南京航空航天大學
NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Gains: Fine-grained Federated Domain Adaptation in Open Set

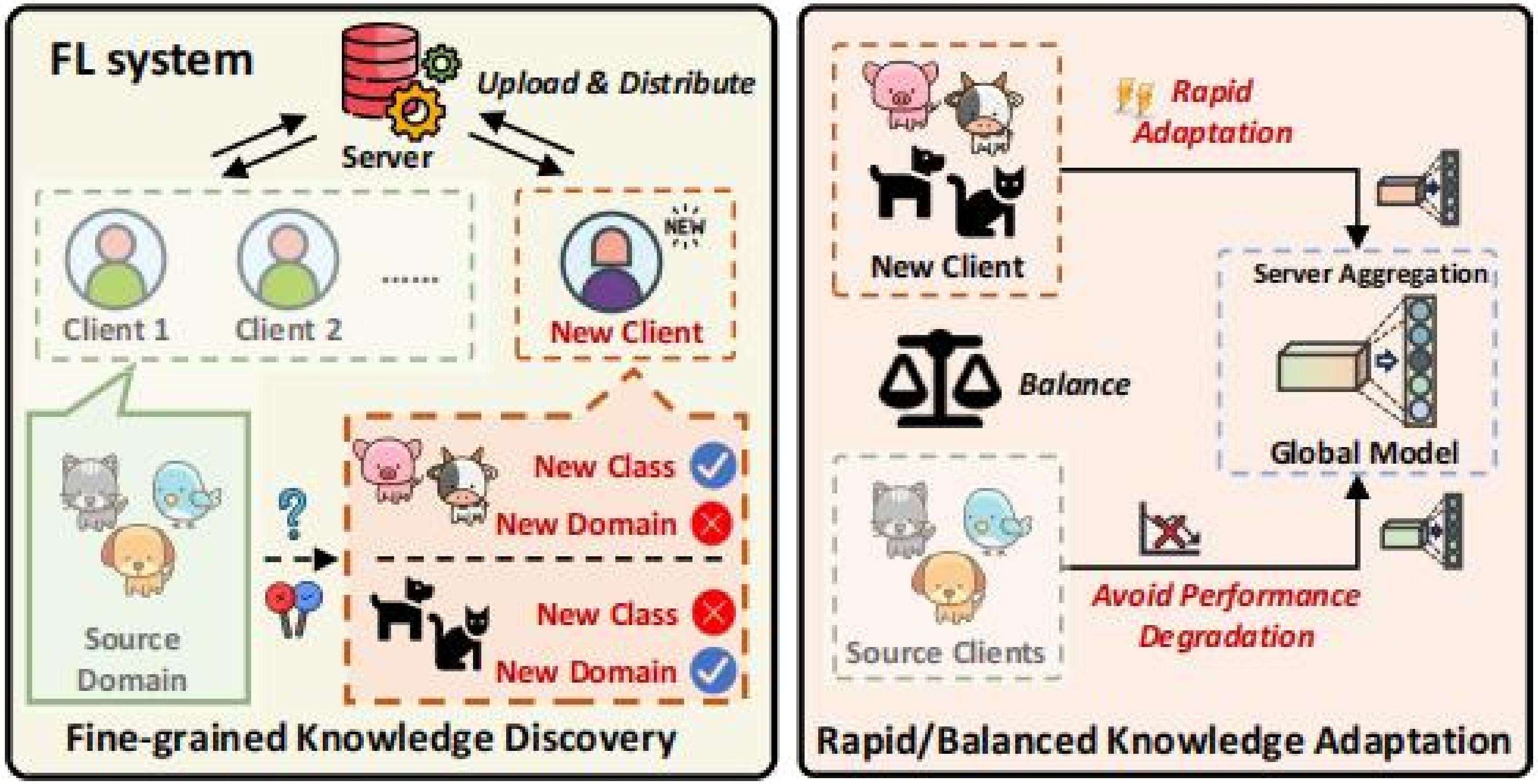
Zhengyi Zhong[1,3,*], Wenzheng Jiang[1,*], Weidong Bao[1], Ji Wang[1,†], Qi Wang[2], Guanbo Wang[3], Yongheng Deng[3], Ju Ren[3]

[1]Laboratory for Big Data and Decision, National University of Defense Technology
[2]College of Science, National University of Defense Technology
[3]Department of Computer Science and Technology, Tsinghua University

NeurIPS 2025

Background

- Conventional FL is often studied in a setup with a fixed number of clients.
  → New clients should be allowed to join the learning process
- How to deal with heterogeneous or evolving client data distributions ?



key challenges:

1) **Not** applicable to both of the class-incremental scenarios and domain-incremental scenarios.

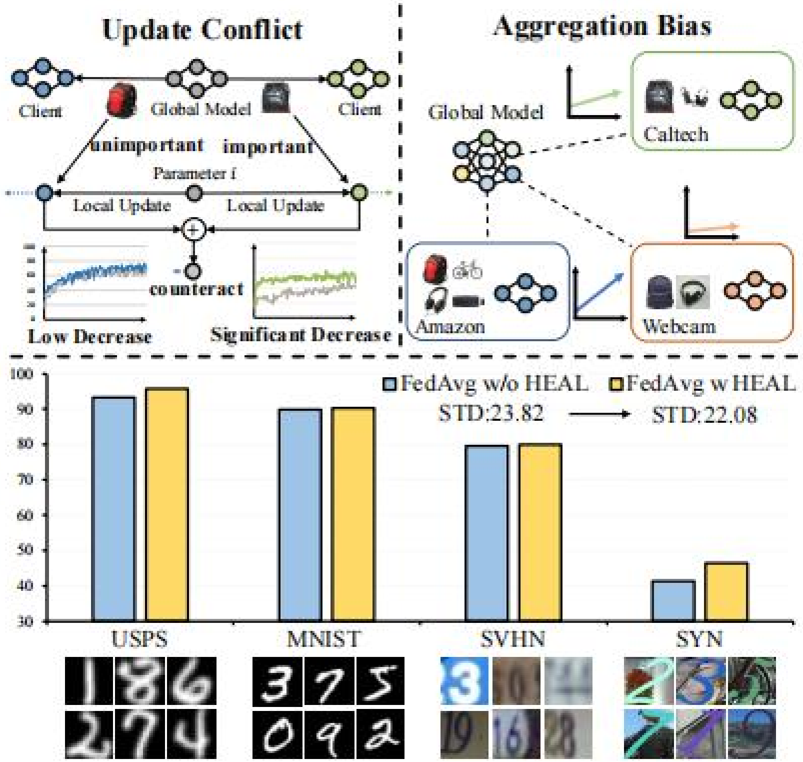
2) **Performance degradation** on the source domain.

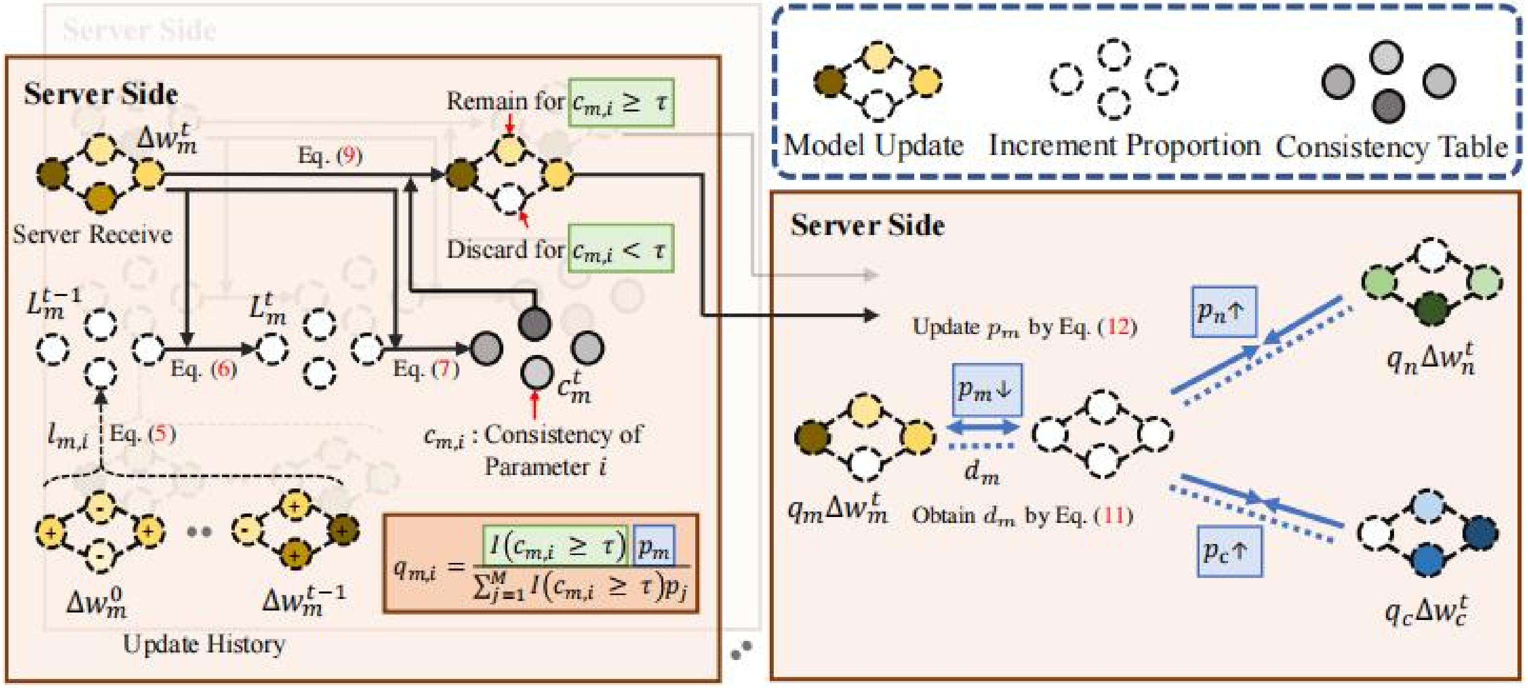
Related work

- FedHEAL(CVPR 2024)

$$l_{m,i} \leftarrow \frac{l_{m,i} * (t-1) + I(\Delta w^t_{m,i} \geq 0)}{t}.$$

$$c_{m,i} = PUC(w_{m,i}) = \begin{cases} l_{m,i} & \text{if } \Delta w^t_{m,i} \geq 0, \\ 1 - l_{m,i} & \text{otherwise} \end{cases}.$$





(a) Federated Parameter-Harmonized Learning

(b) Federated Aggregation-Equalized Learning

$$\mathcal{W}_i^{t+1} = \mathcal{W}_i^t + \sum_{m=1}^{M} q_{m,i}^t \Delta w_{m,i}^t, \qquad d_m = ||\sum_{i=1}^{G} I(c_{m,i} \geq \tau) \cdot \Delta w_{m,i}^t||_2^2,$$

$$\Delta p_m^t = (1-\beta)\Delta p_m^{t-1} + \beta \frac{d_m}{\sum_{j=1}^{M} d_j}, \qquad p_m^t = p_m^{t-1} + \Delta p_m^t, \quad p_m^t = \frac{p_m^t}{\sum_{j=1}^{M} p_j^t}.$$

Experiments

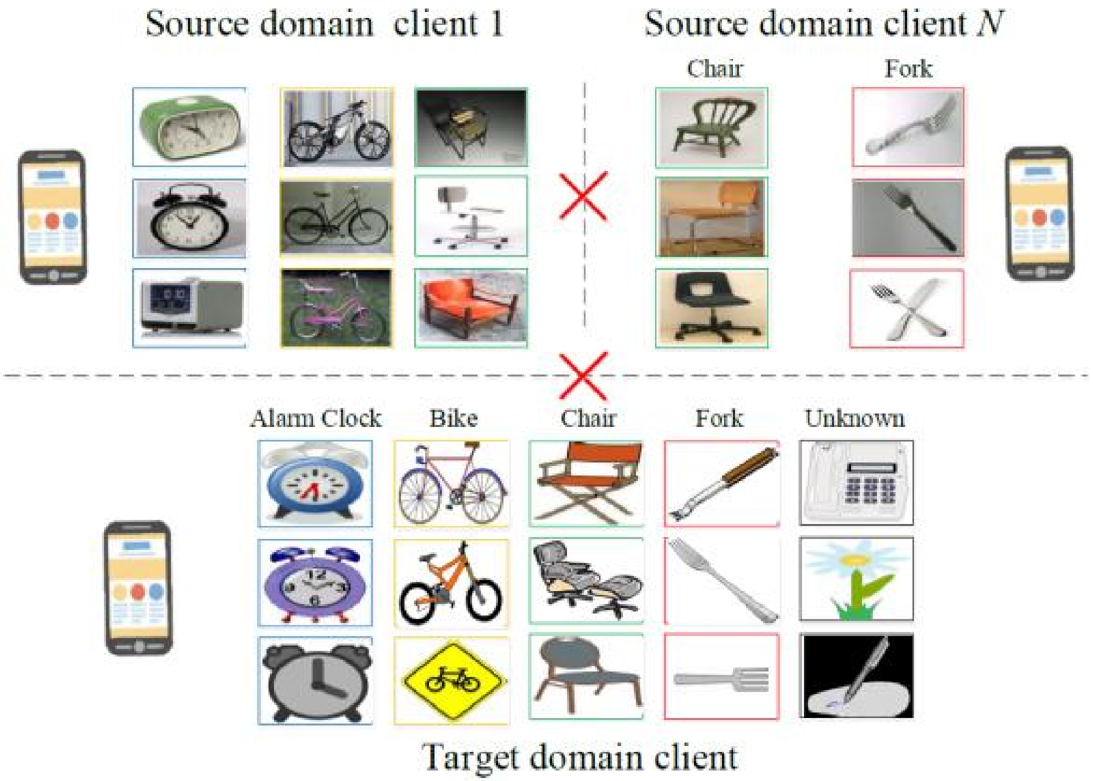
南京航空航天大學
NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

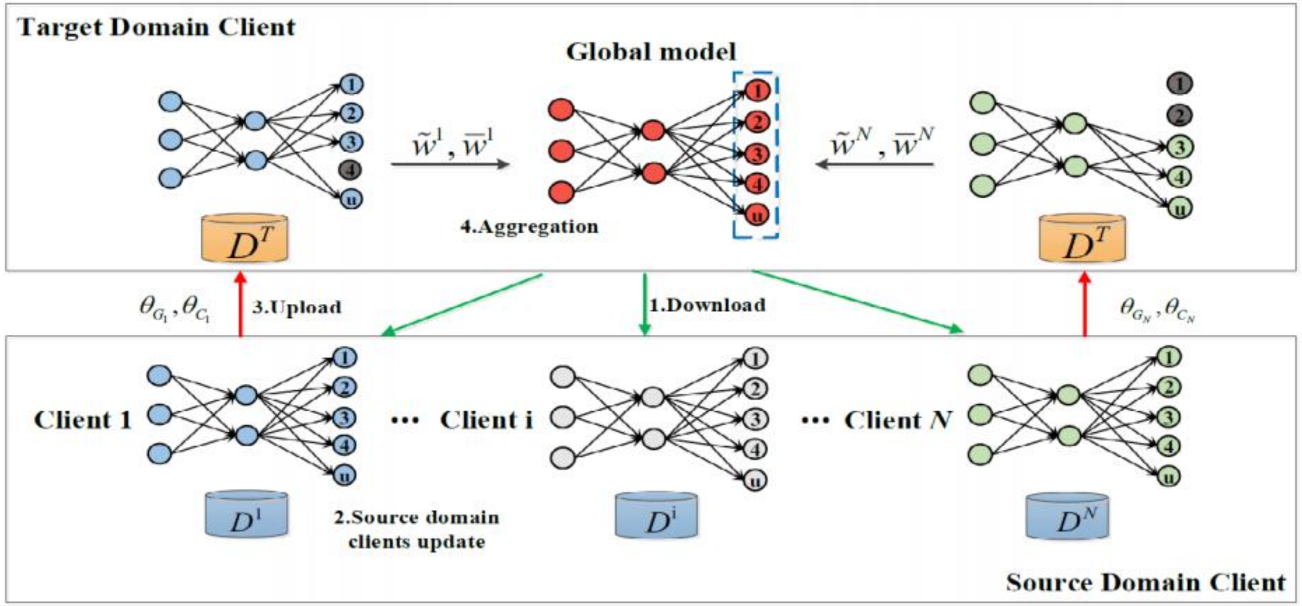
Methods	Digits						Office-Caltech					
	MNIST	USPS	SVHN	SYN	AVG↑	STD↓	Amazon	DSLR	Caltech	Webcam	AVG↑	STD↓
FedAvg [36]	89.84	93.25	79.54	41.35	76.00	23.82	72.63	56.67	58.57	45.52	58.35	11.13
+AFL [38]	90.59	95.83	75.13	44.42	76.49	23.12	64.21	65.37	57.50	48.28	58.83	7.84
+q-FFL [27]	91.44	94.10	76.33	44.48	76.59	22.79	60.00	64.01	53.39	51.72	57.28	5.73
+FedHEAL	90.27	95.69	79.94	46.45	**78.09**	**22.08**	67.90	66.00	59.28	66.21	**64.85**	**3.80**
FedProx [29]	90.27	93.93	80.04	42.82	76.76	23.38	69.90	58.00	60.27	45.52	58.42	10.03
+AFL [38]	92.86	96.17	74.47	42.22	76.43	24.72	68.10	62.67	59.29	52.41	60.62	6.57
+q-FFL [27]	88.58	93.49	75.58	44.23	75.47	22.15	61.37	72.66	54.91	55.52	61.11	8.23
+FedHEAL	89.06	95.52	79.44	46.67	**77.67**	**21.70**	66.11	72.67	57.50	67.59	**65.97**	**6.30**
Scaffold [22]	94.15	94.44	76.87	44.22	77.42	23.61	69.37	59.33	59.55	46.21	58.62	9.50
+AFL [38]	91.77	96.05	78.60	46.39	78.20	22.47	66.42	63.33	59.11	49.31	59.54	7.45
+q-FFL [27]	87.73	94.59	74.00	43.76	75.02	22.53	61.79	73.33	55.18	55.86	61.54	8.40
+FedHEAL	92.68	96.25	78.54	47.72	**78.80**	**22.08**	64.11	67.99	55.18	62.41	**62.42**	**5.37**
MOON [26]	90.46	92.65	80.48	40.58	76.04	24.23	74.00	59.33	60.63	46.90	60.21	11.08
+AFL [38]	91.25	96.03	75.31	44.34	76.73	23.34	66.74	67.33	60.80	55.17	62.51	5.71
+q-FFL [27]	90.43	94.84	76.48	43.95	76.42	23.02	64.32	65.33	54.28	61.03	61.24	4.99
+FedHEAL	91.34	94.94	81.32	44.96	**78.14**	**22.86**	67.68	65.33	59.11	64.14	**64.07**	**3.62**
FedDyn [1]	91.23	92.36	80.15	41.55	76.32	23.83	71.16	62.00	59.20	48.62	60.24	9.28
+AFL [38]	92.11	96.10	71.46	41.52	75.30	24.97	70.10	58.67	59.82	51.03	59.91	7.84
+q-FFL [27]	92.53	95.17	76.37	44.75	77.20	23.18	62.10	67.33	54.82	56.21	60.12	5.76
+FedHEAL	89.87	95.00	80.18	44.23	**77.32**	**22.90**	67.47	60.66	59.02	54.83	**60.50**	**5.26**
FedProc [39]	91.86	91.16	78.54	39.87	75.36	24.44	60.21	46.00	55.98	46.90	52.27	6.95
+AFL [38]	87.85	94.28	78.52	41.54	75.55	23.58	52.63	52.67	55.09	43.45	50.96	5.14
+q-FFL [27]	92.09	92.09	74.97	45.21	76.15	22.17	65.79	42.01	55.80	50.69	53.57	9.94
+FedHEAL	94.23	92.93	81.43	48.67	**79.31**	**21.22**	67.58	66.00	56.79	61.38	**62.94**	**4.87**
FedProto [48]	89.99	92.90	81.09	40.93	76.23	24.06	71.48	42.67	62.23	60.34	59.18	12.04
+AFL [38]	85.27	92.90	67.16	42.36	71.92	22.47	70.74	56.67	57.77	79.65	66.21	11.01
+q-FFL [27]	93.35	94.92	77.08	46.31	77.91	22.56	72.74	54.67	64.20	82.76	68.59	11.99
+FedHEAL	88.49	94.62	81.39	48.46	**78.24**	**20.58**	75.68	76.00	65.18	80.34	**74.30**	**6.44**

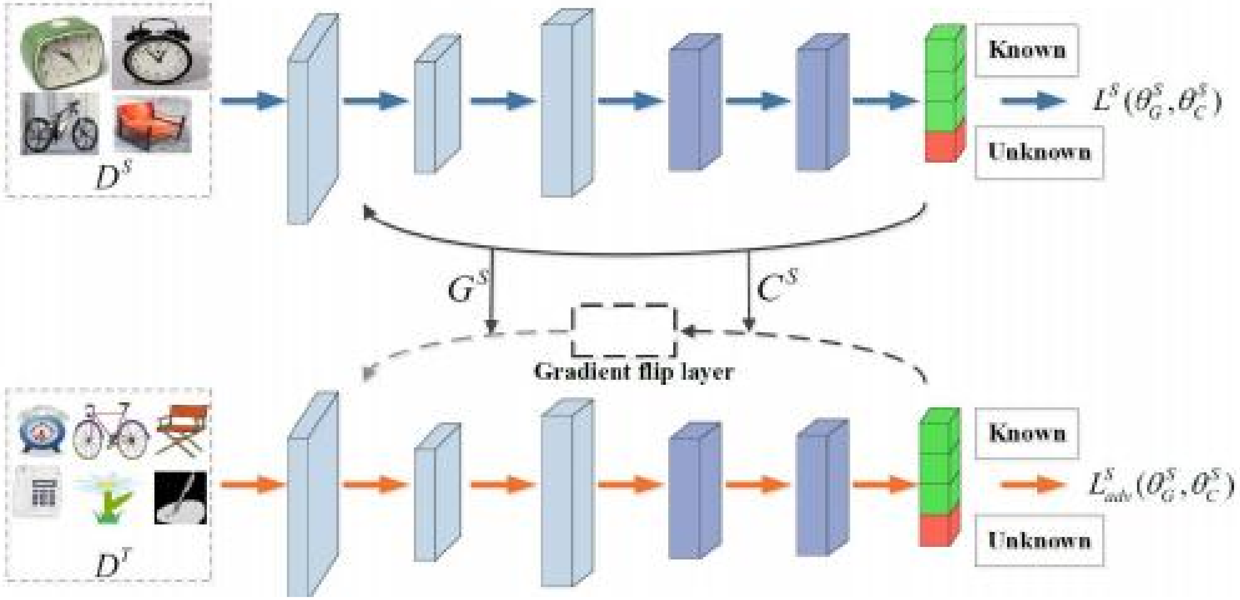
Related work

$$(\tilde{w}^s)^\tau = |\tilde{\mu}^{\tau-1} - (\mu^s)^\tau|. \qquad (\bar{w}^s)^\tau = |\bar{\mu}^{\tau} - (\mu^s)^\tau|.$$

- FOSDA(TNNLS 2024)







$$L^s(\theta_G^s, \theta_C^s) = \frac{1}{n_s}\sum_{i=1}^{n_s} \ell_{ce}\left(C^s\left(G^s\left(\mathbf{x}_i^s; \theta_G^s\right); \theta_C^s\right), y_i^s\right)$$

$$L^s_{\text{adv}}(\theta_G^s, \theta_C^s) = \frac{1}{n^t}\sum_{j=1}^{n^t} \ell_{ce}\left(C^s\left(G^s\left(\mathbf{x}_j^t; \theta_G^s\right), \theta_C^s\right), y^t_{j|K+1}\right).$$

$$w^s = \alpha\tilde{w}^s + (1-\alpha)\bar{w}^s \quad \tilde{\theta}_{C_k} = \sum_{s=1}^{N} w^s \operatorname{sign}(\theta^s_{C_k})\theta^s_{C_k}, \quad k = 1, \ldots, K+1.$$

Experiments

CLASSIFICATION ACCURACY (%) OF OSDA TASKS ON DIGITS

Compared Methods	$S \rightarrow M$		$U \rightarrow M$		$M \rightarrow U$		AVG	
	OS	OS*	OS	OS*	OS	OS*	OS	OS*
IID								
FedAvg [23]	56.5	67.5	76.0	84.7	36.5	43.3	56.3	65.2
CMFL [24]	47.8	55.7	78.3	81.9	49.1	56.5	58.4	64.7
MOON [33]	49.7	58.9	40.8	50.0	43.9	50.9	44.8	54.2
FedEuc [31]	51.2	60.6	65.0	62.8	56.0	64.4	57.4	62.6
FedCos	51.2	60.8	67.4	72.7	52.4	61.5	57.0	65.0
SHOT [21]	**68.5**	**80.4**	70.2	79.5	74.2	85.0	63.4	72.3
IVC [22]	53.2	62.6	<u>78.4</u>	<u>91.0</u>	<u>76.9</u>	<u>90.3</u>	<u>69.5</u>	<u>81.3</u>
FOSDA	<u>65.1</u>	<u>77.0</u>	**81.4**	**93.1**	**83.0**	**92.4**	**76.5**	**87.5**
Non-IID								
FedAvg [23]	52.9	61.4	64.4	70.8	64.4	70.8	60.6	67.7
CMFL [24]	50.8	59.9	64.5	62.2	36.1	38.4	50.5	53.5
MOON [33]	34.9	46.0	62.2	71.0	48.9	56.1	48.7	57.7
FedEuc [31]	39.3	45.0	<u>75.5</u>	<u>82.5</u>	32.7	37.1	49.2	54.9
FedCos	48.1	57.2	63.7	72.0	60.6	71.6	57.5	66.9
SHOT [21]	<u>63.3</u>	<u>73.0</u>	68.9	78.7	**77.6**	**89.6**	<u>69.9</u>	<u>80.4</u>
IVC [22]	35.5	38.0	41.1	49.8	49.0	53.8	41.9	47.2
FOSDA	**65.2**	**75.1**	**80.1**	**92.6**	<u>75.0</u>	<u>83.2</u>	**73.4**	**83.6**

Motivation

Observation:
1) The variation of the encoder does not show a clear fluctuation trend no matter in class or domain incremental scenarios.
2) The changes in the classifier parameters are more pronounced in the class-incremental scenario.
3) While both new classes and domain will bring obvious changes to the feature values, it is more significant in the domain-incremental scenario.

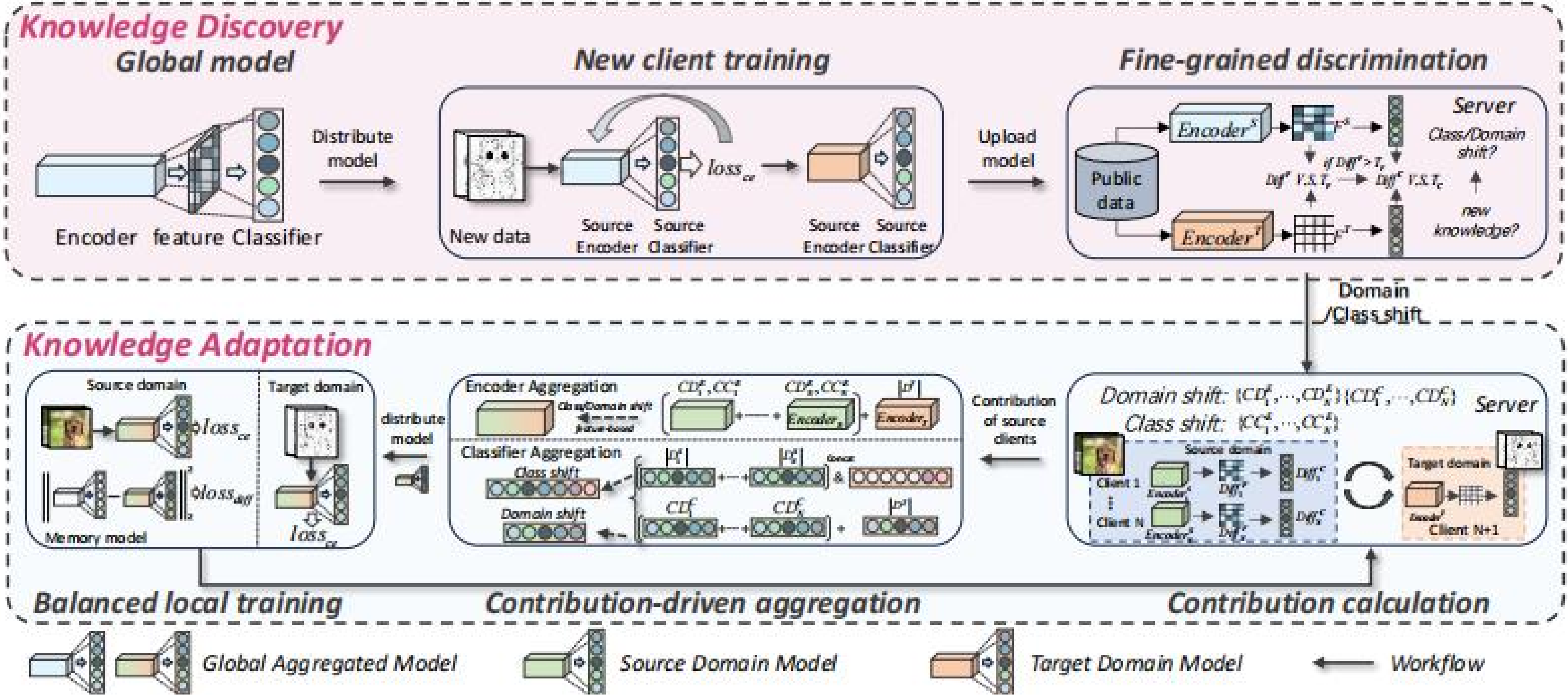
Method

$$\mathcal{W}^T(q+1) = \mathcal{W}^T(q) - \eta \nabla \mathcal{L}(\mathcal{W}^T(q), \mathcal{D}^T), q = 0, \cdots, Q-1,$$

$$Diff^F(i) = \frac{1}{P \times d} \sum_{p=1}^{P} \sum_{k=1}^{d} \left| F^T(i)_{p,k} - F_{src,k}(i) \right|$$

$Diff^F > T_F$ New knowledge

$Diff^C > T_C$ class shift



Method

南京航空航天大學
NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Domain-incremental contribution-driven aggregation

encoder contribution

$$\mathcal{CD}_n^E(i) = \frac{1}{(1 + Diff_n^E(i)) \times \sum_{n=1}^{N} \left({}^{1}\!/\!(1 + Diff_n^F(i)) \right)} \times \frac{\sum_{n=1}^{N} |\mathcal{D}_n^S|}{|\mathcal{D}^T| + \sum_{n=1}^{N} |\mathcal{D}_n^S|},$$

classifier contribution

$$\mathcal{CD}_n^C(i) = \frac{1}{(1 + Diff_n^C(i)) \times \sum_{n=1}^{N} \left({}^{1}\!/\!(1 + Diff_n^C(i)) \right)} \times \frac{\sum_{n=1}^{N} |\mathcal{D}_n^S|}{|\mathcal{D}^T| + \sum_{n=1}^{N} |\mathcal{D}_n^S|},$$

aggregation

$$E(i) = \sum_{n=1}^{N} \mathcal{CD}_n^E(i) \times E_n^S(i) + \frac{|\mathcal{D}^T|}{|\mathcal{D}^T| + \sum_{n=1}^{N} |\mathcal{D}_n^S|} \times E^T(i),$$

$$C(i) = \sum_{n=1}^{N} \mathcal{CD}_n^C(i) \times C_n^S(i) + \frac{|\mathcal{D}^T|}{|\mathcal{D}^T| + \sum_{n=1}^{N} |\mathcal{D}_n^S|} \times C^T(i).$$

Method

南京航空航天大學
NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Class-incremental contribution-driven aggregation

encoder aggregation

$$E(i) = \sum_{n=1}^{N} \mathcal{C}C_n^E(i) \times E_n^S(i) + \frac{|\mathcal{D}^T|}{|\mathcal{D}^T| + \sum_{n=1}^{N} |\mathcal{D}_n^S|} \times E^T(i).$$

classifier aggregation

$$C^S(i) = \sum_{n=1}^{\mathcal{N}} \frac{|D_n^S|}{\sum_{n=1}^{N} |D_n^S|} \times C_n^S(i)$$

$$C^T(i) = [Channel_1^T, \cdots, Channel_{K^s}^T, Channel_{K^s+1}^T, \cdots, Channel_{K^s+K^T}^T]$$

$$C(i) = \Big[\underbrace{Channel_1^S, \cdots, Channel_{K^S}^S}_{SourceDomain}, \underbrace{Channel_{K^S+1}^T, \cdots, Channel_{K^S+K^T}^T}_{TargetDomain} \Big]$$

Anti-forgetting mechanism

$$\mathcal{L}(\mathcal{W}_n^S(i,r), \mathcal{D}_n^S) = -\frac{1}{|\mathcal{D}_{\mathrm{n}}^S|} \sum_{j=1}^{|\mathcal{D}_{\mathrm{n}}^S|} \sum_{c=1}^{K^S+K^T} y_{j,c}^n \log(\hat{y}_{j,c}^n) + \lambda \left\| \mathcal{W}_n^S(i,r) - \mathcal{W}_n^S(0,0) \right\|_2^2$$

Method

Algorithm 1: *Gains*

Input: Number of source clients N; original source global model $\mathcal{W}^S$ and client model $\{\mathcal{W}_1^S(0,0), \mathcal{W}_2^S(0,0), \cdots, \mathcal{W}_N^S(0,0)\}$; number of iteration I; number of local training R; public data $\mathcal{D}^P = \{(x^p, y^p)\}$

Output: Global model $\mathcal{W}$

Distribute original source model $\mathcal{W}^S$ to target client
$\mathcal{W}^T \leftarrow$ Target client performs local updating based on $\mathcal{W}^S$
Target client Uploads $\mathcal{W}^T$ to the server
//Knowledge Discovery
Split the $\mathcal{W}^S$ into encoder E^S and classifier C^S, split the $\mathcal{W}^T$ into E^T and C^T
$F^S \leftarrow E^S(x^p)$, $F^T \leftarrow E^T(x^p)$
Calculating $Diff^C$ and $Diff^F$
if $Diff^F > T_F$ **then**
  Target client brings new knowledge
  if $Diff^C > T_C$ **then**
    $Class\ Increment$=True
  else
    $Domain\ Increment$=True
  //Knowledge Adaptation
  for *iteration* $i = 0, \cdots, I$ **do**
    if $Domain\ Increment$=*True* **then**
      Calculating encoder contributions $\{\mathcal{CD}_1^E, \mathcal{CD}_2^E, \cdots, \mathcal{CD}_N^E\}$ based on Eq. (2)
      Calculating classifier contributions $\{\mathcal{CD}_1^C, \mathcal{CD}_2^C, \cdots, \mathcal{CD}_N^C\}$ based on Eq. (3)
      Aggregating all clients' parameters using Eq.(4) and Eq. (5)
    if $Class\ Increment$=*True* **then**
      Calculating encoder contributions $\{\mathcal{CC}_1^E, \mathcal{CC}_2^E, \cdots, \mathcal{CC}_N^E\}$ based on Eq. (2)
      Aggregating all clients' parameters using Eq.(6) and Eq. (7)
    Server distributes the aggregated model to all clients
    for *client* $n = 1, \cdots, N$ **do**
      Locally update model R rounds using Eq.(8)
      Upload $\mathcal{W}_n^S(i, R)$ to the server
    Target client locally update model R rounds and upload to the server
else
  Apply the original model to newly joined clients for inference tasks without training

Experiments

NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Table 1: Main results. The bold font represents the optimal result.

Scenario	Metric	Federated Domain Adaptation					Heter-FL		
		Ours	FOSDA [TNNLS'24]	SemiFDA [ICDM'24]	AutoFedGP [ICLR'24]	FedHEAL [CVPR'24]	FedAVG [AISTATS'17]	FedProx [MLSys'20]	FedProto [AAAI'22]
DigitFive									
Mild	*T-Acc*	**99.34**	0.00	0.00	68.11	22.60	55.73	72.35	77.61
	S-Acc	93.21	12.72	13.53	0.00	99.29	0.36	**99.53**	0.16
	G-Acc	**94.44**	10.18	10.83	13.62	83.95	11.44	94.09	62.12
Medium	*T-Acc*	**97.91**	11.29	7.91	9.78	93.68	90.79	94.88	45.66
	S-Acc	**90.09**	19.46	19.44	6.22	88.71	76.20	86.50	33.56
	G-Acc	**91.65**	17.82	17.14	6.93	89.70	79.12	88.18	43.23
Strong	*T-Acc*	**98.98**	11.29	31.14	10.37	96.98	85.80	85.29	31.28
	S-Acc	**93.18**	13.60	14.21	11.60	83.32	43.90	43.32	62.23
	G-Acc	**94.34**	13.13	17.60	11.35	86.05	52.28	51.72	37.47
Amazon Review									
Medium	*T-Acc*	**84.60**	49.55	50.45	50.50	50.56	66.74	74.55	50.11
	S-Acc	**82.81**	49.55	49.33	50.50	50.56	67.19	74.44	50.11
	G-Acc	**83.09**	49.82	49.82	50.58	50.48	67.38	74.12	50.01
Strong	*T-Acc*	**80.54**	50.48	55.41	50.03	83.34	51.20	53.73	50.10
	S-Acc	**84.95**	50.27	59.25	50.02	86.54	51.36	53.95	50.11
	G-Acc	**83.85**	50.33	58.29	50.02	85.74	51.32	53.89	50.10

Mild: MNIST targat domain: {1, 5} source domain: {0, 2, 3, 4, 6, 7, 8, 9}
Medium: target domain: {SVHN} source domain: {MNIST}
Strong: target domain: {MNIST} source domain:{MNIST-M, SVHN, USPS, SynthDigits}

Experiments

南京航空航天大學
NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Table 2: Generalization verification.

		{1,5}	{6,9}	{0,1,5}
Mi-DF	NA	99.34	94.42	99.59
	OA	93.21	96.03	87.16
	GA	94.44	95.71	89.64
		SV-MT	MT-MTM	SYN-MTM
Me-DF	NA	97.91	94.46	88.48
	OA	90.09	99.56	97.76
	GA	9165	98.54	95.90
		MT	SV	MTM
St-DF	NA	98.98	91.67	93.94
	OA	93.18	97.58	96.20
	GA	94.34	96.40	95.75
		DD-BK	BK-DD	ET-KC
Me-AR	NA	84.6	82.01	86.59
	OA	82.81	86.85	89.93
	GA	83.09	85.88	89.26
		BK	DD	KC
St-AR	NA	80.54	78.22	85.38
	OA	84.95	88.90	87.73
	GA	83.85	86.23	87.14

Table 3: The performance of sequential FDA.

		{4,5}	{6,7}	{8,9}
Mi	NA	99.88	91.35	96.89
	OA	93.53	99.43	99.35
	GA	96.82	98.08	99.00
		MNIST	MNISTM	SYN
Me	NA	95.27	83.53	93.53
	OA	87.91	90.05	89.66
	GA	89.38	88.96	90.21

Table 5: Convergence Comparison of Different Methods.

Method	Converge Round	Time
Gains	**5**	**807.45**
FedHEAL	40	1368.4
FedAVG	20	1977.20
FedProx	40	6880.80
FedProto	32	9519.68

Thanks