Combating Data Imbalances in Federated Semi-Supervised Learning with Dual Regulators - Paper Review

Devere Anthony Weaver
Towson University
Towson, USA
dweave8@students.towson.edu

I. Introduction

Federated learning (FL) has emerged as a pivotal privacy-preserving technique for decentralized machine learning, enabling multiple clients to collaboratively train a model without directly sharing their data. This paradigm is increasingly relevant in sensitive applications, such as healthcare and consumer devices, where data privacy is paramount. However, real-world federated learning systems face several challenges, particularly when data are heterogeneous and insufficiently labeled. Federated Semi-Supervised Learning (FSSL) attempts to address this issue by leveraging the abundant unlabeled data available on decentralized devices alongside limited labeled data. Although promising, FSSL encounters critical obstacles when data distributions are nonuniform. Specifically, it suffers from external imbalance, where the data distributions vary across clients, and internal imbalance, where the labeled and unlabeled data within a single client exhibit different distributions. These imbalances result in suboptimal global models and reduced performance.

The paper, Combating Data Imbalances in Federated Semi-supervised Learning with Dual Regulators, introduces a novel framework called FedDure, designed to address the dual imbalances that plague FSSL systems. FedDure incorporates two adaptive components, the coarse-grid regulator (C-reg) and the fine-grid regulator (F-reg), to improve the robustness of local and global model training. The C-reg tracks overall learning progress on labeled data, while the F-reg assigns dynamic weights to individual unlabeled data instances. These regulators enable FedDure to rectify the imbalances inherent in FSSL by balancing the contributions of labeled and unlabeled data, both across and within clients. Through theoretical analysis, the authors prove the convergence of the proposed method, and empirical results demonstrate its superior performance across multiple datasets and experimental configurations. This report examines in detail the technical details, contributions, and findings of this innovative work.

II. Related Work

Federated learning (FL) has become a widely adopted paradigm for privacy-preserving machine learning, enabling decentralized data utilization. A foundational approach in this field is FedAvg by McMahan et al. (2017), which introduced communication-efficient aggregation of local model updates. However, FL systems often face statistical heterogeneity, where non-IID data distributions lead to client drift—a divergence between local and global models. To address this, methods such as FedProx and Scaffold (Karimireddy et al., 2020) incorporate constraints or corrections to stabilize model updates. While these methods are effective for supervised FL, they do not generalize well to semi-supervised scenarios.

In the domain of semi-supervised learning (SSL), two dominant techniques are pseudo-labeling and consistency regularization. Pseudo-labeling, first popularized by Lee (2013), assigns artificial labels to unlabeled data based on model predictions, using high-confidence samples to improve learning. This approach is simple but sensitive to pseudo-label accuracy, which can degrade under class imbalances. Consistency regularization, proposed by Tarvainen and Valpola (2017), enforces prediction consistency across perturbed inputs, enhancing the model's robustness and representation learning. SSL methods have proven effective in centralized settings but face additional challenges when applied in federated environments.

Federated Semi-Supervised Learning (FSSL) combines FL and SSL to address scenarios where labeled data is scarce and distributed non-uniformly among clients. Existing FSSL methods, such as FedMatch (Jeong et al., 2021), extend SSL principles by incorporating inter-client consistency regularization to handle external imbalances. However, these approaches rely on strong IID assumptions and fail to address internal imbalances, where labeled and unlabeled data within a client exhibit mismatched distributions. Similarly, SSL methods like Fixmatch (Sohn et al., 2020) and UDA perform poorly in federated settings due to their dependence on centralized assumptions and their inability to adapt to client-specific data distributions.

FedDure builds upon these foundations by introducing dual regulators that adaptively address both external and internal imbalances, overcoming the limitations of prior work. By leveraging insights from FL, SSL, and FSSL, FedDure establishes itself as a robust framework tailored for real-world non-IID federated environments.

III. Problem Definition

The authors define the problem within the context of FSSL, focusing on two primary imbalances: external imbalance and internal imbalance. External imbalance arises when data distributions vary across clients, as often seen in real-world applications where individual clients have domain-specific data (e.g., a client containing only a subset of classes). Internal imbalance refers to the mismatch between the labeled and unlabeled data distributions within a client, which is particularly problematic when pseudo-labeling methods are employed. For example, a client may have labeled images predominantly belonging to one class but unlabeled images spanning multiple classes, leading to biased training and poor generalization.

In this setup, each client C_k possesses a private dataset D_k, which is divided into a labeled subset D_s^k and an unlabeled subset D_u^k. Formally: $D_k = D_s^k \cup D_u^k \quad \text{with} \quad D_s^k \cap D_u^k = \emptyset$. Here, the labeled data D_s^k consists of pairs of samples and labels: $D_s^k = \{(x_i, y_i)\}_{i=1}^{N_s^k}$, while the unlabeled data contains only feature samples: $D_u^k = \{u_i\}_{i=1}^{N_u^k}, \quad \text{where} \quad N_s^k \ll N_u^k$.

The global model, denoted f_g, is parameterized by θ. It is trained collaboratively by aggregating updates from local models, f_l^k, trained on each client. The parameters of a local model are represented as θ_l^k. The goal of FSSL is to optimize the global model f_g such that it generalizes well across all clients despite the presence of external and internal imbalances. This is expressed as: $\theta_g^* = \arg\min_{\theta_g} \sum_{k=1}^{K} w_k \mathcal{L}(f_l^k; \theta_g)$, where w_k is the weight assigned to client k based on its dataset size, and L is the aggregate loss function combining supervised and unsupervised losses.

Traditional methods fail in this scenario due to their reliance on uniform data distribution assumptions or static weighting schemes, which are ineffective under the dual imbalance challenges of FSSL. FedDure addresses this issue through adaptive regulators that dynamically adjust the contributions of labeled and unlabeled data during training, ensuring balanced and effective learning outcomes.

IV. Proposed Framework

A. Coarse-Grained Regulator (C-reg)

The C-reg focuses on addressing internal imbalances by tracking the overall learning progress on labeled data. It calculates the cross-entropy loss difference between iterations to quantify the effectiveness of model updates. This feedback mechanism enables the local model to dynamically adjust the importance assigned to labeled and unlabeled data, preventing pseudo-label corruption. Additionally, the C-reg incorporates global model parameters received from the server, providing a means to counteract client drift caused by external imbalances.

Mathematically, the C-reg optimizes its parameters ϕ by minimizing the supervised loss on labeled data. At iteration t the learning effect is quantified as the difference in cross-entropy losses:

$$d_{t+1} = \mathbb{E}_{(x,y)\in D_s^k} \left[L_{\text{CE}}(y, f_d(x;\phi_t)) - L_{\text{CE}}(y, f_d(x;\phi_{t+1})) \right].$$

B. Fine-Grained Regulator (F-reg)

The F-reg addresses internal imbalances by assigning dynamic weights to individual unlabeled samples based on their relevance to the model's learning objectives. Unlike fixed thresholding techniques in traditional pseudo-labeling, the F-reg learns an adaptive weighting scheme tailored to each client's data distribution. This ensures that high-confidence pseudo-labels contribute more significantly to the training process, while low-confidence pseudo-labels are downweighted. The weight for each unlabeled sample u_i is computed as:

$$m_i = f_w(f_l(T_s(u_i), \theta_l); w).$$

C. Bi-level Optimization

The dual regulators are trained alternately in a bi-level optimization framework, ensuring that both regulators dynamically adapt to the data distributions. In each local training iteration, the F-reg updates its parameters first, refining instance weights based on feedback from the labeled data. The C-reg then adjusts its parameters, incorporating global feedback to regularize the local model's updates. The final local model update combines supervised loss, pseudo-label loss, and feedback from the regulators:

$$\theta_l^{t+1} = \theta_l^t - \eta \left(g_s + g_u + g_d \right).$$

V. Experiments

The experimental evaluation of FedDure was conducted across multiple benchmarks to test its efficacy in addressing both external and internal data imbalances. The authors utilized three publicly available datasets: CIFAR-10, Fashion-MNIST, and CINIC-10. These datasets offer varying levels of complexity and class diversity, providing robust scenarios for evaluating the framework's performance.

To simulate realistic non-IID scenarios, three distinct data distribution settings were employed:

(IID, IID): Both labeled and unlabeled data are independently and identically distributed (IID) across all clients, representing the idealized scenario. (IID, DIR): Labeled data remains IID, while unlabeled data is distributed according to a Dirichlet distribution, introducing inter-client variability. (DIR, DIR): Both labeled and unlabeled data are non-IID, with significant skewness across clients, reflecting the most challenging real-world scenario. Each experiment involved a simulated federated environment with 20 clients. A subset of these clients had access to labeled data, with the remaining clients relying solely on unlabeled data. The labeled-to-unlabeled data ratio varied to analyze the impact of imbalance severity. The backbone model for local training was a ResNet-9, while the fine-grained regulator (F-reg) used a multi-layer perceptron (MLP). All models were trained using the Adam optimizer with a batch size of 10 and a learning rate of 0.0005.

Baseline methods included supervised FL frameworks (FedAvg, FedProx) and FSSL adaptations like FedMatch, which employs inter-client consistency loss. Additionally, the authors incorporated SSL techniques such as Fixmatch and UDA into the baselines to ensure a fair comparison.

VI. Results

The results of the experiments clearly demonstrate the superiority of FedDure over baseline methods, particularly in challenging non-IID settings.

A. Performance Under (DIR, DIR) Setting

The (DIR, DIR) configuration represents the most realistic and difficult scenario, where both labeled and unlabeled data are non-IID. In this setting, FedDure achieved a significant improvement in classification accuracy compared to baseline methods. On the CIFAR-10 dataset, FedDure outperformed FedMatch by over 11%, with an accuracy of 79.1% versus 67.8%. Similarly, for the CINIC-10 dataset, FedDure achieved an accuracy of 61.2%, marking a 12% improvement over the best-performing baseline.

B. Robustness Across Labeled-to-Unlabeled Ratios

To evaluate robustness, the authors varied the ratio of labeled to unlabeled data across clients. FedDure consistently maintained high accuracy even as the proportion of labeled data decreased. In contrast, baseline methods exhibited sharp performance degradation under the same conditions, highlighting their reliance on sufficient labeled data. This robustness is attributed to the adaptive weighting mechanism of the F-reg, which ensures that high-quality pseudo-labels contribute more significantly to training.

C. Scalability Across Clients

The framework's scalability was tested by increasing the number of clients in the federation. FedDure demonstrated stable performance as the number of clients increased, maintaining an accuracy drop of less than 3% with 50 clients compared to 20 clients. Baseline methods, on the other hand, exhibited significant performance degradation due to the amplified effects of external imbalance.

D. Impact of Dual Regulators

Ablation studies were conducted to isolate the contributions of the C-reg and F-reg. The absence of the C-reg resulted in a 7% drop in accuracy under the (DIR, DIR) setting, emphasizing its role in mitigating client drift and pseudo-label corruption. Similarly, removing the F-reg led to a 10% performance drop, underscoring its importance in handling internal imbalances. These results validate the complementary roles of the dual regulators in achieving robust and balanced learning.

VII. Discussion

The results of the experiments underscore several critical findings that highlight the strengths and practical utility of FedDure.

First, FedDure excels in scenarios characterized by significant data imbalance, achieving state-of-the-art accuracy across all datasets and settings. This is particularly noteworthy in the (DIR, DIR) configuration, where traditional methods falter. The dual-regulator design ensures that both external and internal imbalances are addressed simultaneously, a feature absent in existing FSSL frameworks. The C-reg provides coarse-grained adjustments to prevent global model degradation due to skewed client updates, while the F-reg enables fine-grained control over unlabeled data contributions.

Second, the framework's robustness to varying labeled-to-unlabeled ratios is a testament to its adaptability. Many real-world applications of FL involve minimal labeled data, as labeling is both time-intensive and expensive. FedDure's ability to maintain high performance under such conditions makes it an attractive choice for deployment in these scenarios.

Third, scalability is another strong suit of FedDure. As the number of clients increases, traditional methods suffer from amplified client drift and model divergence. The dual regulators in FedDure mitigate these effects, ensuring consistent performance across a wide range of federation sizes. This scalability positions FedDure as a viable solution for large-scale federated systems involving hundreds or thousands of clients.

Despite these strengths, there are areas for improvement. The bi-level optimization framework, while effective, introduces additional computational overhead compared to simpler approaches like FedAvg. This may pose challenges in resource-constrained environments where clients have limited computational capabilities. Additionally, the framework's reliance on gradient updates for the C-reg may be less effective when data distributions change dynamically over time. Future work could explore adaptive mechanisms that account for evolving distributions and reduce computational complexity.

Overall, FedDure addresses fundamental challenges in FSSL and sets a new benchmark for handling real-world data imbalances. Its innovative dual-regulator approach demonstrates the potential of adaptive frameworks in overcoming the limitations of traditional federated learning systems.

VIII. Conclusion

The paper "Combating Data Imbalances in Federated Semi-Supervised Learning with Dual Regulators" presents a groundbreaking solution to the persistent challenges of external and internal imbalances in Federated Semi-Supervised Learning (FSSL). By introducing the FedDure framework with its dual-regulator design, the authors successfully bridge the gap between theoretical innovation and practical applicability in FSSL. The Coarse-Grained Regulator (C-reg) addresses external imbalances by providing global feedback and mitigating client drift, while the Fine-Grained Regulator (F-reg) dynamically adapts to client-specific data distributions to handle internal imbalances.

Empirical results demonstrate that FedDure not only outperforms existing methods under challenging non-IID conditions but also maintains robustness across varying client numbers and labeled-to-unlabeled data ratios. Theoretical convergence guarantees further solidify its contributions, establishing FedDure as a state-of-the-art solution for FSSL.

From a broader perspective, the paper highlights the importance of adaptive mechanisms in addressing real-world data heterogeneity and scarcity in federated learning systems. Its bilevel optimization strategy and dual-regulator design introduce a level of flexibility and precision previously unseen in this domain. However, the increased computational overhead and potential limitations in dynamically changing environments indicate areas for future exploration.

In my assessment, this paper represents a significant step forward in the evolution of federated learning frameworks. Its meticulous design, backed by theoretical rigor and comprehensive experimentation, provides a robust foundation for further advancements in this field. Future work building upon FedDure could explore scalability to even larger and more diverse datasets, reduce computational complexity, and address dynamically evolving data distributions. These directions would enhance the framework's applicability across an even broader range of real-world use cases. Overall, FedDure exemplifies the potential of combining theory and practice to solve critical challenges in federated learning.

REFERENCES

[1] S. Bai, S. Li, W. Zhuang, J. Zhang, K. Yang, J. Hou, S. Zhang, S. Yi, and J. Gao, "Combating Data Imbalances in Federated Semi-supervised Learning with Dual Regulators," AAAI-24, 2024.

[2] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in Proc. 20th Int. Conf. Artif. Intell. Stat. (AISTATS), 2017, pp. 1273–1282.

[3] S. P. Karimireddy, S. Kale, M. Mohri, S. J. Reddi, S. U. Stich, and A. T. Suresh, "Scaffold: Stochastic controlled averaging for federated learning," in Proc. Int. Conf. Mach. Learn. (ICML), 2020, pp. 5132–5143.

[4] K. Sohn, D. Berthelot, C.-L. Li, Z. Zhang, N. Carlini, E. D. Cubuk, A. Kurakin, Z. Zhang, H. Raffel, "Fixmatch: Simplifying semi-supervised learning with consistency and confidence," in Advances in Neural Information Processing Systems (NeurIPS), 2020, pp. 10735–10747.

[5] D.-H. Lee, "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," in Workshop on Challenges in Representation Learning, ICML, 2013, pp. 1–8.

[6] A. Tarvainen and H. Valpola, "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results," in Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 4748–4757.

[7] W. Jeong, Y. Cho, and J. Shin, "Federated semi-supervised learning with inter-client consistency & disjoint learning," in Proc. Int. Conf. Learn. Representations (ICLR), 2021.