Zhaohui WANG, Hongjiao LI, Jinguo LI, Renhao Hu, Baojin WANG, 2024.
Federated learning on non-IID and long-tailed data via dual-decoupling.
Frontiers of Information Technology & Electronic Engineering, 25(5):728-741.
https://doi.org/10.1631/FITEE.2300284

Federated learning on non-IID and long-tailed data via dual-decoupling

Key words: Federated learning; Non-IID; Long-tailed data; Decoupling learning; Knowledge distillation

Corresponding author: Hongjiao LI
E-mail: hjli@shiep.edu.cn
ORCID: https://orcid.org/0000-0003-0642-9046

Motivation

- When both non-independent and identically distributed (non-IID) and long-tailed distributions are presented, the accuracy of the federated averaging (FedAvg) algorithm significantly deteriorates, and the trained models tend to favor the head classes. However, existing federated learning (FL) methods seldom take into account heterogeneous scenarios where non-IID and long-tailed distributions coexist.
- Inspired by the principle of decoupling learning, we introduce a new solution to address the combined challenge of non-IID and long-tailed data distributions in FL, through a federated dual-decoupling via model and logit calibration (FedDDC).

Main idea

- We propose an FL framework, the FedDDC, which studies FL problems based on non-IID and long-tailed data from the perspectives of decoupling model and knowledge.
- We propose a novel confidence-weighting method that efficiently reassigns the weights for each client through the loss of the client model and the logits of the target class.
- We validate the proposed approach in a variety of contexts with different degrees of non-IID-ness and imbalance factors. According to the results of the experiments, FedDDC routinely outperforms state-of-the-art FL methods in terms of performance.

Framework

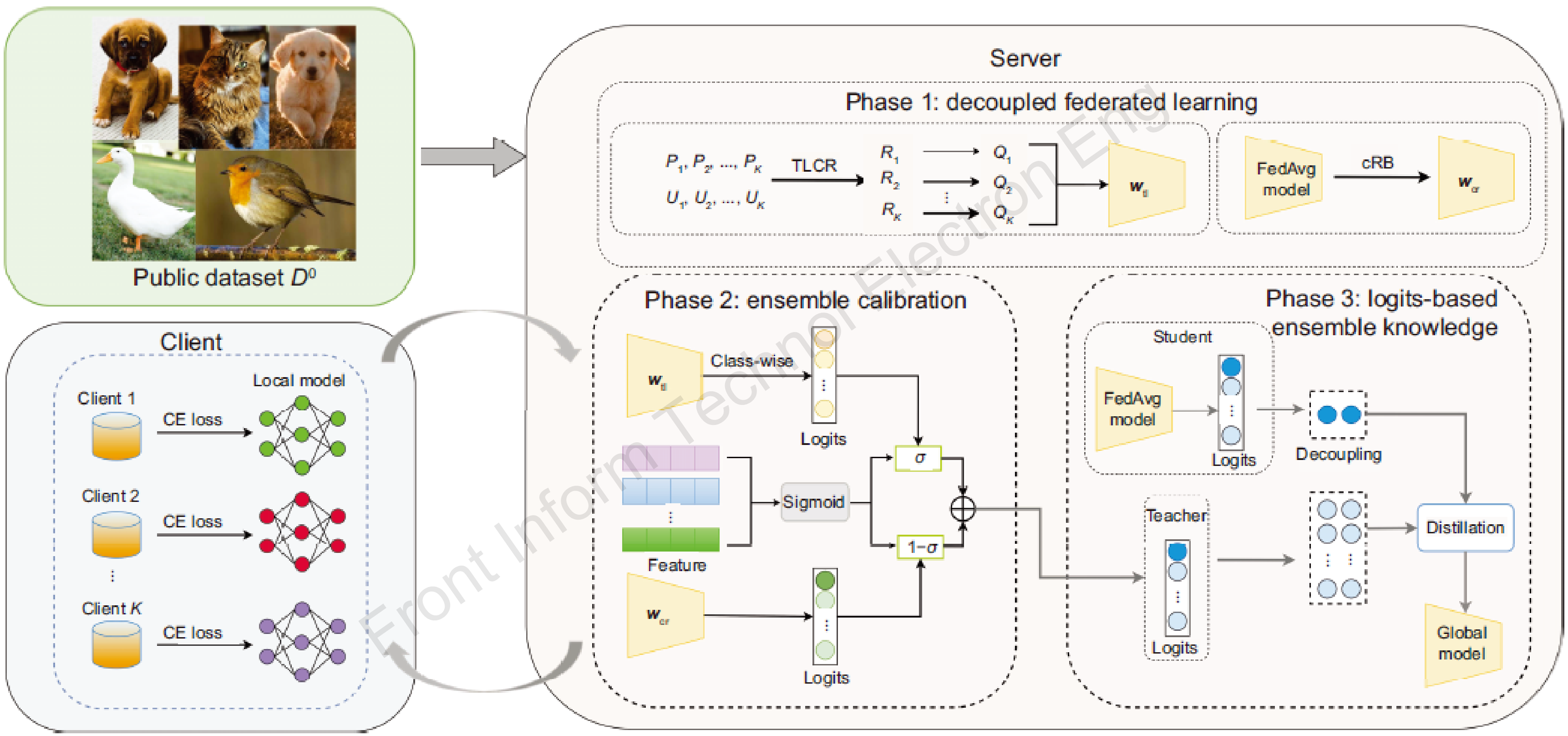

Fig. 1 Framework of FedDDC

In each round, clients send updated local models to the server, and the server sends the aggregated global model to the clients. In the decoupled federated learning, target-loss confidence re-weighting (TLCR) and classifier re-balancing (cRB) are adopted to calibrate the biased feature extractor and the biased classifier, respectively. Here, w_{tl} is the model with the unbiased feature extractor. w_{cr} indicates the model with the unbiased classifier

Method

- We propose the decoupled FL to decouple the global model into the feature extractor and the classifier. We specifically suggest a client confidence re-weighting method to calibrate the biased feature extractor. At the same time, in the other perspective, we use a classifier re-balancing (cRB) scheme to adjust the biased classifier.
- However, the generalization performance of the aggregated global model on the tail class is poor when the number of tail classes is small. Therefore, we propose an ensemble calibration method to weigh the output of the two perspectives.
- Finally, we use logits-based ensemble distillation to distill unbiased knowledge from the calibrated ensemble model to the global model.

Conclusions

In this paper, we propose a dual-decoupling FL framework, FedDDC, for tackling the problem of non-IID and long-tailed distribution. FedDDC concentrates on decoupling the global model and knowledge while decreasing the negative effects of non-IID and long-tailed data through the class-wise logit adjustment and a calibration function. In decoupled FL, we particularly propose an innovative confidence re-weighting approach and a classifier rebalancing method to calibrate the biased feature extractor and the biased classifier, respectively. To extract unbiased knowledge from the calibrated ensemble model, the logits-based ensemble distillation method is adopted to transfer knowledge. Extensive experiments prove the superior performance of our approach to cutting-edge FL methods in the setting of non-IID and long-tailed distribution. Moreover, numerous studies verify the effectiveness of each component included in our approach.

Zhaohui WANG received the B.Eng degree from Anhui Sanlian University, in 2021. He is currently working toward the MS degree in computer science and engineering with Shanghai University of Electric Power. His research interests include applied federated learning, knowledge distillation, and privacy-preserving deep learning.

Hongjiao LI received the PhD degree in the School of Electronic Information and Electrical Engineering from Shanghai Jiao Tong University, Shanghai, China, in 2002. She is currently an associate professor with the College of Computer Science and Technology, Shanghai University of Electric Power, Shanghai, China. Her research interests include federated learning, differential privacy, and information security.

Jinguo LI (Member, IEEE) received the BS degree in information security and the PhD degree in computer science and technology from Hunan University, Changsha, China, in 2007 and 2014, respectively. He is currently an associate professor with the College of Computer Science and Technology, Shanghai University of Electric Power, Shanghai, China. His research interests include information security and privacy, applied cryptography, and cloud computing.

Renhao HU received his B.Eng degree from Hefei Normal University, in 2021. He is currently working toward the MS degree with the College of Computer Science and Technology, Shanghai University of Electric Power. His research interests include information security, differential privacy, and cloud computing.

Baojin WANG received the B.Eng degree from the Shanghai University of Electric Power, in 2021. He is currently working toward the MS degree in power information technology with Shanghai University of Electric Power. His research interests include federated learning and efficient communication.