Navigating Data Heterogeneity in Federated Learning: A Semi-Supervised Approach for Object Detection

Taehyeon Kim[1], Eric Lin[2], Junu Lee[3], Christian Lau[3], Vaikkunth Mugunthan[3]

KAIST AI[1], Seoul, South Korea

DynamoFL[2], San Fransisco, US

The Wharton School[3], Pennsylvania, US

Motivation

- Federated Learning (FL)'s reliance on fully labeled data limits its effectiveness and raises privacy concerns, as it often requires transferring data to central servers for labeling.
- In autonomous driving, a novel approach is required to bridge the knowledge gap between labeled and unlabeled data without the need for direct data exchange.

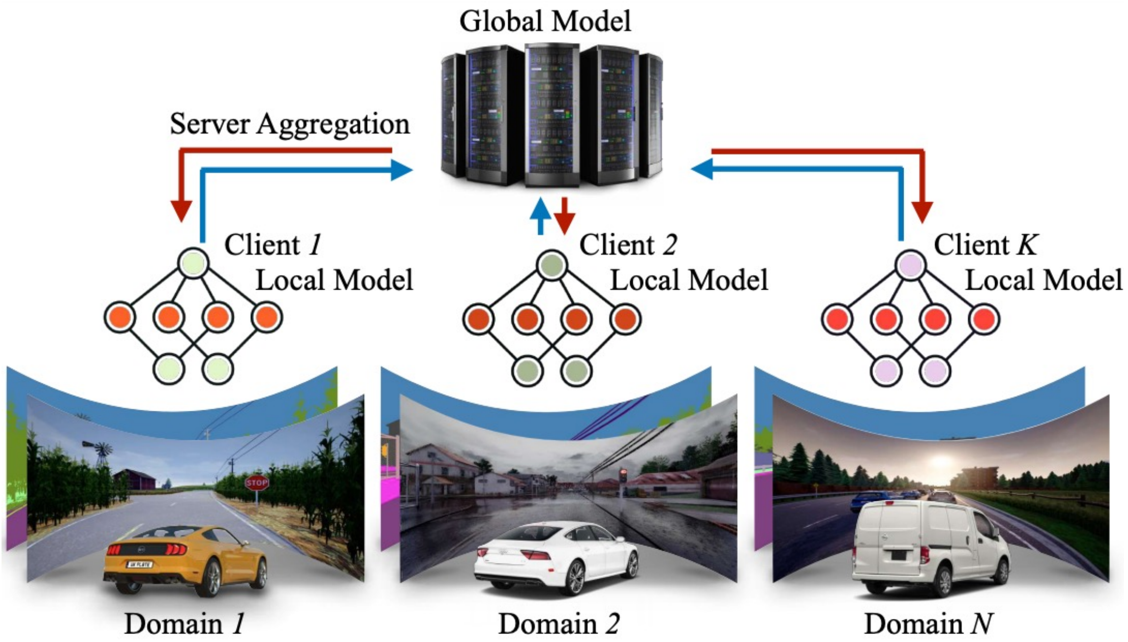
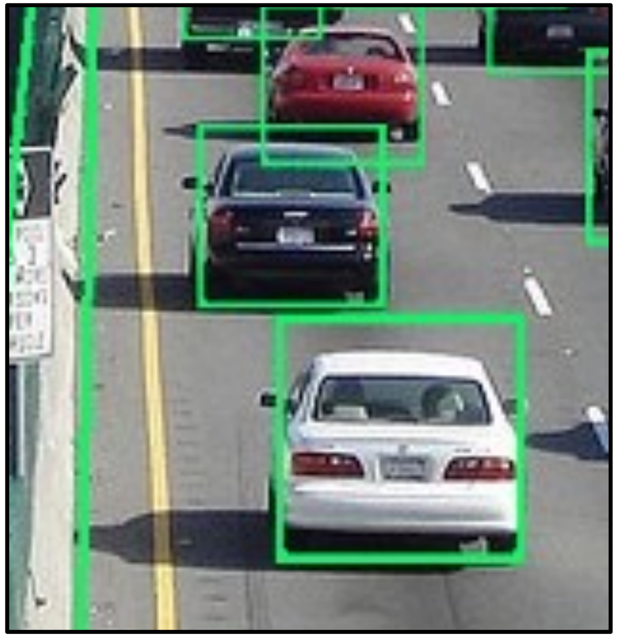
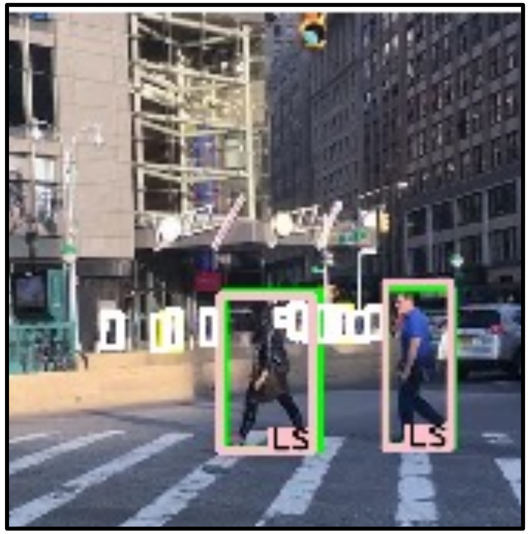


Image Source: https://feddrive.github.io/

Different Annotations from General Image Classification

Federated object detection framework struggles with the complexity of YOLO-like annotations, where object detection requires detailed information like object ID, bounding box coordinates, and confidence scores, beyond simple image category labels.





Challenges: Data Heterogeneity

Federated object detection faces big challenges due to unique dataset heterogeneity from (1) weather-induced feature skew, (2) class distribution imbalance, and (3) label density variability, each impacting model performance and training efficiency in different ways.

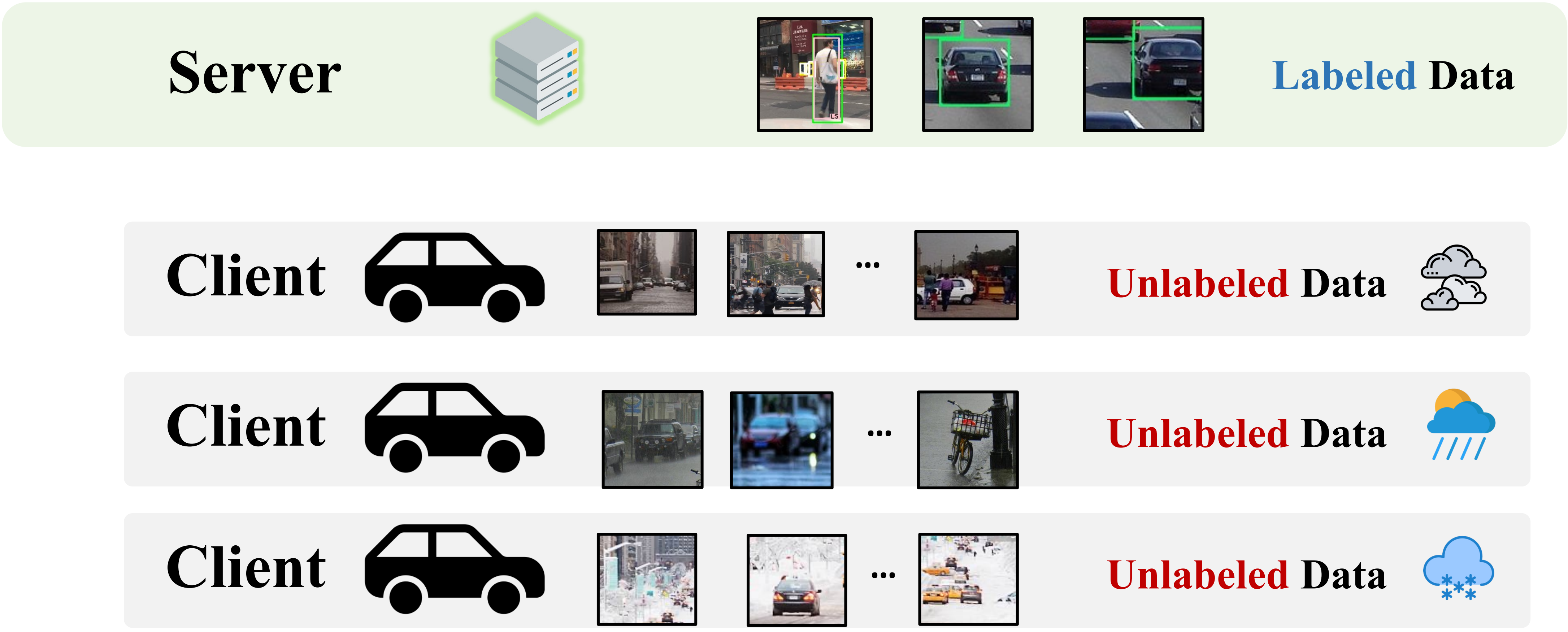
Overcast

Rainy

Snowy

Semi-Supervised Federated Object Detection (SSFOD)

We present a **novel** SSFOD framework, designed for scenarios where labeled data reside only at the server while clients possess unlabeled data.



Method: FedSTO

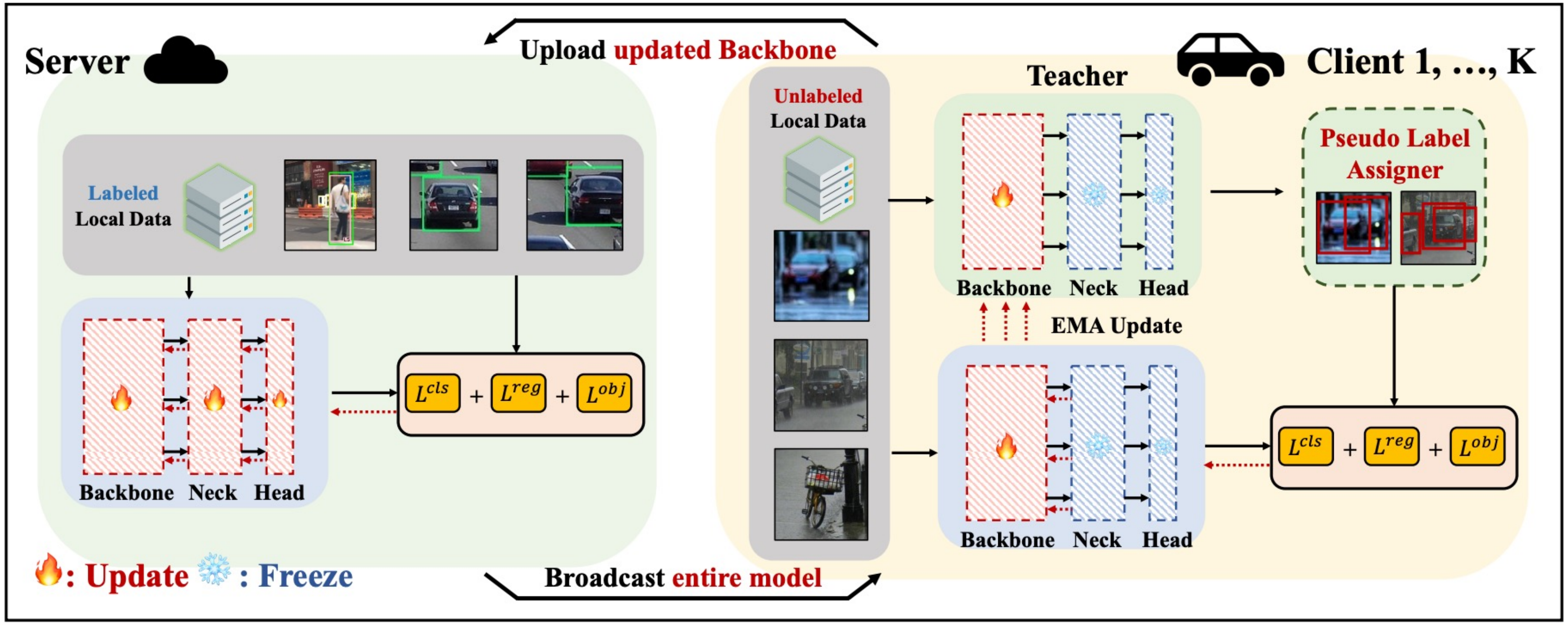
FedSTO is a two-stage method. It begins with a warmup stage (i.e., training only with labeled dataset on server) focused on (1) robust pretraining using labeled server data, followed by (2) a full parameter training phase that further enhances the model's capabilities, thereby improving the performance close to the centralized supervised one.

Training only with labeled dataset on server

Method	Non-IID					IID				
	Cloudy	Overcast	Rainy	Snowy	Total	Cloudy	Overcast	Rainy	Snowy	Total
Partially Supervised	0.540	0.545	0.484	0.474	0.511	0.528	0.545	0.533	0.510	0.529
+ SSFL [5] with Local EMA Model	0.560	0.566	0.553	0.553	0.558	0.572	0.588	0.593	**0.610**	0.591
+ Selective Training	0.571	0.583	0.557	0.556	0.567	0.576	0.578	0.594	0.599	0.587
+ FPT with Orthogonal Enhancement [16]	**0.596**	**0.607**	**0.590**	**0.580**	**0.593**	**0.591**	**0.634**	**0.614**	0.595	**0.609**

Method: FedSTO – Stage1

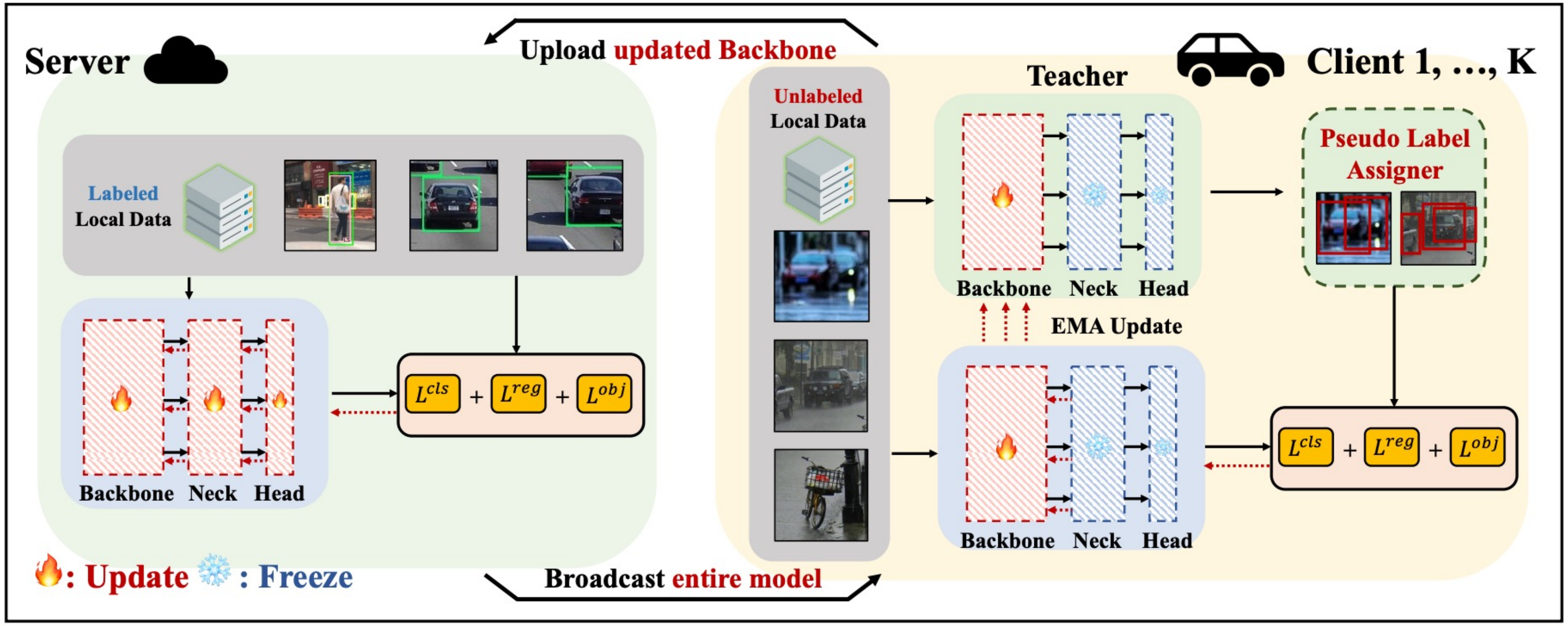
Selective Training (ST) involves three steps: (1) initial training (i.e., warm-up) with a labeled dataset, (2) client-side training on unlabeled data while only updating the backbone, and (3) server-side aggregation of backbone parameters to synthesize information from diverse datasets, with this cycle repeating until performance convergence.



1. Representation Learning with Selective Training

Method: FedSTO – Stage1

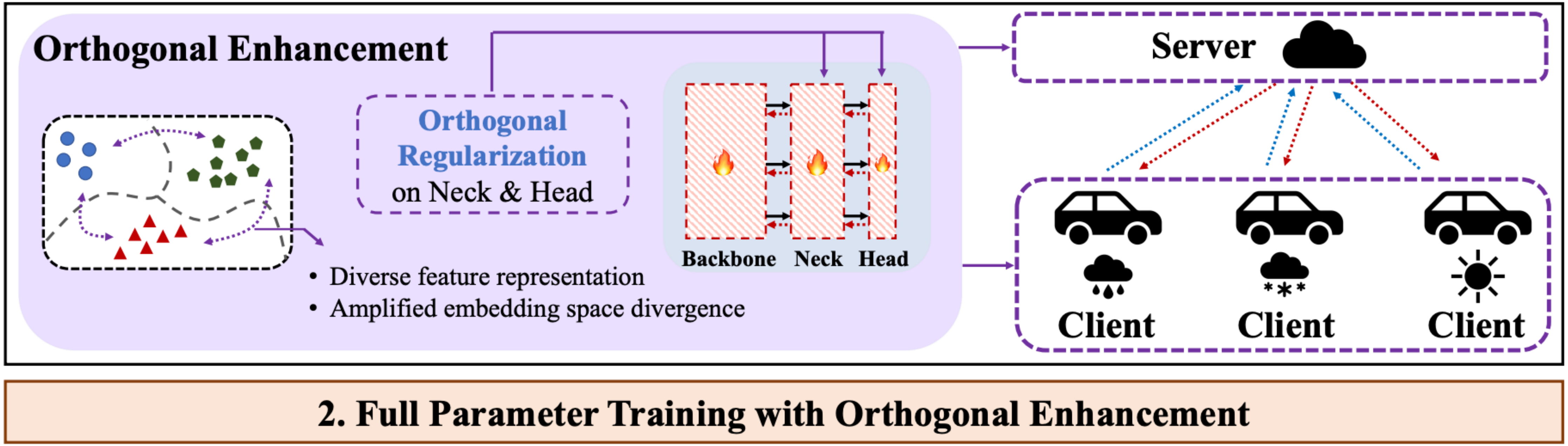
Our approach features a local pseudo labeler which is an Exponential Moving Average (EMA) model, specifically designed to adapt and respond effectively to dynamic local data conditions.



1. Representation Learning with Selective Training

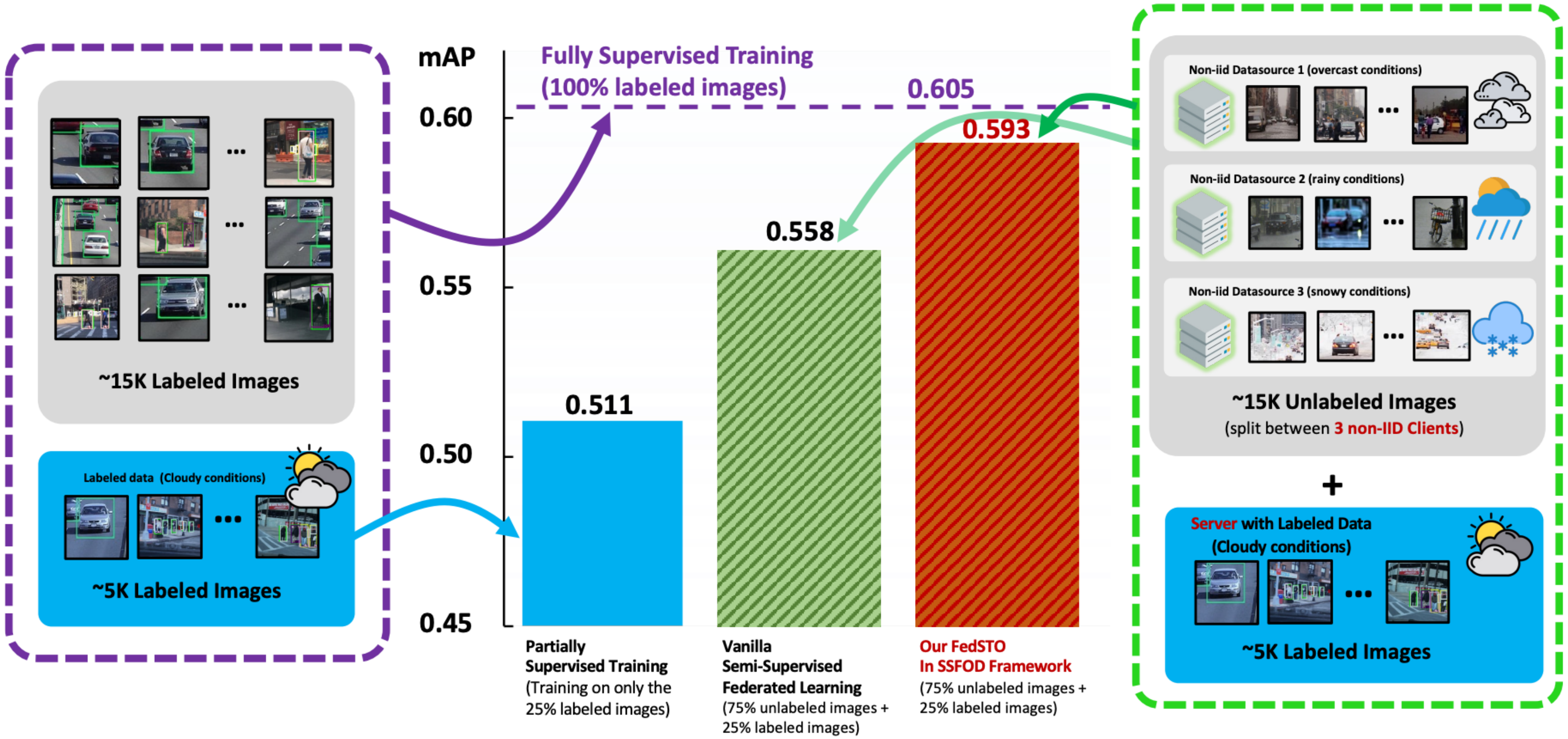
Method: FedSTO – Stage2

In next stage, FedSTO improves model robustness across domains by incorporating orthogonality regularization. This approach, focusing on non-backbone parts like the neck and head, balances training, reduces biases related to weather and object categories, and consequently enhances overall performance.



Experiments

FedSTO outperforms partially supervised and SSFL baselines in object detection by achieving 0.082 and 0.035 higher mAP@0.5 respectively, and nearly matches the performance of fully supervised models with just a 0.012 gap, using only 25% labeled data on BDD100K dataset.



Experiments

FedSTO showcases superior performance over other techniques in both IID and Non-IID weather conditions on the BDD100K dataset, even demonstrating competitive centralized results against fully supervised approaches and SSL methods like EMA Teacher.

Type	Algorithm	Method	Non-IID					IID				
			Cloudy	Overcast	Rainy	Snowy	Total	Cloudy	Overcast	Rainy	Snowy	Total
Centralized	SL	Fully Supervised	0.600	0.604	0.617	0.597	0.605	0.600	0.604	0.617	0.597	0.605
		Partially Supervised	0.540	0.545	0.484	0.474	0.511	0.528	0.545	0.533	0.510	0.529
	SSL	Unbiased Teacher [25]	0.551	0.550	0.502	0.503	0.527	0.546	0.557	0.541	0.533	0.544
		EMA Teacher [38]	0.598	0.59	0.568	0.568	0.581	0.586	0.570	0.571	0.573	0.575
Federated	SFL	Fully Supervised	0.627	0.614	0.607	0.585	0.608	0.635	0.612	0.608	0.595	0.613
	SSFL†	FedAvg [27]	0.560	0.566	0.553	0.553	0.558	0.572	0.588	0.593	**0.610**	0.591
		FedDyn [1]	0.508	0.569	0.541	0.522	0.535	0.355	0.414	0.420	0.397	0.400
		FedOpt [33]	0.561	0.572	0.565	0.566	0.566	0.591	0.587	0.588	0.577	0.586
		FedPAC [39]	0.514	0.532	0.496	0.489	0.508	0.510	0.549	0.547	0.554	0.540
		FedSTO	**0.596**	**0.607**	**0.590**	**0.580**	**0.593**	**0.591**	**0.634**	**0.614**	0.595	**0.609**

Experiments

Compared to other methods, FedSTO consistently demonstrates improved generalization across most object categories, both for labeled and unlabeled data.

Type	Algorithm	Method	Labeled					Unlabeled				
			Categories									
			Person	Car	Bus	Truck	Traffic Sign	Person	Car	Bus	Truck	Traffic Sign
Centralized	SL	Fully Supervised	0.569	0.778	0.530	0.307	0.500	0.560	0.788	0.571	0.283	0.510
		Partially Supervised	0.380	0.683	0.193	0.302	0.246	0.358	0.648	0.343	0.138	0.255
	SSL	Unbiased Teacher [25]	0.391	0.695	0.225	0.320	0.297	0.410	0.689	0.373	0.129	0.354
		EMA Teacher [38]	0.475	0.711	0.354	0.347	0.379	0.460	0.727	0.436	0.144	0.378
Federated	SFL	Fully Supervised	0.498	0.715	0.357	0.289	0.410	0.492	0.714	0.451	0.251	0.425
	SSFL†	FedAvg [27]	0.450	0.697	0.310	**0.304**	0.356	0.482	0.725	0.425	**0.247**	0.397
		FedBN [22]	0.488	0.709	0.325	0.285	0.411	0.375	0.618	0.046	0.031	0.286
		FedSTO	**0.504**	**0.720**	**0.342**	0.261	**0.415**	**0.487**	**0.740**	**0.460**	0.181	**0.437**

Ablation Studies – Sampling Ratio

With 100 clients, FedSTO still performs well on the BDD100k dataset even at a lower sampling ratio of 0.1, effectively handling Non-IID scenarios and demonstrating the efficiency of FL.

Method	Labeled					Unlabeled				
	Categories									
	Person	Car	Bus	Truck	Traffic Sign	Person	Car	Bus	Truck	Traffic Sign
Server Only (i.e., client sampling ratio 0.0)	0.378	0.710	0.141	0.425	0.490	0.337	0.707	0.160	0.338	0.491
FedSTO with client sampling ratio 0.1	0.393	0.714	0.442	0.510	0.540	0.487	**0.738**	**0.573**	**0.589**	**0.617**
FedSTO with client sampling ratio 0.2	**0.458**	**0.747**	**0.476**	**0.521**	**0.571**	0.440	0.731	0.378	0.525	0.573
FedSTO with client sampling ratio 0.5	0.444	0.745	0.437	0.502	0.550	**0.489**	0.730	0.438	0.512	0.538

Network Bandwidth

FedSTO achieves a 20.52% reduction in network bandwidth to 2,166.23 GB compared to traditional FL methods, by reducing the Yolov5L model size from 181.7MB to 107.13MB over 350 training rounds with 100 clients.

Method	Warm-up (50 rounds)	Phase 1 (150 rounds)	Phase 2 (150 rounds)	Total	Reduction
FedAvg FedProx	0	100 * 0.50 * 150 * 181.7 = 1,362.75 GB	100 * 0.50 * 150 * 181.7 = 1,362.75 GB	2,725.50 GB	-
FedBN	0	100 * 0.50 * 150 * 181.24 = 1359.30 GB	100 * 0.50 * 150 * 181.24 = 1359.30 GB	2,718.60 GB	0.25 %
FedSTO	0	100 * 0.50 * 150 * 107.13 = 803.48 GB	100 * 0.50 * 150 * 181.7 = 1,362.75 GB	**2,166.23 GB**	**20.52 %**

Conclusion

- **Two-Stage Training Strategy**: FedSTO introduces a novel approach in Semi-Supervised Federated Object Detection, focusing on heterogeneous, unlabeled data using a two-stage training method.
- **Enhanced Feature Learning**: It leverages selective training, orthogonality regularization, and personalized pseudo labeling to enhance object detection performance across diverse conditions and data distributions.
- **High Performance in Non-IID Settings**: FedSTO achieves results comparable to fully supervised models, even with non-IID clients lacking labels, demonstrating significant progress in efficient, privacy-preserving learning in federated learning environments.

Paper Link

Personal Blog

Thank you

Contact: potter32@kaist.ac.kr, kimtaehyeon610@gmail.com (permanent)