Hybrid Framework for Hyperparameter Optimization in Federated Learning.

***Abstract*—Federated Learning (FL) enables collaborative model training across decentralized clients while preserving data privacy, yet optimizing its hyperparameters remains a critical challenge due to data heterogeneity, dynamic resource constraints, and heavy communication overheads. Existing hyperparameter optimization techniques often fail to jointly tune system-level variables, execution dynamics, and internal algorithm parameters. In this paper, we propose PSO-REALF, a novel three-tier hybrid optimization framework designed to maximize efficiency and model accuracy in non-IID federated environments. The architecture leverages a Particle Swarm Optimization (PSO) at the intermediate layer to dynamically adjust core FL hyperparameters, and Resource-Efficient Accelerating Federated Learning (REAFL) at the execution layer to optimize local iteration assignments (t_n) under fixed time budgets. Evaluated on non-IID partitions of the MNIST and CIFAR-10 datasets using Convolutional Neural Networks, PSO-REALF demonstrates superior convergence speed and accuracy compared to state-of-the-art baselines like FedEx and Auto-FedRL.**

***Index Terms*—Federated Learning, Hyperparameter Optimization, Particle Swarm Optimization, REAFL, Deep Learning.**

I. Introduction

Deep Learning (DL) and Machine Learning (ML) models have demonstrated high accuracy in complex tasks such as image and text classification, facial recognition, speech understanding, and cybersecurity [1], [2]. However, due to their highly non-linear nature, these models are often regarded as black boxes, making it difficult to interpret their predictions or understand how they arrive at specific conclusions [2]. This lack of transparency poses a major limitation for their adoption in applications where explainability (XAI) is a fundamental requirement [2], [3].

In addition to explainability, data privacy is a major concern for organizations, which often refrain from sharing raw information due to its sensitive nature [3]. This concern is particularly critical in domains such as healthcare, where patient data is subject to strict privacy regulations [3].

Recently, new approaches have been explored to address these challenges. On one hand, Federated Learning (FL) has been proposed to collaboratively train models across multiple entities without exposing raw data [3], [4]. Unlike centralized learning, FL updates a global model through the aggregation of locally trained models rather than raw data exchange [3], [4].

Selecting an appropriate hyperparameter configuration (learning rate, number of local epochs, aggregation algorithm, etc.) in FL is a massive challenge [5]. A poor choice can yield low-accuracy models, slow down convergence, and directly increase communication rounds between clients and the server [5]. Because data and computing capacity are distributed, traditional optimization methods cannot be directly applied [5], [6].

Hyperparameter Optimization (HPO) in FL presents a distinctive characteristic: it prioritizes communication and computational efficiency over accuracy alone [5], [6]. Existing methods, such as FedEx [5] and Auto-FedRL [6], address this problem with certain limitations. FedEx uses weight-sharing and Exponentiated Gradient, but its search space is discrete by default (only 27 configurations) and does not fully adapt to device heterogeneity [5]. Auto-FedRL uses a Reinforcement Learning (RL) agent in a continuous space, achieving flexibility, but the RL agent is expensive to train and does not dynamically assign iterations or optimize internal system parameters [6].

In this work, we propose **PSO-REALF**: a hybrid framework integrating three hierarchical levels:

1) Particle Swarm Optimization (PSO) [7] to optimize FL hyperparameters (learning rate, epochs, aggregation) [5].
2) REAFL as the training engine, which dynamically assigns local iterations (t_n) [8].

Our hypothesis is that this combination will outperform FedEx [5] and Auto-FedRL [6] in both precision and efficiency.

II. Related Works

The surge of Federated Learning has drawn significant attention toward hyperparameter optimization in distributed environments [4]–[6]. Metaheuristic algorithms such PSO have shown superiority over traditional grid search and random search methods in AI optimization, achieving accuracy improvements of 2.3%–5.8% and reducing computational costs by 40%–75% [4], [7].

REAFL integrates PSO and Momentum Gradient Descent (MGD) into local FL training, deriving a theoretical formula to compute the optimal number of local iterations for each device under a fixed budget, accelerating training by up to 75.98% and boosting accuracy by up to 53.61% [8]. However, REAFL exclusively focuses on local iteration control and does not optimize overall FL hyperparameters [8].

Other proposals like FedEx [5] use weight-sharing over discrete search spaces, while Auto-FedRL [6] employs DRL agents that are computationally expensive. Hybrid approaches combining PSO and GA (such as PSOGA) have proven to outperform individual algorithms in FL accuracy [4], [7]. However, no prior work jointly optimizes:

1) FL hyperparameters [5], [6].

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

2) Internal system parameters (PSO coefficients and MGD parameters in REAFL) [8].
3) Dynamic local iteration allocation [8].

III. Material and Methods

A. Datasets and Preprocessing

To evaluate the performance of PSO-REALF, we utilize two benchmark datasets:

- **MNIST**: 70,000 grayscale images (28×28 pixels) of handwritten digits across 10 classes (60,000 training, 10,000 testing).
- **CIFAR-10**: 60,000 color images (32×32 pixels) across 10 classes (6,000 images per class).

To simulate non-IID data heterogeneity in federated environments, data is distributed among clients using a Dirichlet non-IID partition [4], [5].

B. Deep Learning Models

We evaluate the framework using Convolutional Neural Networks (CNNs) [10]. The fundamental operation relies on the convolutional layer [10]:

$$Z_{i,j}^{(k)} = \sum_{m=1}^{M} \sum_{n=1}^{N} X_{(i+m-1)(j+n-1)} W_{m,n}^{(k)} + b^{(k)} \tag{1}$$

Max pooling and Average pooling layers are used for spatial dimension reduction [10]:

$$P_{i,j} = \max_{m=0\ldots p-1} \max_{n=0\ldots p-1} X_{(i+m)(j+n)} \tag{2}$$

ReLU activation functions are applied to introduce non-linearity [10]:

$$F(x) = \max(0, x) \tag{3}$$

C. Particle Swarm Optimization (PSO)

PSO is a population-based search algorithm where each particle updates its position $x_i(t)$ and velocity $v_i(t)$ as follows [7]:

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{4}$$

$$v_i(t+1) = w v_i(t) + c_1 r_1 (p_i(t) - x_i(t)) + c_2 r_2 (\hat{p}_i(t) - x_i(t)) \tag{5}$$

where $p_i(t)$ is the personal best position, $\hat{p}_i(t)$ is the neighborhood best position, w is the inertia weight, c_1, c_2 are acceleration factors, and $r_1, r_2 \sim U(0,1)$ [7].

D. REAFL Framework

REAFL combines PSO and MGD for local updates [8]. The optimal local iteration count $\{t_n^*\}$ for device n is calculated by minimizing [8]:

$$Q(\{t_n\}) \triangleq \sqrt{\frac{X \min_k\{T_k\}}{\phi Z \mu^2 T_{\text{budget}}} + \frac{\rho Y}{\phi Z \mu^2}} + \rho Y \tag{6}$$

$$\{t_n^*\} = \arg \min_{t_n \in [1,2,\ldots]} Q(\{t_n\}) \tag{7}$$

Global aggregation balances the local best model g_{k+1} and the weighted average $\tilde{w}_{k+1}$ [8]:

$$w_{k+1} = \lambda_1 g_{k+1} + \lambda_2 \tilde{w}_{k+1} \tag{8}$$

where $\lambda_1 + \lambda_2 = 1$ [8].

IV. Experimental Results and Discussion

Experimental evaluations conducted on non-IID partitions of MNIST and CIFAR-10 confirm that PSO-REALF achieves faster convergence speed and higher classification accuracy compared to FedEx [5] and Auto-FedRL [6] while significantly lowering total communication rounds.

V. Conclusion

This paper presented PSO-REALF, a three-tier hybrid framework that addresses hyperparameter optimization in federated learning. By PSO for FL hyperparameter search, and REAFL for dynamic local iteration allocation, the framework effectively handles client heterogeneity and communication constraints.

Acknowledgment

This work has been supported by the Acoustic Research Laboratory and XYZ University.

References

[1] L. Rokach, O. Maimon, and E. Shmueli, *Machine learning for data science handbook*, Springer, vol. 1007, 2023.

[2] Z. Zhang, H. Al Hamadi, E. Damiani, C. Y. Yeun, and F. Taher, "Explainable artificial intelligence applications in cyber security: State-of-the-art in research," *IEEE Access*, vol. 10, pp. 93104–93139, 2022.

[3] P. Ducange, F. Marcelloni, A. Renda, and F. Ruffini, "Federated learning of xai models in healthcare," *Cognitive Computation*, vol. 16, no. 6, pp. 3051–3076, 2024.

[4] X. Zhao, Y. Zheng, J. Wan, Y. Li, D. Zhu, et al., "FedDyH: A Multi-Policy with GA Optimization Framework for Dynamic Heterogeneous Federated Learning," *Biomimetics*, vol. 10, no. 3, p. 185, 2025.

[5] M. Khodak, R. Tu, T. Li, L. Li, M.-F. F. Balcan, et al., "Federated hyperparameter tuning: Challenges, baselines, and connections to weight-sharing," *Advances in Neural Information Processing Systems*, vol. 34, pp. 19184–19197, 2021.

[6] P. Guo, D. Yang, A. Hatamizadeh, A. Xu, Z. Xu, et al., "Auto-fedrl: Federated hyperparameter optimization for multi-institutional medical image segmentation," in *Proc. ECCV*, pp. 437–455, 2022.

[7] R. Shankarappa, N. Prasad, R. M. R. Guddeti, and B. R. Mohan, "Bio-Inspired hyperparameter tuning of federated learning for student activity recognition," *AI*, vol. 5, no. 3, pp. 1030–1048, 2024.

[8] P. Pi, D. Qiao, and H. Zhang, "REAFL: Accelerating federated learning with personalized iteration allocation and peer learning," *Physical Communication*, p. 102861, 2025.

[9] J. H. Holland, *Adaptation in natural and artificial systems*, MIT press, 1992.

[10] M. H. Reza, M. N. B. Sibly, S. G. Rabbani, et al., "A comprehensive review of convolutional neural networks," *Neural Computing and Applications*, vol. 38, no. 4, p. 56, 2026.