ARTICLE

Federated Semi-Supervised Learning Based on Feature Space Fusion

Zhe Ding[1,2], **Hao Yi**[3,4,*], **Wenrui Xie**[3,4], **Ming Zhang**[1], **Yuxuan Xiao**[1], **Qixu Wang**[1,2], **Qing Chen**[5], **Zhiguang Qin**[1] **and Dajiang Chen**[1]

[1]School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu, China
[2]School of Cybersecurity, Chengdu University of Information Technology, Chengdu, China
[3]China Electronic Products Reliability and Environmental Testing Research Institute, Guangzhou, China
[4]Key Laboratory of the Ministry of Industry and Information Technology for Performance and Reliability Evaluation of Software and Hardware for Information Technology Application Innovation Foundation, Guangzhou, China
[5]Accelink Technologies Co., Ltd., Wuhan, China
*Corresponding Author: Hao Yi. Email: yihao@ceprei.com

ABSTRACT: Federated semi-supervised learning (FSSL) has garnered substantial attention for enabling collaborative global model training across multiple clients to address the scarcity of labeled data and to preserve data privacy. However, FSSL is plagued by formidable challenges stemming from cross-client data heterogeneity, as existing methods fail to achieve effective fusion of feature subspaces across distinct clients. To address this issue, we propose a novel FSSL framework, named FedSPQR, which is explicitly tailored for the label-at-server scenario. On the server side, FedSPQR adopts subspace clustering and fusion method based on the Grassmann manifold to construct a unified global feature space, which is further leveraged to refine the global model. On the client side, the pre-established global feature space acts as a benchmark for aligning the local feature subspaces. Based on the aligned local feature subspaces, integrating self-supervised learning with knowledge distillation facilitates effective local learning to alleviate local bias caused by data heterogeneity. Extensive experiments on two standard public benchmarks confirm that FedSPQR outperforms state-of-the-art (SOTA) baselines by a significant margin.

KEYWORDS: Federated semi-supervised learning; feature space fusion; knowledge distillation

1 Introduction

Federated Learning (FL) has been successfully applied across various fields, including finance and healthcare [1]. Compared with traditional machine learning, FL can train a global model without sharing training data [2]. Therefore, it ensures the security of training data during the training process. However, a major practical challenge is the scarcity of labeled data, which significantly impedes the training effectiveness of FL [3].

FSSL integrates FL with semi-supervised learning (SSL) to train global models by leveraging both the limited labeled and the abundant unlabeled data, all while maintaining data privacy [4]. FSSL scenarios can be divided into two categories: label-at-server and label-at-client scenarios [5,6]. In the label-at-server scenario, the clients possess only unlabeled data and therefore perform local unsupervised learning. The server, which holds the labeled data, conducts supervised learning and aggregates the local models received from the clients [7]. Conversely, in the label-at-client scenario, the clients with both labeled and unlabeled data perform semi-supervised learning, and the server's role is limited to model aggregation [8]. Training

a global model is more complex in the label-at-server scenario than in the label-at-client scenario [9]. Therefore, this paper focuses on the label-at-server scenario. Existing FSSL methods tailored to this setting are confronted with two primary challenges.

- **Data heterogeneity among clients:**Traditional FL models, such as FedAvg [10], are built on the fundamental assumption that data across clients is independent and identically distributed (IID). However, this assumption often fails in practice due to minimal restrictions on participating clients, leading to data heterogeneity (non-IID). Such non-IID data not only causes global model bias and convergence difficulties, but also makes the model prone to getting stuck in local optima [11,12].
- **Training inconsistency and low aggregation efficiency:**In the label-at-server scenario, the server uses labeled data for supervised learning, while the clients employ consistency regularization to extract knowledge from their local unlabeled data. The disparity between these two distinct learning tasks creates a significant performance gap, which reduces the overall training efficiency. Although existing methods leverage semi-supervised learning techniques, such as pseudo-labeling, to improve learning consistency, the aggregation process remains inefficient due to the global model's limited performance in the early training stages [13,14].

Although recent FSSL methods have attempted to address these issues, their approaches often remain inadequate. To cope with data heterogeneity, many methods simply adapt regularization techniques. However, these methods operate primarily on the parameter space of the model and do not explicitly reconcile the underlying feature distribution drift across clients. Consequently, the aggregated global model lacks a coherent and unified feature representation. Regarding training inconsistency, common solutions involve applying SSL techniques like pseudo-labeling on the clients. The global model is aggregated from local models on client sides and updated with the limited labeled data on the server side. So, in the early stages of training, the global model tends to produce low-quality pseudo-labels for unlabeled data. This initiates a vicious cycle: poor global guidance induces noisy local updates, which impair subsequent global aggregation quality and further lead to slow convergence as well as suboptimal performance.

In this paper, we propose a novel federated semi-supervised learning framework, denoted as FedSPQR, to address the aforementioned challenges. To tackle the first challenge, FedSPQR employs subspace clustering and fusion techniques based on Grassmann manifold theory to aggregate feature subspaces from different clients to generate a unified global feature space. This global feature space is then incorporated into the local training on clients to alleviate data heterogeneity. For the second challenge, FedSPQR leverages this global feature space alongside a knowledge distillation mechanism and self-supervised learning based on SimSiam to integrate global knowledge into the local training process, thereby enhancing training consistency. Specifically, the knowledge distillation mechanism addresses heterogeneity in the models' prediction layers, while the fused feature space mitigates heterogeneity in the representation layers. Overall, the main contributions of this paper are as follows:

- We utilize a feature space alignment technique to integrate feature subspaces from all clients, thereby generating a unified global feature space. This global feature space is then incorporated into the local training process on the client sides to mitigate the challenges posed by non-IID data.
- We leverage knowledge distillation mechanism and SimSiam-based self-supervised learning to incorporate global knowledge into the local training on client sides, thereby improving training consistency and aggregation efficiency.
- We conduct extensive evaluations on the performance of our proposed method using the standard CIFAR-10 and Fashion-MNIST datasets. Experimental results clearly demonstrate its superiority over SOTA baselines across key performance metrics.

The remainder of this paper is organized as follows. Section 2 provides an overview of related work. The details of FedSPQR are described in Section 3. In Section 4, two public datasets are used to evaluate the efficiency of FedSPQR.

2 Related Work

In this section, we introduce the related work for federated semi-supervised learning.

2.1 Federated Learning

FL is a distributed machine learning approach that enables model training across multiple nodes (e.g., institutions or devices) while keeping raw data decentralized, thus preserving data privacy and security. In 2017, McMahan et al. proposed FedAvg, which enhances communication efficiency by allowing clients to conduct multiple local SGD updates before aggregation, significantly reducing communication rounds [10]. FedAvg has since become the baseline for numerous follow-up studies aiming to optimize its efficiency further. ConfShield employs attested Confidential Virtual Machines (CVMs) to construct a secure enclave, which prevents cloud service providers from accessing core federated learning (FL) processes and guarantees model confidentiality [14]. In ref. [15], Cao et al. organically integrate data distillation-based initialization, divergence metric strategy and adaptive optimization into the federated learning framework to achieve the decentralized fault diagnosis of pumping units.

Statistical heterogeneity, as a well-recognized obstacle in FL, causes client drift and further degrades model performance by creating a divergence between local and global objectives [16]. To mitigate this issue, several algorithms have been proposed. For example, the FedProx algorithm [11] introduced a proximal term into the local loss function, which mitigates client drift by constraining local updates to remain close to the global model and thus stabilizing training. Similarly, the SCAFFOLD algorithm [17] employs control variables (variance reduction terms) that are exchanged between clients and the server to correct the direction of local updates. This approach has demonstrated significant performance improvements on highly non-IID data. In ref. [18], Durmus et al. proposed a dynamic regularization to fuse the global solution among client sides for reducing communication overhead. In order to optimize the global model in FL, some researchers focus on optimizing parameter aggregation in server side. Beyond these, other approaches focus on optimizing the aggregation process. For example, Chen et al. [19] proposed an aggregation weighting scheme based on theoretical upper bounds of each client's convergence rate for efficient global update.

While these methods improve robustness under non-IID conditions in standard supervised FL, they are not directly designed for scenarios where labeled data is scarce, which is the focus of FSSL. This gap motivates the need for specialized techniques that can jointly address data heterogeneity and the challenges arising from limited supervision.

2.2 Semi-Supervised Learning

The core assumption of SSL is that the data distribution of unlabeled data contains information that can enhance a model's generalization capability [20]. Therefore, SSL leverages a large amount of unlabeled data, in addition to a small set of labeled data, to improve model performance. Classical SSL methods have inspired many FSSL approaches. The Mean-Teacher model [21] employs a teacher-student framework, where the teacher's predictions guide the updates of the student model. Tri-training [22] is a committee-based method that utilizes three models. In this approach, one model is trained on unlabeled data instances where the other two models agree in their predictions.

Building upon traditional SSL, FSSL has yielded substantial achievements. For example, CBAFed [23] maintains a fixed set of information from unlabeled clients and applies class-balanced adaptive thresholds, considering the empirical data distribution to steer training toward class balance. It also incorporates a residual weight connection mechanism during aggregation to improve efficiency. Similarly, pFSSL [24] utilizes helper agents to assist partial-labeled and unlabeled clients and designs an uncertainty-based data-relation metric to improve pseudo-label quality.

In the label-at-server scenario we focus on, these methods still face two difficulties. Firstly, severe data heterogeneity remains problematic because existing alignment or regularization techniques do not explicitly construct a unified feature representation space across clients; Secondly, the inconsistency between super-vised learning on the server and unsupervised learning on clients leads to inefficient aggregation, especially in early training stages when the global model provides low-quality guidance. Our work, FedSPQR, aims to address these specific limitations by proposing a unified global feature space and a consistency-enhanced training framework.

3 Proposed Method

In this section, we will introduce the details of the proposed federated semi-supervised learning model, denoted as FedSPQR.

3.1 Problem Definition

This paper focuses on FSSL in the label-at-server scenario, where limited labeled data is centrally located on the server while vast volumes of unlabeled data are distributed across clients. Under the premise of non-IID among client data and a lack of supervision information at the client side, the core objective is to break the heterogeneity barrier between the global and local feature spaces. The proposed framework, FedSPQR, aims to train a global model using all available data without sharing any training samples between the server and clients, thereby achieving efficient transmission and precise guidance of supervision signals to ultimately enhance the generalization performance and training stability of the global model.

3.2 The Framework of FedSPQR

Fig. 1 illustrates the FedSPQR framework, which employs a teacher-student architecture on both server and clients. The framework consists of two main modules, described below. Clients utilize the teacher model as a repository of global knowledge to supervise the student model's learning from local data. The server, in turn, uses the teacher model to aggregate client models and the student model to distill this aggregated knowledge.

- **Collaborative Learning:** On the server side, FedSPQR first constructs a global feature space from the feature subspaces provided by the clients. Leveraging this global feature space, it then aggregates the local models from the clients to generate a new global model. This global model updates the teacher model, from which global knowledge is distilled into the student model. The student model is further optimized through supervised learning. Finally, the server distributes the updated global feature space and the student model parameters back to the clients.
- **Local Knowledge Learning:** On the client side, FedSPQR first updates the local teacher model with the global model received from the server. The student model then learns local representations through self-supervised learning mechanism, while simultaneously receiving global knowledge from the teacher model via knowledge distillation. To enhance training efficiency, the local feature subspace is aligned with the global feature space provided by the server. Upon completion of local training, the client

generates a feature subspace using Principal Component Analysis (PCA). Finally, the updated student model and the generated feature subspace are returned to the server.
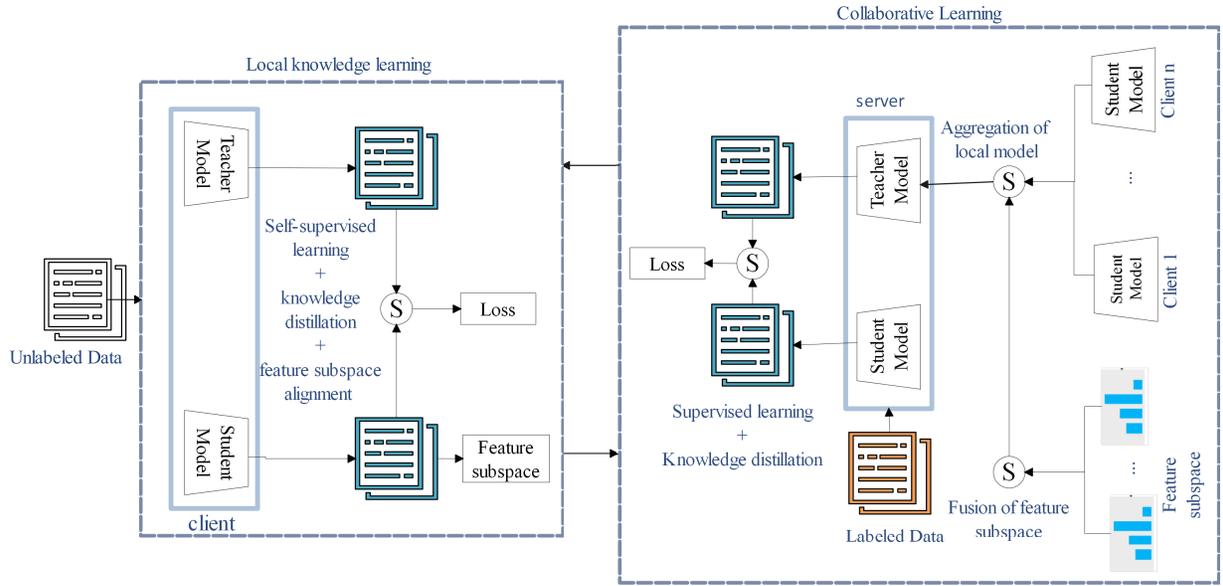

Figure 1: The framework of FedSPQR.

In FedSPQR, the backbone network $f(\cdot)$ for both student and teacher models is ResNet-18. A fully connected layer is used on both client and server sides as the projection head $z(\cdot)$ and prediction head $p(\cdot)$, respectively.

3.3 Collaborative Learning

FedSPQR employs PCA and Grassmann manifold-based alignment technique to aggregate the local models from clients and update the global model. Subsequently, global knowledge is transferred from the teacher model to the student model and refined through a knowledge distillation mechanism.

3.3.1 Extraction of Feature Space for Labeled Data

Traditional FSSL faces the challenge of global model semantic drift. To address this issue, inspired by FedPCA [25], FedSPQR extracts a feature space from the server's labeled data using PCA, which serves as an anchor to provide a stable and accurate guiding signal for both feature space alignment and the knowledge distillation mechanism. Let $F^G \in \mathbb{R}^{S \times d_f}$ denote the feature matrix obtained from the server-side student model's backbone output for the labeled data. The extraction process is detailed below.

$$C^G = \frac{1}{S}(\hat{F}^G)^T \hat{F}^G \tag{1}$$

$$C^G = U^G \Lambda (U^G)^T \tag{2}$$

where $\hat{F}^G$ is centralization of F^G. Finally, FedSPQR selects the first k eigenvectors $U^G = \{u_1, u_2, \cdots, u_k\}$ as the basis of the feature subspace for labeled data.

3.3.2 Generation of Global Feature Space

The global feature space is a low-dimensional feature manifold that is constructed on the server and shared across all clients to align their feature distributions. FedSPQR generates this global feature space by aggregating the client feature subspaces using a K-means clustering algorithm based on Grassmann distance. The key advantages of this approach are as follows:

- FedSPQR enhances the precision of feature fusion by employing Grassmann distance to accurately quantify the geometric discrepancies between client feature subspaces. This approach mitigates the mutual interference of heterogeneous feature subspaces, a common issue in traditional fusion methods. Consequently, FedSPQR ensures that the resulting global feature space more consistently reflects the core feature structures prevalent across the client population.
- FedSPQR aims to adapt to non-IID data and mitigate feature subspace fragmentation. In FL, the non-IID client data can lead to feature subspace fragmentation where the feature subspaces of different clients become scattered. The Grassmann distance-based K-means clusters clients with similar feature distributions into groups. After fusing the feature subspaces within each group, local-global subspaces are generated. These generated local-global subspaces are used to generate the final global feature space. This approach hierarchically unifies feature distributions, gradually reducing the differences between heterogeneous subspaces.

The details of K-means based on Grassmann distance are shown as follow:

- **Initialization:** FedSPQR randomly selects C client subspaces as the initial cluster centers.
- **Assignment step:** FedSPQR calculates the Grassmann distance from each client's feature subspace to each cluster. Subsequently, each client is assigned to the cluster with the smallest distance. The Grassmann distance is defined as Eq. (3).

$$d_G(X, Y) = \sqrt{k - tr(X^T Y Y^T X)} \tag{3}$$

where $X, Y \in \mathbb{R}^{D \times k}$ are the bases of feature subspaces received from two client sides.

- **Updation:** For each group, FedSPQR calculates the Grassmann mean of the feature subspace within the cluster as new clustering center.
- **Generation of global feature space:** FedSPQR calculates the Grassmann mean of all clustering centers as global feature space U_g.

3.3.3 Aggregation of Local Model

FedSPQR employs an aggregation mechanism based on projection distance to generate the global model on the server. The core idea is to quantify the projection differences between the local models from the clients and the global model, which enables the dynamic assignment of aggregation weights. Let $\{U_i\}_{i=1}^N$ and $\{\theta_i\}_{i=1}^N$ be the feature subspaces and local models received from the N clients, respectively, and let U_g denote the global feature space. The aggregation mechanism is formalized in Eq. (6).

$$\omega_i = \frac{\exp\{-\alpha_t \cdot d_G(U_i, U_g)\}}{\sum_j \exp(-\alpha \cdot d_G(U_j, U_g))} \tag{4}$$

$$\alpha_t = \alpha_0 \cdot (1 - e^{(-\beta \cdot t)}) \tag{5}$$

$$\theta_{global} = \sum_j \omega_j \cdot \theta_j \tag{6}$$

where β is a hyperparameter, and t denotes the communication round. The resulting global model θ_{global} is then used to update the teacher model on the server side.

3.3.4 Server Fine-Tuning

FedSPQR refines the server-side teacher model using the labeled data and the global feature space to calibrate the non-IID bias introduced during client model aggregation. This calibration is crucial because the teacher model, aggregated from client parameters, can be skewed by clients with larger datasets. Let $(x_i, y_i)_{i=1}^M$ be the labeled dataset on the server. The refinement involves minimizing a joint loss function, comprising a feature alignment loss (Eq. (7)) and a Cross-Entropy loss (Eq. (8)) on the labeled data.

$$\mathcal{L}_{align} = 1 - \frac{1}{B} \sum_{j=1}^{B} \frac{f(x_j; \theta_{global}) \cdot (U_g U_g^T \cdot f(x_j; \theta_{global}))}{\| f(x_j; \theta_{global}) \|_2 \cdot \| U_g U_g^T \cdot f(x_j; \theta_{global}) \|_2} \tag{7}$$

$$\mathcal{L}_{CE} = -\frac{1}{B} \sum_{i=1}^{B} y_i \cdot \log(\hat{y}_i) \tag{8}$$

where $f(\cdot)$ is the output of teacher model in server side. B is batch size. The total loss function is shown as Eq. (9). y_i and $\hat{y}_i$ are the true label and prediction result of x_i, respectively. Finally FedSPQR employs knowledge distillation mechanism to transfer global knowledge from teacher model to student model. The loss of knowledge distillation is defined as Eq. (10):

$$\mathcal{L}_{finetune} = \mathcal{L}_{CE} + \lambda \cdot \mathcal{L}_{align} \tag{9}$$

$$\mathcal{L}_{KD}1 = \frac{1}{B} \cdot \sum_{j=1}^{B} KL\left(p_{stu}\left(\frac{f_{stu}(x_j)}{T}\right) \| p_{tea}\left(\frac{f_{tea}(x_j)}{T}\right)\right) \times T^2 \tag{10}$$

where $f_{tea}(\cdot)$ and f_{stu} are the backbone output of teacher model and student model in server side, respectively. $z_{tea}(\cdot)$ and $z_{stu}(\cdot)$ are projection output of teacher model and student model, respectively.

3.4 Local Knowledge Learning

FedSPQR enhances local client learning through a unified approach that integrates self-supervised learning, knowledge distillation, and feature subspace alignment. Specifically, it employs self-supervised learning to derive robust features from unlabeled data, leverages knowledge distillation to infuse global insights from the teacher model, and utilizes subspace alignment to optimize the local models.

3.4.1 Representative Learning for Unlabeled Data

FedSPQR employs SimSiam [26], a lightweight self-supervised learning method, to learn features from unlabeled data without requiring negative sample pairs. Its framework is naturally compatible with the server-client architecture of federated learning, making it well-suited for the label-at-server scenario. Additionally, SimSiam helps mitigate overfitting. We adopt its loss function [26], which is defined in Eq. (11).

$$\mathcal{L}_{simsiam} = -\frac{1}{B} \cdot \sum_{j=1}^{B} \left[\frac{p_{stu}(\text{Aug}_w(x_j)) \cdot (z_{tea}(\text{Aug}_s(x_j)))}{\|p_{stu}(\text{Aug}_w(x_j))\| \cdot \|(z_{tea}(\text{Aug}_s(x_j)))\|} \| \right.$$
$$\left. + \frac{p_{tea}(\text{Aug}_s(x_j)) \cdot (z_{stu}(\text{Aug}_w(x_j)))}{\|p_{tea}(\text{Aug}_s(x_j))\| \cdot \|(z_{stu}(\text{Aug}_w(x_j)))\|} \right] \tag{11}$$

where $\text{Aug}_s(x)$ and $\text{Aug}_w(x)$ are weakness and strong augument of unlabeled data x. B is batch size.

On the client side, FedSPQR uses the teacher model to hold global knowledge and the student model to acquire local knowledge. Knowledge distillation is then applied to transfer global insights from the teacher to the student, compensating for client data heterogeneity (non-IID). The distillation loss is given in Eq. (12).

$$\mathcal{L}_{KD} = \frac{1}{B} \cdot \sum_{j=1}^{B} KL\left(p_{stu}\left(\frac{f_{stu}(x_j)}{T}\right) \parallel p_{tea}\left(\frac{f_{tea}(x_j)}{T}\right)\right) \times T^2 \tag{12}$$

where T is a hyperparameter, specifically the temperature parameter.

3.4.2 The Alignment of Local Feature Subspace

Although FedSPQR employs knowledge distillation and SimSiam to mitigate non-IID bias in client data, two key issues continue to impede effective model training.

- The knowledge distillation mechanism only aligns the output layers between the teacher model and the student model while ignoring the heterogeneity of intermediate feature spaces. In non-IID scenarios, the data distributions on the client sides differ significantly from that of the global model. The intermediate features of client models may only capture local patterns of their own data, whereas the intermediate features of the global model need to cover universal patterns across all classes. Even if the knowledge distillation aligns the outputs of client models with those of the global model, the intermediate feature spaces remain disconnected. It means that client models do not truly acquire the ability to extract globally universal features, and only adapt to the global model by mimicking outputs, resulting in poor generalization.
- SimSiam only learns local self-supervised representations and is prone to falling into the local feature trap. The core of SimSiam resides in learning invariant representations of local data through feature prediction across strong and weak augmented views. But this process relies entirely on client-side data and lacks global-level constraints.

The above two issues lead to model drift. Due to the disconnection between their intermediate feature spaces and the global counterpart, client models converge to local optima during each local training round. And the discrepancies between client models become increasingly significant.

FedSPQR addresses these challenges through a feature subspace fusion mechanism that unifies intermediate feature dimensions and imposes global constraints to mitigate model drift. This integrates scattered client models into a coherent global feature framework, thereby endowing the knowledge distillation and SimSiam-learned features with consistent global validity.

For FedSPQR, on the client side, the student model and the teacher model are used to learn local knowledge and global knowledge, respectively. Therefore, the subspace alignment mechanism primarily targets the student model, ensuring that the feature subspace of its output is consistent with the global feature subspace. The alignment loss function is defined in Eq. (13).

$$\mathcal{L}_{align} = 1 - \frac{1}{B} \sum_{j=1}^{B} \frac{U_c^j \cdot (U_g U_g^T \cdot U_c^j)}{\parallel U_c^j \parallel_2 \cdot \parallel U_g U_g^T \cdot U_c^j \parallel_2} \tag{13}$$

where U_g is global feature space information received from the server side. U_c is local feature space obtained in previous round. The loss function in client is defined as Eq. (14).

$$\mathcal{L} = \lambda_1 \cdot \mathcal{L}_{simsiam} + \lambda_2 \cdot \mathcal{L}_{KD} + \lambda_3 \cdot \mathcal{L}_{align} \tag{14}$$

where λ_1, λ_2 and λ_3 are hyperparameters.

3.4.3 Generation of Feature Subspace

Effective extraction of feature subspaces on the client side is crucial in FSSL. Traditional frameworks typically generate these subspaces on the server using the aggregated local models and its labeled data, which overlooks the distribution shift between client and server data. Inspired by FedPCA [25], FedSPQR adopts a PCA-based method for on-client subspace generation. The detailed procedure is described in Section 3.3.1.

4 Experiment

4.1 Experiment Setup

In our experiment, we apply CIFAR-10 [27] and Fashion-MNIST dataset [28] to verify the efficiency of FedSPQR.

- **Fashion-MNIST:** The Fashion-MNIST dataset is a classic benchmark comprising single-channel grayscale images of 28×28 pixels with pixel values ranging from 0 to 255. It contains 60,000 training and 10,000 test examples, evenly distributed across 10 categories.
- **CIFAR-10:** The CIFAR-10 dataset is a classic color image classification dataset. It consists of 3-channel color images (RGB mode), with each image having a fixed size of 32×32 pixels and pixel values ranging from 0 to 255. The dataset contains a total of 60,000 samples, covering 10 categories. Among them, the training set includes 50,000 samples, and the test set includes 10,000 samples.

During training, we use the Adam optimizer with an initial learning rate of 0.001. The model is trained for 15 warm-up epochs and 10 local epochs. We simulate a system with 100 clients in total, randomly selecting 10 clients in each communication round. To simulate non-IID data distributions across clients, we employ Dirichlet distributions with concentration parameters of $\alpha = 1.0$ and $\alpha = 0.1$. For the IID setting, we use $\alpha = 100$. The server holds 500 labeled data samples. The data distributions of 10 randomly selected clients are visualized in Fig. 2.
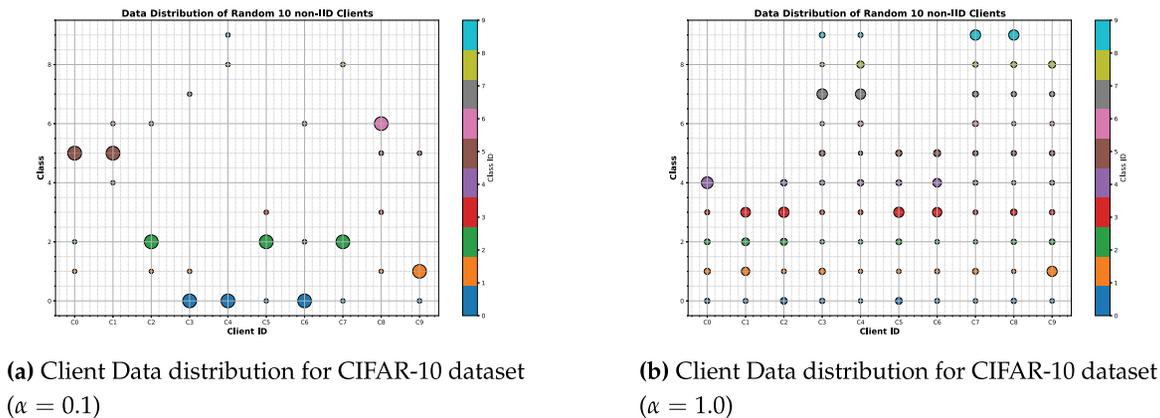

(a) Client Data distribution for CIFAR-10 dataset ($\alpha = 0.1$)

(b) Client Data distribution for CIFAR-10 dataset ($\alpha = 1.0$)

Figure 2: Client data distribution.

The loss function on server side utilizes a weighted combination of classification loss and subspace alignment loss, defined Eq. (9). The alignment loss weight λ follows a progressive growth strategy. The initial value of λ is 0.05. Its value increases continuously as interactions proceed. The upper limit of λ is set at 0.2. In the early stage of training, λ is small to prioritize the optimization of classification tasks. It helps quickly build basic discriminative capabilities. As training rounds progress, λ increases gradually. It strengthens the alignment constraint between the local feature space and the global subspace step by step. It promotes the consistency of feature representations among clients. In this way, it alleviates the model drift problem caused

by non-IID data. The upper limit of 0.2 prevents the alignment loss from dominating excessively. It avoids weakening classification performance.

For the loss function on the client side, defined as Eq. (14), the weight of SimSiam loss $\lambda_1 = 0.5$ ensures that self-supervised learning provides a stable function for feature representation. The weight of knowledge distillation $\lambda_2 = 1.0$ to transfer global knowledge from the teacher model to the student model. The weight of subspace alignment λ_3 adopts adaptive mechanism, which is dynamically adjusted according to the Grassmann distance between the client's local subspace and the global subspace. This design enables clients with data distributions that differ significantly from the global distribution to obtain stronger alignment constraints to alleviate the client drift problem caused by non-IID data. At the same time, it preserves a personalized space for clients whose data distributions are similar to the global one. The synergistic effect of hyperparameters forms a dual balance mechanism during training. The self-supervised loss guarantees the quality of feature representation. The loss of knowledge distillation fuses global knowledge during local learning. And the dynamic alignment loss adaptively regulates the trade-off between personalization and consistency.

4.2 Baseline Method

We chose three models as baselines to verify the efficiency of FedSPQR.

- **Only Server:** This model only utilizes labeled data to train the global model on the server side. Thus, this model serves as a lower bound for federated semi-supervised learning.
- **FedMatch [29]:** This model integrates consistency regularization among clients with parameter decomposition to implement separate learning for labeled and unlabeled data.
- **FedMix [6]:** This model integrates parameter mixing strategy for separate learning with dynamic weight allocation for client weights to increase data availability and alleviate the non-IID problem in a federated semi-supervised learning environment.
- **FedRGD [30]:** This model is a grouping-based model averaging method. It integrates group normalization with a consistency regularization loss to improve training efficiency for federated semi-supervised learning.

4.3 Overall Result

In order to verify the efficiency, we compare FedSPQR with other baselines on CIFAR-10 and Fashion-MNIST datasets. After each round of communication, the global model evaluates its accuracy using the test dataset. The evaluation results are shown in Fig. 3.
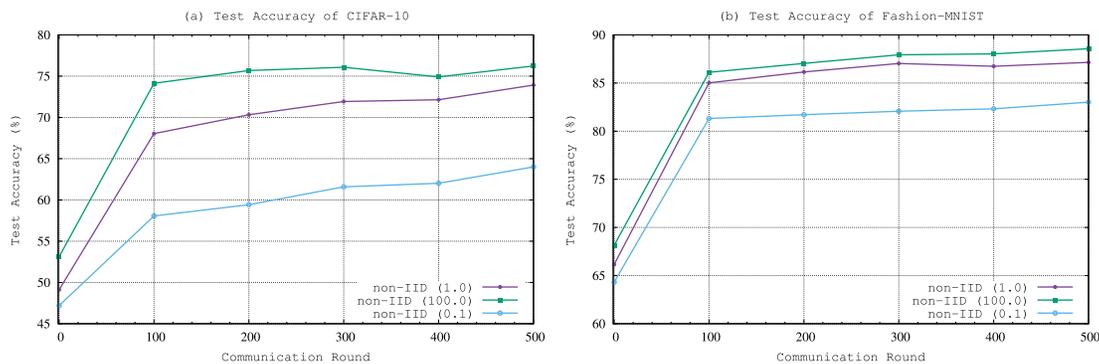

Figure 3: Test accuracy variation with increasing of number of commnuncation.

Fig. 3a,b shows the variation of test accuracy with the increase in the number of communications for CIFAR-10 and Fashion-MNIST datasets under IID and non-IID settings, respectively. As can be shown from Fig. 3, the test accuracy continuously improves with the increase in the number of communications. The improvement in test accuracy is most significant during the first 0 to 100 communications. This is because supervised learning based on labeled data on the server side effectively enhances the overall model performance. Therefore, the subsequent improvement in accuracy mainly relies on the learning of unlabeled data on the client side. It can be seen from Fig. 3 that when the number of communications between the client and server exceeds 300, the global model begins to converge. Table 1 shows the highest accuracy of FedSPQR and other baselines in Fig. 3. For the CIFAR-10 dataset, under non-IID and IID data distributions, the highest accuracy of FedSPQR is 73.12% and 76.98%, respectively. This accuracy exceeds that of other baselines by at least 20%. For the Fashion-MNIST dataset, under non-IID and IID data distributions, the highest accuracy of our model is 87.03% and 88.82%, respectively. This accuracy exceeds that of other baselines by at least 3%. As a result, the efficiency of FedSPQR is better than that of other baselines.

Table 1: Compared FedSPQR with the others baselines for the highest accuracy where the number of labeled data is 500.

Model	Highest Accuracy (%)			
	CIFAR-10		Fashion-MNIST	
	IID ($\alpha = 100$)	Non-IID ($\alpha = 1.0$)	IID ($\alpha = 100$)	Non-IID ($\alpha = 1.0$)
Only Server	51.82	51.82	80.17	80.17
FedMatch	53.62	52.76	81.92	81.59
FedMix	57.82	56.93	83.57	82.61
FedRGD	48.98	47.53	83.72	83.13
FedSPQR (our)	76.98	73.12	88.82	87.03

In order to verify the impact of the amount of labeled data on the efficiency of FedSPQR, we randomly select 500 and 1000 samples as labeled data to test the model's performance on the CIFAR-10 dataset. The variation of test accuracy with the increase in the number of communications is shown in Fig. 4. As shown in Fig. 4, with the increase in the amount of labeled data on the server side, the test accuracy of FedSPQR improves. Table 2 presents the highest accuracy of FedSPQR and other baselines when 500 and 1000 samples are used as labeled data on the server side. When the number of labeled samples is only 500, the highest accuracy of FedSPQR is 74.13% and 68.04% under IID and non-IID scenarios, respectively. However, when the number of labeled samples reaches 1000, the test accuracy of our model increases to 78.12% and 77.54% under IID and non-IID scenarios, respectively. Furthermore, the highest accuracy of FedSPQR is also better than that of other baselines.
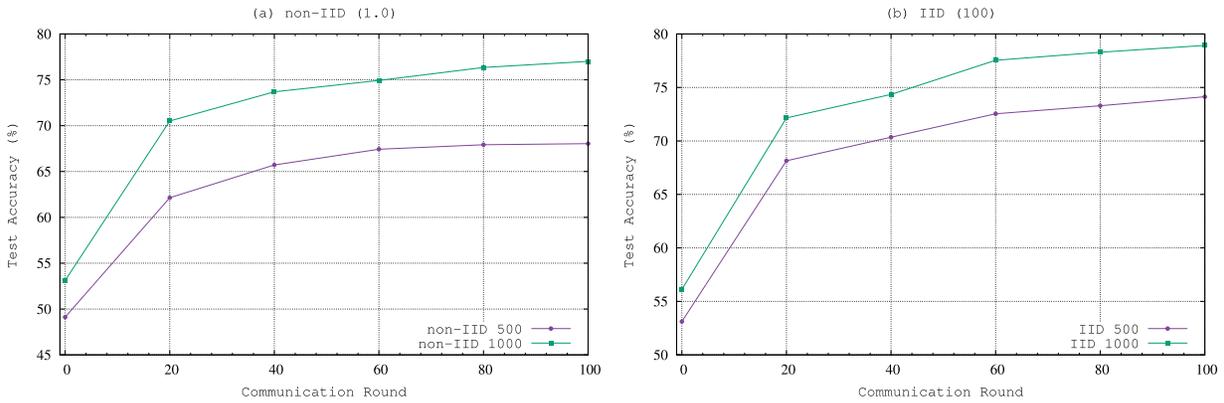
Figure 4: Test accuracy variation with different amounts of labeled data on server side for CIFAR-10.

Table 2: Compared FedSPQR with the others baselines for the highest accuracy for CIFAR-10, where the amount of labeled data is 500 or 1000.

Model	Highest Accuracy (%)			
	500		**1000**	
	IID (α = 100)	**Non-IID (α = 1.0)**	**IID (α = 100)**	**Non-IID (α = 1.0)**
Only Server	51.82	51.82	67.13	67.13
FedMatch	53.62	52.76	68.22	67.14
FedMix	57.82	56.93	69.43	68.72
FedRGD	48.98	47.53	61.53	60.42
FedSPQR (our)	74.13	68.04	78.93	77.54

4.4 Analysis and Ablation Study

In this section, we analyze the contribution of feature space alignment (-sa) and self-supervised learning (-ss) to FedSPQR. After 100 communication rounds, the experimental results of comparing FedSPQR with its variants without feature space alignment and self-supervised learning are shown in Figs. 5 and 6, and Table 3.
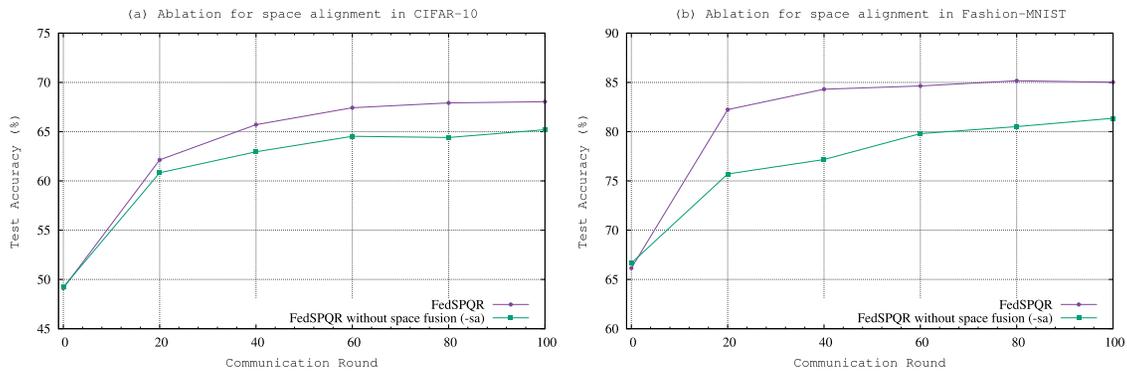

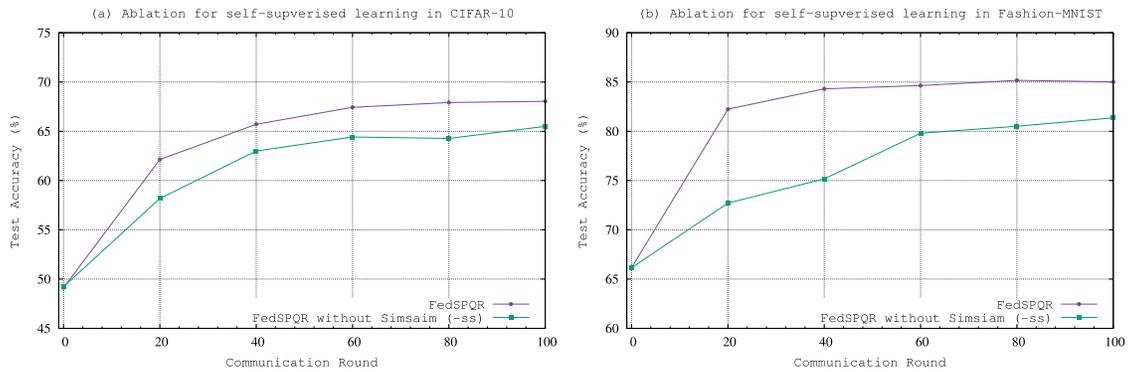**Figure 5:** The ablation experimental results for space alignment.

Figure 6: The ablation experimental results for self-supervised learning.

Table 3: The ablation Results.

| Labeled data | Setting | Highest Accuracy (%) | | | | | |
| | | CIFAR-10 | | | Fashion-MNIST | | |
		FedSPQR	-sa	-ss	FedSPQR	-sa	-ss
500	IID (α = 100)	74.13	65.72	66.03	88.82	85.92	86.72
	non-IID (α = 1.0)	68.04	63.01	62.17	87.03	82.67	84.67
1000	IID (α = 100)	78.93	69.31	71.92	89.12	86.83	87.92
	non-IID (α = 1.0)	77.54	66.35	69.63	87.52	84.93	85.43

Regarding feature space alignment, as shown in Fig. 5, with the increase in the number of interactions, although the test accuracy of both models in Fig. 5 have improved, the improvement of the model without feature space alignment is limited. This indicates that the feature space alignment can effectively assist the model in achieving more effective training. When comparing the experimental results of the CIFAR-10 dataset and the Fashion-MNIST dataset, it is found that on CIFAR-10, the performance improvement of FedSPQR over its variant without feature space alignment is more significant. This indicates that for datasets with greater classification difficulty, feature space alignment can more effectively enhance the model's performance.

For self-supervised learning, as shown in Fig. 6, the change in its test accuracy is similar to that of the subspace alignment model. Furthermore, compared with the results of the ablation experiment on the Fashion-MNIST dataset, the results of the ablation experiment on the CIFAR-10 dataset are more pronounced. This indicates that the self-supervised model effectively facilitates model training and can more effectively extract features when dealing with datasets that are more difficult to classify.

5 Conclusion

In this paper, we propose a novel federated semi-supervised learning method, denoted as FedSPQR, for the label-at-server framework. Compared with traditional federated semi-supervised learning methods, FedSPQR utilizes the subspace alignment technology to generate the global feature subspace in federated learning. This global feature subspace alleviates data heterogeneity among clients. Furthermore, the knowledge distillation mechanism is applied to inject global knowledge into local training on the client side. Finally,

we utilize CIFAR-10 and Fashion-MNIST data to verify the performance of FedSPQR. The experimental results show that the performance of FedSPQR is better than that of the traditional models.

Despite the promising results, several avenues remain for further improvement and exploration. Firstly, the processes of subspace alignment and knowledge distillation introduce additional computational overhead on both the server and client sides. Future work will focus on optimizing these components, potentially through more efficient clustering algorithms, lightweight knowledge distillation techniques, or the development of sparse feature subspace representations, to reduce the computational burden and enhance scalability. Secondly, communication efficiency is paramount in federated systems. The current framework requires transmitting feature subspace information between rounds. A key direction for future research is for the server to evaluate the client-side local feature spaces using the locally uploaded models and its own data, thereby reducing communication overhead.

Author Contributions: Zhe Ding, Hao Yi and Zhiguang Qin: Model Design, Writing—Original Draft; Zhe Ding, Qixu Wang, Wenrui Xie: Experiment; Ming Zhang, Yuxuan Xiao, Qing Chen, Dajiang Chen: Editing. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics Approval: Not applicable.

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

References

1. Zhang C, Xie Y, Bai H, Yu B, Li W, Gao Y. A survey on federated learning. Knowl Based Syst. 2021;216(1):106775. doi:10.1016/j.knosys.2021.106775.
2. 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(2):513–35. doi:10.1007/s13042-022-01647-y.
3. Jin Y, Wei X, Liu Y, Yang Q. Towards utilizing unlabeled data in federated learning: a survey and prospective. arXiv:2002.11545. 2020.
4. Jin Y, Liu Y, Chen K, Yang Q. Federated learning without full labels: a survey. arXiv:2303.14453. 2023.
5. Zhang Z, Zhang F, Xiong Z, Zhang K, Chen D. LsiA3CS: deep-reinforcement-learning-based cloud-edge collaborative task scheduling in large-scale IIoT. IEEE Internet Things J. 2024;11(13):23917–30. doi:10.1109/jiot.2024.3386888.
6. Zhang Z, Ma S, Yang Z, Xiong Z, Kang J, Wu Y, et al. Robust semisupervised federated learning for images automatic recognition in Internet of drones. IEEE Internet Things J. 2023;10(7):5733–46. doi:10.1109/JIOT.2022.3151945.
7. Liu H, Mi Y, Tang Y, Guan J, Zhou S. Boosting semi-supervised federated learning by effectively exploiting server-side knowledge and client-side unconfident samples. Neural Netw. 2025;188:107440. doi:10.1016/j.neunet.2025.107440.
8. Naeem F, Ali M, Kaddoum G. Federated-learning-empowered semi-supervised active learning framework for intrusion detection in ZSM. IEEE Commun Mag. 2023;61(2):88–94. doi:10.1109/MCOM.001.2200533.
9. Bian J, Fu Z, Xu J. FedSEAL: semi-supervised federated learning with self-ensemble learning and negative learning. arXiv:2110.07829. 2021.

10. Brendan M, Moore E, Ramage D, Hampson S, Aguera Y, Arcas B. Communication-efficient learning of deep networks from decentralized data. In: Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017; 2017 Apr 20–22; Fort Lauderdale, FL, USA. p. 1273–82.

11. 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;135:244–58. doi:10.1016/j.future.2022.05.003.

12. Mammen PM. Federated learning: opportunities and challenges. arXiv:2101.05428. 2021.

13. Li T, Sahu AK, Zaheer M, Sanjabi M, Talwalkar A, Smith V. Federated optimization in heterogeneous networks. In: Proceedings of the 3rd MLSys Conference; 2020 Mar 2–4; Austin, TX, USA. p. 429–50.

14. Cao Y, Zhang J, Zhao Y, Liu D, Han Y, Jiang W, et al. ConfShield: a dual-stage fusion framework for robust and privacy-enhancing federated learning in cloud environments. Inf Fusion. 2026;129:104037. doi:10.1016/j.inffus.2025.104037.

15. Gao ZW, Xiang YH, Lu SX, Liu YH. An optimized updating adaptive federated learning for pumping units collaborative diagnosis with label heterogeneity and communication redundancy. Eng Appl Artif Intell. 2025;152:110724. doi:10.1016/j.engappai.2025.110724.

16. Zhu H, Xu J, Liu S, Jin Y. Federated learning on non-IID data: a survey. Neurocomputing. 2021;465:371–90. doi:10.1016/j.neucom.2021.07.098.

17. Karimireddy SP, Kale S, Mohri M, Reddi S, Stich S, Suresh AT. Scaffold: stochastic controlled averaging for federated learning. In: Proceedings of the 37th International Conference on Machine Learning, Virtual Event; 2020 Jul 13–18; Online. p. 5132–43.

18. Saligrama V, Acar DA, Whatmough PN, Matas R, Mattina M, Zhao Y. Federated learning based on dynamic regularization. In: 2021 International Conference on Learning Representations; 2021 May 3–7; Vienna, Austria.

19. Chen S, Shen C, Zhang L, Tang Y. Dynamic aggregation for heterogeneous quantization in federated learning. IEEE Trans Wirel Commun. 2021;20(10):6804–19. doi:10.1109/TWC.2021.3076613.

20. Yang X, Song Z, King I, Xu Z. A survey on deep semi-supervised learning. IEEE Trans Knowl Data Eng. 2023;35(9):8934–54. doi:10.1109/TKDE.2022.3220219.

21. Tarvainen A, Valpola H. Mean teachers are better role models: weight-averaged consistency targets improve semi-supervised deep learning results. In: Proceedings of the 31st International Conference on Neural Information Processing Systems; 2017 Dec 4–9; Long Beach, CA, USA. p. 1195–204.

22. Zhou ZH, Li M. Tri-training: exploiting unlabeled data using three classifiers. IEEE Trans Knowl Data Eng. 2005;17(11):1529–41. doi:10.1109/TKDE.2005.186.

23. Li M, Li Q, Wang Y. Class balanced adaptive pseudo labeling for federated semi-supervised learning. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2023 Jun 17–24; Vancouver, BC, Canada. p. 16292–301. doi:10.1109/CVPR52729.2023.01563.

24. Shi Y, Chen S, Zhang H. Uncertainty minimization for personalized federated semi-supervised learning. IEEE Trans Netw Sci Eng. 2023;10(2):1060–73. doi:10.1109/TNSE.2022.3226574.

25. Nguyen TA, He J, Le LT, Bao W, Tran NH. Federated PCA on Grassmann manifold for anomaly detection in IoT networks. In: IEEE INFOCOM 2023-IEEE Conference on Computer Communications; 2023 May 17–20; New York, NY, USA. p. 1–10. doi:10.1109/INFOCOM53939.2023.10229026.

26. Chen X, He K. Exploring simple Siamese representation learning. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2021 Jun 20–25; Nashville, TN, USA. p. 15745–53. doi:10.1109/cvpr46437.2021.01549.

27. Krizhevsky A, Hinton. Geoffrey Learning multiple layers of features from tiny images. Toronto, ON, Canada: University of Toronto; 2009.

28. Xiao H, Rasul K, Vollgraf R. Fashion-MNIST: a novel image dataset for benchmarking machine learning algorithms. arXiv:1708.07747. 2017.

29. Jeong W, Yoon J, Yang E, Hwang SJ. Federated semi-supervised learning with inter-client consistency & disjoint learning. arXiv:2006.12097. 2020.

30. Zhang Z, Yang Y, Yao Z, Yan Y, Gonzalez JE, Ramchandran K, et al. Improving semi-supervised federated learning by reducing the gradient diversity of models. In: 2021 IEEE International Conference on Big Data (Big Data); 2021 Dec 15–18; Orlando, FL, USA. p. 1214–25. doi:10.1109/bigdata52589.2021.9671693.