EAST CHINA NORMAL UNIVERSITY

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Class Balanced Adaptive Pseudo Labeling for Federated Semi-Supervised Learning

Ming Li, Qingli Li, Yan Wang

Shanghai Key Laboratory of Multidimensional Information Processing,
East China Normal University

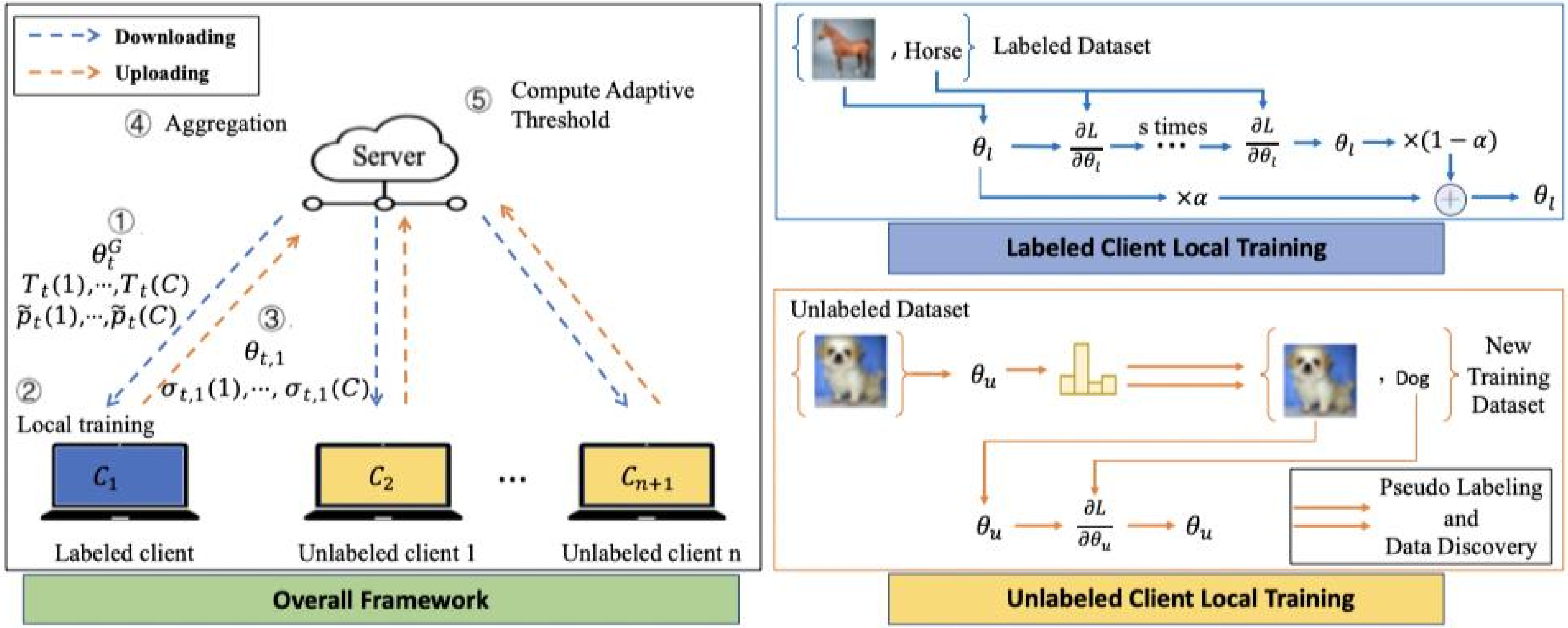
Paper Tag: WED-PM-375

Overview of Method

We present CBAFed, a new method to deal with federated semi supervised learning task.

- For labeled clients, we introduce residual weight connection to make training more stable and reach better optimum.

- For unlabeled clients, we propose fixed pseudo labeling, class balanced adaptive thresholds and tain class data discovery to deal with Non-IID problem in federated learning



Experimental Results and Analysis

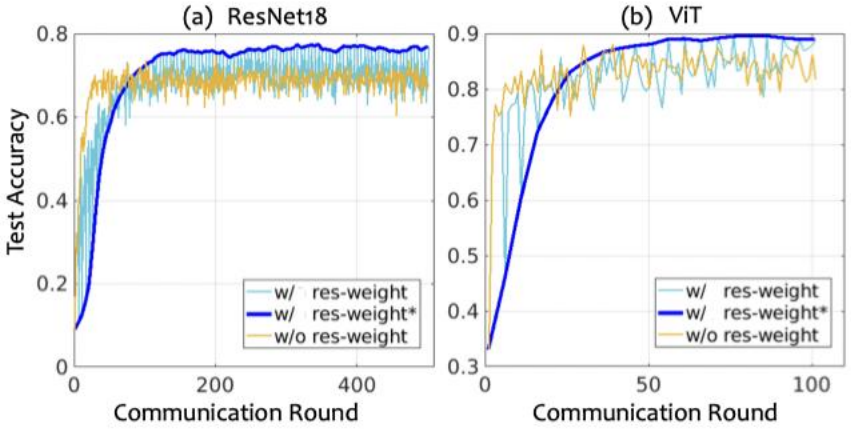
We conduct experiments on five datasets and do lots of analysis

Table 1. Results on SVHN, CIFAR-10/100, Fashion MNIST and ISIC 2018 datasets under heterogeneous data partition with ResNet18. FedAVG$^+$ means FedAvg [19] trained with all one labeled clients using our residual weight connection. Fed-consist$^+$ means Fed-Consist [31] using our proposed fixed pseudo labeling without enlarging the weight of labeled client.

Labeling Strategy	Method	Client Num.		Dataset				
		labeled	unlabeled	SVHN	CIFAR10	CIFAR100	Fashion-MNIST	ISIC 2018
Fully supervised	FedAvg [19](upper-bound)	10	0	91.83	80.89	51.38	90.14	81.32
	FedAvg [19](lower-bound)	1	0	67.71	54.66	20.49	74.87	65.13
	FedAvg$^+$ [19]	1	0	76.98	58.21	24.84	78.26	66.69
Semi supervised	FedIRM [18]	1	9	69.22	52.84	20.20	76.83	64.85
	Fed-Consist [31]	1	9	70.56	54.23	21.81	76.57	65.20
	Fed-Consist$^+$ [31]	1	9	86.57	56.35	23.25	78.35	65.50
	RSCFed [14]	1	9	76.74	57.07	28.46	78.40	67.21
	CBAFed(ours)	1	9	**88.07**	**67.08**	**30.18**	**85.49**	**68.29**

Table 2. Comparison of our method against RSCFed [14], Fed-Consist [31] and FedAVG [19] in SVHN dataset on ViT [5] as the backbone, with one labeled and nine unlabeled clients.

Method	Client Num.		Accuracy
	labeled	unlabeled	
FedAVG [19](upper bound)	10	0	96.81
FedAVG [19](lower bound)	1	0	81.68
FedAVG$^+$ [19]	1	0	88.93
FedIRM [18]	1	9	79.44
Fed-Consist [31]	1	9	85.91
Fed-Consist$^+$ [31]	1	9	93.21
RSCFed [14]	1	9	89.43
CBAFed(ours)	1	9	**95.09**



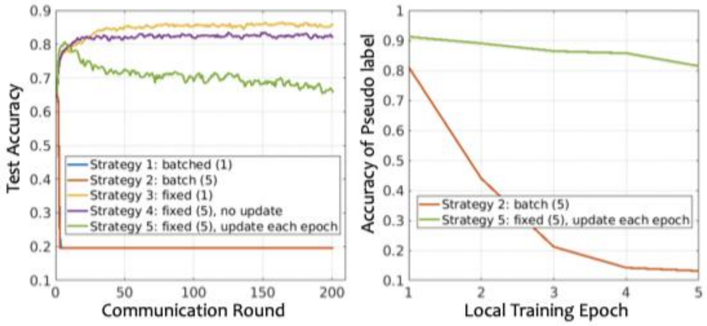


Table 3. Comparison of our method against RSCFed [14], Fed-Consist [31], FedIRM [18] and FedAVG [19] with the number of labeled and unlabeled client set to 2 and 8.

Method	Client Num.		Accuracy
	labeled	unlabeled	
FedAVG [19](upper bound)	10	0	80.89
FedAVG [19](lower bound)	2	0	61.85
FedAVG$^+$ [19]	2	0	66.55
FedIRM [18]	2	8	62.62
Fed-Consist [31]	2	8	61.67
Fed-Consist$^+$ [31]	2	8	68.04
RSCFed [14]	2	8	64.25
CBAFed(ours)	2	8	**72.01**

Introduction

Semi-supervised learning aims to effectively utilize both small size labeled data and large size unlabeled data.

In recent years, with the development of federated learning, federated semi-supervised learning(FSSL) becomes popular.

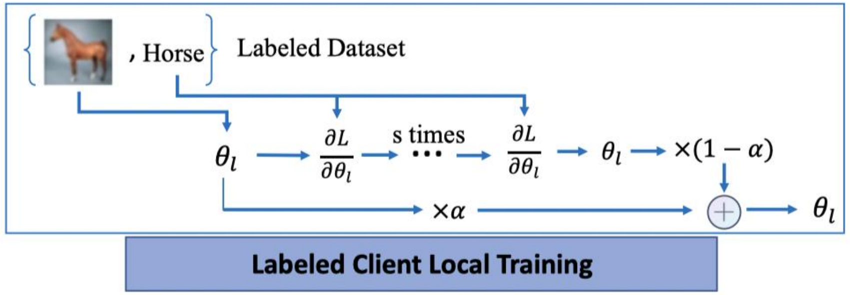
Compared with standard semi-supervised learning, FSSL is more challenging:

- There are no labeled data in unlabeled clients.
- Due to Non-IID setting, class distributions of labeled and unlabeled clients are divergent.
- Catastrophic forgetting problems may harm pseudo labeling process.

To deal with above problems, we propose Class Balanced Adaptive Pseudo Labeling method.

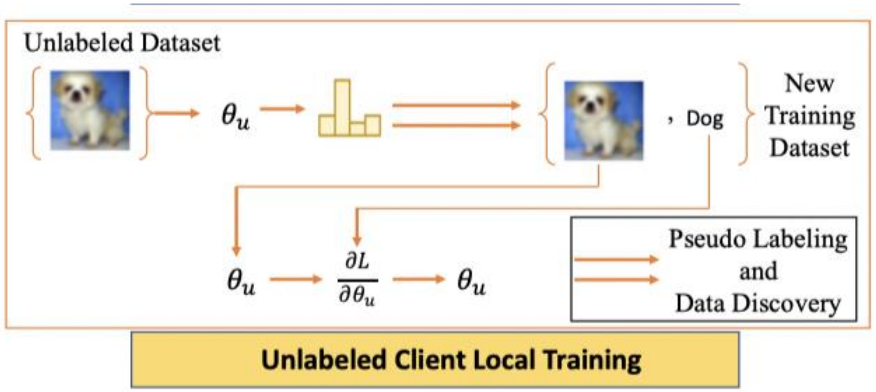
Method

For labeled clients, we propose residual weight connection



We periorically recall model from previous epoch, and use exponentail moving average to obtain new model

For unlabeled clients, we propose fixed pseudo labeling, class balanced adaptive threshold and tail class data discovery



- Compared with traditional batch-based pseudo labeling, fixed pseudo labeling obtain fixed pseudo label training set at the begining of training.
- We design class balanced adaptive thresholds via considering the empirical distribution of training data. Analysis shows that our method can set a reasonably high threshold for scarce classes.
- To discover unlabeled data from tail classes, we propose to leverage information from so-called “not informative” unlabeled data by considering their second largest confidence.

Experimental Results

We conduct experiments on five datasets to show the superiority of our method.

Table 1. Results on SVHN, CIFAR-10/100, Fashion MNIST and ISIC 2018 datasets under heterogeneous data partition with ResNet18. FedAVG$^+$ means FedAvg [19] trained with all one labeled clients using our residual weight connection. Fed-consist$^+$ means Fed-Consist [31] using our proposed fixed pseudo labeling without enlarging the weight of labeled client.

Labeling Strategy	Method	Client Num.		Dataset				
		labeled	unlabeled	SVHN	CIFAR10	CIFAR100	Fashion-MNIST	ISIC 2018
Fully supervised	FedAvg [19](upper-bound)	10	0	91.83	80.89	51.38	90.14	81.32
	FedAvg [19](lower-bound)	1	0	67.71	54.66	20.49	74.87	65.13
	FedAvg$^+$ [19]	1	0	76.98	58.21	24.84	78.26	66.69
Semi supervised	FedIRM [18]	1	9	69.22	52.84	20.20	76.83	64.85
	Fed-Consist [31]	1	9	70.56	54.23	21.81	76.57	65.20
	Fed-Consist$^+$ [31]	1	9	86.57	56.35	23.25	78.35	65.50
	RSCFed [14]	1	9	76.74	57.07	28.46	78.40	67.21
	CBAFed(ours)	1	9	**88.07**	**67.08**	**30.18**	**85.49**	**68.29**

Results on ViT backbone

Table 2. Comparison of our method against RSCFed [14], Fed-Consist [31] and FedAVG [19] in SVHN dataset on ViT [5] as the backbone, with one labeled and nine unlabeled clients.

Method	Client Num.		Accuracy
	labeled	unlabeled	
FedAVG [19](upper bound)	10	0	96.81
FedAVG [19](lower bound)	1	0	81.68
FedAVG$^+$ [19]	1	0	88.93
FedIRM [18]	1	9	79.44
Fed-Consist [31]	1	9	85.91
Fed-Consist$^+$ [31]	1	9	93.21
RSCFed [14]	1	9	89.43
CBAFed(ours)	1	9	**95.09**

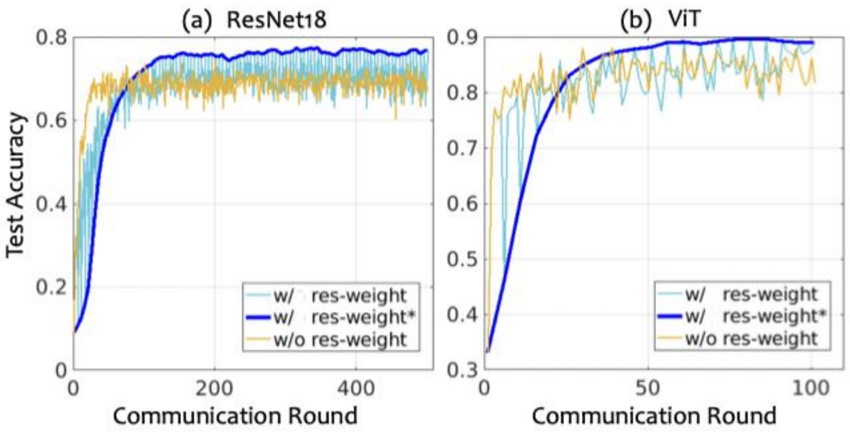
Results on 2 labeled clients

Table 3. Comparison of our method against RSCFed [14], Fed-Consist [31], FedIRM [18] and FedAVG [19] with the number of labeled and unlabeled client set to 2 and 8.

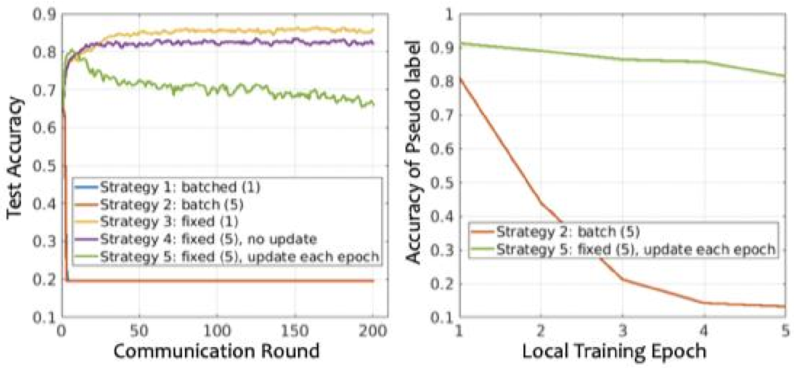
Method	Client Num.		Accuracy
	labeled	unlabeled	
FedAVG [19](upper bound)	10	0	80.89
FedAVG [19](lower bound)	2	0	61.85
FedAVG$^+$ [19]	2	0	66.55
FedIRM [18]	2	8	62.62
Fed-Consist [31]	2	8	61.67
Fed-Consist$^+$ [31]	2	8	68.04
RSCFed [14]	2	8	64.25
CBAFed(ours)	2	8	**72.01**

Analysis

Analysis of residual weight connection



Residual weight connection can make training more stable and finally reach better optimum, and hence increase prediction accuracy in pseudo labeling process.



We study different pseudo labeling strategies for FSSL, experiments show that fixed pseudo labeling with epoch=1 is best.

Thank you for listening!