

# Personalizing Federated Medical Image Segmentation via Local Calibration

Jiacheng Wang<sup>1,†</sup>, Yueming Jin<sup>2,†</sup>, Liansheng Wang<sup>1,\*</sup>

<sup>1</sup> Department of Computer Science at School of Informatics, Xiamen University,

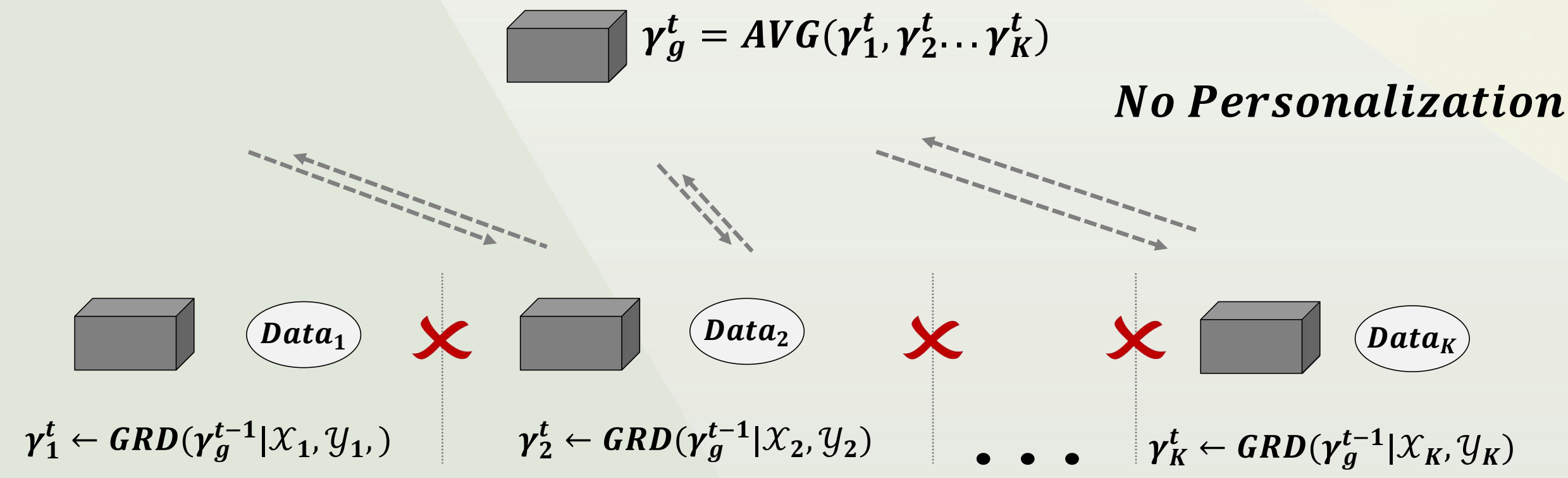
<sup>2</sup> Wellcome/EPSCRC Centre for Interventional and Surgical Sciences (WEISS) and Department of Computer Science, University College London,

<sup>†</sup> They contributed equally to this work.

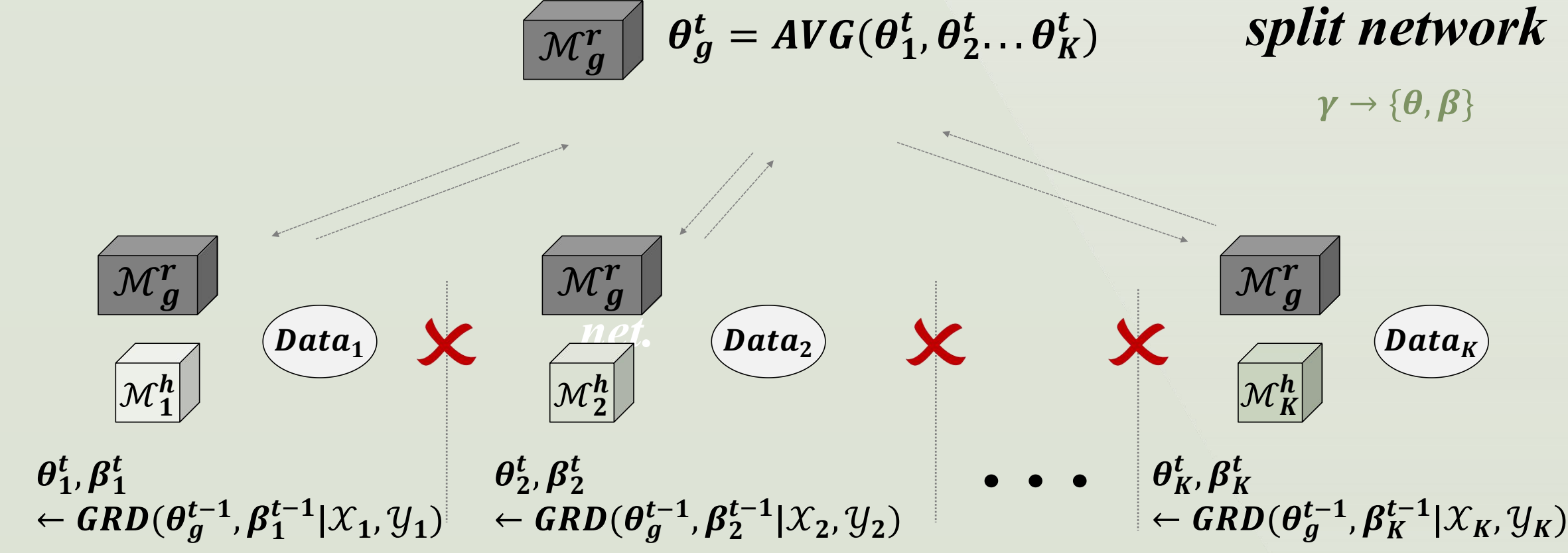


## INTRODUCTION

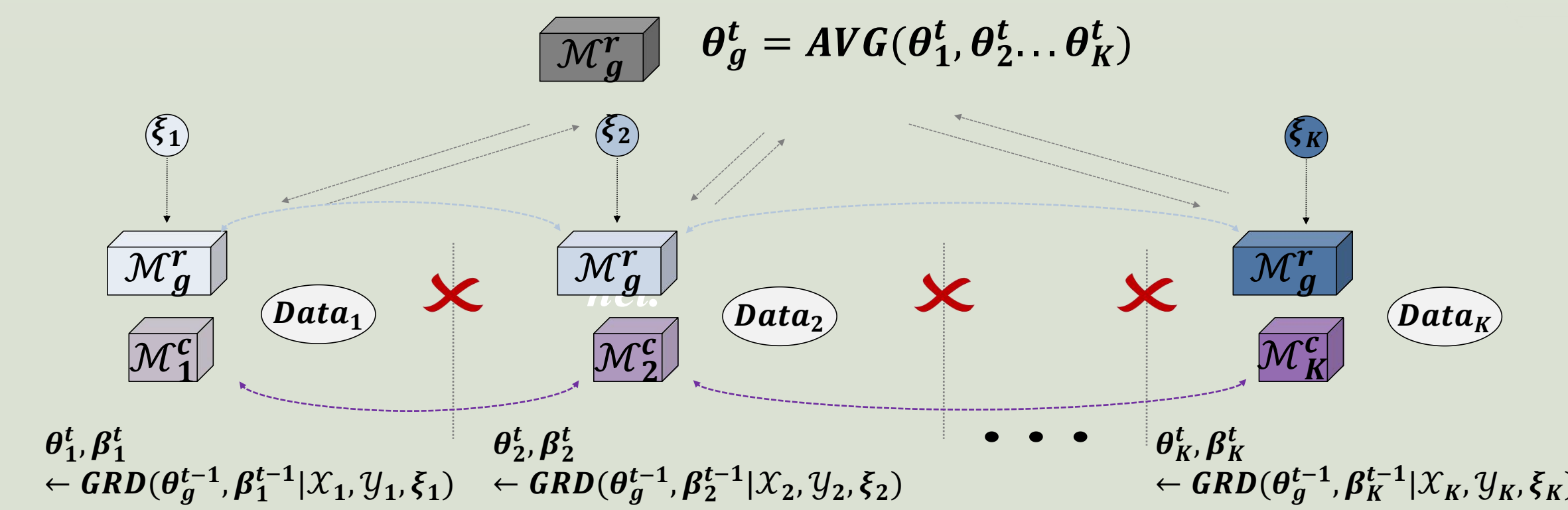
**FedAVG:** Federated Learning (FL) lacks personalized solutions.



**Model Decoupling (FedRep):** Personalized Federated Learning (FL) fails to explore the inter-site inconsistencies.



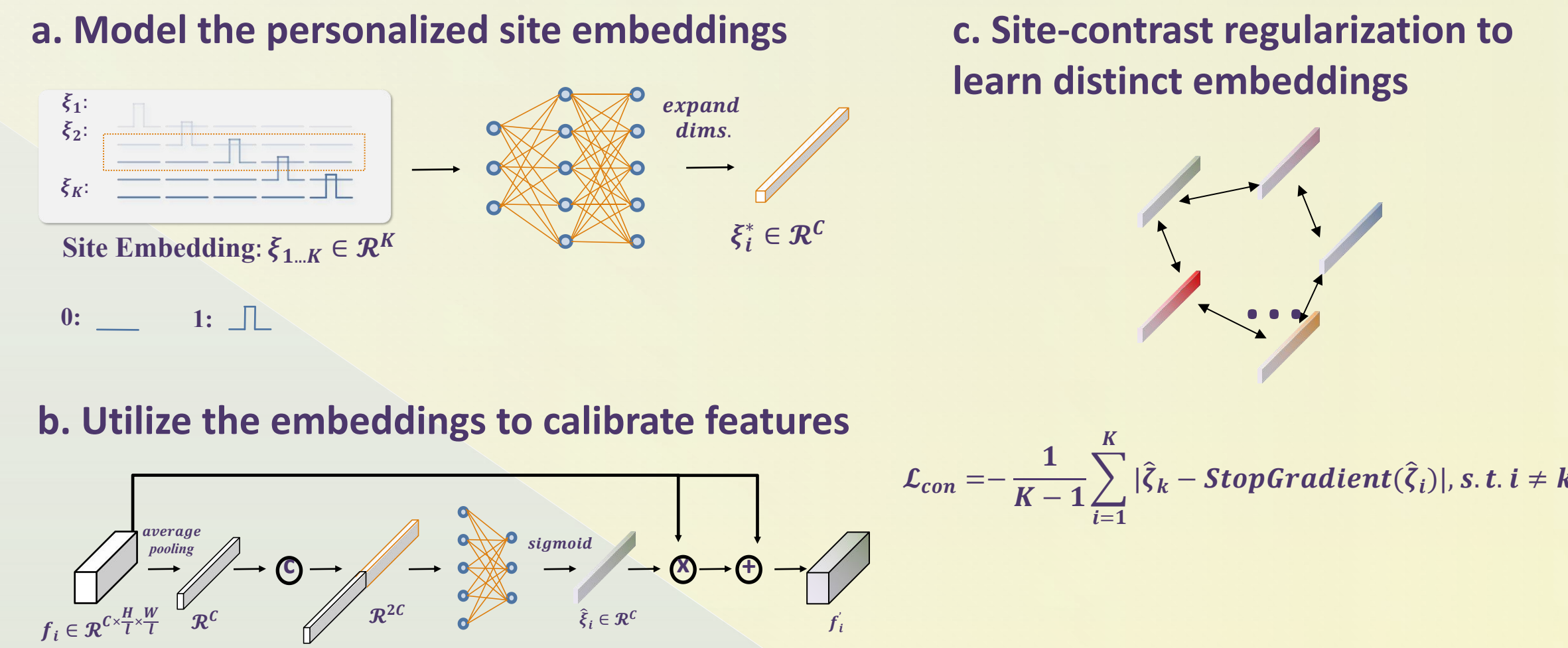
**Ours (LC-Fed):** to explore inconsistencies at **feature-** and **prediction-** levels



## Method

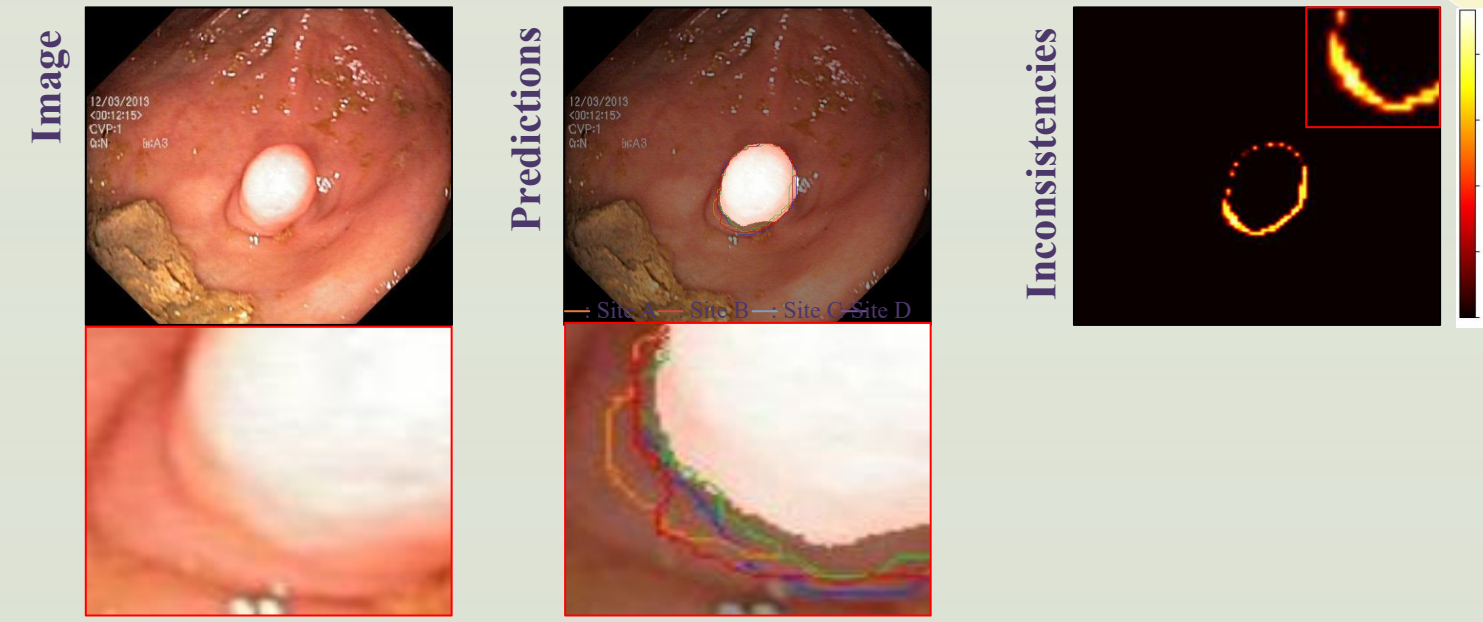
**Solutions:** Model the inconsistencies and utilize them to refine the features and the predictions.

### Feature: Channel-level inconsistencies

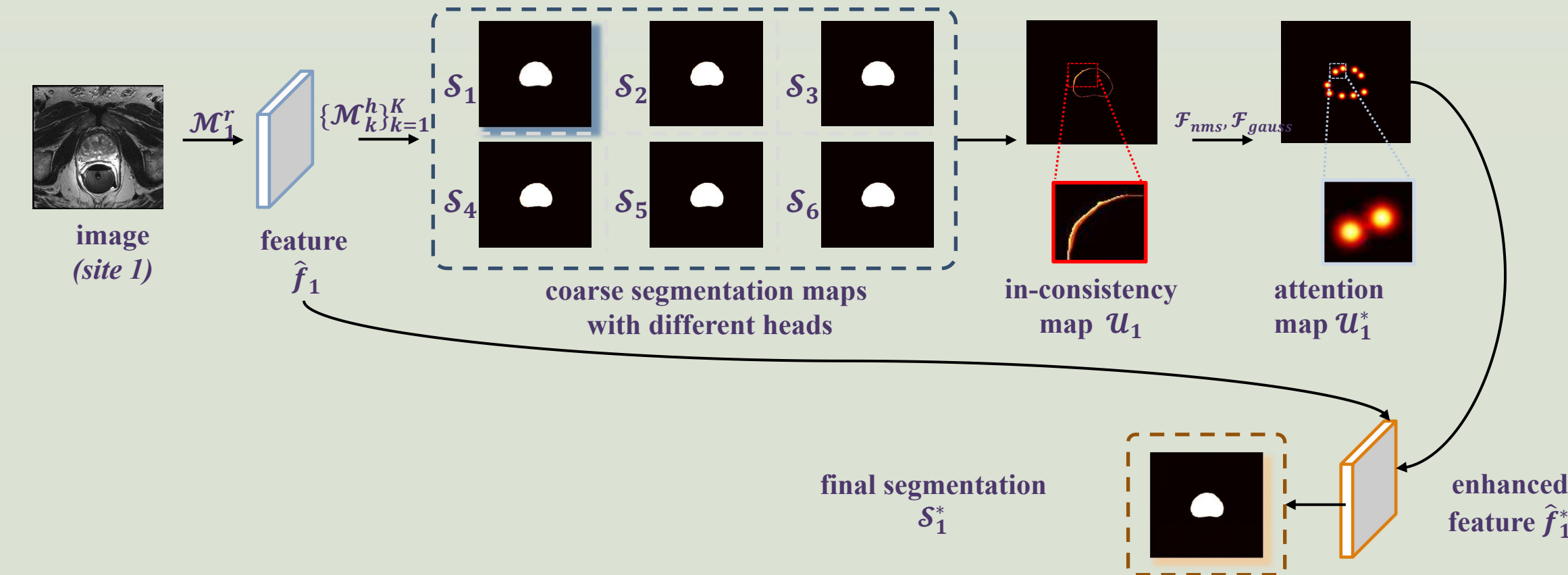


### Prediction: Pixel-level inconsistencies

a. Model the prediction-level inconsistencies



b. Calibrate the predictions



## EXPERIMENTS

**SOTA Comparison:** The comparison is finished on three public datasets.

Prostate segmentation in MR images collected from six sites.

Sites	A	B	C	D	E	F	Avg.	A	B	C	D	E	F	Avg.
Local Train	77.01	74.15	77.8	76.73	79.67	45.39	71.79	0.6	0.93	0.53	0.84	0.86	12.4	2.69
FedAVG	79.74	80.89	84.91	82.59	83.69	73.27	80.85	0.47	0.52	0.32	0.59	0.36	2.02	0.71
FT*	81.66	81.51	83.04	80.93	82.09	72.76	80.33	0.41	0.59	0.36	0.68	0.39	1	0.57
PRR-FL*	75.1	68.67	79.67	79.71	69.99	51.05	70.7	0.79	1.6	0.46	0.79	1.34	12.69	2.95
FedBN*	78.91	52.3	77.15	61.75	77.58	64.14	68.64	0.54	19.95	0.49	13.8	0.6	1.67	6.17
FedRep*	81.09	81.41	84.7	83.46	82.31	73.81	81.13	0.41	0.49	0.32	0.58	0.39	0.64	0.47
LC-Fed	85.91	82.27	86.28	85.31	86.08	79.47	84.22	0.35	0.48	0.25	0.49	0.32	0.75	0.44

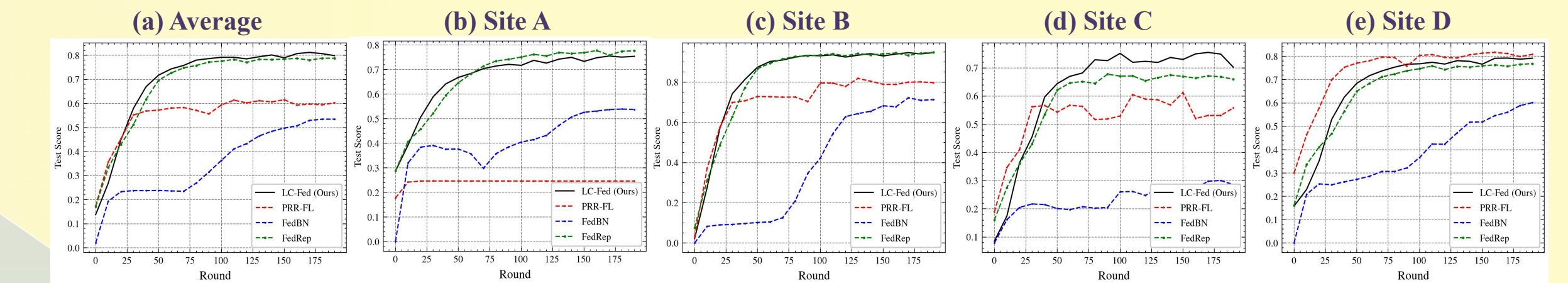
Polyp segmentation in Endoscopic images collected from four sites.

Sites	A	B	C	D	E	A	B	C	D	Avg.
Local Train	48.27	55.26	38.37	62.74	51.16	27.89	21.14	36.66	22.62	27.08
FedAVG	64.56	86.76	61.28	65.93	69.63	18.12	2.85	15.33	17.69	13.5
FT*	65.95	87.45	60.63	69.04	70.77	17.43	2.62	16.42	13.65	12.53
PRR-FL*	15.29	71.69	43.37	73.39	50.93	116.87	13.18	31.14	13.15	43.58
FedBN*	51.76	78.23	31.21	60.55	55.44	30.59	15.16	105.4	28.41	44.89
FedRep*	67.23	88.94	61.17	69.56	71.73	16.36	2.11	18.69	16.77	13.48
LC-Fed	69.21	88.51	68.1	76.68	75.63	15.59	2.64	11.6	12	10.46

Polyp segmentation in Endoscopic images collected from four sites.

Sites	A	B	C	D	E	A	B	C	D	Avg.
Local Train	82.8	78.55	84.8	85.58	82.93	5.88	5.12	3.72	2.76	4.37
FedAVG	84.81	77.88	83.91	84.51	82.77	5.19	5.68	3.99	3	4.46
FT*	85.84	80.21	84.58	85.2	83.96	4.62	4.56	3.82	2.85	3.96
PRR-FL*	81.24	78.49	83.75	83.19	81.67	7.57	5.09	4.1	3.34	5.02
FedBN*	84.7	78.01	85.01	85.13	83.21	5.05	5.32	3.65	2.84	4.22
FedRep*	85.33	79.81	83.95	83.53	83.15	4.84	4.93	3.95	3.14	4.21
LC-Fed	86.33	81.91	85.15	86.81	85.05	4.54	4.29	3.62	2.51	3.74

**Learning Curve:** The comparison is finished on three public datasets.



## CONCLUSIONS

- We propose to unify the feature- and prediction-level personalization by local calibration.
- The PCS module aims to calculate contrastive site embedding and couple it with channel selection operation.
- The HC module is designed to explore the inter-site inconsistency at prediction-level as a disagreement map.