Federated Cross Learning for Medical Image Segmentation

Xuanang Xu[1], Hannah H. Deng[2], Tianyi Chen[3], Tianshu Kuang[2], Joshua C. Barber[2], Daeseung Kim[2], Jaime Gateno[2,4], James J. Xia[2,4 ✉], and Pingkun Yan[1 ✉]

[1] Department of Biomedical Engineering and the Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
[2] Department of Oral and Maxillofacial Surgery, Houston Methodist Research Institute, Houston, TX 77030, USA
[3] Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
[4] Department of Surgery (Oral and Maxillofacial Surgery), Weill Medical College, Cornell University, New York, NY 10065, USA

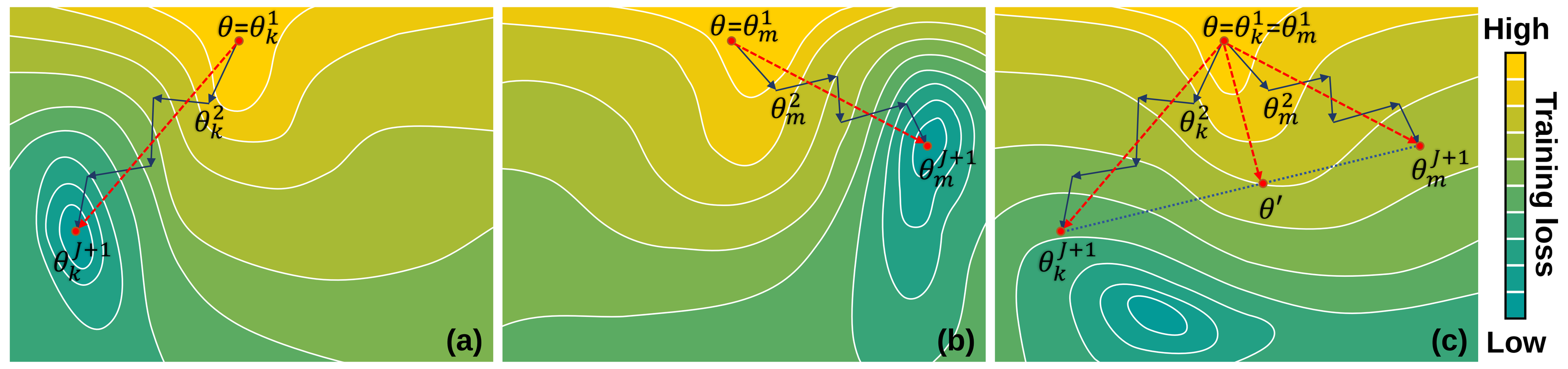
Presenting author:
Xuanang Xu (xux12@rpi.edu)
Corresponding authors:
James J. Xia (jxia@houstonmethodist.org)
Pingkun Yan (yanp2@rpi.edu)

Paper PDF | Source code

Introduction

- **Federated learning (FL)**: A decentralized learning paradigm for collaboratively training deep learning (DL) models without sharing raw data across different data owners, which is especially apt for resolving the tension between maintaining privacy for medical image data and satisfying the substantial data needs of DL models.
- **Non-iid issue of FL**: Performance degrades when handling the data that are not independently and identically distributed (non-iid data).
- **Cause of non-iid issue**: Model averaging process in the FL methods may lead to sub-optimal solutions in the parameter space due to the non-convex nature of the training objective of deep neural networks.

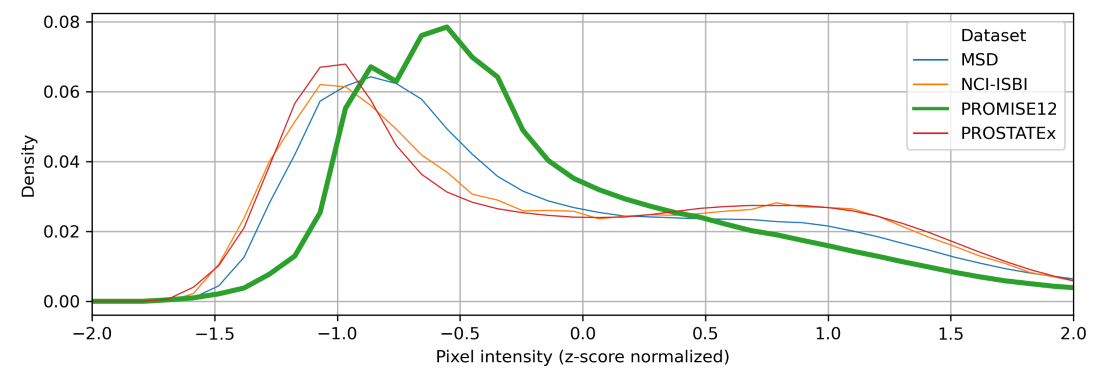
$$\theta' = \theta - \sum_{k=1}^{K} w_k \alpha_1 \nabla l(\theta, D_k) - \sum_{k=1}^{K} w_k \sum_{j=2}^{J} \alpha_j \nabla l(\theta_k^j, D_k)$$

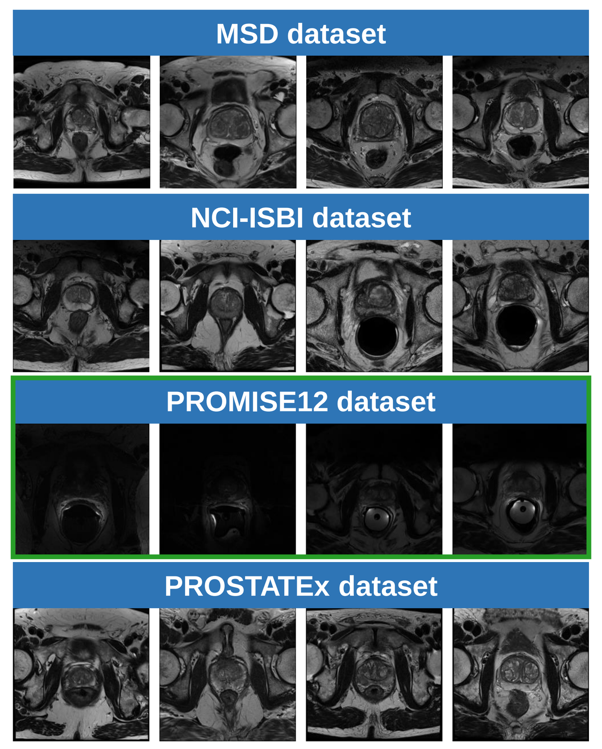


Datasets

- **Four public datasets** for MRI prostate segmentation: *MSD-Prostate*, *NCI-ISBI-Prostate*, *PROMISE12*, and *PROSTATEx*.
- **FL with non-iid data**: Each dataset mimics a local site with different imaging devices and protocols.
- **PROMISE12 dataset** exhibits significantly skewed distribution than other three datasets.

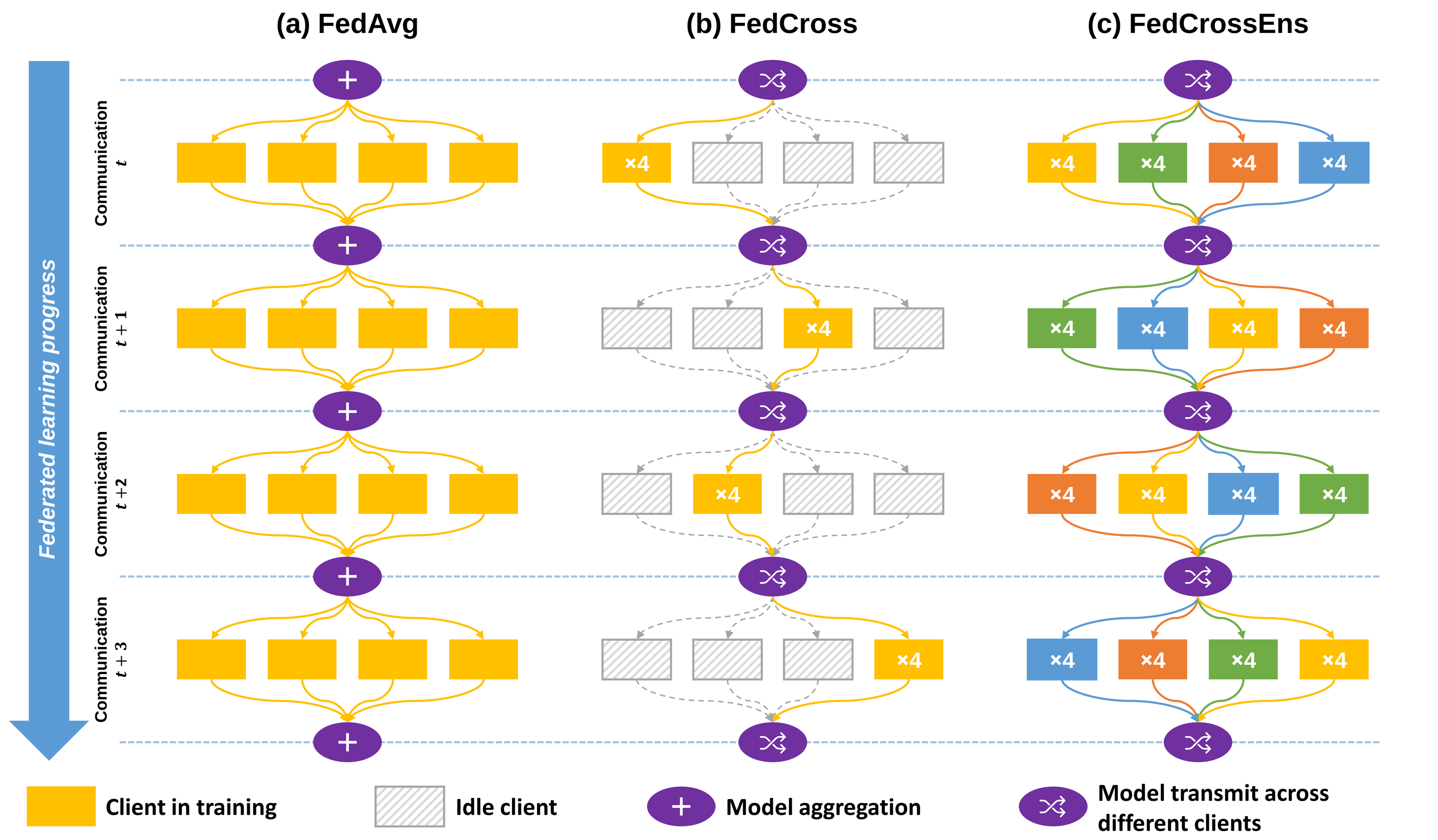
Datasets	Training set (60%)	Validation set (10%)	Testing set (30%)	Total
MSD-Prostate	19	3	10	32
NCI-ISBI-Prostate	48	8	24	80
PROMISE12	30	5	15	50
PROSTATEx	122	20	62	204
Centralized	**219**	**36**	**111**	**366**





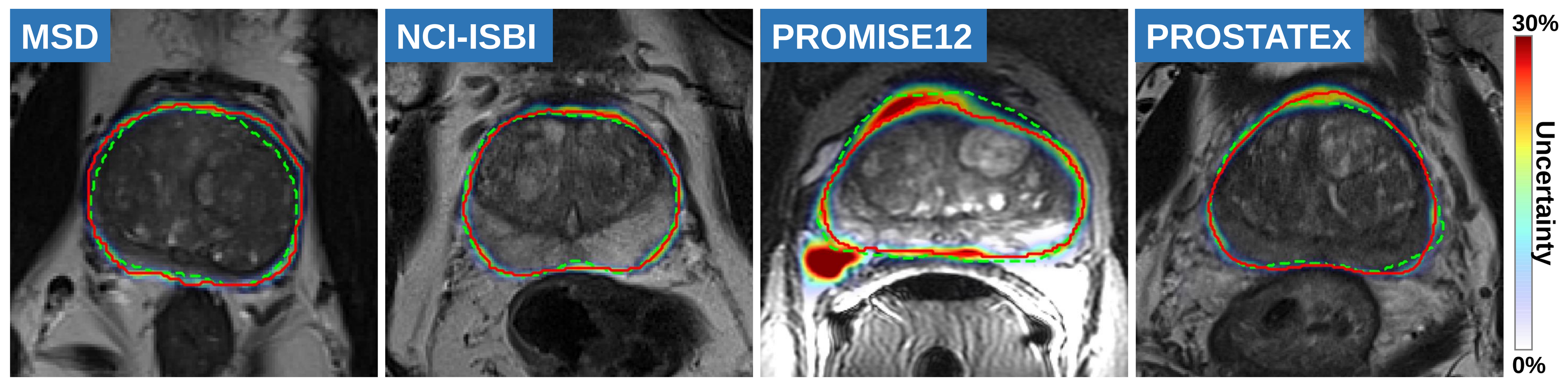
Method

- **FedAvg**[1]: Local sites simultaneously train the copy of global model and average the locally trained models after each communication round.
- **FedCross**: Local sites sequentially train global model in a round-robin manner, which avoids model parameter averaging process in FL.
- **FedCrossEns**: Multiple global model copies are independently trained by FedCross method. The output of these global model copies are assembled to get more accurate results with uncertainty estimation.



- **Uncertainty estimation** with FedCrossEns:

  By combining with ensemble mechanism, FedCrossEns can further refine the segmentation results of FedCross and estimate uncertainties.



References

1. Brendan McMahan et al., "Communication-efficient learning of deep networks from decentralized data," 2017.
2. Tian Li et al., "Federated optimization in heterogeneous networks," 2020.
3. Xiaoxiao Li et al., "FedBN: Federated learning on non-iid features via local batch normalization," 2021.

Results

- **Comparison with benchmarking training strategies**:

  Localized training: Each local site individually train a model with local data.

  Centralized training: All local sites share and gather their data to collaboratively train a model.

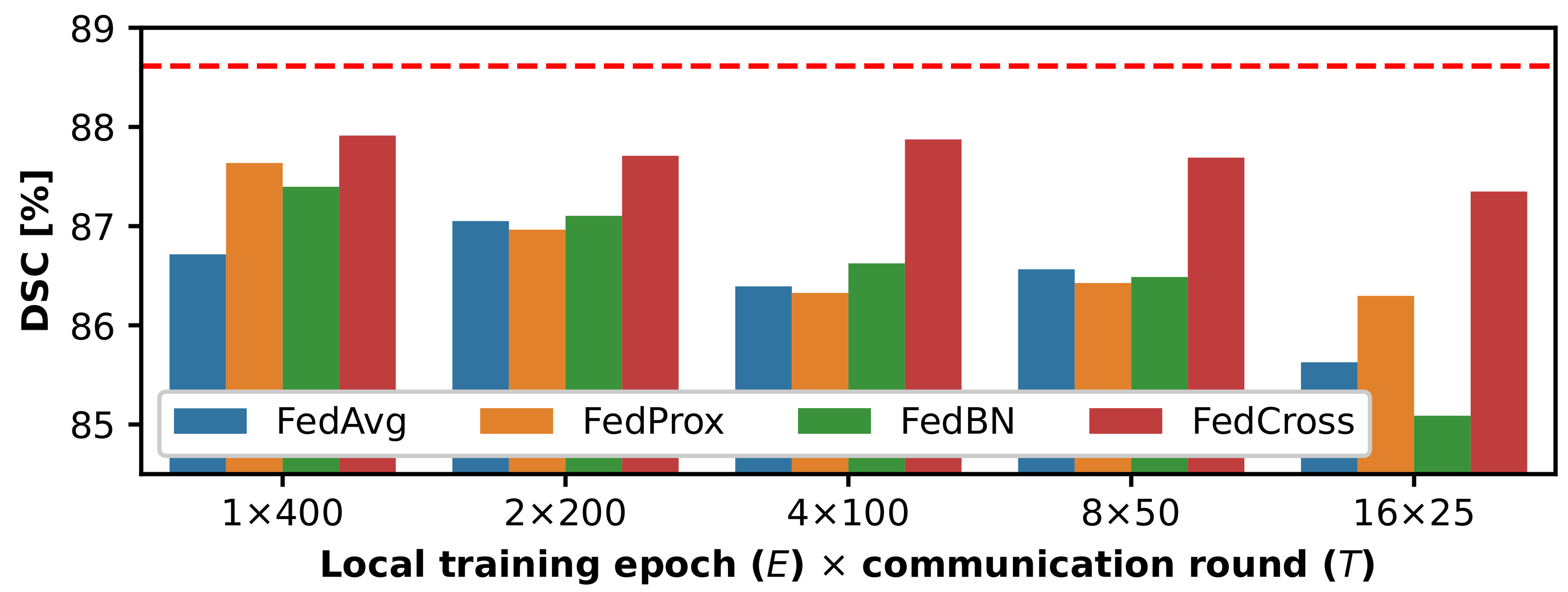
  Federated training: FedAvg[1], FedProx[2], FedBN[3], and our method.

Datasets	MSD	NCI-ISBI	PROMISE12	PROSTATEx	Global	
	DSC [Mean(SD)%]				DSC [%]	ASD [mm]
Localized	83.97(11.92)	84.04(4.87)	81.64(14.05)	**90.67(2.79)**	85.08	2.33
Centralized	90.68(2.40)	**87.19(4.68)**	86.15(5.56)	90.44(2.74)	88.61	1.43
FedAvg[1]	89.96(2.85)	84.93(7.59)	81.64(9.94)	90.34(2.96)	86.72	1.57
FedProx[2]	90.16(2.40)	86.27(5.02)	83.53(8.48)	90.59(2.84)	87.64	1.54
FedBN[3]	90.06(2.99)	86.06(6.33)	83.08(7.80)	90.38(2.96)	87.40	1.58
FedCross	90.31(2.36)	85.96(6.87)	85.09(5.68)	90.29(2.85)	87.91	1.57
FedCrossEns	**90.77(2.47)**	86.78(6.60)	**86.72(5.54)**	90.66(2.80)	**88.73**	**1.22**

Note: Underlined numbers indicate a result with statistical significance compared with the bottom row ($p<0.05$).

- **Impact of local training epoch number** in FL:

  The risk of getting sub-optimal model via model averaging can be amplified when the local training step number increases.



Conclusion

- **FedCross** can effectively address the non-iid data issue in FL-based medical image segmentation by the **aggregation-free design**.
- **FedCrossEns** can further boost the segmentation accuracy of FedCross by utilizing ensemble mechanism, which also enables **uncertainty estimation**.
- **Catastrophic forgetting** could be a potential limitation of the proposed method.