The 43rd International Conference on Machine Learning (ICML)
July 6-11, 2026 | Seoul, South Korea
The Coex Convention & Exhibition Center

Decoupled Training with Local Reinforcement Fine-Tuning in Federated Learning

Yuting Ma[1], Lechao Cheng[2], Xiaohua Xu[1]
[1]University of Science and Technology of China
[2] Hefei University of Technology

ICML 2026 Main Conference Poster

Challenge & Motivation

Federated VLM

- Federated learning
- Pre-trained vision-language models
- Parameter-efficient fine-tuning techniques
- Heterogeneous data distributions

Inter-Client Optimization Inconsistency

- Representation-level client drift and misaligned local gradients
- Decrease global task adaptation in heterogeneous data settings

Intra-Client Over-Specialization

- Compromised generalization to unseen classes or shifted domains
- Reduce generalization ability in local fine-tuning

Few-Shot and Full-Data Regimes

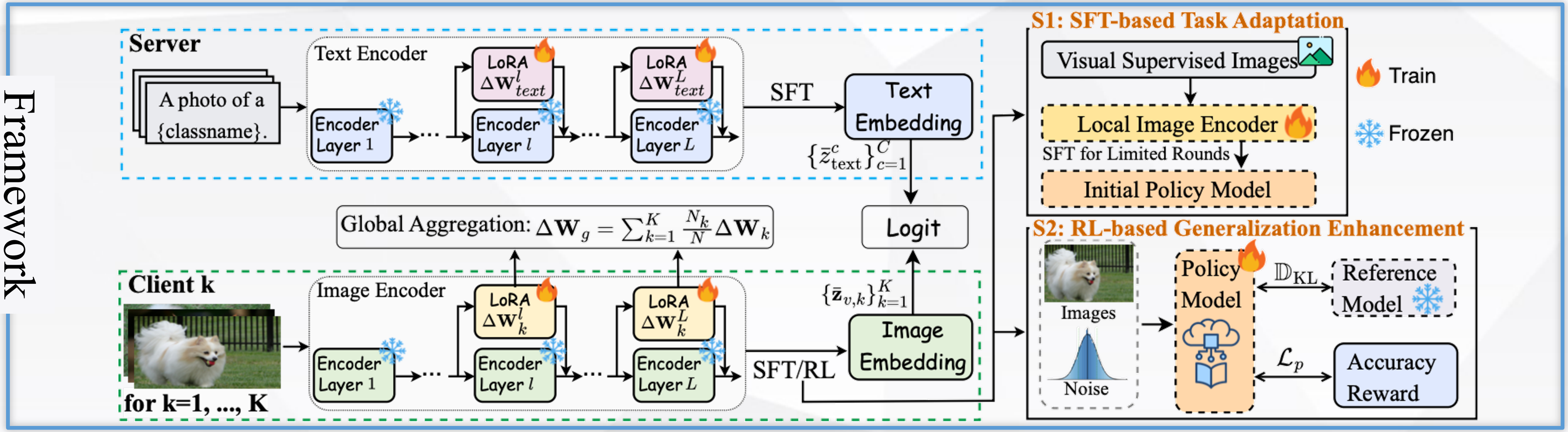
- Balance global task adaptation and generalization
- Diverse data settings in federated learning

We propose a novel federated VLM framework (FedDTL) to solve these challenges

Method

FedDTL

Framework



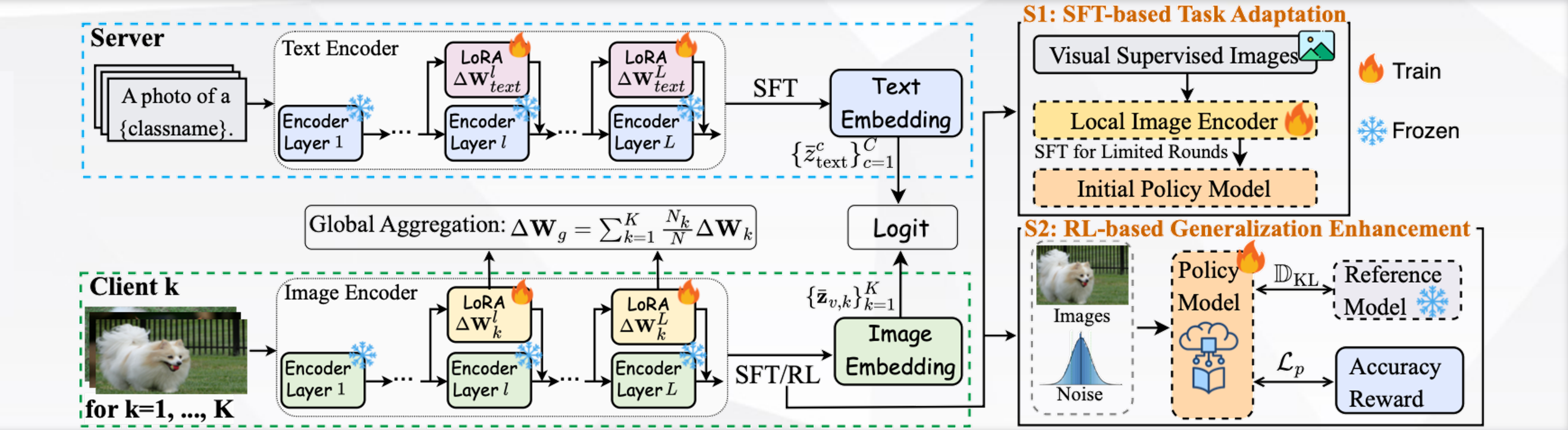
1 Decoupled Encoder Training

- ▶ Modality decoupling paradigm in CLIP
  - Client: privacy-preserving image encoding
  - Server: global text encoding
- ▶ Freeze pre-trained CLIP backbone
- ▶ Low-rank adaptation (LoRA) tuning
  - Optimize parameters of local image encoders
  - Optimize parameters of the global text encoder
- ▶ Embedding alignment between clients and server
  - Global text encoder as a semantic constraint
  - Align local image embeddings with text embeddings

Method

FedDTL

Framework



2 Two-Stage Local Fine-Tuning

- Stage 1: SFT-based task adaptation
  - A rapid end efficient optimization warm-start

$$\min_{\Delta W_k} \mathcal{L}_{ce}([W_0, \Delta W_k]; \{\bar{z}^c_{\text{text}}\}^C_{c=1}, \mathcal{D}_k)$$

- Stage 2: RL-based generalization enhancement
  - A novel GRPO-inspired RL stage with a controlled stochasticity mechanism

$$\min_{\Delta W_k} \mathcal{L}_{rl}([W_0, \Delta W_k]; \{\bar{z}^c_{\text{text}}\}^C_{c=1}, \mathcal{D}_k, \varepsilon)$$

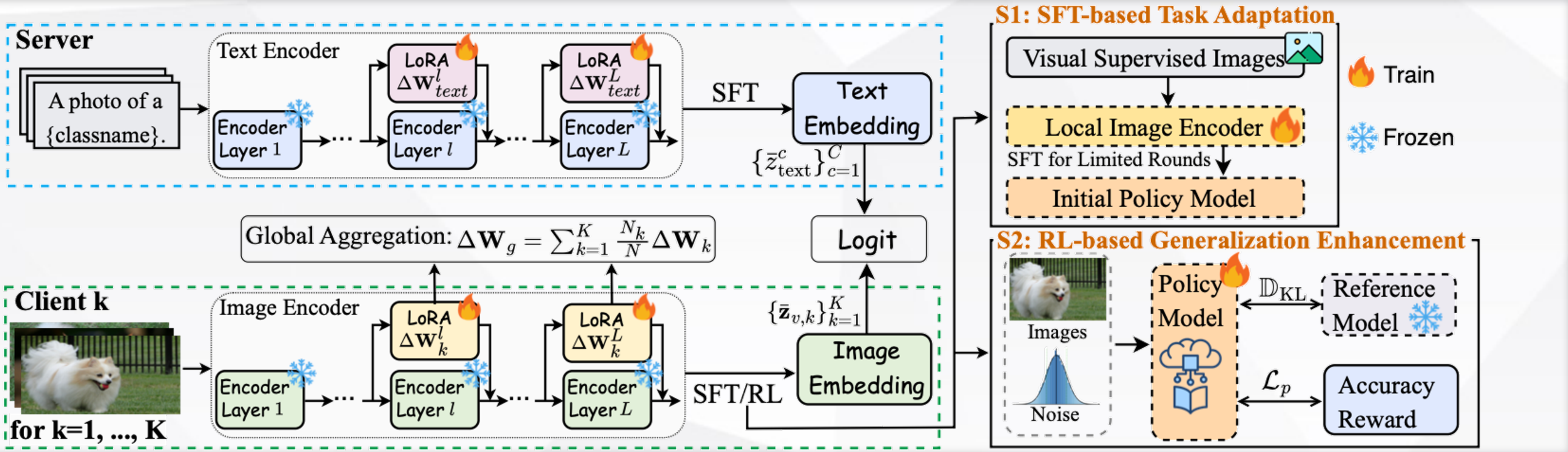
  - 1) sampling process: gaussian noise for action sampling

$$a_{i,j} \sim \pi_{\theta_{k,\text{old}}}(a_i | x_i, \varepsilon), \quad j = 1, 2, \cdots, G$$

Method

FedDTL

Framework



2 Two-Stage Local Fine-Tuning

- 2) reward estimation: accuracy-based reward

$$A_{i,j} = \frac{r_{i,j} - \text{mean}_j(r_{i,j})}{\text{std}_j(r_{i,j})}, \ j = 1, 2, \cdots, G.$$

- 3) policy update: a ϵ-clipping policy gradient term and a KL regularization term

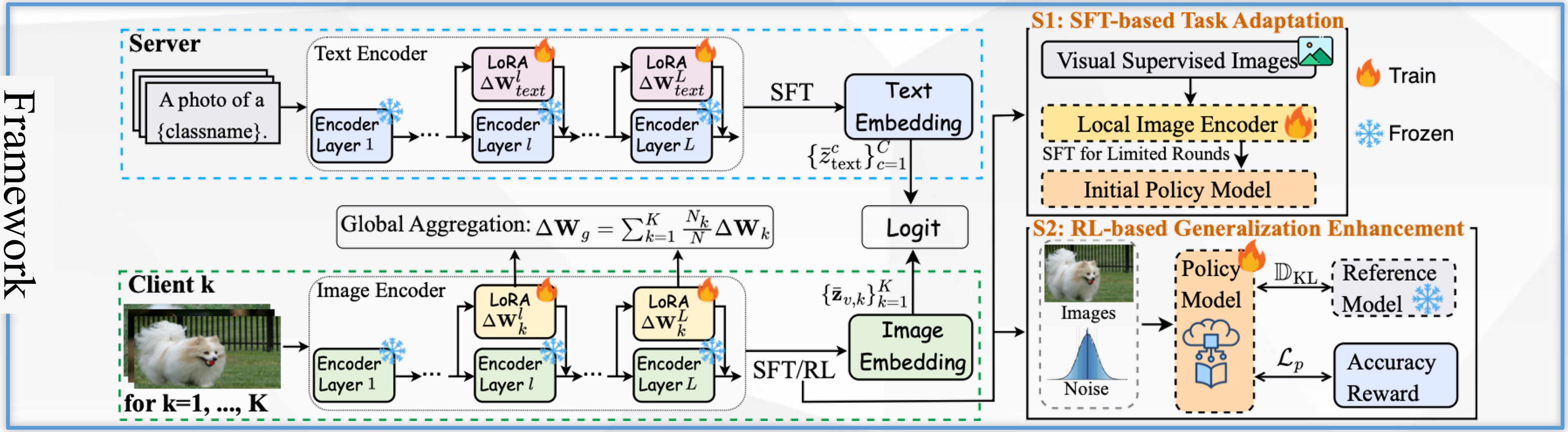
$$\mathcal{L}_{rl} = -\frac{1}{G}\sum_{j=1}^{G}\frac{1}{bs}\sum_{i=1}^{bs}(\mathcal{L}_p - \beta \mathbb{D}_{\text{KL}})$$

$$\mathcal{L}_p = \min\left[\rho_{i,j} A_{i,j}, \ \text{clip}(\rho_{i,j}, 1-\epsilon, 1+\epsilon)\, A_{i,j}\right]$$

$$\mathbb{D}_{\text{KL}} = \frac{\pi_{k,\text{ref}}(a_{i,j}|x_i)}{\pi_{\theta_k}(a_{i,j}|x_i)} - \log \frac{\pi_{k,\text{ref}}(a_{i,j}|x_i)}{\pi_{\theta_k}(a_{i,j}|x_i)} - 1$$

Method

FedDTL



3 Global Coordination and Optimization

- Parameter aggregation
  - Clients upload visual LoRA parameters and image embeddings
  - Weighted averaging $\Delta \mathbf{W}_g = \sum_{k=1}^{K} \frac{N_k}{N} \Delta \mathbf{W}_k$
- Global optimization
  - Perform global textual training at the server to promote consistent knowledge learning

$$\min_{\Delta \mathbf{W}_{\text{text}}} \mathcal{L}_{ce}([\mathbf{W}_{0,\text{text}}, \Delta \mathbf{W}_{\text{text}}]; \{\bar{\mathbf{z}}_{v,k}\}_{k=1}^{K}, text_c)$$

  - Server broadcasts the global text embeddings of all classes and global visual LoRA parameters

Experiments

Performance Exploration

Label Skew

Few-shot	Non-IID			Dirichlet(0.1)			Dirichlet(0.3)			Dirichlet(0.5)			IID		
	Local	Base	Novel	Local	Base	Novel	Local	Base	Novel	Local	Base	Novel	Local	Base	Novel
CLIP	80.22	79.62	81.99	79.44	79.76	82.12	79.83	79.76	82.12	79.77	79.76	82.12	79.76	79.76	82.12
pFedDC	96.28	37.55	54.58	87.36	69.96	69.54	85.84	80.34	74.97	88.02	85.77	75.52	89.66	89.66	76.70
FedPGP	82.17	80.66	78.87	81.43	79.75	78.48	83.35	83.02	79.72	84.40	84.31	80.42	85.09	85.09	79.04
pFedMMA	**97.23**	56.55	72.47	87.13	72.77	74.91	86.33	81.45	76.69	88.52	86.39	76.04	90.68	90.68	76.25
PromptFL	82.69	82.23	79.84	82.02	80.41	76.09	83.79	83.40	80.06	85.85	85.61	79.76	86.24	86.24	80.52
FedMaPLe	83.89	83.63	77.56	85.70	84.05	77.69	88.47	88.18	80.40	90.40	90.18	80.01	90.85	90.85	81.65
FedDTL	89.76	**89.58**	**83.01**	**91.66**	**90.95**	**82.64**	**92.18**	**91.94**	**83.61**	**92.17**	**92.02**	**83.25**	**92.58**	**92.58**	**82.53**

Full-data	Non-IID			Dirichlet(0.1)			Dirichlet(0.3)			Dirichlet(0.5)			IID		
	Local	Base	Novel	Local	Base	Novel	Local	Base	Novel	Local	Base	Novel	Local	Base	Novel
CLIP	80.22	79.62	**81.99**	79.44	79.76	**82.12**	79.83	79.76	**82.12**	79.77	79.76	**82.12**	79.76	79.76	**82.12**
pFedDC	97.44	28.75	40.16	80.12	60.08	63.55	83.37	77.51	74.37	86.25	83.87	77.46	91.51	91.51	79.82
FedPGP	83.74	48.26	49.02	74.62	65.23	57.61	78.32	76.38	66.17	82.77	82.04	66.37	87.76	87.76	78.10
pFedMMA	**98.29**	31.54	45.26	84.07	62.56	65.56	84.93	78.34	71.42	87.31	84.55	73.32	92.83	92.83	73.67
PromptFL	62.92	62.98	48.48	77.52	75.82	57.98	81.62	81.00	63.06	85.32	84.93	66.48	87.29	87.29	77.39
FedMaPLe	79.77	80.56	69.41	90.49	89.27	70.10	92.28	91.70	74.39	92.87	92.57	76.67	93.97	93.97	78.77
FedDTL	91.52	**91.64**	77.72	**92.69**	**92.40**	76.59	**93.05**	**92.93**	79.87	**93.70**	**93.61**	79.23	**94.07**	**94.07**	79.72

Feature Shift

Method	Office-Caltech10				DomainNet			
	Few-one	Few-Dir(0.1)	Full-one	Full-Dir(0.1)	Few-one	Few-Dir(0.1)	Full-one	Full-Dir(0.1)
CLIP	92.88	92.88	92.88	92.88	87.68	87.68	87.68	87.68
pFedDC	96.54	72.01	97.40	70.71	85.69	69.31	87.11	59.87
FedPGP	97.89	97.75	98.64	96.12	88.98	89.01	89.08	87.70
pFedMMA	98.22	82.25	98.23	78.56	88.94	78.67	86.51	67.24
PromptFL	98.24	97.03	98.49	97.87	88.88	88.79	89.77	86.99
FedMaPLe	**98.44**	98.07	98.02	96.55	88.99	88.81	91.94	90.51
FedDTL	98.33	**98.20**	**98.65**	**98.65**	**90.88**	**91.06**	**93.38**	**93.47**

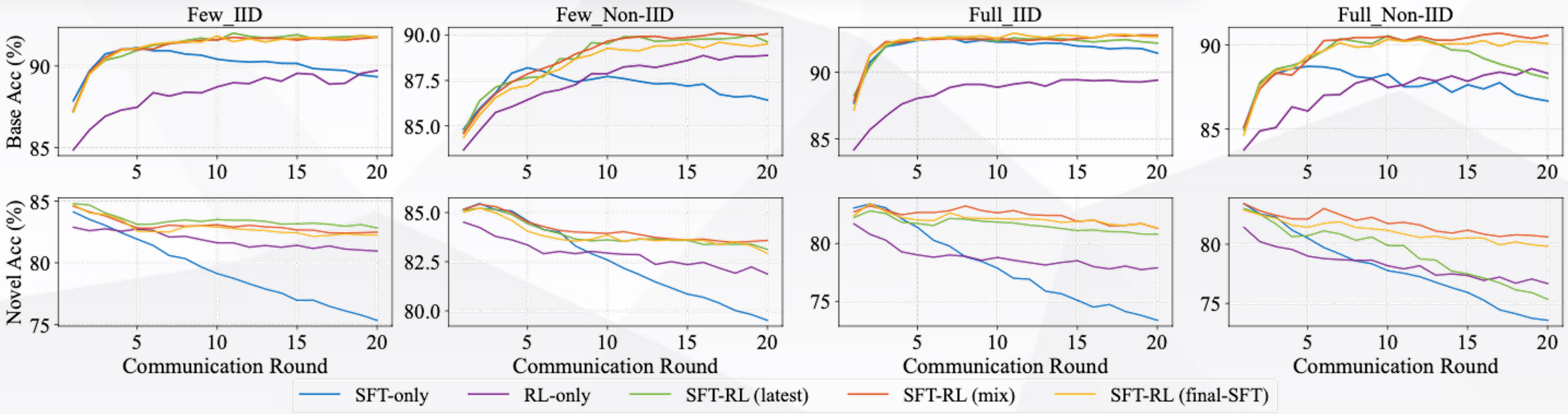
Experiments

Ablation Study

Core Module

Core Module	Few_IID Base	Few_IID Novel	Few_IID HM	Few_Non-IID Base	Few_Non-IID Novel	Few_Non-IID HM	Full_IID Base	Full_IID Novel	Full_IID HM	Full_Non-IID Base	Full_Non-IID Novel	Full_Non-IID HM
FedLoRA	89.27	77.44	82.68	78.32	78.86	78.56	91.75	75.80	82.58	58.11	70.51	63.12
+ Decoupled Encoder Training	89.34	75.35	81.33	86.42	79.52	82.60	91.42	73.36	80.93	86.68	73.57	79.20
+ Two-Stage Local Fine-Tuning	90.48	**83.11**	86.46	79.46	**83.84**	81.47	91.71	**81.86**	86.25	47.91	76.43	57.86
FedDTL	**91.76**	82.48	**86.59**	**90.06**	83.58	**86.51**	**92.75**	81.29	**86.35**	**90.58**	**80.62**	**85.03**

Local fine-tuning strategy and choice of reference model



THANKS!

E-mail: ytma@mail.ustc.edu.cn