ConFREE: Conflict-free Client Update Aggregation for Personalized Federated Learning

Hao Zheng[1,2], **Zhigang Hu**[1], **Liu Yang**[1*], **Meiguang Zheng**[1], **Aikun Xu**[1], **Boyu Wang**[2*]

[1]School of Computer Science and Engineering, Central South University, Changsha, China
[2]Department of Computer Science, University of Western Ontario, London, Canada
{zhenghao, zghu, yangliu, zhengmeiguang, aikunxu}@csu.edu.cn; bwang@csd.uwo.ca

Abstract

Negative transfer (NF) is a critical challenge in personalized federated learning (pFL). Existing methods primarily focus on adapting the local data distribution on the client side, which can only resist NF, rather than avoiding NF itself. To tackle NF at its root, we investigate its mechanism through the lens of the global model, and argue that it is caused by update conflicts among clients during server aggregation. In light of this, we propose a conflict-free client update aggregation strategy (ConFREE), which enables us to avoid NF in pFL. Specifically, ConFREE guides the global update direction by constructing a conflict-free guidance vector through projection and utilizes the optimal local improvements of the worst-performing clients near the guidance vector to regularize server aggregation. This prevents the conflicting components of the updates from transferring, achieving balanced updates between different clients. Notably, ConFREE is model-agnostic and can be straightforwardly adopted as a complement to enhance various existing NF-resistance methods implemented on the client side. Extensive experiments demonstrate significantly enhances existing pFL algorithms when integrated with ConFREE.

Introduction

Federated Learning (FL) is a paradigm that enables multiple distributed clients to collaboratively train a global model without sharing their private data (Wen et al. 2023). With the increasing implementation of data privacy protection legislation (Regulation 2016; Goldman 2020), FL has found success in various artificial intelligence domains (Liu et al. 2020; Tan et al. 2020). However, the influx of vast amounts of non-independent and identically distributed (non-IID) and imbalanced data generated by different devices poses a challenge. Consequently, the traditional approaches of employing a single global model in FL often lead to significant deviations from local data distributions on individual clients, resulting in slow convergence and poor inference performance. To address the statistical heterogeneity in FL, personalized Federated Learning (pFL) has been proposed (Lee et al. 2024), which aims to learn a tailored model for each client while still benefiting from the global model to enhance their models and address data scarcity issues.

[*]Corresponding authors: Liu Yang, Boyu Wang

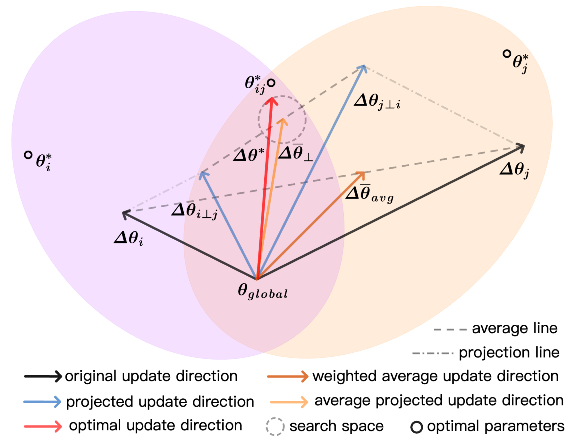

Figure 1: Weighted average aggregation vs. Conflict-free aggregation method. In non-IID settings, where client data varies widely and is imbalanced, updates $\Delta\theta_i$ and $\Delta\theta_j$ after training the global model θ_{global} on clients i and j may conflict when $\Delta\theta_i \cdot \Delta\theta_j < 0$. The classical FedAvg $\Delta\bar{\theta}_{\text{avg}}$ tends to be dominated by the $\Delta\theta_j$ from clients with more data, leading to insufficient optimization for client i. ConFREE addresses this by projecting conflicting updates onto a normal plane, creating a conflict-free guiding vector $\Delta\bar{\theta}_\perp$. The optimal $\Delta\theta^*$ is then found within a ball centered around $\Delta\bar{\theta}_\perp$, which maximizes the local improvement of the worst performing client in the neighborhood. This ensures the updated global model is closer to the global optimum θ_{ij}^* and balances the update across clients.

Although there has been a significant amount of research, negative transfer is still a critical issue in pFL (Kulkarni, Kulkarni, and Pant 2020; Mansour et al. 2020). Negative transfer refers to the situation where models are shared between different clients and the performance is degraded because some client updates negatively affect the shared model (Wu et al. 2023). Existing works have been studied from the client side (Collins et al. 2021; Zhang et al. 2023d) and from the server side (Sattler, Müller, and Samek 2020; Ye et al. 2023) to mitigate the impact of negative transfer. The former aims to preserve more local features while incorporating a small amount of global information through techniques such as specialized layers (Zhang et al. 2023d), regularization (Huang et al. 2021), or selective mixing (Zhang et al. 2023c). The latter focuses on collaborative training by grouping clients with similar update directions, e.g., clustering (Ye et al. 2023; Carrillo et al. 2024), which helps to avoid the influence of clients from different update directions. However, this limits the availability of global information only within a single cluster. Therefore, reducing the bias of the global model during server aggregation is important to help clients obtain more effective global information.

As an illustrative example, the most commonly used weighted average aggregation strategy in practical pFL can be dominated by the client with larger datasets, resulting in insufficient optimization for another client, which motivates us to seek an effective way to eliminate the effects of global model bias in the server aggregation. In this work, we argue that the model bias problem in pFL arises from conflicts in client updates. When two updates are far away from each other i.e., satisfy $\Delta\theta_i \cdot \Delta\theta_j < 0$, global model aggregation is hindered by these conflicting updates and lead to degraded performance of local models.

To reduce the client update conflicts during server aggregation, we propose a conflict-free client update aggregation (ConFREE) strategy for pFL. Specifically, we construct a global conflict-free guidance vector through projection and maximize the local improvement of the worst-performing client near the guidance vector to regularize server aggregation. As shown in Fig.1, we project each conflicting update onto the hyperplane orthogonal to the update of another client. The constructed global conflict-free guidance vector prevents the interference component of the updates from affecting the global aggregation. Furthermore, to balance the optimization objectives of different clients, we search within the neighborhood of the guidance vector for the optimal update vector that maximizes the worst local improvement for any client. To evaluate our method's effectiveness and scalability, we combined ConFREE with eight state-of-the-art (SOTA) methods and conducted experiments on four typical image classification tasks in two common FL heterogeneous environments. Extensive experimental results show that ConFREE can be seamlessly integrated with a range of personalization techniques to improve the accuracy of SOTA models. Additionally, we varied client numbers and participation rates to enrich statistical heterogeneity. The contributions of the paper are summarized as follows:

- To the best of our knowledge, ConFREE is the first method to address client update conflicts in pFL. By optimizing global model aggregation, it provides each client with more effective and comprehensive global information.
- ConFREE is model-agnostic, which requires a single modification to the server aggregation program that can be easily applied to enhance various client-side personalization techniques.
- Through enhancing different SOTA methods and conducting extensive evaluations on real-world datasets with different data heterogeneity, ConFREE has been proven to significantly boost their overall performance.

Related Work

PFL has garnered widespread attention for its ability to effectively address statistical heterogeneity by tailoring models to individual clients. In this section, we briefly review the research on pFL from two perspectives: client-side personalization and server-side personalization.

Client-Side Personalization

In pFL, client-side personalization is achieved through techniques such as designing specific layers (Ma et al. 2022), regularization (Zhang et al. 2023a) and selective mixtures (Chen et al. 2023) to better adapt to local data distributions. For example, early methods like FedPer (Arivazhagan et al. 2019), LG-FedAvg (Liang et al. 2020), FedRep (Collins et al. 2021) and GPFL (Zhang et al. 2023b) decouple model parameters, achieving effective personalization and global representation learning by sharing a feature extractor and fine-tuning a private classifier locally. FedCP (Zhang et al. 2023d) goes a step further by separating global features and personalized features at the data level, processing them with a global head and a personalized head, respectively. FedProx (Li et al. 2020), FedAMP (Huang et al. 2021), L2GD (Hanzely and Richtárik 2020), and pFedMe (T Dinh, Tran, and Nguyen 2020) incorporate a regularization term into the global model aggregation process to extract global information from global model parameters. Ditto (Li et al. 2021a) leverages a proximity term to locally train personalized models for each client, while FedROD (Chen and Chao 2021) aligns the distribution of client models with the global model using optimal transport. Selective mixing of local and global models can better balance generalization and personalization. APFL (Deng, Kamani, and Mahdavi 2020) and FedPHP (Li et al. 2021b) combines local and global models linearly, and FedALA (Zhang et al. 2023c) performs a weighted fusion of the global model from each round with the local model from the previous round to retain local personalized information.

Server-Side Personalization

Server-side personalization in pFL primarily involves grouping clients based on data distribution similarity and aggregating models within these groups (Ghosh et al. 2020; Long et al. 2023; Cai et al. 2023), while some works are also focused on designing highly generalizable global models (Achituve et al. 2021; Marfoq et al. 2022; Zhang et al. 2022; Singhal et al. 2021). For example, CFL (Sattler, Müller, and Samek 2020), FL+HC (Briggs, Fan, and Andras 2020), and PFedSA (Ye et al. 2023) adjust clustering by calculating the similarity between clients, where each cluster of clients undergoes model training and updates separately, and the characteristics of each cluster are weighted during global aggregation. Per-FedAvg (Fallah, Mokhtari, and Ozdaglar 2020) and (Acar et al. 2021) learns a shared meta-model that satisfies all clients, ensuring it has good generalization performance across all clients.

Due to the model bias problems caused by the non-IID nature of client data, client-side personalization literature focuses on decoupling local and global feature extractors to minimize the negative transfer caused by the global model. On the other hand, server-side personalization literature leverages clients with similar biases to train together, thereby incorporating as much global feature information as possible. Thus, the global model's feature information is crucial. However, these approaches overlook the issue of update bias during global model aggregation. This paper aims to reduce the parameter update conflicts during global aggregation, generating a global model that balances updates across different clients, thereby providing more effective global information to each client.

Method

Preliminaries

Consider a typical personalized federated learning system setup consisting of a server and N clients. Assume that N clients have local private datasets $\mathcal{D}_1, \ldots, \mathcal{D}_N$, which are non-IID and imbalanced. Under the collaborative training of the server and clients, the objective is to train a global representation network $\theta_g \in \Theta$ that contains the shared knowledge of all clients, as well as client-specific head $h_i \in \mathcal{H}$ for each client $i \in N$ that adapt to their data distributions. Therefore, the objective function can be described as:

$$\min_{\theta_g \in \Theta} \sum_{i=1}^{N} p_i \min_{h_i \in \mathcal{H}} \mathcal{L}_i \left(h_i, \theta_g; D_i\right) \tag{1}$$

where p_i is the proportional aggregation weight relative to the size of the local dataset $\mathcal{D}_i$ of client i, with $p_i = |D_i| / \sum_{j=1}^{N} |D_j|$. Additionally, $\mathcal{L}_i(h_i, \theta_g; D_i)$ is the loss function, which measures the empirical risk of the personalized model that combines the client-specific head h_i and the global shared network θ_g on the local dataset $\mathcal{D}_i$. Clearly, this objective function is a bilevel optimization problem, where the optimal solutions to the inner and outer problems are found through iterative client updates and server aggregations, respectively.

Client Update. For a given outer shared parameter θ_g, each client i updates its specific head h_i by minimizing its local loss function to optimize the inner problem. The update of h_i for the t-th global iteration is:

$$h_i^{t+1} \leftarrow h_i^t - \eta \nabla_{h_i^t} \mathcal{L}_i \left(h_i^t, \theta_g^t; D_i\right) \tag{2}$$

where η is the learning rate for the local client optimization. $\nabla_{h_i^t} \mathcal{L}_i \left(h_i^t, \theta_g^t; D_i\right)$ is the gradient of the loss function $\mathcal{L}_i$ with respect to the local parameter h_i^t. The goal is to reduce the empirical risk on the local dataset $\mathcal{D}_i$ by adjusting the client-specific head h_i while keeping the global shared parameter θ_g^t fixed. This iterative process ensures that each client's head is fine-tuned to its local data distribution, thereby optimizing the inner problem.

Once the optimal head h_i is found, the next step is to perform a gradient update on the shared parameter θ_g while keeping h_i fixed. The update step for the shared parameter θ_g^t in the t-th global iteration is given by:

$$\theta_i^t \leftarrow \theta_g^t - \alpha \nabla_{\theta_g^t} \mathcal{L}_i \left(h_i^{t+1}, \theta_g^t; D_i\right) \tag{3}$$

where α is the learning rate for the outer optimization, θ_i^t is a shared parameter that matches the data distribution $\mathcal{D}_i$ of client i. This process ensures that the global model incorporates the personalized adjustments from each client, enhancing the overall performance. After a round of gradient updates, the updated shared parameters θ_i^t are uploaded to the server for global aggregation.

Server Aggregation. At the end of the t-th communication round, the server receives the global shared parameters θ_i^t uploaded by all clients and updates the global parameters θ_g through Federated Averaging (FedAvg). The aggregation process is given by:

$$\theta_g^{t+1} \leftarrow \sum_{i=1}^{N} p_i^t \theta_i^t = \theta_g^t + \sum_{i=1}^{N} p_i^t \Delta\theta_i^t \tag{4}$$

where p_i^t represents the aggregation weight for client i, $\Delta\theta_i^t = -\alpha \nabla_{\theta_g^t} \mathcal{L}_i \left(h_i^{t+1}, \theta_g^t; D_i\right)$ is the change of shared parameter in client i in the t-th communication round, $\Delta\theta_{avg}^t = \sum_{i=1}^{N} p_i^t \Delta\theta_i^t$ is the classical weighted average aggregation. This step ensures that the global model θ_g^{t+1} integrates the updates from all clients, leading to global knowledge fusion.

Client Update Conflicts

Negative transfer refers to the situation when models are shared between different clients and the performance is degraded because some client updates negatively affect the shared model. Specifically, when the shared parameter θ_g is trained on the non-IID dataset D_i at local client i, the parameter update $\Delta\theta_i$ impacts the global shared parameters θ_g is $\theta_g \leftarrow \theta_g + \Delta\theta_i$. This minor adjustment to the θ_g can influence the performance of another client j as follows:

$$\Delta\mathcal{L}_j = \mathcal{L}_j(h_j, \theta_g + \Delta\theta_i) - \mathcal{L}_j(h_j, \theta_g) \tag{5}$$

by applying the first order Taylor approximation around θ, we expand $\mathcal{L}_j(h_j, \theta_g + \Delta\theta_i)$ as:

$$\begin{aligned} \mathcal{L}_j(h_j, \theta_g + \Delta\theta_i) \approx \mathcal{L}_j(h_j, \theta_g) + \nabla_{\theta_g} \mathcal{L}_j(h_j, \theta_g) \cdot \Delta\theta_i \\ + o(\|\Delta\theta_i\|) \end{aligned} \tag{6}$$

where $\nabla_{\theta_g} \mathcal{L}_j(h_j, \theta_g)$ is the gradient of $\mathcal{L}_j$ over θ_g, and $o(\|\Delta\theta_i\|)$ is a higher-order small quantity and thus can be considered negligible in the context of the equation. Using the definition of the parameter update, $\nabla_{\theta_g} \mathcal{L}_j(h_j, \theta_g)$ can be represented as $-\frac{1}{\alpha}\Delta\theta_j$. Substituting these expressions into the original formula, we get:

$$\Delta\mathcal{L}_j = -\frac{1}{\alpha} \Delta\theta_i \cdot \Delta\theta_j \tag{7}$$

from the above formula, it follows that when $\Delta\theta_i \cdot \Delta\theta_j < 0$, the parameter update from client i has a negative impact on client j because it leads to an increase in the loss of client j, and vice versa. Consequently, the essence of negative transfer is ***client update conflict***.

Furthermore, according to the classical weighted average aggregation $\Delta\theta_g \leftarrow \Delta\theta_{avg} = \sum_{i=1}^{N} p_i \Delta\theta_i$, the total average loss impact on all clients due to the global parameter update can be expressed as:

$$\Delta\mathcal{L}_{\text{total}} = -\frac{1}{\alpha}\sum_{i=1}^{N}\sum_{j=1}^{N} p_i p_j \Delta\theta_i \cdot \Delta\theta_j \tag{8}$$

where p_i represents the influence of each client's updates on the global model, typically determined by the client's data volume. When the parameter updates from dominant clients conflict with those of clients with fewer samples, the global loss can rise instead of decrease. Over time, these less influential clients remain under-optimized, further accumulating conflicts and exacerbating imbalances within the FL system.

ConFREE:Conflict-free Client Update Aggregation

Our first goal is to construct a conflict-free guidance vector that minimizes negative conflicts with parameter updates from other clients. Most importantly, based on this guidance vector, we seek an optimal update vector in its vicinity to maximize the worst local improvement among all clients in the neighborhood. The ConFREE method not only resolves conflicts in parameter updates but also ensures the best possible improvement for the worst-performing clients. This approach prevents the dominance of updates by clients with the largest sample sizes and promotes balanced optimization for all clients.

Guidance Vector. The first step is to determine if a conflict occurs by calculating the cosine similarity between each pair of client updates $\Delta\theta_i$ and $\Delta\theta_j$ (where $i, j \in N$ and $i \neq j$). If the cosine similarity is negative, resolve the conflict using the projection method; otherwise, retain the original update. Use the average of all conflict-free client updates as the guidance vector $\Delta\overline{\theta}_\perp$:

$$\Delta\overline{\theta}_\perp = \frac{1}{N}\sum_{i=1}^{N}\left(\Delta\theta_i - \sum_{j\neq i, \cos\theta_{ij}<0} \frac{\Delta\theta_i \cdot \Delta\theta_j}{\|\Delta\theta_j\|^2}\Delta\theta_j\right) \tag{9}$$

Optimization Objective. To mitigate the dominant influence of clients with large data volumes during aggregation and minimize the overall loss change $\Delta\mathcal{L}_{\text{total}}$, we aim to find an update vector $\mathbf{d}$ to replace the average parameter update $\Delta\theta_{avg}$. This vector should not only reduce the average loss across all clients but also specifically decrease the loss for each individual client. In other words, we are looking for the client with the worst decrease in loss during the minimization process and then trying to maximize the loss reduction for that client as much as possible. The objective function can thus be described as follows:

$$\begin{aligned} \min \Delta\mathcal{L}_{\text{total}} &\approx \max_{\mathbf{d}\in\boldsymbol{\Theta}} \frac{1}{N}\sum_{j=1}^{N}\min_{j\in[N]} \mathbf{d}\cdot\Delta\theta_j \\ \text{s.t.} \quad & \left\|\mathbf{d} - \Delta\overline{\theta}_\perp\right\| \leq c\|\Delta\overline{\theta}_\perp\| \end{aligned} \tag{10}$$

Algorithm 1: ConFREE

Input: N clients, local datasets $\{\mathcal{D}_1, \dots, \mathcal{D}_N\}$, learning rates $\{\eta, \alpha\}$, total communication rounds T.
Output: Conflict-free global aggregation parameter θ_g^{t+1}.

```
Initialize global parameter θ_g.
for each global iteration t = 1, 2, ..., T do
    for client i = 1, 2, ..., N in parallel do
        Update local head h_i^{t+1} via Eq. 2.
        Update shared parameter θ_i^t via Eq. 3.
        Compute shared parameter update Δθ_i^t = θ_i^t − θ_g^t.
    end for
    for each client update Δθ_i^t do
        Adjust Δθ_i^t using projection for all Δθ_j^t (where
        j ≠ i) with negative cosine similarity.
    end for
    Compute the conflict-free guidance vector by averag-
    ing the adjusted updates using Eq. 9.
    Solve optimization problem to find w* using Eq. 13.
    Compute λ and d* based on the solution.
    Server Aggregation: θ_g^{t+1} ← θ_g^t + d*.
end for
return θ_g^{t+1}.
```

where the constraint ensures that the chosen direction $\mathbf{d}$ stays within a range determined by the constant $c \in [0, 1)$ relative to the conflict-free guidance vector $\Delta\overline{\theta}_\perp$. This helps prevent the negative effects of update conflicts on the optimization process. Additionally, the update vector $\mathbf{d}$ is aligned as closely as possible with the update directions of all clients, especially those with the least contribution. This alignment is essential to ensure effective optimization for even the least contributing clients, thereby preventing any single client from dominating the update direction.

Lagrangian Function. To avoid the high costs associated with optimizing d in high-dimensional spaces, weights w are introduced to transform the constrained optimization problem into a more solvable dual problem. This is expressed as $\frac{1}{N}\sum_{j=1}^{N}\min_{j\in[N]} \mathbf{d}\cdot\Delta\theta_j = \min_{w\in\mathcal{W}} \mathbf{d}\cdot\sum_j w_j\Delta\theta_j$, where $w = (w_1, \dots, w_N)$ and $\mathcal{W} = \left\{w : \sum_j w_j = 1 \text{ and } w_j \geq 0\right\}$. Here, $\mathcal{W}$ represents the probability simplex, and $\Delta\theta_w = \sum_j w_j\Delta\theta_j$ is defined. By adjusting the values of w, the influence of each client on the overall optimization direction can be controlled. Specifically, the Lagrangian function can be written as:

$$\mathcal{L}(\mathbf{d}, \lambda, w) = \mathbf{d}\cdot\Delta\theta_w - \frac{\lambda}{2}\left(\|\mathbf{d} - \Delta\overline{\theta}_\perp\|^2 - c^2\|\Delta\overline{\theta}_\perp\|^2\right) \tag{11}$$

Dual Problem. Each term $\mathbf{d}\cdot\Delta\theta_j$ is linear with respect to $\mathbf{d}$, making $\min_{j\in[N]} \mathbf{d}\cdot\Delta\theta_j$ a concave function. Moreover, the inequality constraints define a convex set and satisfy the Slater condition. Consequently, we can interchange the min and max operations to solve the dual problem, as strong duality ensures that the solutions of the dual and primal prob-

Methods	Pathological heterogeneous		Practical heterogeneous($\beta = 0.1$)				#S.	#T.
	Flowers102	CIFAR100	CIFAR10	CIFAR100	TINY	Flowers102		
Per-FedAvg	67.04	61.77	87.76	41.49	25.62	55.33	151.72	66.86
+ ConFREE	67.64 (↑ 0.60)	62.45 (↑ 0.68)	88.23 (↑ 0.47)	41.76 (↑ 0.27)	27.42 (↑ 1.80)	55.96 (↑ 0.63)	148.19	67.38
LG-FedAvg	63.89	63.32	89.23	47.63	33.75	51.48	148.39	46.71
+ ConFREE	64.09 (↑ 0.20)	63.48 (↑ 0.16)	89.52 (↑ 0.29)	47.88 (↑ 0.25)	33.96 (↑ 0.21)	52.72(↑ 1.24)	144.86	46.94
APFL	68.87	63.53	89.45	54.18	41.16	63.65	292.85	56.27
+ ConFREE	70.79 (↑ 1.92)	63.71 (↑ 0.18)	89.64 (↑ 0.19)	55.44 (↑ 1.26)	41.97 (↑ 0.81)	65.11 (↑ 1.46)	289.32	56.66
FedROD	69.21	63.79	89.93	51.02	39.74	61.02	159.55	54.30
+ ConFREE	69.56 (↑ 0.35)	64.02 (↑ 0.23)	90.80 (↑ 0.87)	52.57 (↑ 1.55)	40.98 (↑ 1.24)	61.52(↑ 0.50)	156.02	54.81
FedRep	68.62	67.41	90.40	51.33	41.72	57.52	151.52	56.90
+ ConFREE	68.92 (↑ 0.30)	67.92 (↑ 0.51)	90.61 (↑ 0.21)	51.41 (↑ 0.08)	42.31 (↑ 0.59)	57.81(↑ 0.29)	147.99	56.97
FedCP	71.38	70.47	91.30	55.33	44.87	60.73	539.17	99.49
+ ConFREE	71.73 (↑ 0.35)	70.98 (↑ 0.51)	91.72 (↑ 0.42)	55.68 (↑ 0.35)	45.62 (↑ 0.75)	61.51(↑ 0.78)	535.64	99.59
FedALA	68.82	67.33	90.67	57.24	44.56	62.92	683.23	55.91
+ ConFREE	70.64 (↑ 1.82)	67.52 (↑ 0.19)	91.32 (↑ 0.65)	57.70 (↑ 0.46)	45.01 (↑ 0.45)	63.94 (↑ 1.02)	679.72	56.39
FedPAC	74.48	72.60	90.86	61.94	46.63	68.56	160.01	110.67
+ ConFREE	74.68 (↑ 0.20)	73.01 (↑ 0.41)	91.42 (↑ 0.56)	62.85 (↑ 0.91)	46.83 (↑ 0.20)	69.10 (↑ 0.54)	156.48	147.20

Table 1: Average test accuracies (%) for eight State-of-The-Art pFL methods combined with the ConFREE aggregation strategy in the Pathological heterogeneous setting and Practical heterogeneous setting. All experiments are repeated over 3 random seeds and the mean values are reported. The ($\uparrow \Delta$) denotes the corresponding improvement is marked in green. The maximum improvement for the SOTA methods across different datasets is marked in red. Using the CIFAR100 dataset as an example, $\#\mathbf{S}$ denotes the communication overhead, i.e., the model parameters size (MB) transmitted in each iteration, and $\#\mathbf{T}$ represents the computation overhead, i.e., the total time (s) taken by 20 clients in one iteration.

lems are equivalent:

$$\min_{w\in\mathcal{W},\lambda\geq 0}\ \max_{\mathbf{d}\in\Theta}\quad \mathbf{d}\cdot\Delta\theta_w - \frac{\lambda}{2}\left(\|\mathbf{d}-\Delta\overline{\theta}_\perp\|^2 - c^2\|\Delta\overline{\theta}_\perp\|^2\right) \tag{12}$$

differentiate with respect to $\mathbf{d}$ and λ, and solve to obtain $d^* = \Delta\overline{\theta}_\perp + \frac{c\|\Delta\overline{\theta}_\perp\|}{\|\Delta\theta_{w^*}\|}\Delta\theta_{w^*}$ and $\lambda = \frac{\|\Delta\theta_{w^*}\|}{c\|\Delta\overline{\theta}_\perp\|}$. Consequently, the final objective function is:

$$w^* = \arg\min_{w\in\mathcal{W}}\quad \mathbf{d}\cdot\Delta\theta_w + c\|\Delta\overline{\theta}_\perp\|\|\Delta\theta_w\| \tag{13}$$

the optimization problem for w can be solved using the Minimize function provided by the scientific computing library Scipy, which then computes the final update vector d^*, and c is a hyperparameter with a value range of $(0, 1)$. The dimension of w is equal to the number of participating clients, which is much smaller than directly solving the dimension of the aggregated update vector d, thereby simplifying the computational complexity.

Experiments

In this section, we conduct extensive experiments to evaluate our method ConFREE for personalized federated learning and demonstrate its effectiveness, efficiency, and versatility.

Implementation Detail

Datasets and Non-IID. We evaluated the effectiveness of ConFREE on four public datasets: CIFAR10 and CIFAR100(Krizhevsky et al. 2009), Tiny-ImageNet (Chrabaszcz, Loshchilov, and Hutter 2017) and Flowers102 (Nilsback and Zisserman 2008). Additionally, we simulated two widely used Non-IID scenarios: the Pathological heterogeneous setting (Shamsian et al. 2021) and the Practical heterogeneous setting (Li, He, and Song 2021). In the first scenario, each client is assigned only a small subset of the data categories, such as sampling 5 classes from the total of 102 classes in the Flowers102 dataset for each client, with non-overlapping data samples. In the second scenario, the Dirichlet distribution $Dir(\beta)$ is used to simulate the class distribution of each client's data, with the default parameter $\beta = 0.1$ controlling the distribution difference. In both scenarios, the sample quantities for each client are imbalanced, with the second scenario being applicable to a wider range of practical applications. All data is split into 70% training and 30% testing sets. The testing and training sets have the same data distribution for all clients.

Baselines. We combined the proposed ConFREE with several advanced pFL methods focused on learning client-specific models, including two classic baselines, Per-FedAvg (Fallah, Mokhtari, and Ozdaglar 2020) and LG-FedAvg (Liang et al. 2020), as well as six state-of-the-art methods. The APFL (Deng, Kamani, and Mahdavi 2020) method adjusts the balance between the global model and the personalized model through adaptive weighting. The FedROD (Chen and Chao 2021) method enhances the similarity of model parameters across clients using regularization techniques. The FedRep (Collins et al. 2021) method

Methods	CIFAR100				Flowers102			
	C = 20		**C = 60**		**C = 20**		**C = 60**	
	$P = 30\%$	$P = 60\%$	$P = 30\%$	$P = 60\%$	$P = 30\%$	$P = 60\%$	$P = 30\%$	$P = 60\%$
APFL + ConFREE	52.85 53.49 (↑ 0.64)	53.64 54.11 (↑ 0.47)	43.56 44.90 (↑ 1.34)	44.02 44.46 (↑ 0.44)	62.92 64.14 (↑ 1.22)	63.94 64.48 (↑ 0.54)	45.88 49.30 (↑ 3.42)	50.31 52.53 (↑ 2.22)
FedROD + ConFREE	53.91 54.18 (↑ 0.27)	53.06 53.29 (↑ 0.23)	52.49 53.23 (↑ 0.74)	50.59 52.29 (↑ 1.70)	61.41 62.41 (↑ 1.00)	61.46 62.69 (↑ 1.23)	56.87 57.74 (↑ 0.87)	55.28 56.34 (↑ 1.06)
FedRep + ConFREE	50.95 51.11 (↑ 0.16)	51.03 51.39 (↑ 0.36)	44.36 44.66 (↑ 0.30)	44.04 44.37 (↑ 0.33)	58.44 58.93 (↑ 0.49)	57.42 58.10 (↑ 0.68)	49.40 49.93 (↑ 0.53)	48.72 49.16 (↑ 0.44)
FedCP + ConFREE	56.25 56.38 (↑ 0.13)	56.10 56.20 (↑ 0.10)	50.20 50.29 (↑ 0.09)	48.95 49.02 (↑ 0.07)	62.19 63.03 (↑ 0.84)	61.31 61.80 (↑ 0.49)	53.30 53.69 (↑ 0.39)	51.81 52.29 (↑ 0.48)
FedALA + ConFREE	56.62 56.98 (↑ 0.36)	56.97 57.45 (↑ 0.48)	58.01 58.40 (↑ 0.39)	58.15 58.60 (↑ 0.45)	62.19 63.50 (↑ 1.31)	62.48 63.70 (↑ 1.22)	56.54 59.33 (↑ 2.79)	58.71 60.88 (↑ 2.17)
FedPAC + ConFREE	61.45 62.09 (↑ 0.64)	61.61 62.23 (↑ 0.62)	61.93 62.63 (↑ 0.70)	62.84 63.34 (↑ 0.50)	67.49 67.83 (↑ 0.34)	68.08 68.86 (↑ 0.78)	58.95 60.20 (↑ 1.25)	59.04 60.78 (↑ 1.74)

Table 2: Average test accuracies $(\%)$ of 20 clients and 60 clients with participation rates of 30% and 60% on the CIFAR100 and Flowers102 datasets. All experiments are repeated over 3 random seeds and the mean values are reported. **C** represents the number of clients, and P represents the client participation rate.

optimizes each client's classifier based on local data by separating the training of representations and classifiers. The FedCP (Zhang et al. 2023d) method balances global sharing and local optimization through a collaborative personalization strategy. The FedALA (Zhang et al. 2023c) method allows clients to autonomously choose the appropriate fusion ratio between the global model and the local model during training. Finally, FedPAC (Xu, Tong, and Huang 2023) uses clustering algorithms to enable each cluster group to have its own personalized model.

Training Details. Based on the training settings of FedALA, FedCP, and FedPAC, we use a 4-layer CNN model as the base model and SGD as the local optimizer for all methods. The local learning rate is set to 0.005, the batch size to 10, and the number of local training epochs to 1. The global communication rounds are set to 500 across all datasets until all methods empirically converge. The default number of clients is 20, with a 100% client participation rate.

Implementation. We run all experiments on a workstation equipped with two Intel Xeon Gold 6248R CPUs (48 cores), 128GB of memory and two NVIDIA 3090 Ti GPUs. All methods are reproduced in PyTorch, and each method is run 3 times to report the average accuracy.

Performance Evaluation with the SOTA

ConFREE improves the performance of all SOTA models. The main results under pathological heterogeneous setting and practical heterogeneous setting on four typical datasets are presented in Table.1. ConFREE further improves the performance of all SOTA methods and performs even better on FedALA and APFL, with maximum enhancements of 1.82% and 1.92%, respectively. We observed that the performance gain on datasets with more categories, such as TINY and Flowers102, is higher than on datasets with fewer categories like CIFAR10. This is because a larger number of data categories leads to increased heterogeneity, which in turn intensifies the conflicts between client updates. Thus, this also demonstrates the effectiveness of our approach to resolve client update conflicts.

The overall overhead of ConFREE is acceptable. We report the communication overhead and the computation overhead of each method in Table.1. The conflict optimization in ConFREE does not significantly increase the computation time in each iteration, remaining almost on par with the original aggregation methods. Additionally, in terms of communication overhead, our aggregation strategy only computes the client parameter updates each time, thereby reducing the size of model parameters transmitted during each communication. Overall, ConFREE effectively improves accuracy with minimal additional time cost and reduced communication overhead.

ConFREE is scalable across different client amounts and participation rates. We evaluate the stability of ConFREE on the CIFAR100 and Flowers102 datasets, which have more categories, by varying the number of clients and their participation rates. It is worth noting that the total dataset size is fixed and increasing the number of clients leads to higher heterogeneity per client. In Table.2, the highest performance improvements achieved with ConFREE for APFL, FedROD, FedRep, FedCP, FedALA and FedPAC are 3.42%,1.70%, 0.68%, 0.84%, 2.79% and 1.74% respectively. These results show substantial gains compared to the default experimental settings reported in Table.1. The observed improvements are attributed to the increased heterogeneity of the client data distributions caused by increasing the number of clients and reducing the participation rate.

ConFREE performs better in the strong heterogeneity setting. Flowers102 has more categories and fewer data than CIFAR100, making it more heterogeneous in the same experimental settings. From the experiments in Table.2, we can see that the maximum gains of most SOTA models are concentrated in the Flowers102, and the Top-2 boosts are all in the most heterogeneous C=60, P=30% setting. This proves that our method is very effective in solving client

Heterogeneous setting	Methods	Base	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
C=60, P=30%	APFL	45.88	47.13	48.63	48.87	**49.30**	48.14	47.13	46.16	45.15	44.38
	FedROD	56.87	57.12	57.31	57.31	57.26	57.69	**57.74**	57.36	57.26	56.78
	FedRep	49.40	**49.93**	49.59	49.59	49.73	49.93	49.69	49.59	49.78	49.20
	FedCP	53.30	53.35	53.50	53.64	**53.69**	53.64	53.59	53.35	53.35	53.55
	FedALA	56.54	57.79	58.51	58.61	58.85	59.04	**59.33**	59.09	59.00	58.90
	FedPAC	58.95	59.43	**60.20**	59.91	60.11	59.77	59.82	59.86	59.53	58.19
C=60, P=60%	APFL	50.31	51.62	52.39	**52.53**	51.81	50.27	48.67	46.64	45.30	44.72
	FedROD	55.28	55.57	55.76	56.01	56.05	55.91	56.05	55.48	**56.34**	55.19
	FedRep	48.72	**49.16**	48.82	49.11	49.06	49.11	48.67	48.91	48.63	48.67
	FedCP	51.81	52.05	51.95	52.05	51.95	52.19	**52.29**	52.00	51.76	51.91
	FedALA	58.71	58.66	59.77	59.29	59.72	59.82	60.35	60.64	**60.88**	60.15
	FedPAC	59.04	59.58	59.91	**60.78**	60.35	59.91	60.20	59.72	59.48	59.36

Table 3: Average model accuracy of ConFREE combined with the SOTA method on Flowers102 for different hyperparameters c. Bold numbers indicate the optimal value for the corresponding method.

conflicts caused by strong heterogeneity. Therefore, in scenarios with strong heterogeneity, addressing update conflicts during global model aggregation can significantly improve the performance of local clients.

ConFREE can provide more global information. From Table.2, we observe that ConFREE provides significantly more gains for APFL, FedPAC, and FedALA compared to FedROD, FedRep, and FedCP, using the Flowers102 dataset with 60 clients as an example. This is because the former three methods are based on selectively aggregating more global knowledge, whereas the latter three methods focus on designing specific layers to retain more local information, thereby overlooking some global information. After the global model resolves conflicts using the ConFREE strategy, it incorporates more effective global information that benefits each client. As a result, methods that focus on acquiring more global information see higher gains.

ConFREE improves different SOTA methods differently. When we further observe Table.2, we find that when the dataset and the number of clients are fixed, the performance of different SOTA methods varies as the client participation rate rises from 30% to 100%. Using CIFAR100 as an example, the accuracy of APFL, FedPAC, and FedALA methods gradually improves as the participation rate increases. This is because more clients bring in more data, allowing the global model to learn more knowledge. Therefore, the performance of the three methods focusing on aggregating more global knowledge is improved. However, since the total amount of data remains fixed, the increase in client participation reduces the heterogeneity of data distribution, leading to a decrease in the performance gains provided by ConFREE for these methods. In contrast to this phenomenon, the accuracy of FedROD, FedRep, and FedCP methods either remains stable or gradually decreases as the participation rate increases. This is because these three methods retain more local information, while the transfer of more global knowledge gradually interferes with their local model learning. However, ConFREE can eliminate this conflicting component in updates, resulting in increasing gains for these three methods. In summary, ConFREE is a model-agnostic approach that minimizes the impact of negative transfer by resolving parameter update conflicts in the global model. Furthermore, the stronger the heterogeneity of the client data distribution, the greater the impact of negative transfer and the better the performance of ConFREE in such scenarios.

Effect of the Hyperparameter c

To find the optimal update vector within the conflict-free guidance vector space, the hyperparameter c is used to control the size of the search range. As shown in Table.3, the accuracy of the three models first increases and then decreases as the hyperparameter increases, which is consistent between settings with different levels of heterogeneity. In the heterogeneous setting with $C = 60$ and $P = 30\%$, the optimal hyperparameter values for APFL, FedROD, FedRep, FedCP, FedALA, and FedPAC are 0.4, 0.6, 0.1, 0.4, 0.6 and 0.2, respectively. In another heterogeneous setting, the optimal values are 0.3, 0.8, 0.1, 0.6, 0.8 and 0.3, respectively. The optimal hyperparameters for the models may vary slightly across different datasets and heterogeneous settings, and typically, we start searching from 0.5 in both directions to find the optimal value.

Conclusion

We have proposed an effective ConFREE method to reduce conflicting client updates in personalized federated learning. By constructing a global conflict-free guidance vector and maximizing the local improvement of the worst-performing client, ConFREE can achieve automatic balancing of different clients' update objectives. Our method is model-agnostic and can be easily applied to enhance various client-side personalization techniques. Extensive experimental results demonstrate that ConFREE can effectively improve the performance of SOTA methods. Currently, we primarily use the projected vector as a conflict-free guidance vector to achieve global loss optimization, which could be replaced by other main objectives.

Acknowledgements

We appreciate constructive feedback from anonymous reviewers and meta-reviewers. This work was supported by the National Natural Science Foundation of China (62172442, 62172451), China Scholarship Council, and High Performance Computing Center of Central South University.

References

Acar, D. A. E.; Zhao, Y.; Zhu, R.; Matas, R.; Mattina, M.; Whatmough, P.; and Saligrama, V. 2021. Debiasing model updates for improving personalized federated training. In *International conference on machine learning*, 21–31. PMLR.

Achituve, I.; Shamsian, A.; Navon, A.; Chechik, G.; and Fetaya, E. 2021. Personalized federated learning with gaussian processes. *Advances in Neural Information Processing Systems*, 34: 8392–8406.

Arivazhagan, M. G.; Aggarwal, V.; Singh, A. K.; and Choudhary, S. 2019. Federated learning with personalization layers. *arXiv preprint arXiv:1912.00818*.

Briggs, C.; Fan, Z.; and Andras, P. 2020. Federated learning with hierarchical clustering of local updates to improve training on non-IID data. In *2020 international joint conference on neural networks (IJCNN)*, 1–9. IEEE.

Cai, L.; Chen, N.; Cao, Y.; He, J.; and Li, Y. 2023. FedCE: personalized federated learning method based on clustering ensembles. In *Proceedings of the 31st ACM international conference on multimedia*, 1625–1633.

Carrillo, J. A.; Trillos, N. G.; Li, S.; and Zhu, Y. 2024. FedCBO: Reaching group consensus in clustered federated learning through consensus-based optimization. *Journal of machine learning research*, 25(214): 1–51.

Chen, H.-Y.; and Chao, W.-L. 2021. On bridging generic and personalized federated learning for image classification. *arXiv preprint arXiv:2107.00778*.

Chen, M.; Jiang, M.; Dou, Q.; Wang, Z.; and Li, X. 2023. FedSoup: improving generalization and personalization in federated learning via selective model interpolation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 318–328. Springer.

Chrabaszcz, P.; Loshchilov, I.; and Hutter, F. 2017. A downsampled variant of imagenet as an alternative to the cifar datasets. *arXiv preprint arXiv:1707.08819*.

Collins, L.; Hassani, H.; Mokhtari, A.; and Shakkottai, S. 2021. Exploiting shared representations for personalized federated learning. In *International conference on machine learning*, 2089–2099. PMLR.

Deng, Y.; Kamani, M. M.; and Mahdavi, M. 2020. Adaptive personalized federated learning. *arXiv preprint arXiv:2003.13461*.

Fallah, A.; Mokhtari, A.; and Ozdaglar, A. 2020. Personalized federated learning with theoretical guarantees: A model-agnostic meta-learning approach. *Advances in neural information processing systems*, 33: 3557–3568.

Ghosh, A.; Chung, J.; Yin, D.; and Ramchandran, K. 2020. An efficient framework for clustered federated learning. *Advances in Neural Information Processing Systems*, 33: 19586–19597.

Goldman, E. 2020. An introduction to the california consumer privacy act (ccpa). *Santa Clara Univ. Legal Studies Research Paper*.

Hanzely, F.; and Richtárik, P. 2020. Federated learning of a mixture of global and local models. *arXiv preprint arXiv:2002.05516*.

Huang, Y.; Chu, L.; Zhou, Z.; Wang, L.; Liu, J.; Pei, J.; and Zhang, Y. 2021. Personalized cross-silo federated learning on non-iid data. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, 7865–7873.

Krizhevsky, A.; et al. 2009. Learning multiple layers of features from tiny images. *Citeseer*.

Kulkarni, V.; Kulkarni, M.; and Pant, A. 2020. Survey of personalization techniques for federated learning. In *2020 fourth world conference on smart trends in systems, security and sustainability (WorldS4)*, 794–797. IEEE.

Lee, R.; Kim, M.; Li, D.; Qiu, X.; Hospedales, T.; Huszár, F.; and Lane, N. 2024. Fedl2p: Federated learning to personalize. *Advances in Neural Information Processing Systems*, 36.

Li, Q.; He, B.; and Song, D. 2021. Model-contrastive federated learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10713–10722.

Li, T.; Hu, S.; Beirami, A.; and Smith, V. 2021a. Ditto: Fair and robust federated learning through personalization. In *International conference on machine learning*, 6357–6368. PMLR.

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

Li, X.-C.; Zhan, D.-C.; Shao, Y.; Li, B.; and Song, S. 2021b. Fedphp: Federated personalization with inherited private models. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 587–602. Springer.

Liang, P. P.; Liu, T.; Ziyin, L.; Allen, N. B.; Auerbach, R. P.; Brent, D.; Salakhutdinov, R.; and Morency, L.-P. 2020. Think locally, act globally: Federated learning with local and global representations. *arXiv preprint arXiv:2001.01523*.

Liu, Y.; Huang, A.; Luo, Y.; Huang, H.; Liu, Y.; Chen, Y.; Feng, L.; Chen, T.; Yu, H.; and Yang, Q. 2020. Fedvision: An online visual object detection platform powered by federated learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, 13172–13179.

Long, G.; Xie, M.; Shen, T.; Zhou, T.; Wang, X.; and Jiang, J. 2023. Multi-center federated learning: clients clustering for better personalization. *World Wide Web*, 26(1): 481–500.

Ma, X.; Zhang, J.; Guo, S.; and Xu, W. 2022. Layer-wised model aggregation for personalized federated learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10092–10101.

Mansour, Y.; Mohri, M.; Ro, J.; and Suresh, A. T. 2020. Three approaches for personalization with applications to federated learning. *arXiv preprint arXiv:2002.10619*.

Marfoq, O.; Neglia, G.; Vidal, R.; and Kameni, L. 2022. Personalized federated learning through local memorization. In *International Conference on Machine Learning*, 15070–15092. PMLR.

Nilsback, M.-E.; and Zisserman, A. 2008. Automated flower classification over a large number of classes. In *2008 Sixth Indian conference on computer vision, graphics & image processing*, 722–729. IEEE.

Regulation, P. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council. *Regulation (eu)*, 679: 2016.

Sattler, F.; Müller, K.-R.; and Samek, W. 2020. Clustered federated learning: Model-agnostic distributed multitask optimization under privacy constraints. *IEEE transactions on neural networks and learning systems*, 32(8): 3710–3722.

Shamsian, A.; Navon, A.; Fetaya, E.; and Chechik, G. 2021. Personalized federated learning using hypernetworks. In *International Conference on Machine Learning*, 9489–9502. PMLR.

Singhal, K.; Sidahmed, H.; Garrett, Z.; Wu, S.; Rush, J.; and Prakash, S. 2021. Federated reconstruction: Partially local federated learning. *Advances in Neural Information Processing Systems*, 34: 11220–11232.

T Dinh, C.; Tran, N.; and Nguyen, J. 2020. Personalized federated learning with moreau envelopes. *Advances in neural information processing systems*, 33: 21394–21405.

Tan, B.; Liu, B.; Zheng, V.; and Yang, Q. 2020. A federated recommender system for online services. In *Proceedings of the 14th ACM conference on recommender systems*, 579–581.

Wen, J.; Zhang, Z.; Lan, Y.; Cui, Z.; Cai, J.; and Zhang, W. 2023. A survey on federated learning: challenges and applications. *International Journal of Machine Learning and Cybernetics*, 14(2): 513–535.

Wu, J.; Bao, W.; Ainsworth, E.; and He, J. 2023. Personalized federated learning with parameter propagation. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2594–2605.

Xu, J.; Tong, X.; and Huang, S.-L. 2023. Personalized federated learning with feature alignment and classifier collaboration. *arXiv preprint arXiv:2306.11867*.

Ye, C.; Zheng, H.; Hu, Z.; and Zheng, M. 2023. PFedSA: Personalized federated multi-task learning via similarity awareness. In *2023 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, 480–488. IEEE.

Zhang, H.; Li, C.; Dai, W.; Zou, J.; and Xiong, H. 2023a. FedCR: Personalized federated learning based on across-client common representation with conditional mutual information regularization. In *International Conference on Machine Learning*, 41314–41330. PMLR.

Zhang, J.; Hua, Y.; Wang, H.; Song, T.; Xue, Z.; Ma, R.; Cao, J.; and Guan, H. 2023b. Gpfl: Simultaneously learning global and personalized feature information for personalized federated learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 5041–5051.

Zhang, J.; Hua, Y.; Wang, H.; Song, T.; Xue, Z.; Ma, R.; and Guan, H. 2023c. Fedala: Adaptive local aggregation for personalized federated learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 11237–11244.

Zhang, J.; Hua, Y.; Wang, H.; Song, T.; Xue, Z.; Ma, R.; and Guan, H. 2023d. Fedcp: Separating feature information for personalized federated learning via conditional policy. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 3249–3261.

Zhang, X.; Li, Y.; Li, W.; Guo, K.; and Shao, Y. 2022. Personalized federated learning via variational bayesian inference. In *International Conference on Machine Learning*, 26293–26310. PMLR.