Supplementary Material: Federated Incremental Semantic Segmentation

This supplementary material introduces detailed optimization procedure of our proposed FBL model (Section A) and qualitative ablation studies (Section B). More importantly, the source code of this paper is publicly available at https://github.com/JiahuaDong/FISS.

A. Optimization Procedure

The optimization pipeline of our FBL model to address the FISS problem is presented in **Algorithm 1**. Starting from the first segmentation task, all local clients employ Eq. (11) to calculate the average entropy $\mathcal{I}_l^{r,t}$ of local training data $\mathcal{T}_l^t$ at the beginning of each global round, and then some of local clients are randomly selected by global server $\mathcal{S}_g$ to perform local training for each global round. After these selected local clients utilize task transition monitor to accurately recognize new classes, they automatically store the global model learned at the last global round as the old model Θ^{t-1} to generate confident pseudo labels for old classes via Eq. (2), and optimize local model $\Theta_l^{r,t}$ via $\mathcal{L}_{\mathrm{obj}}$ in Eq. (10) at the r-th global round. Finally, the updated local models $\Theta_l^{r,t}$ of selected local clients are aggregated as global model $\Theta^{r+1,t}$ by global server $\mathcal{S}_g$, and $\Theta^{r+1,t}$ will be distributed to local clients for the next round training.

Algorithm 1: Optimization of The FBL Model.

Input: In the t-th ($t \geq 2$) task, global server $\mathcal{S}_g$ randomly select w local clients $\{\mathcal{S}_{l_1}, \mathcal{S}_{l_2}, \cdots, \mathcal{S}_{l_w}\}$ with their local datasets as $\{\mathcal{T}_{l_1}^t, \mathcal{T}_{l_2}^t, \cdots, \mathcal{T}_{l_w}^t\}$ at the r-th global round; The global server $\mathcal{S}_g$ transmits the latest global model $\Theta^{r,t}$ to selected local clients;

All Local Clients:

for $\mathcal{S}_l$ *in* $\{\mathcal{S}_1, \mathcal{S}_2, \cdots, \mathcal{S}_L\}$ **do**
- Calculate averaged entropy $\mathcal{I}_l^{r,t}$ of local training data $\mathcal{T}_l^t$ via Eq. (11);

Selected Local Clients:

Obtain $\Theta^{r,t}$ from $\mathcal{S}_g$ as the local segmentation model $\Theta_l^{r,t}$;

for $\mathcal{S}_l$ *in* $\{\mathcal{S}_{l_1}, \mathcal{S}_{l_2}, \cdots, \mathcal{S}_{l_w}\}$ **do**
- *Task* = False;
- **if** $\mathcal{I}_l^{r,t} - \mathcal{I}_l^{r-1,t} \geq \tau$ **then**
  - *Task* = True;
- **if** *Task* = True **then**
  - Store the latest global model $\Theta^{r,t}$ as old model Θ^{t-1} for local client $\mathcal{S}_l$;
- **for** $\{\mathbf{x}_{li}^t, \mathbf{y}_{li}^t\}_{i=1}^B$ *in* $\mathcal{T}_l^t$ **do**
  - Generate confident pseduo labels via Eq. (2);
  - Update local model $\Theta_l^{r,t}$ via Eq. (10);

Global Server:

$\mathcal{S}_g$ aggregates the parameters of all local models $\Theta_l^{r,t}$ as $\Theta^{r+1,t}$ for the training of next global round.

B. Ablation Studies

In this subsection, we present qualitative ablation studies to verify the effectiveness and superiority of our proposed modules. As shown in Table 1, when removing one of the designed modules, the performance in terms of mIoU heavily degrades about $3.9\% \sim 11.6\%$. Specifically, when compared with Ours, Ours-w/oAPL decreases $3.9\% \sim 9.0\%$ mIoU, which validates the effectiveness of the proposed adaptive class-balanced pseudo labeling to mine confident pseudo labels of old classes. These pseudo labels provide strong guidance for two forgetting-balanced losses to address intra-client heterogeneous forgetting on old classes. Moreover, Ours significantly outperforms Ours-w/oFSC by a large margin of $3.9\% \sim 6.5\%$ mIoU. This significant performance improvement verifies that our FBL model could effectively tackle forgetting heterogeneity of different old classes within each local client via the forgetting-balanced semantic compensation loss. In addition, Ours-w/oFRC degrades the segmentation performance of $7.0\% \sim 11.6\%$ mIoU, compared with Ours. This phenomenon illustrates the effectiveness and superiority of the proposed forgetting-balanced relation consistency loss to compensate heterogeneous relation distillation gains. More importantly, the performance degradation illustrates that all designed modules are effective to collaboratively learn a global incremental segmentation model under the practical FISS settings.

Table 1. Ablation studies on Pascal-VOC 2012 dataset [12] under the 4-4 and 8-2 settings with overlapped foregrounds.

Settings	Variants			VOC Overlapped 4-4 [12]						VOC Overlapped 8-2 [12]							
	APL	FSC	FRC	t=1 (Base)	t=2	t=3	t=4	t=5	Imp.	t=1 (Base)	t=2	t=3	t=4	t=5	t=6	t=7	Imp.
Our-w/oAPL	✗	✓	✓	70.4	**67.6**	**53.3**	**49.8**	**40.0**	⇑ 3.9	80.4	**65.1**	54.5	39.5	**41.5**	32.9	26.7	⇑ 9.0
Our-w/oFSC	✓	✗	✓	70.4	**67.3**	52.2	45.8	**40.0**	⇑ 3.9	80.4	**65.8**	55.0	40.8	40.6	32.5	**29.2**	⇑ 6.5
Our-w/oFRC	✓	✓	✗	70.4	61.4	43.2	41.4	32.3	⇑ 11.6	80.4	**65.1**	**57.9**	**43.5**	41.0	**33.8**	28.7	⇑ 7.0
FBL (Ours)	✓	✓	✓	70.4	66.6	**53.6**	**49.6**	**43.9**	–	80.4	65.0	**58.1**	**47.3**	**45.8**	**39.4**	**35.7**	–

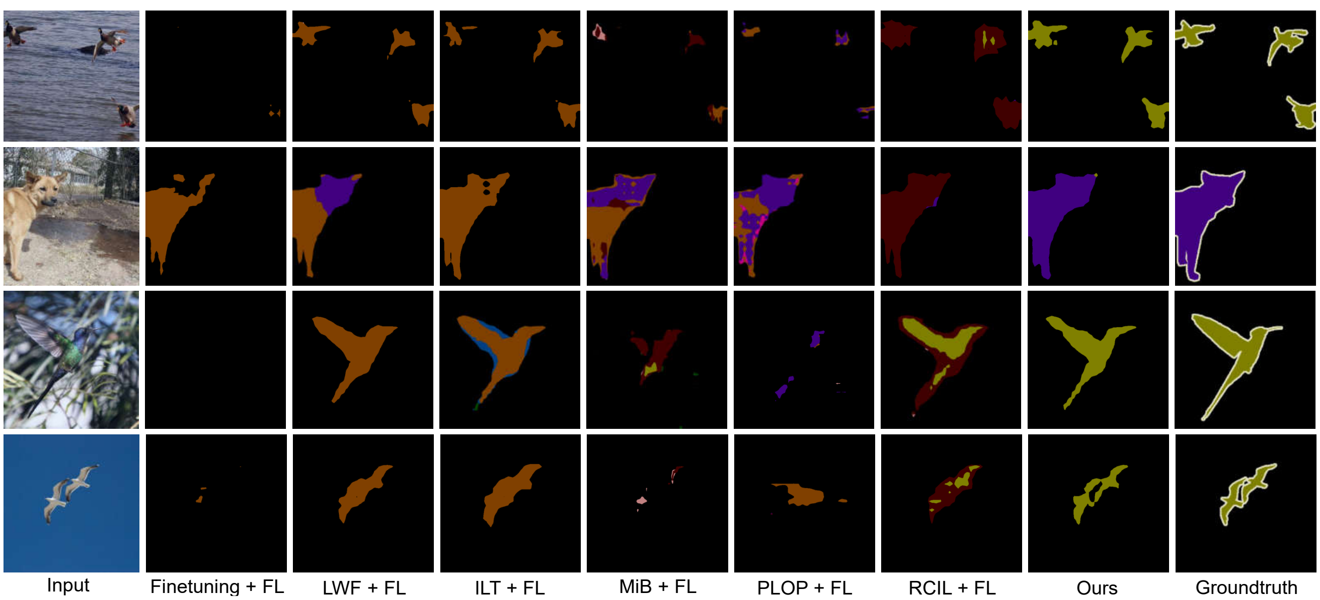

Figure 1. Visualization of some qualitative comparison results on Pascal-VOC 2012 [12] under the overlapped 4-4 setting of the FISS.