FedPall: Prototype-based Adversarial and Collaborative Learning for Federated Learning with Feature Drift

Yong Zhang, Feng Liang*, Guanghu Yuan, Min Yang, Chengming Li, Xiping Hu*

ICCV OCT 19-23, 2025 HONOLULU HAWAII

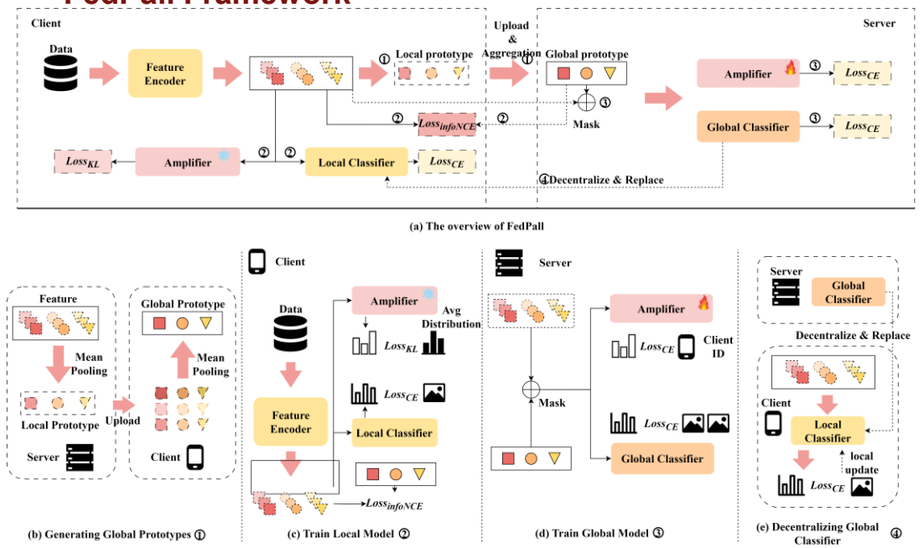
Introduction

Feature Drift in Federated Learning: Feature drift in federated learning arises when data collected across clients exhibit discrepancies in the feature space. Such variations often result from differences in sensing devices, task requirements, or environmental conditions.

Core Design of FedPall: FedPall leverages prototype-based adversarial learning to align heterogeneous feature spaces. Meanwhile, it employs collaborative learning to preserve class-specific information, ensuring that discriminative features remain intact.

Effectiveness: By jointly addressing feature alignment and class consistency, FedPall provides an effective solution to the problem of feature drift in federated learning.

FedPall Framework



① Local class prototypes are generated through client collaboration, and the server aggregates these local prototypes to form global prototypes.

② The feature encoder is trained using the adversarial learning method, and the optimized loss function is used to force the local feature encoder to enhance the client-agnostic features while generating category information.

③ The server enhances the expressive power of the amplifier through encryption hybrid features, forming an adversarial relationship with the client, and retraining the global classifier at the same time.

④ We replace local classifiers with a global classifier to construct a more generalizable classification model, while leveraging local data to strengthen its personalization capability.

Loss Function

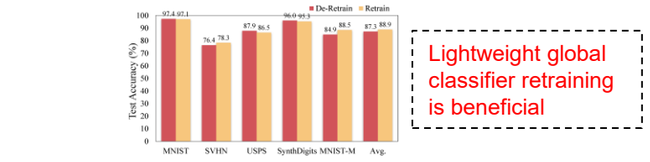
- $Loss_{KL}$ improves the generalization of each client's feature encoder while minimizing client-specific heterogeneity in feature representations.
- $Loss_{infoNCE}$ minimizes the distance between local features and their corresponding global prototypes, while maximizing the difference between local features and global prototypes of other categories.
- $Loss_{CE}$ improves the classification ability of local models in common scenarios.

Main Experiment

		SingleSet	FedAvg	FedProx	PerfedAvg	FedRep	FedBN	MOON	FedProto	ADCOL	RUCR	FedHEAL	ours(FedPall)
Office-10	amazon	74.0(2.7)	56.9(2.5)	56.6(2.6)	57.1(2.2)	45.3(1.9)	40.8(15.8)	51.7(16.1)	69.4(2.1)	73.3(4.4)	52.1(8.5)	65.1(3.3)	72.9(1.4)
	caltech	44.7(3.2)	46.5(4.6)	51.0(5.2)	50.8(1.6)	38.4(4.9)	33.9(6.5)	41.3(13.6)	39.4(6.3)	37.2(1.7)	44.3(1.0)	44.6(3.2)	44.7(8.7)
	dslr	60.2(6.7)	30.1(4.9)	33.3(10.4)	31.2(4.9)	34.4(4.9)	38.7(3.2)	24.7(1.9)	65.6(4.9)	76.3(4.9)	30.1(6.7)	67.7(1.9)	77.4(3.2)
	webcam	71.3(2.6)	37.9(6.2)	43.7(7.2)	47.1(7.8)	55.8(2.6)	30.5(6.1)	33.3(12.7)	71.3(4.3)	71.3(2.6)	37.4(5.0)	60.9(1.0)	74.7(1.0)
	avg	62.5(0.4)	42.9(1.2)	46.1(2.6)	46.6(2.9)	43.5(1.3)	36.0(6.5)	37.8(10.9)	61.4(1.7)	64.5(1.8)	41.0(0.6)	59.6(0.5)	**67.5(2.7)**
Digits	MNIST	95.5(0.2)	92.9(2.2)	91.8(3.0)	90.1(4.8)	86.5(6.1)	96.7(0.1)	93.4(1.1)	96.4(0.5)	96.3(0.4)	92.6(2.0)	93.6(0.6)	97.2(0.4)
	SVHN	71.1(0.9)	77.4(0.2)	76.9(0.3)	75.6(0.4)	67.2(1.7)	79.4(0.3)	79.6(0.8)	72.5(0.3)	75.1(2.1)	77.9(0.3)	68.3(2.0)	78.0(0.4)
	USPS	86.4(0.3)	89.3(0.9)	89.2(1.4)	88.7(0.7)	90.0(3.0)	90.1(0.5)	81.8(0.7)	87.0(0.8)	86.7(1.3)	88.9(2.4)	87.0(0.5)	87.3(1.3)
	SynthDigits	95.2(0.1)	95.5(0.1)	95.4(0.1)	95.0(0.2)	94.2(0.8)	95.6(0.1)	96.6(0.2)	95.3(0.6)	96.4(0.3)	96.0(0.2)	89.3(1.6)	95.3(0.4)
	MNIST-M	76.6(0.4)	73.8(1.5)	74.0(1.5)	73.2(0.8)	69.1(0.9)	76.3(0.4)	72.2(0.9)	78.3(1.2)	78.3(4.4)	72.7(0.4)	67.8(1.9)	85.9(1.4)
	avg	84.9(0.1)	85.8(0.9)	85.5(1.1)	84.5(1.3)	81.4(2.5)	87.6(0.1)	84.7(0.6)	85.9(0.2)	86.6(1.3)	85.6(0.9)	81.2(1.3)	**88.7(0.2)**
PACS	art_painting	33.6(0.8)	25.8(1.9)	24.3(4.1)	26.5(2.2)	26.9(3.3)	36.7(1.8)	30.6(2.0)	32.7(0.7)	34.9(1.2)	24.7(1.1)	31.2(1.2)	35.6(0.6)
	cartoon	58.5(2.5)	45.4(2.3)	51.4(0.6)	48.3(1.2)	44.4(2.1)	55.6(2.0)	51.5(1.8)	57.3(1.5)	57.2(0.8)	47.5(3.3)	50.8(0.4)	59.7(2.3)
	photo	63.0(1.9)	48.7(3.1)	49.6(2.0)	46.9(2.6)	41.9(2.8)	66.1(1.0)	53.0(3.3)	64.0(1.3)	62.1(2.0)	47.5(6.2)	61.1(2.0)	64.7(1.3)
	sketch	79.7(0.1)	49.0(2.0)	40.7(1.5)	44.4(3.8)	40.5(1.3)	79.6(1.7)	55.1(1.4)	79.6(0.8)	80.1(1.0)	42.2(2.4)	73.8(0.1)	82.2(0.7)
	avg	58.7(1.2)	42.2(1.6)	41.5(1.8)	41.5(1.9)	38.4(1.1)	59.5(1.4)	47.6(0.9)	58.4(0.3)	58.6(0.6)	40.5(2.0)	54.2(0.5)	**60.6(0.4)**

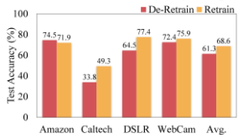
Our method outperforms other personalized federated learning algorithms on three benchmark datasets, achieving SOTA performance in feature drift scenarios.

Ablation Study

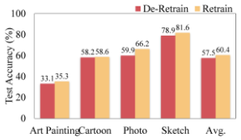


(a) Digits Datasets

Lightweight global classifier retraining is beneficial

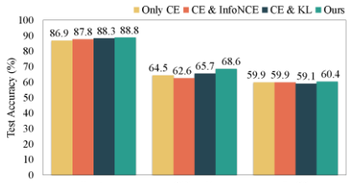


(b) Office-10 Datasets

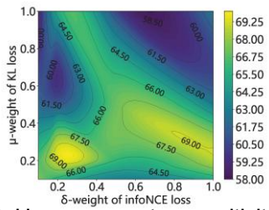


(c) PACS Datasets

Without retraining the global classifier (De-Retrain), classification performance declines to varying degrees across all datasets.



1. Effect of loss combination



2. Hyperparameter sensitivity

1. The results of different loss function combinations indicate that our loss function design can effectively address the feature drift problem.
2. b) When μ is set to around 0.2, the performance remains relatively stable, and it reaches its best when δ is around 0.2.

Acknowledgement

This work was supported in part by the Innovation Team Project of Guangdong Province of China (2024KCXTD017) and Guangdong-Hong Kong-Macau University Joint Laboratory (2023LSYS005).