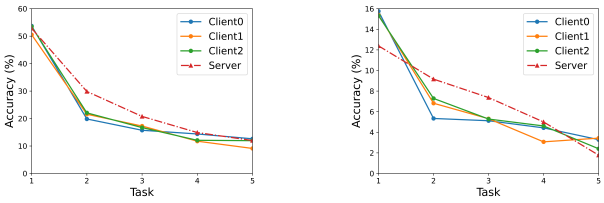
FedProK: Trustworthy Federated Class-Incremental Learning via Prototypical Feature Knowledge Transfer

Supplementary Material

Preliminary Experiments

To further investigate the issues of asynchronous FCIL, we conduct two preliminary experiments and disclose the averaged results of three repeated experiments in Fig. 1. In the first preliminary experiment, the setting is synchronous FCIL, where the task sequence of three clients is identical, to investigate the local catastrophic forgetting of previous knowledge when the asynchronous issue is not involved. In the second experiment, we set the task sequences of three clients differently to investigate the challenges of asynchronous FCIL and figure out whether asynchronous issue exacerbates catastrophic forgetting of previous knowledge. In both settings, we adopt FedAvg as the FL protocol without utilizing any CL techniques. Specifically, the dataset is CIFAR-100 and the backbone is ResNet-18. The incremental state occurs every five communication rounds simultaneously on three clients. We can find that: (1) In both synchronous and asynchronous settings, the local models suffer severe catastrophic forgetting of previous knowledge with the arrival of new tasks, and thus cause significant performance degradation of the global model, which is clearly illustrated by the sudden decline of accuracy in every incremental state. (2) The asynchronous issue has a significant impact on performance, even in the initial incremental state. The global model accuracy, with quantity-based label imbalance among clients Fig. 1b (asynchronous FCIL), is much lower compared to that with distribution-based label imbalance among clients Fig. 1a (synchronous FCIL).



(a) Synchronous FCIL. (b) Asynchronous FCIL.

Figure 1. Catastrophic forgetting of previous tasks when the task sequences are synchronous (Fig. 1a) and asynchronous (Fig. 1b) among clients.

Implementation Details

In this paper, we utilized two RGB-colored visual classification datasets, i.e. CIFAR-10 and CIFAR-100 to validate the effectiveness of our method, as shown in Tab. 1.

Dataset	Train	Test	μ(train)	σ(train)	Size
CIFAR-10	50000	10000	500	0.0	32×32
CIFAR-100	50000	10000	500	0.0	32×32

Table 1. Summary of datasets. μ is the mean number of train images per class and σ is the standard deviation

And we adopt two types of data partitions to simulate the dynamic task streams and data heterogeneity in real-world scenarios: synchronous and asynchronous FCIL. In the synchronous FCIL, clients share a common task order but suffer a heterogeneity of Dirichlet Distribution. To better illustrate our setting, we show the settings of synchronous FCIL on CIFAR-100 with 5 incremental tasks in Tab. 2 and asynchronous FCIL on CIFAR-100 with 5 incremental tasks in Tab. 3. We conduct experiments using random seeds

Incre State	Class Index
Task 0	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19
Task 1	20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39
Task 2	40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59
Task 3	60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79
Task 4	80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99

Table 2. The data partition in the synchronous FCIL. In this setting, clients share the same task order and each task has the same class indices. Data heterogeneity is under Dirichlet Distribution.

$\{42, 1999, 2024\}$ and compute the average to improve the robustness of our results. The learning rate is 0.001.

Compared methods include static FL methods (FedAvg, FedProx and FedNova), hybrid methods (FedEWC and FediCaRL) and FCIL methods (GLFC, FedSpace and TARGET). **FedAvg** is the vanilla algorithm of FL without any rectification to handle catastrophic forgetting and data heterogeneity. **FedProx** is a classical FL method dealing with data heterogeneity, which adds a proximal term in the loss function to effectively limit the impact of variable local updates. **FedNova** (Federated Normalized Averaging) is a further modification of FedAvg, it normalizes the local updates before aggregation to weaken the deviation of local optimization. **FedEWC** is the combination of FedAvg and EWC(Elastic Weight Consolidation). EWC is a typical regularization-based CL method, which alleviates forgetting by adding a regularization term in the loss function to penalize the changes in the model's parameters that are important to the previous tasks. **FediCaRL** is the combination of FedAvg and iCaRL (Incremental Classifier and Rep-

Incre State	Client	Class Index
Task 0	Client 1	$59, 2, 7, 27, 91, 64, 29, 88, 0, 54$
	Client 2	$17, 28, 42, 12, 78, 70, 97, 23, 3, 54$
	Client 3	$57, 11, 24, 71, 28, 20, 86, 38, 27, 31$
Task 1	Client 1	$39, 86, 80, 4, 35, 41, 77, 36, 22, 14$
	Client 2	$66, 99, 29, 35, 13, 85, 63, 77, 15, 75$
	Client 3	$69, 45, 58, 13, 3, 53, 51, 72, 81, 82$
Task 2	Client 1	$97, 69, 40, 56, 3, 11, 8, 95, 73, 68$
	Client 2	$62, 27, 84, 64, 32, 71, 87, 69, 48, 86$
	Client 3	$76, 46, 55, 75, 4, 19, 37, 92, 9, 54, 1$
Task 3	Client 1	$74, 94, 28, 75, 89, 17, 50, 31, 65, 84$
	Client 2	$31, 11, 88, 14, 79, 49, 18, 6, 21, 44$
	Client 3	$1, 61, 60, 83, 14, 17, 5, 94, 35, 77$
Task 4	Client 1	$90, 47, 71, 30, 33, 25, 13, 81, 67, 26$
	Client 2	$94, 52, 81, 25, 96, 89, 10, 16, 4, 93$
	Client 3	$98, 29, 25, 34, 97, 89, 88, 43, 84, 64$

Table 3. The data partition in the asynchronous FCIL with $\gamma = 0.5$. In this setting, clients have different local tasks.

resentation Learning). iCaRL maintains a memory buffer to restore the exemplars of previous tasks and replay them when training on new tasks, which shows excellent performance even in distributed settings but breaks the underlying privacy protocol of FL.

GLFC (Global Local Forgetting Compensation) is an FCIL method that addresses global forgetting caused by data heterogeneity and local forgetting caused by dynamic task streams respectively. To address local forgetting, it utilizes a class-aware gradient compensation loss and a class-semantic relation distillation loss to maintain previous knowledge and distill consistent inter-class relations across tasks. To address global forgetting, it implements a proxy server to select the optimal previous global model, assisting the relation distillation on the client side. **FedSpace** (Federated learning System with Prototype Aggregation for Continual rEpresentation) is an asynchronous FCIL model, which can also be employed in synchronous FCIL. It addresses catastrophic forgetting and data heterogeneity with prototype-based learning, a representation loss, fractal pretraining, and a modified aggregation policy. **TARGET** generates pseudo samples of the previous classes on the server according to the global data distribution and sends them to the clients to address FCIL issues.

Stability-Plasticity Trade-Off

The stability-plasticity dilemma is a challenge in CL, where a model needs to adapt to new information while retaining knowledge from previously learned tasks. This dilemma exists in FCIL as well. Therefore, we need to strike a balance between the two points to enhance the trustworthiness of FCIL models. In this paper, we measured stability with the test accuracy of the global model on current tasks $\mathcal{A}_G^{t-1}$, and plasticity with the test accuracy of the global model on previous tasks $\mathcal{A}_G^t$, and combined them in Eq. (1) to strike a balance.

$$U = \lambda \times \mathcal{A}_G^{t-1} + (1 - \lambda) \times \mathcal{A}_G^t, \tag{1}$$

where λ is the hyperparameter. A larger λ indicates a higher emphasis on stability and a lower emphasis on plasticity. The results in **??** clearly demonstrate that our method FedProK achieves state-of-the-art continual utility under different values of λ.

The concepts of stability and plasticity can also be measured using other metrics. However, existing works typically utilize two distinct metrics, making it challenging to unify them into a single formula.

Future Work

Further research efforts will be dedicated to empirically analyzing the effectiveness of privacy preservation against other forms of attacks. Besides, we also wish to explore a more comprehensive trustworthy FCIL framework with regard to interpretability, robustness and fairness within the context of asynchronous FCIL.