top. Signal Process.* **2022**, *16*, 369–380. [[CrossRef](#)]
159. Chen, C.; Xu, H.; Wang, W.; Li, B.; Chen, L.; Zhang, G. Communication-efficient federated learning with adaptive parameter freezing. In Proceedings of the 2021 IEEE 41st International Conference on Distributed Computing Systems (ICDCS), Washington, DC, USA, 7–10 July 2021; pp. 1–11.
160. Yang, Y.; Zhang, Z.; Yang, Q. Communication-efficient federated learning with binary neural networks. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 3836–3850. [[CrossRef](#)]
161. Zhou, Y.; Ye, Q.; Lv, J. Communication-efficient federated learning with compensated overlap-fedavg. *IEEE Trans. Parallel Distrib. Syst.* **2021**, *33*, 192–205. [[CrossRef](#)]
162. Krouka, M.; Elgabli, A.; Issaid, C.B.; Bennis, M. Communication-efficient federated learning: A second order newton-type method with analog over-the-air aggregation. *IEEE Trans. Green Commun. Netw.* **2022**, *6*, 1862–1874. [[CrossRef](#)]
163. Shi, Z.; Eryilmaz, A. Communication-efficient subspace methods for highdimensional federated learning. In Proceedings of the 2021 17th International Conference on Mobility, Sensing and Networking (MSN), Exeter, UK, 13–15 December 2021; pp. 543–550.
164. Zhang, X.; Liu, Y.; Liu, J.; Argyriou, A.; Han, Y. D2d-assisted federated learning in mobile edge computing networks. In Proceedings of the 2021 IEEE Wireless Communications and Networking Conference (WCNC), Nanjing, China, 29 March–1 April 2021; pp. 1–7.
165. Liu, Y.; Garg, S.; Nie, J.; Zhang, Y.; Xiong, Z.; Kang, J.; Hossain, M.S. Deep anomaly detection for time-series data in industrial iot: A communication-efficient on-device federated learning approach. *IEEE Internet Things J.* **2020**, *8*, 6348–6358. [[CrossRef](#)]
166. Zhang, W.; Zhou, T.; Lu, Q.; Wang, X.; Zhu, C.; Sun, H.; Wang, Z.; Lo, S.K.; Wang, F.-Y. Dynamic-fusion-based federated learning for covid-19 detection. *IEEE Internet Things J.* **2021**, *8*, 15884–15891. [[CrossRef](#)]
167. Chai, Z.; Chen, Y.; Anwar, A.; Zhao, L.; Cheng, Y.; Rangwala, H. Fedat: A high-performance and communication-efficient federated learning system with asynchronous tiers. In Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, St. Louis, MO, USA, 14–19 November 2021; pp. 1–16.
168. Qu, L.; Song, S.; Tsui, C.-Y. Feddq: Communication-efficient federated learning with descending quantization. In Proceedings of the GLOBECOM 2022–2022 IEEE Global Communications Conference, Rio de Janeiro, Brazil, 4–8 December 2022; pp. 281–286.
169. Mu, Y.; Garg, N.; Ratnarajah, T. Federated learning in massive mimo 6g networks: Convergence analysis and communication-efficient design. *IEEE Trans. Netw. Sci. Eng.* **2022**, *9*, 4220–4234. [[CrossRef](#)]
170. Chen, H.; Huang, S.; Zhang, D.; Xiao, M.; Skoglund, M.; Poor, H.V. Federated learning over wireless iot networks with optimized communication and resources. *IEEE Internet Things J.* **2022**, *9*, 16592–16605. [[CrossRef](#)]
171. Chan, Y.H.; Ngai, E.C. Fedhe: Heterogeneous models and communication-efficient federated learning. In Proceedings of the 2021 17th International Conference on Mobility, Sensing and Networking (MSN), Exeter, UK, 13–15 December 2021; pp. 207–214.
172. Lee, M.-L.; Chou, H.-C.; Chen, Y.-A. Fedsauc: A similarity-aware update control for communication-efficient federated learning in edge computing. In Proceedings of the 2021 Thirteenth International Conference on Mobile Computing and Ubiquitous Network (ICMU), Tokyo, Japan, 17–19 November 2021; pp. 1–6.
173. Chen, D.; Hong, C.S.; Zha, Y.; Zhang, Y.; Liu, X.; Han, Z. Fedsvrg based communication efficient scheme for federated learning in mec networks. *IEEE Trans. Veh. Technol.* **2021**, *70*, 7300–7304. [[CrossRef](#)]
174. Prakash, P.; Ding, J.; Chen, R.; Qin, X.; Shu, M.; Cui, Q.; Guo, Y.; Pan, M. Iot device friendly and communication-efficient federated learning via joint model pruning and quantization. *IEEE Internet Things J.* **2022**, *9*, 13638–13650. [[CrossRef](#)]
175. Ng, J.S.; Lim, W.Y.B.; Dai, H.-N.; Xiong, Z.; Huang, J.; Niyato, D.; Hua, X.-S.; Leung, C.; Miao, C. Joint auction-coalition formation framework for communication-efficient federated learning in uav-enabled internet of vehicles. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 2326–2344. [[CrossRef](#)]
176. Sun, J.; Chen, T.; Giannakis, G.B.; Yang, Q.; Yang, Z. Lazily aggregated quantized gradient innovation for communication-efficient federated learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *44*, 2031–2044. [[CrossRef](#)] [[PubMed](#)]
177. Chu, D.; Jaafar, W.; Yanikomeroglu, H. On the design of communication-efficient federated learning for health monitoring. In Proceedings of the GLOBECOM 2022–2022 IEEE Global Communications Conference, Rio de Janeiro, Brazil, 4–8 December 2022; pp. 1128–1133.

178. Yang, P.; Jiang, Y.; Wang, T.; Zhou, Y.; Shi, Y.; Jones, C.N. Over-the-air federated learning via second-order optimization. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 10560–10575. [\[CrossRef\]](#)
179. Xu, C.; Liu, S.; Huang, Y.; Huang, C.; Zhang, Z. Over-the-air learning rate optimization for federated learning. In Proceedings of the 2021 IEEE International Conference on Communications Workshops (ICC Workshops), Montreal, QC, Canada, 14–23 June 2021; pp. 1–7.
180. Qu, Z.; Guo, S.; Wang, H.; Ye, B.; Wang, Y.; Zomaya, A.Y.; Tang, B. Partial synchronization to accelerate federated learning over relay-assisted edge networks. *IEEE Trans. Mob. Comput.* **2021**, *21*, 4502–4516. [\[CrossRef\]](#)
181. Huang, T.; Ye, B.; Qu, Z.; Tang, B.; Xie, L.; Lu, S. Physical-layer arithmetic for federated learning in uplink mu-mimo enabled wireless networks. In Proceedings of the IEEE INFOCOM 2020-IEEE Conference on Computer Communications, Toronto, ON, Canada, 6–9 July 2020; pp. 1221–1230.
182. Deng, Y.; Lyu, F.; Ren, J.; Zhang, Y.; Zhou, Y.; Yang, Y. Share: Shaping data distribution at edge for communication-efficient hierarchical federated learning. In Proceedings of the 2021 IEEE 41st International Conference on Distributed Computing Systems (ICDCS), Washington, DC, USA, 7–10 July 2021; pp. 24–34.
183. Prakash, P.; Ding, J.; Shu, M.; Wang, J.; Xu, W.; Pan, M. Squafl: Sketch quantization inspired communication efficient federated learning. In Proceedings of the 2021 IEEE/ACM Symposium on Edge Computing (SEC), San Jose, CA, USA, 14–17 December 2021; pp. 350–354.
184. Xu, J.; Du, W.; Jin, Y.; He, W.; Cheng, R. Ternary compression for communication efficient federated learning. *IEEE Trans. Neural. Netw. Learn. Syst.* **2022**, *33*, 1162–1176. [\[CrossRef\]](#)
185. Asad, M.; Moustafa, A.; Rabhi, F.A.; Aslam, M. Thf: 3-way hierarchical framework for efficient client selection and resource management in federated learning. *IEEE Internet Things J.* **2021**, *9*, 11085–11097. [\[CrossRef\]](#)
186. Ozfatura, E.; Ozfatura, K.; Gündüz, D. Time-correlated sparsification for communication-efficient federated learning. In Proceedings of the 2021 IEEE International Symposium on Information Theory (ISIT), Melbourne, Australia, 12–20 July 2021; pp. 461–466.
187. Zhou, X.; Deng, Y.; Xia, H.; Wu, S.; Bennis, M. Time-triggered federated learning over wireless networks. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 11066–11079. [\[CrossRef\]](#)
188. Oh, Y.; Jeon, Y.-S.; Chen, M.; Saad, W. Vector quantized compressed sensing for communication-efficient federated learning. In Proceedings of the 2022 IEEE Globecom Workshops (GC Wkshps), Rio de Janeiro, Brazil, 4–8 December 2022; pp. 365–370.
189. Asad, M.; Moustafa, A.; Aslam, M. Ceep-fl: A comprehensive approach for communication efficiency and enhanced privacy in federated learning. *Appl. Soft Comput.* **2021**, *104*, 107235. [\[CrossRef\]](#)
190. Liu, J.; Xu, H.; Xu, Y.; Ma, Z.; Wang, Z.; Qian, C.; Huang, H. Communication-efficient asynchronous federated learning in resource-constrained edge computing. *Comput. Netw.* **2021**, *199*, 108429. [\[CrossRef\]](#)
191. Chen, M.; Shlezinger, N.; Poor, H.V.; Eldar, Y.C.; Cui, S. Communication efficient federated learning. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2024789118. [\[CrossRef\]](#) [\[PubMed\]](#)
192. Jia, N.; Qu, Z.; Ye, B. Communication-efficient federated learning via quantized clipped sgd. In Proceedings of the Wireless Algorithms, Systems, and Applications: 16th International Conference, WASA 2021, Nanjing, China, 25–27 June 2021; Springer: Berlin/Heidelberg, Germany, 2021. Part I 16. pp. 559–571.
193. Mao, Y.; Zhao, Z.; Yan, G.; Liu, Y.; Lan, T.; Song, L.; Ding, W. Communication-efficient federated learning with adaptive quantization. *ACM Trans. Intell. Syst. Technol.* **2022**, *13*, 1–26. [\[CrossRef\]](#)
194. Cui, Z.; Wen, J.; Lan, Y.; Zhang, Z.; Cai, J. Communication-efficient federated recommendation model based on many-objective evolutionary algorithm. *Expert Syst. Appl.* **2022**, *201*, 116963. [\[CrossRef\]](#)
195. Ji, S.; Jiang, W.; Walid, A.; Li, X. Dynamic sampling and selective masking for communication-efficient federated learning. *IEEE Intell. Syst.* **2021**, *37*, 27–34. [\[CrossRef\]](#)
196. Paragliola, G. Evaluation of the trade-off between performance and communication costs in federated learning scenario. *Future Gener. Comput. Syst.* **2022**, *136*, 282–293. [\[CrossRef\]](#)
197. Yang, K.; Jiang, T.; Shi, Y.; Ding, Z. Federated learning via over-the-air computation. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 2022–2035. [\[CrossRef\]](#)
198. Reiszadeh, 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 twenty third International Conference on Artificial Intelligence and Statistics, PMLR, Online, 26–28 August 2020; pp. 2021–2031.
199. Gorbunov, E.; Burlachenko, K.P.; Li, Z.; Richt, P. Marina: Faster nonconvex distributed learning with compression. In Proceedings of the 38th International Conference on Machine Learning, PMLR, Virtual, 18–24 July 2021; pp. 3788–3798.
200. Gao, H.; Xu, A.; Huang, H. On the convergence of communication-efficient local sgd for federated learning. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 2–9 February 2021; Volume 35, pp. 7510–7518.
201. Ma, L.; Liao, Y.; Zhou, B.; Xi, W. Perhefed: A general framework of personalized federated learning for heterogeneous convolutional neural networks. *World Wide Web* **2022**, *26*, 2027–2049. [\[CrossRef\]](#)
202. Huang, A.; Chen, Y.; Liu, Y.; Chen, T.; Yang, Q. Rpn: A residual pooling network for efficient federated learning. In Proceedings of the ECAI 2020: 24th European Conference on Artificial Intelligence, Santiago de Compostela, Spain, 29 August–8 September 2020; IOS Press: Amsterdam, The Netherlands, 2020; pp. 1223–1229.

203. Chen, Z.; Liao, W.; Hua, K.; Lu, C.; Yu, W. Towards asynchronous federated learning for heterogeneous edge-powered internet of things. *Digit. Commun. Netw.* **2021**, *7*, 317–326. [\[CrossRef\]](#)
204. Huang, W.; Yang, Y.; Chen, M.; Liu, C.; Feng, C.; Poor, H.V. Wireless network optimization for federated learning with model compression in hybrid vlc/rf systems. *Entropy* **2021**, *23*, 1413. [\[CrossRef\]](#)
205. Li, Y.; Wu, C.; Zhong, L.; Yoshinaga, T. A communication-efficient distributed machine learning scheme in vehicular network. In Proceedings of the Conference on Research in Adaptive and Convergent Systems, Virtual Event, Japan, 3–6 October 2022; pp. 92–98.
206. Zhou, P.; Xu, H.; Lee, L.H.; Fang, P.; Hui, P. Are you left out? an efficient and fair federated learning for personalized profiles on wearable devices of inferior networking conditions. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2022**, *6*, 1–25.
207. Deng, Y.; Kamani, M.M.; Mahdavi, M. Distributionally robust federated averaging. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 15111–15122.
208. Chandrasekaran, R.; Ergun, K.; Lee, J.; Nanjunda, D.; Kang, J.; Rosing, T. Fhdnn: Communication efficient and robust federated learning for aiot networks. In Proceedings of the 59th ACM/IEEE Design Automation Conference, San Francisco, CA, USA, 10–14 July 2022; pp. 37–42.
209. Chen, X.; Li, X.; Li, P. Toward communication efficient adaptive gradient method. In Proceedings of the 2020 ACM-IMS on Foundations of Data Science Conference, Virtual Event, USA, 19–20 October 2020; pp. 119–128.
210. Kundu, K.; Jaja, J. Fednet2net: Saving communication and computations in federated learning with model growing. In Proceedings of the International Conference on Artificial Neural Networks, Bristol, UK, 6–9 September 2022; Springer: Berlin/Heidelberg, Germany; pp. 236–247.
211. Yang, J.; Duan, Y.; Qiao, T.; Zhou, H.; Wang, J.; Zhao, W. Prototyping federated learning on edge computing systems. *Front. Comput. Sci.* **2020**, *14*, 146318. [\[CrossRef\]](#)
212. Yang, H.; Liu, J.; Bentley, E.S. Cfedavg: Achieving efficient communication and fast convergence in non-iid federated learning. In Proceedings of the 2021 19th International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks (WiOpt), Philadelphia, PA, USA, 18–21 October 2021; pp. 1–8.
213. Rothchild, D.; Panda, A.; Ullah, E.; Ivkin, N.; Stoica, I.; Braverman, V.; Gonzalez, J.; Arora, R. Fetchsgd: Communication-efficient federated learning with sketching. In Proceedings of the 37th International Conference on Machine Learning, PMLR, Virtual, 13–18 July 2020; pp. 8253–8265.
214. Li, X.; Liu, N.; Chen, C.; Zheng, Z.; Li, H.; Yan, Q. Communication efficient collaborative learning of geo-distributed joint cloud from heterogeneous datasets. In Proceedings of the 2020 IEEE International Conference on Joint Cloud Computing, Oxford, UK, 3–6 August 2020; pp. 22–29.
215. Wen, H.; Wu, Y.; Li, J.; Duan, H. Communication-efficient federated data augmentation on non-iid data. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 3377–3386.
216. Sheikholeslami, S.M.; Rasti-Meymandi, A.; Seyed-Mohammadi, S.J.; Abouei, J.; Plataniotis, K.N. Communication-efficient federated learning for hybrid vlc/rf indoor systems. *IEEE Access* **2022**, *10*, 126479–126493. [\[CrossRef\]](#)
217. Mills, J.; Hu, J.; Min, G. Communication-efficient federated learning for wireless edge intelligence in iot. *IEEE Internet Things J.* **2019**, *7*, 5986–5994. [\[CrossRef\]](#)
218. Zhou, S.; Huo, Y.; Bao, S.; Landman, B.; Gokhale, A. Fedaca: An adaptive communication-efficient asynchronous framework for federated learning. In Proceedings of the 2022 IEEE International Conference on Autonomic Computing and Self Organizing Systems (ACSOS), Virtual, CA, USA, 19–23 September 2022; pp. 71–80.
219. Lit, Z.; Sit, S.; Wang, J.; Xiao, J. Federated split bert for heterogeneous text classification. In Proceedings of the 2022 International Joint Conference on Neural Networks (IJCNN), Padua, Italy, 18–23 July 2022; pp. 1–8.
220. Zhang, X.; Hong, M.; Dhople, S.; Yin, W.; Liu, Y. Fedpd: A federated learning framework with adaptivity to non-iid data. *IEEE Trans. Signal Process.* **2021**, *69*, 6055–6070. [\[CrossRef\]](#)
221. Wu, X.; Yao, X.; Wang, C.-L. Fedscr: Structure-based communication reduction for federated learning. *IEEE Trans. Parallel Distrib. Syst.* **2020**, *32*, 1565–1577. [\[CrossRef\]](#)
222. Li, B.; Jiang, Y.; Sun, W.; Niu, W.; Wang, P. Fedvanet: Efficient federated learning with non-iid data for vehicular ad hoc networks. In Proceedings of the 2021 IEEE Global Communications Conference (GLOBECOM), Madrid, Spain, 7–11 December 2021; pp. 1–6.
223. Shu, J.; Zhang, W.; Zhou, Y.; Cheng, Z.; Yang, L.T. Flas: Computation and communication efficient federated learning via adaptive sampling. *IEEE Trans. Netw. Sci. Eng.* **2021**, *9*, 2003–2014. [\[CrossRef\]](#)
224. Sun, J.; Wang, B.; Duan, L.; Li, S.; Chen, Y.; Li, H. Lotteryfl: Empower edge intelligence with personalized and communication-efficient federated learning. In Proceedings of the 2021 IEEE/ACM Symposium on Edge Computing (SEC), San Jose, CA, USA, 14–17 December 2021; pp. 68–79.
225. Sattler, F.; Wiedemann, S.; Müller, K.-R.; Samek, W. Robust and communication efficient federated learning from non-iid data. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *31*, 3400–3413. [\[CrossRef\]](#)
226. Su, Z.; Wang, Y.; Luan, T.H.; Zhang, N.; Li, F.; Chen, T.; Cao, H. Secure and efficient federated learning for smart grid with edge-cloud collaboration. *IEEE Trans. Ind. Inform.* **2021**, *18*, 1333–1344. [\[CrossRef\]](#)
227. Chen, Z.; Li, D.; Zhao, M.; Zhang, S.; Zhu, J. Semi-federated learning. In Proceedings of the 2020 IEEE Wireless Communications and Networking Conference (WCNC), Seoul, Republic of Korea, 25–28 May 2020; pp. 1–6.

228. Yang, Z.; Sun, Q. A dynamic global backbone updating for communication-efficient personalised federated learning. *Connect. Sci.* **2022**, *34*, 2240–2264. [[CrossRef](#)]
229. Liang, J.; Liu, Z.; Zhou, Z.; Xu, Y. Communication-efficient federated indoor localization with layerwise swapping training-fedavg. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* **2022**, *105*, 1493–1502. [[CrossRef](#)]
230. Abdellatif, A.; Mhaisen, N.; Mohamed, A.; Erbad, A.; Guizani, M.; Dawy, Z.; Nasreddine, W. Communication-efficient hierarchical federated learning for iot heterogeneous systems with imbalanced data. *Future Gener. Comput. Syst.* **2022**, *128*, 406–419. [[CrossRef](#)]
231. Ma, Z.; Zhao, M.; Cai, X.; Jia, Z. Fast-convergent federated learning with class-weighted aggregation. *J. Syst. Archit.* **2021**, *117*, 102125. [[CrossRef](#)]
232. Al-Saedi, A.; Boeva, V.; Casalicchio, E. Fedco: Communication-efficient federated learning via clustering optimization. *Future Internet* **2022**, *14*, 377. [[CrossRef](#)]
233. Mo, Z.; Gao, Z.; Zhao, C.; Lin, Y. Feddq: A communication-efficient federated learning approach for internet of vehicles. *J. Syst. Archit.* **2022**, *131*, 102690. [[CrossRef](#)]
234. Mugunthan, V.; Lin, E.; Gokul, V.; Lau, C.; Kagal, L.; Pieper, S. Fedltn: Federated learning for sparse and personalized lottery ticket networks. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer: Berlin/Heidelberg, Germany; pp. 69–85.
235. Chen, W.; Bhardwaj, K.; Marculescu, R. Fedmax: Mitigating activation divergence for accurate and communication-efficient federated learning. In Proceedings of the Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2020, Ghent, Belgium, 14–18 September 2020; Springer: Berlin/Heidelberg, Germany, 2021. Part II. pp. 348–363.
236. Li, X.; Gong, Y.; Liang, Y.; Wang, L.-E. Personalized federated learning with semisupervised distillation. *Secur. Commun. Netw.* **2021**, *2021*, 3259108. [[CrossRef](#)]
237. Sun, J.; Zeng, X.; Zhang, M.; Li, H.; Chen, Y. Fedmask: Joint computation and communication-efficient personalized federated learning via heterogeneous masking. In Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems, Coimbra, Portugal, 15–17 November 2021; pp. 42–55.
238. Sun, J.; Li, P.; Pu, Y.; Li, H.; Chen, Y. Hermes: An efficient federated learning framework for heterogeneous mobile clients. In Proceedings of the 27th Annual International Conference on Mobile Computing and Networking, New Orleans, LA, USA, 25–29 October 2021; pp. 420–437.
239. Wang, H.; Wang, L. Fedkg: Model-optimized federated learning for local client training with non-iid private data. In Proceedings of the 2021 Ninth International Conference on Advanced Cloud and Big Data (CBD), Xi'an, China, 26–27 March 2022; pp. 51–57.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.