Towards Unbiased Training in Federated Open-world Semi-supervised Learning

Jie Zhang [1] **Xiaosong Ma** [1] **Song Guo** [1] **Wenchao Xu** [1]

Abstract

Federated Semi-supervised Learning (FedSSL) has emerged as a new paradigm for allowing distributed clients to collaboratively train a machine learning model over scarce labeled data and abundant unlabeled data. However, existing works for FedSSL rely on a closed-world assumption that all local training data and global testing data are from seen classes observed in the labeled dataset. It is crucial to go one step further: *adapting FL models to an open-world setting, where unseen classes exist in the unlabeled data.* In this paper, we propose a novel **Fed**erated **o**pen-world **S**emi-**S**upervised **L**earning (**FedoSSL**) framework, which can solve the key challenges in distributed and open-world settings, i.e., the biased training process for heterogeneously distributed unseen classes. Specifically, since the advent of a certain unseen class depends on a client basis, the locally unseen classes (exist in multiple clients) are likely to receive differentiated superior aggregation effects than the globally unseen classes (exist only in one client). We adopt an uncertainty-aware suppressed loss to alleviate the biased training between locally unseen and globally unseen classes. Besides, we enable a calibration module supplementary to the global aggregation to avoid potential conflicting knowledge transfer caused by inconsistent data distribution among different clients. The proposed FedoSSL can be easily adapted to state-of-the-art FL methods, which is also validated via extensive experiments on benchmarks and real-world datasets (CIFAR-10, CIFAR-100 and CINIC-10).

[1]Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China. Correspondence to: Song Guo <song.guo@polyu.edu.hk>.

Proceedings of the 40th International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023. Copyright 2023 by the author(s).

1. Introduction

To tackle the privacy issues in distributed machine learning, Federated Learning (FL) (McMahan et al., 2017; Zhang et al., 2021b) has emerged as a promising paradigm by collaboratively training a shared model among multiple clients without exposing their private raw data. As a common practice in FL, the global model is usually obtained by periodically averaging the updated model parameters from distributed clients in a centralized server. While existing FL methods assume that clients' data is fully labeled so that supervised learning can be conducted for local model update on each client, in some real-world applications, the data labeling process can be prohibitive due to the tremendous overhead and the requirement of corresponding expertise (Ouali et al., 2020) such as for medical diagnoses (Ng et al., 2021) and object detection (Liu et al., 2021).

The scarce labeled data and the abundant unlabeled data give the rise to the emergence of federated semi-supervised learning (FedSSL) (Jeong et al., 2021; Liang et al., 2022), which can simultaneously exploit both the labeled and unlabeled data to optimize a global model in distributed environments. Existing FedSSL schemes have demonstrated to train models based on a small amount of labeled data on both client or server side (Jin et al., 2020; Long et al., 2020). However, these works rely on the closed-world assumption that all local training data and global testing data are from the same set of classes that are included in the labeled dataset (Zhou, 2022; Boult et al., 2019), which is often invalid for practical scenarios. In contrast, the open-world settings allow novel class discovery, and thus are common for data in-the-wild, e.g., in medical image classification tasks, some diseases are naturally scarce and never labeled before, while the model may be required to both classify images into predefined types (seen classes) and discover new unknown diseases (unseen classes). Hence, a new fundamental question arises: *how to collaboratively train models on distributed data to enable classification on both seen and unseen classes under the open-world setting?*

To this end, we first construct a new FedSSL benchmark that extends the conventional closed-world training framework to the open-world setting. Surprisingly, it is demonstrated that there is a significant performance degradation due to the existence of unseen classes in unlabeled data during the

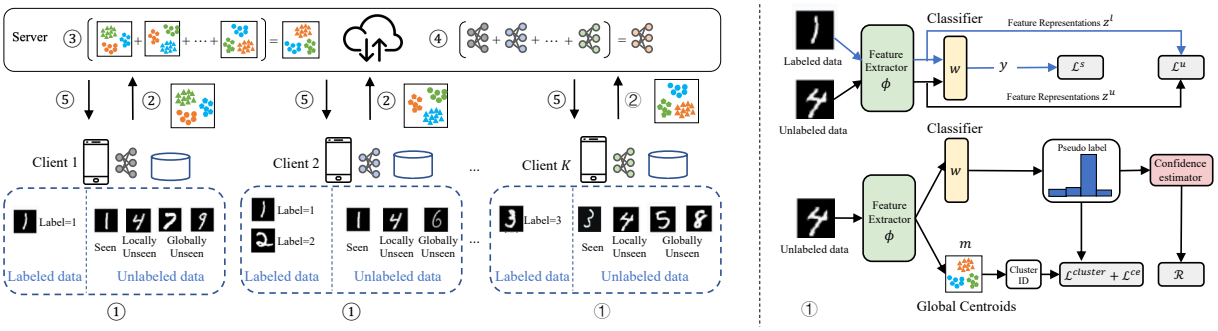


Figure 1: Framework of the proposed FedoSSL algorithm. **Pipeline**: ① **Local Training:** Each client first performs local training on its private dataset for several epochs (i.e., via optimizing loss function in Eq. (1)), and then computes local centroids via a Sinkhorn-Knopp based clustering algorithm (Genevay et al., 2019). ② Upload model parameters and local centroids to the server. ③ The server performs standard model aggregation. ④ The server performs centroids aggregation by again using Sinkhorn-Knopp clustering to obtain global centroids. ⑤ The global model and global centroids are returned to the clients, who use them for local training.

training process. Although a few works explore the unseen class problem to avoid misclassifying unlabeled examples from unseen classes into seen classes, meanwhile alleviating the undesired performance gap between seen and unseen classes (Cao et al., 2022; Guo et al., 2022), there is no literature considers this problem in a distributed environment. With multiple participants, the problem definition is different, i.e., some unseen classes in one client may exist in other clients' side from a global view, and thus requires a novel fine-grained definition on unseen classes as well as the training mechanism for different types of samples/classes. It is worth mentioning that due to the heterogeneous distributed classes across different clients, simply aggregating the parameters following traditional FL mechanism can cause the biased training process for clients possessing different unseen classes.

To tackle the above challenges, in this paper, we propose a brand-new **Fed**erated **o**pen-world **S**emi-**S**upervised **L**earning framework (FedoSSL) that can achieve unbiased training procedure among different types of samples. Specifically, the unseen classes in local clients are first redefined into *locally unseen classes* and *globally unseen classes*, respectively. Then, we design a uncertainty-aware suppressed loss to adaptively control the undesired training divergence between locally and globally unseen classes, i.e. locally unseen classes often have higher training efficiency than globally unseen ones due to the cross-client collaboration. Furthermore, considering the heterogeneously distributed unseen classes across different clients that lead to potential conflicting knowledge transfer during the aggregation phase, we introduce a calibration module to produce corresponding global centroids to perform dedicated local adjusting on each client. We show that our method can significantly improve the model accuracy in a open-world setting when compared with the state-of-the-art baselines over widely used models (i.e., ResNet-18) and downstream tasks (i.e., CIFAR-10, CIFAR-100, CINIC-10). The contributions of the paper are summarized as follows.

- To the best of our knowledge, we are the first to consider the open-world setting in FedSSL, where unseen classes exist in the unlabeled data, which is challenging due to the heterogeneously distributed unseen classes.
- We design a brand-new FedoSSL framework, that can achieve unbiased learning among different types of classes (i.e., locally unseen and globally unseen classes) and calibrated knowledge aggregation given heterogeneous data distributions.
- We conduct extensive experiments on three typical image classification tasks. The empirical evaluation shows the superior performance of FedoSSL over the state-of-the-art approaches.

2. Related Work

2.1. Federated Learning

Federated Learning (FL) has emerged as a promising paradigm to collaboratively train machine learning models using decentralized training data with privacy protection. As a most popular supervised FL training framework, FedAvg (McMahan et al., 2017) periodically aggregates the local updates at the server and transmits the averaged model back to local clients for the next round of training. In FL, Non-IID data is one of the key challenges due to the caused weight divergence and performance drop (Li et al., 2020b). Many solutions have been proposed to solve this problem, i.e., sharing a public dataset among clients to achieve an approximately IID data (Zhao et al., 2018; Yoshida et al., 2019),

Table 1: Comparison between our proposed FedoSSL and other SSL-related methods.

Method	Training Dataset		Testing Dataset		FL Environment?
	Seen classes	Unseen classes	Seen classes	Unseen classes	
SSL	Present	Not Present	Classify	-	×
Open-set SSL	Present	Present	Classify	Detect & Reject	×
Novel Class Discovery	Present	Present	-	Discover & Cluster	×
Open-world SSL	Present	Present	Classify	Discover & Cluster	×
FedSSL	Present	Not Present	Classify	-	√
FedoSSL	Present	Present	Classify	Discover & Cluster	√

data augmentation (Jeong et al., 2018), knowledge distillation (Li & Wang, 2019; Lin et al., 2020), regularization-based methods (Fallah et al., 2020; Li et al., 2020a) and even personalized FL (Collins et al., 2021; Zhang et al., 2021a). However, FL in the semi-supervised setting is less explored.

2.2. Semi-supervised Learning

Semi-Supervised Learning (SSL) refers to the general problem of learning with partially labeled data, especially when the amount of labeled data is much smaller than that of the unlabeled data (Zhou & Li, 2005; Rasmus et al., 2015). The mainstream of SSL can be divided into two types: pseudo-labeling and consistency regularization. Pseudo-labeling methods (Lee et al., 2013; Pham et al., 2021; Zou et al., 2020) usually generate artificial labels of unlabeled data from the model trained on labeled data and apply the filtered high-confidence labels as supervised signals for unlabeled data training. Consistency regularization methods (Tarvainen & Valpola, 2017; Wei et al., 2020; Lee et al., 2022) focus on training models via minimizing the distance among different perturbed outputs of the same input. All the above literature assume a closed-world setting in which labeled and unlabeled data come from the same set of predefined classes. However, this assumption rarely holds for data in real-world applications, namely, unseen classes would occur in unlabeled data. There are two recent lines of works related to different scenarios:

Open-set Semi-supervised Learning. Open-set SSL considers that unseen classes in unlabeled samples only exist in training data, while not exist in testing data (Chen et al., 2020b; Guo et al., 2020; Huang et al., 2021; Saito et al., 2021). The goal is to detect these unseen classes and reject them to ensure no performance degradation on seen classes.

Novel Class Discovery (NCD). Different from Open-set SSL that fails to classify unseen classes, NCD (Han et al., 2019; Hsu et al., 2019; Zhong et al., 2021) aims to classify both seen and unseen classes during the testing phase but assumes all unlabeled instances belonging to unseen classes in training data.

To extend unseen classes into a more practical scenario, open-world SSL (Cao et al., 2022; Guo et al., 2022) focus on solving the class mismatch between labeled and unlabeled data, where each test sample should be either classified into one of existing classes or a new unseen class in the test time.

2.3. Federated Semi-supervised Learning

Learning representations from unlabeled decentralized data while preserving data privacy is still a nascent field. Existing FedSSL frameworks can be categorized into two types according to the location of the labeled data: 1) Labels-at-Server assumes that clients have purely unlabeled data and the server contains a small amount of labeled data (Lin et al., 2021; He et al., 2021; Zhang et al., 2021c; Diao et al., 2022); 2) Labels-at-Client considers labeled data are available at local clients (Jeong et al., 2021; Lin et al., 2021; Liang et al., 2022), which can be further subdivided into two scenarios: a) each client contains both labeled and unlabeled data (Jeong et al., 2021); b) some clients are fully labeled while some clients only contain unlabeled samples (Lin et al., 2021; Liang et al., 2022; Liu et al., 2021). In this paper, we focus on the type 2.a as it has been largely overlooked and is more general in real-world FL scenarios.

2.4. Summary

To sum up, considering the existence of unseen classes in FedSSL is still a vacancy. More importantly, the characteristic of data heterogeneity among multiple clients brings new challenges for the combination of open-world and conventional FedSSL frameworks: *how to tackle heterogeneous unseen classes among the clients in an efficient way?* Inspired by the above observations, we are motivated to develop a novel federated open-world semi-supervised learning framework to achieve high performance on both seen and unseen classes. The detailed comparisons between or proposed FedoSSL and other SSL-related methods are illustrated in Table 1.

3. Methodology

In this section, we first define the problem and elaborate the proposed FedoSSL framework, which contains an

uncertainty-aware suppressed loss for local training and a calibration module for model aggregation. Then, we present the optimization process in Section 3.3.

3.1. Preliminary and Problem Definition

We focus on federated semi-supervised learning in an open-world setting, where the data in each client is partially labeled and the set of classes/categories in labeled data and unlabeled data are not the same, i.e., unseen classes exist in unlabeled data. Assuming K clients that each client i holds a private classification dataset containing a labeled part $\mathcal{D}_i^l = \{(x_j, y_j)\}_{j=1}^{n_i^l}$ and an unlabeled part $\mathcal{D}_i^u = \{(x_j)\}_{j=1}^{n_i^u}$, where $n_i^l \ll n_i^u$, $x \in \mathbb{R}^N$, $1 \leq i \leq K$. The whole label and unlabeled dataset can be represented as $\mathcal{D}^l = \{\mathcal{D}_i^l\}_{i=1}^K$ and $\mathcal{D}^u = \{\mathcal{D}_i^u\}_{i=1}^K$, respectively. We denote the set of classes seen in the full labeled data as $\mathcal{C}^l$ and the set of classes in the unlabeled test data as $\mathcal{C}^u$. Unlike traditional (closed-world) FedSSL that $\mathcal{C}^l = \mathcal{C}^u$, in this paper, we consider $\mathcal{C}^l \neq \mathcal{C}^u$, and denote $\mathcal{C}_{seen} = \mathcal{C}^l \cap \mathcal{C}^u$, $\mathcal{C}_{unseen} = \mathcal{C}^u \setminus \mathcal{C}_{seen}$ as the set of seen classes and unseen classes, respectively.

The goal of FedSSL is to train a generalized global model f with parameter θ from multiple decentralized clients, i.e., $\min_\theta \mathcal{L}(\theta) := \sum_{i=1}^K \frac{n_i}{n} \mathcal{L}_i(\theta)$, where $n_i = n_i^l + n_i^u$ and $n = \sum_{i=1}^K n_i$ is the total data amount. $\mathcal{L}_i(\theta)$ is the training loss function for client i. Specifically, the model f can be decomposed of a feature extractor g with parameter ϕ: $\mathbb{R}^N \rightarrow \mathbb{R}^d$ to learn a low-dimensional feature z and a classifier h with parameter w: $\mathbb{R}^d \rightarrow \mathbb{R}^{|\mathcal{C}_{seen} \cup \mathcal{C}_{unseen}|}$. The training loss of a semi-supervised learning algorithm on each client i usually contains supervised loss $\mathcal{L}_i^s$ and unsupervised loss $\mathcal{L}_i^u$ with weight parameters α, $\alpha > 0$:

$$\mathcal{L}_i = \mathcal{L}_i^s + \alpha \mathcal{L}_i^u \tag{1}$$

Typically, $\mathcal{L}_i^s$ applies the standard cross-entropy loss on labeled instances:

$$\mathcal{L}_i^s = \frac{1}{n_i^l} \sum_{(x_j, y_j) \in \mathcal{D}_i^l} \mathcal{H}(y_j, p(x_j; \theta)) \tag{2}$$

where $p(x; \theta) = \text{Softmax}(f(x; \theta))$ denotes the predicted probabilities produced by the model f for input x, and $\mathcal{H}(\cdot, \cdot)$ is the cross-entropy function.

In terms of unsupervised loss $\mathcal{L}_i^u$, there are two typical forms: **pseudo-labels** (Sohn et al., 2020) based on labeled data and **consistency regularization** (Xie et al., 2020) based on data augmentation. However, in an open-world setting, the existence of unseen classes makes above methods fail to classify seen classes and unseen classes. Therefore, similar to ORCA (Cao et al., 2022) and NACH (Guo et al., 2022), we use pairwise objective as unsupervised loss on unlabeled data to classify unseen classes:

$$\mathcal{L}_i^u = -\frac{1}{n_i^l + n_i^u} \sum_{\substack{z_j, \bar{z}_j \in \\ (Z_i^l \cup Z_i^u, \bar{Z}_i^l \cup \bar{Z}_i^u)}} \mathcal{H}(p(w^\top \cdot z_j), p(w^\top \cdot \bar{z}_j)) \tag{3}$$

where Z_i^l and Z_i^u are the whole set of feature representations for labeled and unlabeled data, respectively. $\bar{Z}_i^l \cup \bar{Z}_i^u$ is the closet set of $Z_i^l \cup Z_i^u$ in a mini-batch by computing the cosine distance between all pairs of feature representations.

3.2. Overview of FedoSSL

Previous FedSSL methods do not consider the existence of unseen classes on each client, which leads to a number of data from unseen classes being misclassified into seen classes. Besides, inconsistent data distribution on different clients raises another new problem: *some unseen classes may exist in more than one client, resulting in biased training among different unseen classes*, e.g., in Figure 1, both client 1, 2 and K have class 4, while class 5, 6, 7, 8 and 9 only exist in one of the clients. In this case, we need a more fine-grained definition on unseen class in a global view.

Definition 1 (locally unseen & globally unseen class)**.** In FedoSSL, the unseen classes $\mathcal{C}_{i,unseen}$ on client i can be divided into two types: locally unseen classes $\mathcal{C}_{i,lu}$, in which $\mathcal{C}_{i,lu} = \mathcal{C}_{1,unseen} \cap \cdots \cap \mathcal{C}_{K,unseen}$; and globally unseen classes $\mathcal{C}_{i,gu}$, in which $\mathcal{C}_{i,gu} = \mathcal{C}_{i,unseen} \setminus \mathcal{C}_{i,lu}$.

When unseen classes exist among multiple clients in FedoSSL, two main challenges need to be considered. Firstly, locally unseen classes may be learned faster than globally unseen classes due to the facilitation of client collaboration on locally unseen classes. The existing unsupervised pairwise loss $\mathcal{L}_i^u$ treats each class equally, while the imbalanced training progresses among unseen classes will result in a big bias on pseudo-label generation and even performance degradation in seen classes. Thus, we propose an uncertainty-aware regularization loss to alleviate training bias among different classes. Moreover, feature-level unsupervised pairwise loss on unlabeled data makes the generated cluster/class id heterogeneous among different clients since both labeled data and unlabeled data are required to feed into the same model classifier. It is critical to design a calibration strategy to align the outputs of the same unseen classes during the model aggregation phase.

3.2.1. Objective

To achieve unbiased training among different unseen classes in FedSSL, we present the proposed FedoSSL method. The overall objective consists of three parts: 1) fundamental semi-supervised loss for all data; 2) an uncertainty-aware regularization loss to reduce the training gap among locally unseen and globally unseen classes; 3) a calibration loss to

achieve efficient model aggregation:

$$\mathcal{L}_i^* = \mathcal{L}_i + \beta\mathcal{R}_i + \gamma\mathcal{L}_i^{cal} \tag{4}$$

where β and γ are trade-off hyper-parameters.

3.2.2. Uncertainty-aware Loss

Considering that different types of unseen classes have various training progress, i.e, locally unseen classes can be facilitated from model collaboration, we seek to add a regularization term to alleviate training divergence between locally unseen and globally unseen classes. Specifically, we use the uncertainty of the generated cluster/class id to reflect the training progress and apply a larger penalty for samples with high uncertainty. Then, the uncertainty-aware loss can be defined as:

$$\mathcal{R}_i = \frac{1}{n_i^u} \sum_{x_j^u \in \mathcal{D}_i^u} |\pi(x_j^u)| \tag{5}$$

where $\pi(\cdot)$ is the data uncertainty function. To precisely explore the uncertainty, we rely on both the confidence of the pseudo-label computed from the output of the softmax function and the proportion of samples belonging to related pseudo-label, i.e.,

$$\pi(x_j^u) = \rho(n^c | \arg\max_c p(x_j^u;\theta))[1 - \max_c p(x_j^u;\theta)] \tag{6}$$

where $\rho(n^c) = -\tau^{1-\frac{n^c}{n_{\max}}}$ is the weight for class c, which can be estimated from the labeled data, $\tau \in (0, 1]$. n^c is the number of training samples of the class c predicted by the model. $n_{\max}$ is the number of samples of the class with the maximal size. Note that $\rho(n^c)$ can be any function inversely proportional to n^c. In this paper, we focus more on unbiased training among locally unseen and globally unseen classes, while the training inconsistency between seen classes and unseen classes is ignored as both ORCA (Cao et al., 2022) and NACH (Guo et al., 2022) have tackled this problem.

3.2.3. Calibration Module

Due to the unseen classes being classified into new clusters by the pair-wise objective, i.e., $\mathcal{L}_i^u$, which only ensures similar samples are classified into one group/cluster, the same unseen class may be classified with different cluster id on different clients. For example, in Figure 2, unseen classes 4, 5 and 6 would be classified with any labels between 4∼9 because of lacking supervision from the labeled data. Such label heterogeneity would significantly degrade the aggregation performance. Therefore, it is required to design a calibration module to align the heterogeneous outputs of local classifiers before the aggregation phase.

Inspired by a clustering-based FL technique (Lubana et al., 2022) that aims to align local clustering performances among different clients via adding a global centroids aggregation mechanism, we extend this technique to our FedoSSL scenario and use the global centroids as a self-supervised signal to guide the classification process on unseen classes. The loss of the calibration module can be represented as

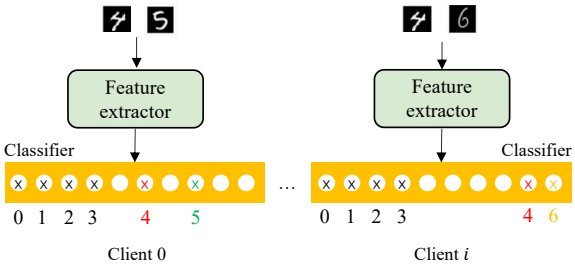


Figure 2: Illustration of label heterogeneity in FedoSSL. In a 10-class classification example, classes {0, 1, 2, 3} are seen classes, while classes {4, 5, 6} are unseen classes. Due to the feature-level pair-wise unsupervised loss (i.e., $\mathcal{L}_i^u$) on unlabeled data, same unseen class would be classified with different label id on different clients, e.g., unseen class 4 would be classified into the sixth position of the client 0's classifier, while in client i class 4 would be classified into the eighth position of the classifier.

$$\mathcal{L}_i^{cal} = \mathcal{L}_i^{ce} + \mathcal{L}_i^{cluster} \tag{7}$$

Specifically, we first aggregate local centroids from all clients to obtain global clusters at the server, i.e, by using Sinkhorn-Knopp (Genevay et al., 2019) clustering. The global centroids are then returned to the clients for further calibration. Taking the global centroids as guidance, the outputs of the classifier on unseen classes can be updated by approaching the cluster assignments with a cross-entropy loss:

$$\mathcal{L}_i^{ce} = \frac{1}{n_i^u} \sum_{z_j \in Z_i^u} \mathcal{H}(q(z_j;m), p(w^\top \cdot z_j)) \tag{8}$$

where m denotes the global centroids. Moreover, to prevent the cluster assignment of one class from changing dramatically during the training process, we additionally design the following loss function for promoting clusterability of feature representations:

$$\mathcal{L}_i^{cluster} = \frac{1}{n_i^u} \sum_{z_j, \bar{z}_j \in Z_i^u, \bar{Z}_i^u} \mathcal{H}(q(z_j;m), q(\bar{z}_j;m)) \tag{9}$$

where $q(z_j;m)$ and $q(\bar{z}_j;m)$ are corresponding cluster assignments, which are computed by matching representations with global centroids.

3.3. Algorithm Workflow

In this subsection, we illustrate the overall learning framework of FedoSSL, which is summarized in Algorithm 1.

Algorithm 1 FedoSSL Algorithm

Input: Number of clients K, learning rate η, local epochs E, labeled data $\{\mathcal{D}_1^l, \mathcal{D}_2^l, \ldots, \mathcal{D}_K^l\}$, unlabeled data $\{\mathcal{D}_1^u, \mathcal{D}_2^u, \ldots, \mathcal{D}_K^u\}$, hyperparameter α, β, γ
Output: Final model θ

```
Initialize the model parameter θ
repeat
    Sample a set of clients S.
    for each client i ∈ S in parallel do
        for j = 1 to E do
            Update local model: θ_i ← θ_i − η∇_θ L_i^*(θ_i)
        end for
        Calculate local centroids m_i
    end for
    Update global model: θ ← (n_i / Σ_{i∈S} n_i n) Σ_{i∈S} θ_i
    Update global centroids m
    Distribute θ and m to all clients
until Model converges
```

Client Update. In each communication round, the clients download the global model and global centroids from the server. In the client update phase, each of them makes several local gradient-based updates (e.g., E epochs) to optimize the local objective in Eq. (4). Then, local centroids are computed via a Sinkhorn-Knopp based clustering algorithm (Genevay et al., 2019).

Server Aggregation. After training the local models at the client-side, both updated models and local centroids will be sent to the server for further aggregation. Specifically, the server first aggregates the local models by taking a weighted average of them. Then, global centroids are calculated by aggregating local centroids (i.e., again using Sinkhorn-Knopp clustering). The above steps will be repeated until the model achieves convergence.

4. Experiments

4.1. Experimental Setup

Dataset. We evaluate the FedoSSL framework over three datasets CIFAR-10, CIFAR-100, and CINIC-10 (Darlow et al., 2018). CINIC-10 is a larger dataset that is constructed from CIFAR-10 and ImageNet. For all datasets, we first divide classes into 60% seen and 40% unseen classes, then select 50% of seen classes as the labeled data and the rest as unlabeled data. For CIFAR-10 and CINIC-10 datasets, one class of unseen classes is selected as the globally unseen class and rest 3 classes are locally unseen classes, each client owns all 6 seen classes, one globally unseen class and one locally unseen class. For CIFAR-100 dataset, 10 classes of unseen classes are selected as the globally unseen class and rest 30 classes are locally unseen classes, each client owns all 60 seen classes, 10 globally unseen classes and 10 locally unseen classes.

Baselines. To compare our FedoSSL with state-of-the-art methods, we conduct experiments from two perspectives: 1) extending existing open-world SSL methods to FL environments; 2) extending existing FedSSL methods to the open-world scenarios.

For the former one, we extend ORCA and NACH to be applicable to federated learning scenarios by implementing two representative FL algorithms with them, i.e., FedAvg (McMahan et al., 2017) and FedRep (Collins et al., 2021). Specifically, the extended baselines are named FedAvg+ORCA (**Fed-AO**), FedRep+ORCA (**Fed-RO**), FedAvg+NACH (**Fed-AN**) and FedRep+NACH (**Fed-RO**). Note that FedRep is a personalized federated learning algorithm that keeps each client's classifier updating locally, while the other parts are aggregated at the server. The reason for implementing FedRep is that the same unseen class may be classified with different labels on different classifiers due to the lack of uniform labels among different clients when extending ORCA and NACH to a distributed environment. In order to avoid the confusion caused by the aggregation of local classifiers, FedRep does not aggregate the classifiers.

For FedSSL methods, we adopt one state-of-the-art method **SemiFL** (Diao et al., 2022) to the open-world setting in the following way: we use SemiFL to classify samples into seen classes and estimate out-of-distribution (OOD) samples based on softmax confidence scores. Note that SemiFL has already shown superiority over FedMatch (Jeong et al., 2021), thus we ignore FedMatch as a baseline in the paper.

To ensure fairness in performance comparison, we also run the open-world SSL methods (i.e., ORCA, NACH) in the centralized (i.e., **Cen-O**, **Cen-N**) and distributed (i.e., **Local-O**, **Local-N**) environment without aggregation, respectively, to indicate the upper bound and lower bound of the performance of our FedoSSL method.

Implementation Details. For all datasets, we use ResNet-18 as the backbone model and train the model using standard SGD with a momentum of 0.9 and a weight decay of 5×10^{-4}. The dimension of the classifier corresponds to the number of classes in each dataset. Unless otherwise explicitly specified, α, β, γ are set to 1. The model is trained for 50 global rounds with 5 local epochs in each round. The batch size is 512 for all experiments. Similar to ORCA (Cao et al., 2022), we only update the parameters of the last block of ResNet in the second training stage to avoid over-fitting. For Sinkhorn-Knopp clustering (Genevay et al., 2019) in our FedoSSL, we compute 32 local centroids and 10 global centroids for CIFAR-10 and CINIC-10 and 128 local centroids and 100 global centroids for CIFAR-100. We evaluate all baselines and our FedoSSL in two settings: 10 clients

Table 2: Classification accuracy of compared methods on seen, unseen and all classes with 10 clients over three benchmark datasets. Asterisk ($*$) in *SemiFL denotes that the original methods cannot classify unseen classes (and we had to extend it). On unseen classes, *LU.* denotes locally unseen classes, while *GU.* denotes globally unseen classes. *AU.* represents the overall accuracy of all unseen classes. Gray rows indicate the upper bound of the model performance of FedoSSL.

#Method	CIFAR-10 (%)					CIFAR-100 (%)					CINIC-10 (%)				
	All	Seen	Unseen			All	Seen	Unseen			All	Seen	Unseen		
			LU.	GU.	AU.			LU.	GU.	AU.			LU.	GU.	AU.
Cen-O	78.26	86.63	-	-	71.95	56.92	73.68	-	-	44.28	69.32	83.18	-	-	58.86
Cen-N	81.02	89.47	-	-	74.64	58.98	75.10	-	-	46.82	71.89	83.82	-	-	62.89
Local-O	65.98	79.57	-	-	45.60	43.10	54.33	-	-	26.25	55.33	65.23	-	-	40.48
Local-N	67.67	83.95	-	-	43.26	45.28	57.24	-	-	27.34	57.31	65.70	-	-	44.73
Fed-AO	69.46	81.01	89.38	42.03	52.15	47.91	59.67	38.07	29.12	30.26	54.85	63.22	71.31	37.88	42.29
Fed-RO	71.72	82.22	89.84	53.43	55.96	47.72	59.79	44.13	28.86	29.62	57.16	62.26	72.24	42.09	49.50
Fed-AN	66.58	84.18	78.76	37.58	40.15	47.25	58.24	42.11	30.44	30.77	53.49	63.61	66.78	36.06	38.32
Fed-RN	68.83	**85.52**	79.84	41.79	43.81	48.02	59.4	**48.77**	30.36	30.96	58.11	65.97	68.81	39.01	46.33
*SemiFL	64.91	81.57	86.33	31.16	39.92	42.28	54.94	31.68	21.46	23.29	52.27	62.72	64.53	37.21	37.34
FedoSSL	**76.26**	84.29	**90.68**	**59.69**	**64.22**	**51.58**	**61.12**	45.76	**33.82**	**31.13**	**63.82**	**68.40**	**79.79**	**47.78**	**56.96**

Table 3: Analysis of Loss function: classification accuracy on CIFAR-10 (the number of clients: 10).

METHOD	SEEN	UNSEEN	ALL
FED-AO	81.01	52.15	69.46
FEDOSSL-$\mathcal{R}_i$-$\mathcal{L}_i^{ce}$	83.53	52.24	71.01
FEDOSSL-$\mathcal{R}_i$	83.13	62.98	75.07
FEDOSSL	84.29	64.22	76.26

Table 4: Analysis of Loss function: classification accuracy on CINIC-10 (the number of clients: 10).

METHOD	SEEN	UNSEEN	ALL
FED-AO	63.22	42.29	54.85
FEDOSSL-$\mathcal{R}_i$-$\mathcal{L}_i^{ce}$	69.10	40.31	57.58
FEDOSSL-$\mathcal{R}_i$	67.59	47.73	59.65
FEDOSSL	68.40	56.69	63.82

with 50% participation ratio and 50 clients with 10% participation ratio. The final average model accuracy of all clients is obtained from the best round among all global rounds. All compared methods are implemented based on the pre-trained model using the contrastive learning algorithm SimCLR (Chen et al., 2020a).

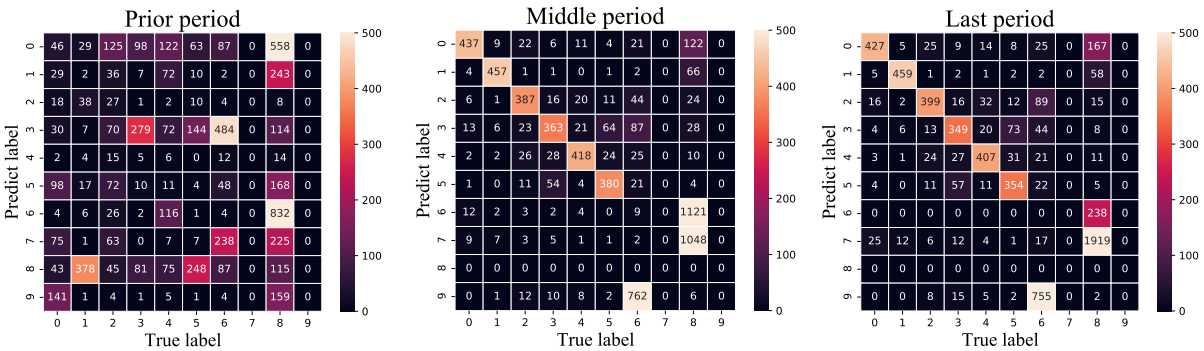
We simulate all clients and the server on a workstation with an RTX 2080Ti GPU, a 3.6-GHZ Intel Core i9-9900KF CPU and 64GB of RAM.

Remark. We are aware that many tricks, e.g., a diversity-constrained regularization loss to avoid trivial solutions (Cao et al., 2022; Guo et al., 2022) or a dynamic threshold to filter samples with lower confidence (Guo et al., 2022), etc., both can be added to our FedoSSL framework. However, to avoid losing focus, we ignore these tricks in this paper.

4.2. Performance Comparison

First, we compare our proposed FedoSSL methods with the baseline methods over three benchmark datasets. The classification accuracy on CIFAR-10, CIFAR-100 and CININ-10 dataset are listed in Table 2. It should be noted that the overall unseen accuracy (AU.) is not always a weighted average of locally unseen accuracy (LU.) and globally unseen accuracy (GU.). Because sometimes most samples of a locally unseen class and most samples of a globally unseen class will be classified in the same label, in this case, when calculating overall unseen accuracy, we have to choose another label (i.e., with the second largest number of samples) for one of those two classes.

From the results, it can be observed that our proposed FedoSSL provides superior performance of overall classification accuracy than baselines and the locally trained versions (i.e., Local-O, Local-N) over all three datasets. In most cases, FedoSSL maintains robust performance on seen classes and locally unseen classes. On globally unseen classes and overall unseen classes, FedoSSL achieves remarkable performance gains. For example, for globally unseen classes, FedoSSL outperforms the best baselines by 11.72% on CIFAR-10, 11.10% on CIFAR-100 and 13.52% on CINIC-10. For overall unseen classes, FedoSSL outperforms the best baselines by 14.76% on CIFAR-10 and 15.07% on CINIC-10. Moreover, we prove that the performance gap between locally and globally unseen classes on FedoSSL has significantly reduced when compared with other methods, i.e., there is a 47% gap between locally and globally unseen classes in Fed-AO, while FedoSSL reduces this gap to 31% on CIFAR-10.



(a) In initial stage. (b) In intermediate stage. (c) In final stage.

Figure 3: Visualization of the predicted clustering assignments in different training stages.

Table 5: Classification accuracy of compared methods on seen, unseen and all classes with 50 clients over three benchmark datasets.

	CIFAR-10 (%)			CIFAR-100 (%)			CINIC-10 (%)		
#Method	All	Seen	Unseen	All	Seen	Unseen	All	Seen	Unseen
Fed-AO	70.22	83.34	50.54	45.63	56.25	29.69	53.81	60.49	43.80
Fed-RO	71.36	84.31	51.93	45.18	56.78	27.79	57.26	61.70	50.61
Fed-AN	69.89	85.36	46.68	45.22	56.30	28.59	53.42	63.62	38.13
Fed-RN	71.49	**86.28**	49.30	45.57	56.79	28.73	57.81	65.29	46.60
FedoSSL	**76.41**	85.71	**62.46**	**47.01**	**58.34**	**30.17**	**64.02**	**69.56**	**55.71**

4.3. Ablation Study

In this subsection, detailed analyses are shown to help understand the superiority of our proposed FedoSSL framework, including analysis on the two additional modules: $\mathcal{R}_i$ and $\mathcal{L}_i^{cal}$, effect of the number of seen classes, robustness on large-scale of FL scenarios, and hyper-parameter sensitivity analysis.

Analysis on Objective Functions. We evaluate our proposed two objective functions $\mathcal{R}_i$ and $\mathcal{L}_i^{cal}$ (i.e., $\mathcal{L}_i^{cal}$ consists of $\mathcal{L}_i^{ce}$ and $\mathcal{L}_i^{cluster}$) of FedoSSL on CIFAR-10 and CINIC-10 dataset. We use the Fed-AO as the basic baseline. First, FedoSSL-$\mathcal{R}_i$-$\mathcal{L}_i^{ce}$ means that only adding $\mathcal{L}_i^{cluster}$ to the baseline, it can be observed that the accuracy of seen classes has improved. Then, FedoSSL-$\mathcal{R}_i$ means that adding both $\mathcal{L}_i^{ce}$ and $\mathcal{L}_i^{cluster}$ could greatly improve the accuracy of unseen classes. Finally, the case of using all two objective functions, i.e., the complete FedoSSL, further improves the accuracy of the unseen classes and achieves the best performance. Results in Table 3 and 4 give a clear ablation study to demonstrate the effectiveness of our proposed objective functions. Besides that, we record the calculated clustering assignments in the initial stage, intermediate stage and final stage, respectively. Figure 3 shows that sharing local centroids can significantly help to calibrate the heterogeneous outputs of the classifier on unseen classes.

Number of Seen Class. To demonstrate the effectiveness of our proposed FedoSSL, we systematically evaluate the performance when changing the ratio of seen classes on CIFAR-10 dataset with 10 clients. Fed-RO is selected as the baseline because it is the best among the 4 baselines. The results in Figure 4 show that FedoSSL always outperforms Fed-RO in all cases. For example, with 40% of seen classes, FedoSSL achieves a 9.8% overall accuracy improvement from Fed-RO. Besides, the performance of FedoSSL is stable for all ratio values and improves significantly with the increasing ratio of seen class.

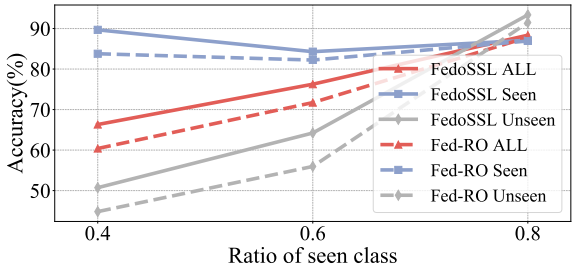


Figure 4: Performance of Fed-RO and FedoSSL with different numbers of seen classes on CIFAR-10.

Analysis on Robustness of FedoSSL. To further demonstrate the robustness of our proposed FedoSSL at different experimental scales, we evaluated it in a 50-client scenario with the same other client settings, as shown in Table 5. The results show that FedoSSL still has the best overall accuracy in all datasets, and the performance advantage of unseen classes remains significant. For CIFAR-10 dataset, FedoSSL outperforms the best baselines by 6.88% on the overall accu-

Table 6: Sensitivity to number of local clusters on CIFAR-10. The number of global centroids is 10.

L	All	Seen	Unseen		
			LU.	GU.	AU.
8	74.28	84.26	88.90	54.09	59.29
16	75.76	84.17	89.28	58.36	63.15
32	76.26	84.29	90.68	59.69	64.22

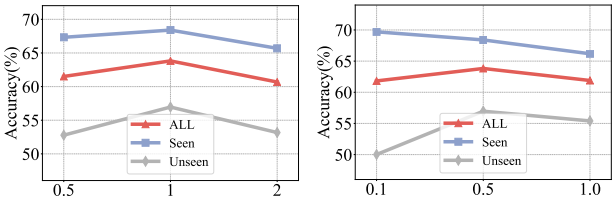
Table 7: Accuracy obtained using different privacy-guarnteed version of FedoSSL on CIFAR-10. 'No Privacy' represents the idealized setting when local representations are shared with the server. The number of global centroids is 10.

	All	Seen	Unseen		
			LU.	GU.	AU.
No Privacy	77.19	85.95	89.76	58.77	64.05
K-anonymity	76.26	84.29	90.68	59.69	64.22

racy and 20.28% on the unseen accuracy, which is 10.74% and 10.08% for the CINIC-10, respectively. Overall, our results demonstrate that FedoSSL has good robustness in large-scale FL scenarios.

Number of Local Centroids. In the calibration module, the key insight is to upload local centroids instead of the representations of all data to enable computation of the global clusters. The local clustering methods can be flexibly chosen, depending on the system's constraints, e.g., for strong privacy guarantees at possibly high loss in utility, locally differentially private (local-DP) clustering methods can be used (i.e., DP-k-Means (Balcan et al., 2017), Local-DP (Chang et al., 2021)); for slightly weaker guarantees but higher utility, K-anonymous clustering methods can be used (i.e., r-Gather clustering (Aggarwal et al., 2010), Sinkhorn-Knopp based clustering (Genevay et al., 2019)). To further demonstrate the effectiveness of FedoSSL, we study the sensitivity of FedoSSL to the number of local centroids (i.e., L). Table 6 shows that FedoSSL is robust to the number of local clusters as long as it even slightly exceeds the number of classes, achieving similar performance in all settings.

Analysis on Privacy. Our proposed FedoSSL uses Sinkhorn-Knopp based clustering algorithm to compute L equally-sized local clusters. This operation enables a n/L-anonymity privacy guarantee across all n samples present on a client (Lubana et al., 2022). Furthermore, it is also feasible to use other local-DP based clustering methods instead of Sinkhorn-Knopp based clustering to provide stronger privacy guarantees. Therefore, we compare our proposed K-anonymity-guaranteed FedoSSL with the no-privacy-guaranteed version (i.e., replace Sinkhorn-Knopp based clustering with general k-means method). As we can see from the Table 7, FedoSSL witnesses only a small performance drop w.r.t K-anonymity.



(a) β (b) γ

Figure 5: Classification accuracy on different setting of β (a) and γ (b).

Analysis on Hyperparameters. We further analyze the impact of hyper-parameters β and γ on CINIC-10 dataset with 10 clients. Figure 5(a) and 5(b) provide the performance of different β and γ respectively. For β, the bigger value means the bigger effect of $\mathcal{R}_i$. When $\beta = 1$, the proposal achieves the best performance on all three accuracies and the performance does not degrade severely with β changes. This demonstrates that FedoSSL is quite robust with the selection of β. For γ, the bigger value means the bigger effect of $\mathcal{L}_i^{cal}$, namely, the model will be more influenced by global centroids. The results show that when $\gamma = 0.5$ the proposal achieves the best overall accuracy. When $\gamma = 0.1$, unseen classes' accuracy degrades severely, which demonstrates that the value of γ should not be too small.

5. Conclusion

In this paper, we have investigated a new federated semi-supervised learning framework under the open-world setting, **FedoSSL**, to provide unbiased training for both labeled and unlabeled data. Specifically, we categorize the unseen classes as locally and globally unseen with the principle of whether the related unseen classes exist in only one client or not. An uncertainty-aware suppressed loss and a calibration module are then designed to balance the training pace among locally unseen and globally unseen classes and supplement global model aggregation phase, respectively. We have proved that FedoSSL is a general framework for FedSSL and can be easily compatible with state-of-the-art FL methods. Extensive empirical experiments have been conducted over various models and datasets to verify the effectiveness and superior performance of FedoSSL framework.

Acknowledgements

This research was supported by fundings from the Key-Area Research and Development Program of Guangdong Province (No. 2021B0101400003), Hong Kong RGC Research Impact Fund (No. R5060-19), General Research Fund (152203/20E, 152244/21E and 152169/22E), Shenzhen Science and Technology Innovation Commission (JCYJ20200109142008673), Areas of Excellence Scheme (AoE/E-601/22-R).

References

Aggarwal, G., Panigrahy, R., Feder, T., Thomas, D., Kenthapadi, K., Khuller, S., and Zhu, A. Achieving anonymity via clustering. *ACM Transactions on Algorithms (TALG)*, 6(3):1–19, 2010.

Balcan, M.-F., Dick, T., Liang, Y., Mou, W., and Zhang, H. Differentially private clustering in high-dimensional euclidean spaces. In *International Conference on Machine Learning*, pp. 322–331. PMLR, 2017.

Boult, T. E., Cruz, S., Dhamija, A. R., Gunther, M., Henrydoss, J., and Scheirer, W. J. Learning and the unknown: Surveying steps toward open world recognition. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pp. 9801–9807, 2019.

Cao, K., Brbic, M., and Leskovec, J. Open-world semi-supervised learning. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=O-r8LOR-CCA.

Chang, A., Ghazi, B., Kumar, R., and Manurangsi, P. Locally private k-means in one round. In *International Conference on Machine Learning*, pp. 1441–1451. PMLR, 2021.

Chen, T., Kornblith, S., Norouzi, M., and Hinton, G. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pp. 1597–1607. PMLR, 2020a.

Chen, Y., Zhu, X., Li, W., and Gong, S. Semi-supervised learning under class distribution mismatch. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pp. 3569–3576, 2020b.

Collins, L., Hassani, H., Mokhtari, A., and Shakkottai, S. Exploiting shared representations for personalized federated learning. In *International Conference on Machine Learning*, pp. 2089–2099. PMLR, 2021.

Darlow, L. N., Crowley, E. J., Antoniou, A., and Storkey, A. J. Cinic-10 is not imagenet or cifar-10. *arXiv preprint arXiv:1810.03505*, 2018.

Diao, E., Ding, J., and Tarokh, V. Semifl: Semi-supervised federated learning for unlabeled clients with alternate training. *Advances in Neural Information Processing Systems*, 35:17871–17884, 2022.

Fallah, A., Mokhtari, A., and Ozdaglar, A. E. Personalized federated learning with theoretical guarantees: A model-agnostic meta-learning approach. In *NeurIPS*, 2020.

Genevay, A., Dulac-Arnold, G., and Vert, J.-P. Differentiable deep clustering with cluster size constraints. *arXiv preprint arXiv:1910.09036*, 2019.

Guo, L.-Z., Zhang, Z.-Y., Jiang, Y., Li, Y.-F., and Zhou, Z.-H. Safe deep semi-supervised learning for unseen-class unlabeled data. In *International Conference on Machine Learning*, pp. 3897–3906. PMLR, 2020.

Guo, L.-Z., Zhang, Y.-G., Wu, Z.-F., Shao, J.-J., and Li, Y.-F. Robust semi-supervised learning when not all classes have labels. In Oh, A. H., Agarwal, A., Belgrave, D., and Cho, K. (eds.), *Advances in Neural Information Processing Systems*, 2022. URL https://openreview.net/forum?id=lDohSFOHr0.

Han, K., Vedaldi, A., and Zisserman, A. Learning to discover novel visual categories via deep transfer clustering. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8401–8409, 2019.

He, C., Yang, Z., Mushtaq, E., Lee, S., Soltanolkotabi, M., and Avestimehr, S. Ssfl: Tackling label deficiency in federated learning via personalized self-supervision. *arXiv preprint arXiv:2110.02470*, 2021.

Hsu, Y.-C., Lv, Z., Schlosser, J., Odom, P., and Kira, Z. Multi-class classification without multi-class labels. *arXiv preprint arXiv:1901.00544*, 2019.

Huang, J., Fang, C., Chen, W., Chai, Z., Wei, X., Wei, P., Lin, L., and Li, G. Trash to treasure: Harvesting ood data with cross-modal matching for open-set semi-supervised learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8310–8319, 2021.

Jeong, E., Oh, S., Kim, H., Park, J., Bennis, M., and Kim, S.-L. Communication-efficient on-device machine learning: Federated distillation and augmentation under non-iid private data. *CoRR*, 2018.

Jeong, W., Yoon, J., Yang, E., and Hwang, S. J. Federated semi-supervised learning with inter-client consistency & disjoint learning. In *International Conference on Learning Representations (ICLR) 2021*. International Conference on Learning Representations (ICLR), 2021.

Jin, Y., Wei, X., Liu, Y., and Yang, Q. Towards utilizing unlabeled data in federated learning: A survey and prospective. *arXiv preprint arXiv:2002.11545*, 2020.

Lee, D., Kim, S., Kim, I., Cheon, Y., Cho, M., and Han, W.-S. Contrastive regularization for semi-supervised learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3911–3920, 2022.

Lee, D.-H. et al. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In *Workshop on challenges in representation learning, ICML*, volume 3, pp. 896, 2013.

Li, D. and Wang, J. Fedmd: Heterogenous federated learning via model distillation. *arXiv preprint arXiv:1910.03581*, 2019.

Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., and Smith, V. Federated optimization in heterogeneous networks. *Proceedings of Machine Learning and Systems*, 2:429–450, 2020a.

Li, X., Huang, K., Yang, W., Wang, S., and Zhang, Z. On the convergence of fedavg on non-iid data. In *Proc. of ICLR*, 2020b.

Liang, X., Lin, Y., Fu, H., Zhu, L., and Li, X. Rscfed: Random sampling consensus federated semi-supervised learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10154–10163, 2022.

Lin, H., Lou, J., Xiong, L., and Shahabi, C. Semifed: Semi-supervised federated learning with consistency and pseudo-labeling. *arXiv preprint arXiv:2108.09412*, 2021.

Lin, T., Kong, L., Stich, S. U., and Jaggi, M. Ensemble distillation for robust model fusion in federated learning. In *NeurIPS*, 2020.

Liu, Q., Yang, H., Dou, Q., and Heng, P.-A. Federated semi-supervised medical image classification via inter-client relation matching. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 325–335. Springer, 2021.

Long, Z., Che, L., Wang, Y., Ye, M., Luo, J., Wu, J., Xiao, H., and Ma, F. Fedsiam: Towards adaptive federated semi-supervised learning. *arXiv preprint arXiv:2012.03292*, 2020.

Lubana, E. S., Tang, C. I., Kawsar, F., Dick, R. P., and Mathur, A. Orchestra: Unsupervised federated learning via globally consistent clustering. *arXiv preprint arXiv:2205.11506*, 2022.

McMahan, B., Moore, E., Ramage, D., Hampson, S., and y Arcas, B. A. Communication-efficient learning of deep networks from decentralized data. In *Proceedings of Artificial Intelligence and Statistics (AISTATS)*, 2017.

Ng, D., Lan, X., Yao, M. M.-S., Chan, W. P., and Feng, M. Federated learning: a collaborative effort to achieve better medical imaging models for individual sites that have small labelled datasets. *Quantitative Imaging in Medicine and Surgery*, 11(2):852, 2021.

Ouali, Y., Hudelot, C., and Tami, M. An overview of deep semi-supervised learning. *arXiv preprint arXiv:2006.05278*, 2020.

Pham, H., Dai, Z., Xie, Q., and Le, Q. V. Meta pseudo labels. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11557–11568, 2021.

Rasmus, A., Berglund, M., Honkala, M., Valpola, H., and Raiko, T. Semi-supervised learning with ladder networks. *Advances in neural information processing systems*, 28, 2015.

Saito, K., Kim, D., and Saenko, K. Openmatch: Open-set semi-supervised learning with open-set consistency regularization. *Advances in Neural Information Processing Systems*, 34:25956–25967, 2021.

Sohn, K., Berthelot, D., Carlini, N., Zhang, Z., Zhang, H., Raffel, C. A., Cubuk, E. D., Kurakin, A., and Li, C.-L. Fixmatch: Simplifying semi-supervised learning with consistency and confidence. *Advances in neural information processing systems*, 33:596–608, 2020.

Tarvainen, A. and Valpola, H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. *Advances in neural information processing systems*, 30, 2017.

Wei, C., Shen, K., Chen, Y., and Ma, T. Theoretical analysis of self-training with deep networks on unlabeled data. In *International Conference on Learning Representations*, 2020.

Xie, Q., Dai, Z., Hovy, E., Luong, T., and Le, Q. Unsupervised data augmentation for consistency training. *Advances in Neural Information Processing Systems*, 33: 6256–6268, 2020.

Yoshida, N., Nishio, T., Morikura, M., Yamamoto, K., and Yonetani, R. Hybrid-fl: Cooperative learning mechanism using non-iid data in wireless networks. *CoRR*, 2019.

Zhang, J., Guo, S., Ma, X., Wang, H., Xu, W., and Wu, F. Parameterized knowledge transfer for personalized federated learning. *Advances in Neural Information Processing Systems*, 34, 2021a.

Zhang, J., Qu, Z., Chen, C., Wang, H., Zhan, Y., Ye, B., and Guo, S. Edge learning: The enabling technology for distributed big data analytics in the edge. *ACM Computing Surveys (CSUR)*, 54(7):1–36, 2021b.

Zhang, Z., Yang, Y., Yao, Z., Yan, Y., Gonzalez, J. E., Ramchandran, K., and Mahoney, M. W. Improving semi-supervised federated learning by reducing the gradient diversity of models. In *2021 IEEE International Conference on Big Data (Big Data)*, pp. 1214–1225. IEEE, 2021c.

Zhao, Y., Li, M., Lai, L., Suda, N., Civin, D., and Chandra, V. Federated learning with non-iid data. *CoRR*, 2018.

Zhong, Z., Fini, E., Roy, S., Luo, Z., Ricci, E., and Sebe, N. Neighborhood contrastive learning for novel class discovery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10867–10875, 2021.

Zhou, Z.-H. Open-environment machine learning. *National Science Review*, 9(8):nwac123, 2022.

Zhou, Z.-H. and Li, M. Tri-training: Exploiting unlabeled data using three classifiers. *IEEE Transactions on knowledge and Data Engineering*, 17(11):1529–1541, 2005.

Zou, Y., Zhang, Z., Zhang, H., Li, C.-L., Bian, X., Huang, J.-B., and Pfister, T. Pseudoseg: Designing pseudo labels for semantic segmentation. *arXiv preprint arXiv:2010.09713*, 2020.