Harnessing Asynchrony to Balance Modalities in Multi-Modal Federated Learning

Yiming Ma[1,5], Boyi Liu[1], Zimu Zhou[2,3], Yanfeng Wang[4,5], and Yongxin Tong[1]

[1]Beihang University, [2]City University of Hong Kong, [3]City University of Hong Kong Shenzhen Research Institute,

[4]Shanghai Jiao Tong University, [5]Shanghai Artificial Intelligence Laboratory

Background

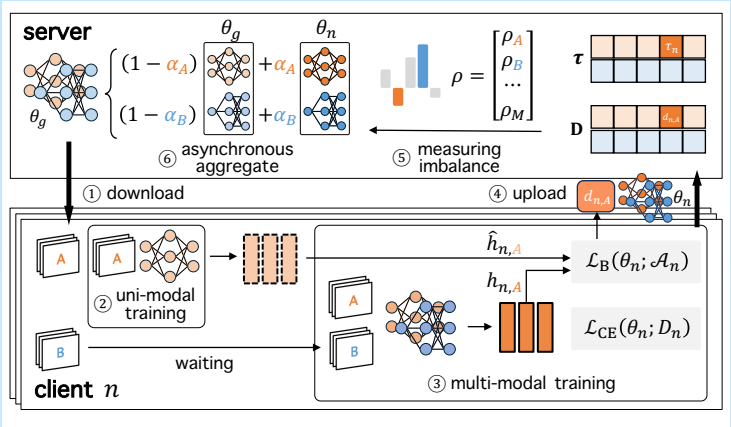
Multi-Modal Federated Learning (MMFL): Enables collaborative training across clients with diverse data types (e.g., images, audio, sensors) without sharing raw data due to strict data privacy regulations (e.g., GDPR). It achieves higher accuracy by leveraging complementary information from multiple modalities.

Challenge

- **Challenge 1: Modality Laziness**
  Dominant modalities overshadow weaker ones during joint optimization, leading to biased representations.
- **Challenge 2: Asynchronous Availability**
  Modalities arrive at clients at different times due to heterogeneous sampling and network jitter.
- **The Conflict:** Waiting for all modalities wastes time; training immediately with partial data exacerbates modality laziness.

Design of MBA

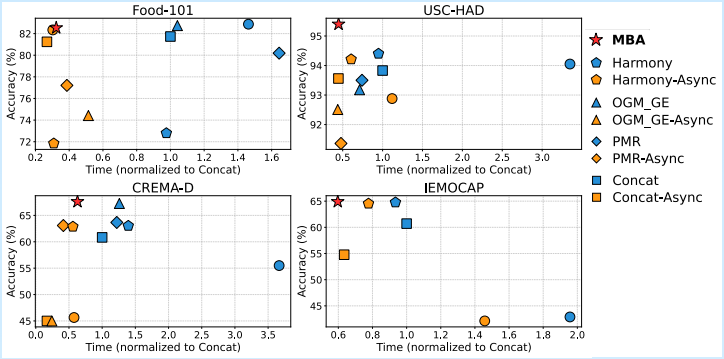
We proposed a two-tire framework of MMFL that harness asynchrony to mitigate modality laziness. We treat asynchrony from a 'bottleneck' into a 'resource'. By utilize idle waiting by performing opportunistic uni-modal training during data arrival gaps. We cache the features of early-arriving modalities as an anchor for joint optimization. And we calculate the distance between the features of the anchors and joint optimization, and the distance is uploaded to server to achieve a global view on estimating the imbalance between modalities. Finally, the server aggregates the uploaded model in a modality-decoupled way to balance the modality laziness.



MBA System Overview

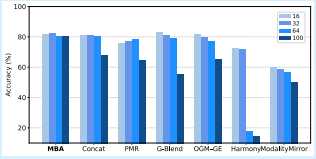
Results

Evaluation on 4 real-world datasets demonstrates that MBA achieves better accuracy and performs high efficiency in other asynchronous baselines.

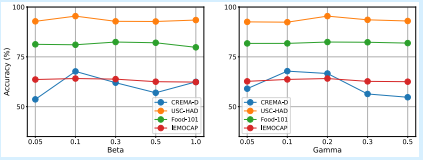


Experiment of Accuracy and Efficiency

We evaluate MBA and all asynchronous baselines under client partitions of 16, 32, 64 and 100. As the number of clients increases, MBA performs strong stability and scalability. MBA maintains stable accuracy across a broad range of settings for beta and gamma with only minor fluctuations.



Scalability Experiment



Hyperparameter Sensitivity

智汇北航 青云传讯