Generalized and Personalized Federated Learning with Black-Box Foundation Models via Orthogonal Transformations

Supplementary Material

Eun Gyung Kong[*] [†] Je Won Yeom[*] Yonghoon Jeon[‡] Taesup Kim[§]
Seoul National University

A. Related Work (Extended)

A.1. Foundation Models

Foundation models [3], exemplified by the GPT series [1, 4, 16] and VLMs[10, 17], are characterized by their substantial scale and extensive pre-training on broad data. CLIP [17] notably showcases the potential of aligning visual and language modalities through contrastive learning. In this study, we concentrate on vision foundation models, including VLMs, and highlight their significant potential for extension to federated learning paradigms.

A.2. Federated Learning

FL facilitates collaborative model training across decentralized data sources [13]. Heterogeneous FL specifically addresses challenges arising from non-IID data distributions, which complicate model aggregation and convergence [8]. Personalized FL (PFL) aims to mitigate these challenges by training personalized local models alongside a generalized global model, adapting to local heterogeneity [15, 18]. The integration of large-scale pre-trained VLMs within the FL framework represents an active and promising area of research [9, 12].

B. FedOT Algorithm Details

B.1. Pseudo-code

The detailed algorithm of FEDOT is presented in the form of a complete pseudo-code in Algorithm 1.

Algorithm 1 FEDOT: Orthogonal Local Updates via Cayley Transform

Global initialisation
Initialise server parameter $w_{\mathrm{g},0}$
for all client $i \in \{1, \dots, N\}$ **in parallel do**
  $X_0^{(i)} \leftarrow I$
  $P_0^{(i)} \leftarrow \frac{1}{2}\big(X_0^{(i)} - X_0^{(i)\top}\big)$
  $w_{\mathrm{l},0}^{(i)} \leftarrow (I + P_0^{(i)})(I - P_0^{(i)})^{-1}$
end for
Server broadcasts $w_{\mathrm{g},0}$ to all clients
for round $t = 0$ **to** $T - 1$ **do**
  for all client i **in parallel do**
    $(w_{\mathrm{g},t+1}^{(i)},\, X_{t+1}^{(i)}) \leftarrow \textsc{LocalUpdate}(i,\, w_{\mathrm{g},t},\, X_t^{(i)})$
    Client sends $w_{\mathrm{g},t+1}^{(i)}$ to server
  end for
  Server aggregation
  $w_{\mathrm{g},t+1} \leftarrow \frac{1}{N}\sum_{i=1}^{N} w_{\mathrm{g},t+1}^{(i)}$
  Server broadcasts $w_{\mathrm{g},t+1}$ to all clients
end for
function $\textsc{LocalUpdate}(i,\, w_{\mathrm{g}},\, X^{(i)})$
  $P^{(i)} \leftarrow \frac{1}{2}\big(X^{(i)} - X^{(i)\top}\big)$
  $w_{\mathrm{l}}^{(i)} \leftarrow (I + P^{(i)})(I - P^{(i)})^{-1}$
  for epoch $e = 1$ **to** E **do**
    for all mini-batch $B \subset D^{(i)}$ **do**
      $w_{\mathrm{g}}^{(i)} \leftarrow w_{\mathrm{g}}^{(i)} - \eta \nabla_{w_{\mathrm{g}}^{(i)}} \ell^{(i)}(w_{\mathrm{g}}^{(i)},\, w_{\mathrm{l}}^{(i)}; B)$
      $X^{(i)} \leftarrow X^{(i)} - \eta \nabla_{X^{(i)}} \ell^{(i)}(w_{\mathrm{g}}^{(i)},\, w_{\mathrm{l}}^{(i)}; B)$
      $P^{(i)} \leftarrow \frac{1}{2}\big(X^{(i)} - X^{(i)\top}\big)$
      $w_{\mathrm{l}}^{(i)} \leftarrow (I + P^{(i)})(I - P^{(i)})^{-1}$ ▷ Cayley transform
    end for
  end for
  return $(w_{\mathrm{g}}^{(i)},\, X^{(i)})$ ▷ Only $w_{\mathrm{g}}^{(i)}$ sent to server
end function

[*]Equal contribution.
[†]Current affiliation: Mobilint, Inc.
[‡]Current affiliation: Kakao Healthcare Corp.
[§]Corresponding author.

B.2. Cayley Transform

We utilize the Cayley transform [6] to enforce strict orthogonality during optimization. The orthogonal matrix Q is derived using $Q = (I + P)(I - P)^{-1}$, where P is a skew-symmetric matrix ($P = -P^\top$). We formulate P as $P = 0.5(X - X^\top)$, where X is an unconstrained transformation matrix. Because P is skew-symmetric, $(I - P)$ is guaranteed to be invertible, ensuring the Cayley transform is well-defined and continuously differentiable. This allows for stable, standard gradient-based optimization directly on the unconstrained matrix X.

C. Theoretical Analysis (Proofs)

C.1. Proof of Theorem 1 (Bounding the Gradient Difference)

We define the cross-entropy loss for client i: $\ell^{(i)} = \mathbb{E}_{(x,y)\in D^{(i)}}\big[\mathrm{CE}(r_x^{(i)}, y)\big]$. The gradient with respect to the global classifier $w_{\mathrm{g}}^{(i)}$ is:

$$\nabla_{w_{\mathrm{g}}^{(i)}} \ell^{(i)} = \mathbb{E}\Big[\tau \left(r_x^{(i)} - y\right) \Big(\frac{w_1^{(i)}\, \mathcal{I}(x)}{\|w_1^{(i)}\, \mathcal{I}(x)\|}\Big)^\top\Big].$$

Let $\Delta_i := r_x^{(i)} - y$. Since $r_x^{(i)}$ is a probability distribution vector and y is a one-hot vector, we know $\|\Delta_i\| \le 2$ (derived from the triangle inequality: $\|r_x^{(i)}\| \le 1$ and $\|y\| \le 1$).

For clients i and j, we analyze the difference between their gradients using the triangle inequality:

$$\begin{aligned}
&\Big\|\nabla_{w_{\mathrm{g}}^{(i)}} \ell^{(i)} - \nabla_{w_{\mathrm{g}}^{(j)}} \ell^{(j)}\Big\|_2 \\
&\le\ \mathbb{E}_{D^{(i)}}\Big\|\tau\, \Delta_i \frac{w_1^{(i)}\, \mathcal{I}(x)}{\|w_1^{(i)}\, \mathcal{I}(x)\|}\Big\| \ +\ \mathbb{E}_{D^{(j)}}\Big\|\tau\, \Delta_j \frac{w_1^{(j)}\, \mathcal{I}(x)}{\|w_1^{(j)}\, \mathcal{I}(x)\|}\Big\| \\
&\le\ \mathbb{E}_{D^{(i)}}\Big\|2\tau \frac{w_1^{(i)}\, \mathcal{I}(x)}{\|w_1^{(i)}\, \mathcal{I}(x)\|}\Big\| \ +\ \mathbb{E}_{D^{(j)}}\Big\|2\tau \frac{w_1^{(j)}\, \mathcal{I}(x)}{\|w_1^{(j)}\, \mathcal{I}(x)\|}\Big\|.
\end{aligned} \tag{1}$$

By definition, the norm of the transformed normalized vector is bounded by the condition number κ of the transformation matrix:

$$\frac{\|w_1^{(i)}\, \mathcal{I}(x)\|}{\|w_1^{(i)}\, \mathcal{I}(x)\|} \le \kappa\big(w_1^{(i)}\big),$$

which leads to the final bound:

$$\Big\|\nabla_{w_{\mathrm{g}}^{(i)}} \ell^{(i)} - \nabla_{w_{\mathrm{g}}^{(j)}} \ell^{(j)}\Big\| \ \le\ 2\tau \Big[\kappa\big(w_1^{(i)}\big) + \kappa\big(w_1^{(j)}\big)\Big].$$

C.2. Condition Number of an Orthogonal Transformation

An orthogonal transformation Q satisfies $Q^\top Q = I$. Consider an eigenvector v such that $Qv = \lambda v$. Taking the inner product: $\langle Qv, Qv\rangle = |\lambda|^2 \langle v, v\rangle$. Alternatively, using the property of orthogonality: $\langle Qv, Qv\rangle = \langle v, Q^\top Qv\rangle = \langle v, Iv\rangle = \langle v, v\rangle$. Equating both expressions yields $|\lambda|^2 = 1$. Since the singular values of Q are the square roots of the eigenvalues of $Q^\top Q = I$, all singular values are exactly 1. Therefore, the condition number, defined as the ratio of the largest to the smallest singular value, is fixed at $\kappa(Q) = 1/1 = 1$.

D. Experimental Setup Details

D.1. Datasets

We utilize five distinct datasets for our evaluation. FEMNIST [5] utilizes digit samples from four specific users (IDs: 0, 25, 26, 33) to simulate cross-domain heterogeneity based on writing style. PACS [11] comprises four domains characterized by stylistic differences (Photo, Art, Cartoon, Sketch). VLCS [7] consists of four sub-datasets representing different contexts (VOC2007, LabelMe, Caltech101, SUN09). Office-Home [19] includes four domains with varying visual appearances (Art, Clipart, Product, Real-world). TerraIncognita [2] utilizes images captured at four distinct geographical locations (L38, L43, L46, L100). For all datasets, we employ a standardized 60/20/20 split for training, validation, and testing sets.

D.2. Implementation Details

We utilize the CLIP pre-trained model with the ViT-B/32 architecture [17]. The local update epoch is set to $E = 1$. The total number of communication rounds is $T = 200$ for PACS, FEMNIST, and VLCS, and $T = 50$ for Office-Home and TerraIncognita, reflecting differences in dataset complexity and convergence speed. We use the SGD optimizer with a weight decay of $5e-4$. The learning rates are tuned for optimal performance: 1×10^{-3} for FEMNIST, and 5×10^{-5} for the other datasets. The optimal block size for FEDOT(+B) for each dataset are as follows: 256 for FEMNIST, 1 for PACS, OfficeHome, and VLCS, and 4 for TerraIncognita.

D.3. Computational Resources and Efficiency Analysis

Experiments were conducted on high-performance computing clusters equipped with NVIDIA A6000 (48GB) and RTX 3090 (24GB) GPUs, utilizing the PyTorch 2.1.0 framework [14]. We provide a detailed analysis of runtime and computational overhead across representative FL baselines on the FEMNIST dataset in Tab. 3.

E. Additional Experimental Results

E.1. Detailed Cross-Device Scalability Results

For the scalability experiments detailed in Sec. 5.4, we utilized the FEMNIST dataset. To ensure stable training conditions while preserving the inherent heterogeneity (writer-specific data partitions), we pre-filtered the dataset to retain only users possessing a sufficient number of data samples (at least 100 images) and randomly sampled participants from this refined pool. This minimum-samples-per-client criterion was employed solely to isolate the effect of the number of participants on performance and does not alter the underlying feature heterogeneity. We designated a fixed set of 5 unseen test clients (User IDs: 1894, 224, 3321, 745, 509) and systematically varied the number of participating clients (from 1 to 75) per communication round.

Tab. 1 presents the detailed generalization results for FEDOT, and Tab. 2 presents the corresponding results for the PromptFL baseline.

Table 1. Detailed Scalability Results for FEDOT on FEMNIST. Generalization accuracy (%) on 5 unseen test clients.

No. of Participants	Test Clients (User ID)					AVG
	1894	**224**	**3321**	**745**	**509**	
0 (=CLIP ZS)	33.33%	26.32%	44.00%	68.75%	47.62%	44.00%
1	29.17%	42.11%	44.00%	50.00%	52.38%	43.53%
2	33.33%	31.58%	52.00%	56.25%	47.62%	44.16%
5	33.33%	31.58%	56.00%	68.75%	47.62%	47.46%
10	37.50%	31.58%	56.00%	81.25%	66.67%	54.60%
15	41.67%	31.58%	60.00%	75.00%	66.67%	54.98%
20	41.67%	36.84%	60.00%	75.00%	66.67%	56.04%
40	54.17%	57.89%	76.00%	75.00%	71.43%	66.90%
60	66.67%	57.89%	76.00%	81.25%	71.43%	70.65%
75	70.83%	57.89%	76.00%	75.00%	76.19%	71.18%

Table 2. Detailed Scalability Results for PromptFL on FEMNIST. Generalization accuracy (%) on 5 unseen test clients.

No. of Participants	Test Clients (User ID)					AVG
	1894	**224**	**3321**	**745**	**509**	
1	33.33%	26.32%	48.00%	68.75%	38.10%	42.90%
2	29.17%	21.05%	44.00%	37.50%	38.10%	33.96%
5	25.00%	21.05%	48.00%	50.00%	47.62%	38.33%
10	37.50%	26.32%	56.00%	62.50%	66.67%	49.80%
15	33.33%	31.58%	48.00%	43.75%	57.14%	42.76%
20	37.50%	26.32%	56.00%	62.50%	66.67%	49.80%
40	58.33%	47.37%	84.00%	87.50%	90.48%	73.54%
60	50.00%	47.37%	72.00%	87.50%	85.71%	68.52%
75	54.17%	42.11%	68.00%	81.25%	80.95%	65.30%

E.2. Computational Analysis

We conduct a computational analysis on FEMNIST. As depicted on the Table 3, our FedOT gives relatively comparable runtime to other methods, and also in parameters.

Table 3. Computational Analysis on FEMNIST. d_{txtp}, d_{imgp}: prompt dimensions; d_{txtemb}, d_{imgemb}: embedding dimensions; K: number of classes; n_{ctx}: context length. Meta-net size for CoCoOp is calculated as $d_{imgemb}^2/16 + d_{imgemb}/16 + d_{imgemb}/16 \times d_{txtp} + d_{txtp}$.

Method	Runtime	Type	Parameters
PromptFL	22s	Server	$d_{txtp} \times n_{ctx}$
		Client	$d_{txtp} \times n_{ctx}$
CoCoOp	2m 9s	Server	$d_{txtp} \times n_{ctx}$
		Client	$d_{txtp} \times n_{ctx}$ + meta-net
VPT	26s	Server	$K \times d_{txtemb}$
		Client	$K \times d_{txtemb} + d_{imgp} \times n_{ctx}$
FedAKT(C)	23s	Server	$2 \times (d_{imgemb}^2 + d_{imgemb})$
		Client	$K \times d_{txtemb} + K$
FedGH(C)	27s	Server	$K \times d_{txtemb} + K$
		Client	$K \times d_{txtemb} + K$
FedCLIP	19s	Server	$2 \times (d_{imgemb}^2 + d_{imgemb})$
		Client	$2 \times (d_{imgemb}^2 + d_{imgemb})$
FedOT(All Global)	28s	Server	$K \times d_{txtemb} + d_{imgemb}^2$
		Client	$K \times d_{txtemb} + d_{imgemb}^2$
FEDOT (Ours)	26s	Server	$K \times d_{txtemb}$
		Client	$K \times d_{txtemb} + d_{imgemb}^2$
FEDOT(+B) (Ours)	24s	Server	$K \times d_{txtemb}$
		Client	$K \times d_{txtemb} + d_{imgemb}^2$

F. Full Results of Main Experiment

F.1. Generalization

Method	33	0	25	26	AVG
Zero Shot	50.88 ± 3.42	48.40 ± 1.47	45.19 ± 2.50	48.24 ± 2.73	48.18 ± 1.31
FedCLIP	68.57 ± 8.25	70.99 ± 5.57	72.12 ± 12.01	70.51 ± 9.15	70.55 ± 8.60
PromptFL ($M = 4$)	94.28 ± 1.78	96.63 ± 0.49	96.80 ± 1.00	97.44 ± 0.74	96.29 ± 0.86
PromptFL ($M = 16$)	95.51 ± 0.41	96.96 ± 0.28	96.64 ± 0.84	$\mathbf{97.76 \pm 0.28}$	96.72 ± 0.30
CoCoOp (All Global)	86.26 ± 10.82	87.02 ± 9.47	87.50 ± 7.51	87.34 ± 7.73	87.03 ± 8.86
CoCoOp (Meta-Net Random Init)	21.22 ± 12.57	20.67 ± 11.81	23.56 ± 10.84	24.04 ± 12.91	22.37 ± 11.79
CoCoOp (Server CoOp)	15.10 ± 1.23	17.31 ± 5.42	17.63 ± 9.50	9.62 ± 2.89	14.91 ± 3.27
VPT-SHALLOW ($p = 2$)	71.43 ± 0.82	56.41 ± 15.22	65.54 ± 6.40	69.55 ± 15.77	65.73 ± 6.97
VPT-SHALLOW ($p = 10$)	49.39 ± 10.30	49.04 ± 6.31	51.76 ± 6.45	52.57 ± 14.01	50.69 ± 8.52
VPT-DEEP ($p = 2, d = 2$)	63.54 ± 6.24	64.42 ± 10.82	49.36 ± 6.56	54.17 ± 6.18	57.87 ± 6.94
VPT-DEEP ($p = 10, d = 9$)	65.17 ± 14.38	51.28 ± 16.01	58.01 ± 20.26	55.13 ± 22.03	57.40 ± 17.56
FedAKT (CLIP)	8.54 ± 3.50	9.44 ± 3.62	10.50 ± 3.55	8.02 ± 3.47	9.13 ± 3.53
FedGH (CLIP)	38.32 ± 3.30	40.17 ± 3.40	39.01 ± 3.35	39.66 ± 3.31	39.29 ± 3.34
FedLT	82.86 ± 0.41	92.15 ± 0.70	91.19 ± 0.43	88.78 ± 0.42	88.74 ± 0.14
FedAdapter	94.42 ± 0.92	96.55 ± 0.56	96.55 ± 0.40	96.87 ± 0.24	96.10 ± 0.47
FedOT (All Global)	92.79 ± 5.48	92.31 ± 3.81	92.47 ± 4.55	95.68 ± 2.93	93.31 ± 3.04
FedOT	93.74 ± 2.50	95.35 ± 0.73	92.31 ± 4.33	95.51 ± 0.28	94.23 ± 0.84
FedOT+(B)	$\mathbf{97.14 \pm 1.52}$	$\mathbf{97.64 \pm 0.29}$	$\mathbf{97.56 \pm 0.58}$	97.70 ± 0.18	$\mathbf{97.51 \pm 0.48}$

Table 4. The generalization accuracy of FEMNIST.

Method	A	C	P	S	AVG
Zero Shot	95.35 ± 0.98	96.94 ± 0.81	99.60 ± 0.46	85.56 ± 0.99	94.36 ± 0.48
FedCLIP	95.36 ± 1.36	96.58 ± 0.57	99.30 ± 0.46	85.73 ± 1.71	94.24 ± 0.71
PromptFL ($M = 4$)	95.27 ± 0.75	97.22 ± 0.57	99.50 ± 0.46	84.25 ± 1.95	94.06 ± 0.61
PromptFL ($M = 16$)	94.05 ± 1.10	97.15 ± 0.65	99.40 ± 0.30	85.94 ± 1.41	94.14 ± 0.57
CoCoOp (All Global)	95.27 ± 2.19	97.08 ± 1.05	98.90 ± 0.62	85.69 ± 0.96	94.24 ± 1.11
CoCoOp (Meta-Net Random Init)	62.10 ± 13.94	77.35 ± 8.04	62.77 ± 15.22	72.57 ± 8.93	68.70 ± 11.29
CoCoOp (Server CoOp)	85.74 ± 0.61	74.58 ± 25.97	89.42 ± 6.81	80.21 ± 11.88	82.49 ± 8.16
VPT-SHALLOW ($p = 2$)	95.84 ± 1.27	97.44 ± 1.07	99.70 ± 0.52	85.01 ± 1.28	94.50 ± 0.59
VPT-SHALLOW ($p = 10$)	95.19 ± 1.48	97.58 ± 0.62	$\mathbf{99.80 \pm 0.35}$	$\mathbf{87.09 \pm 1.12}$	$\mathbf{94.92 \pm 0.57}$
VPT-DEEP ($p = 2, d = 2$)	95.28 ± 1.63	97.58 ± 1.01	99.60 ± 0.46	84.93 ± 1.77	94.35 ± 0.95
VPT-DEEP ($p = 10, d = 9$)	95.11 ± 1.12	97.15 ± 1.23	99.70 ± 0.30	84.63 ± 1.52	94.15 ± 0.78
FedAKT (CLIP)	13.58 ± 1.15	14.39 ± 1.25	15.03 ± 1.20	13.28 ± 1.16	14.07 ± 1.19
FedGH (CLIP)	54.57 ± 1.60	55.99 ± 1.65	55.56 ± 1.62	54.92 ± 1.59	55.23 ± 1.62
FedLT	95.11 ± 0.86	97.08 ± 0.40	99.60 ± 0.26	85.48 ± 0.70	94.32 ± 0.40
FedAdapter	95.60 ± 0.88	97.35 ± 0.47	99.60 ± 0.26	84.95 ± 0.76	94.37 ± 0.45
FedOT (All Global)	96.09 ± 1.12	97.37 ± 0.62	99.60 ± 0.46	85.14 ± 1.58	94.55 ± 0.54
FedOT	$\mathbf{96.33 \pm 0.89}$	$\mathbf{98.15 \pm 0.75}$	99.60 ± 0.46	85.73 ± 1.47	94.53 ± 0.35
FedOT+(B)	$\mathbf{96.33 \pm 0.89}$	$\mathbf{98.15 \pm 0.75}$	99.60 ± 0.46	85.73 ± 1.47	94.53 ± 0.35

Table 5. The generalization accuracy of PACS.

Method	A	C	P	R	AVG
Zero Shot	77.59 ± 1.33	65.29 ± 0.89	86.89 ± 1.61	86.83 ± 0.74	79.15 ± 0.40
FedCLIP	77.32 ± 1.07	65.79 ± 1.56	85.94 ± 1.42	86.68 ± 1.99	78.93 ± 1.06
PromptFL ($M = 4$)	77.39 ± 2.00	68.12 ± 1.00	87.86 ± 0.51	88.33 ± 0.86	80.42 ± 0.80
PromptFL ($M = 16$)	$\mathbf{79.86 \pm 0.43}$	68.50 ± 1.60	88.47 ± 0.46	$\mathbf{89.13 \pm 0.88}$	$\mathbf{81.49 \pm 0.45}$
CoCoOp (All Global)	77.73 ± 1.89	65.56 ± 2.83	88.39 ± 0.85	88.29 ± 1.20	79.99 ± 0.90
CoCoOp (Meta-Net Random Init)	60.96 ± 8.08	50.13 ± 2.75	67.53 ± 4.39	69.00 ± 4.67	61.91 ± 4.84
CoCoOp (Server CoOp)	72.00 ± 1.46	73.00 ± 1.46	73.50 ± 1.46	72.42 ± 1.46	72.73 ± 1.46
VPT-SHALLOW ($p = 2$)	78.42 ± 2.10	67.47 ± 1.00	87.98 ± 1.30	87.56 ± 1.33	80.36 ± 0.74
VPT-SHALLOW ($p = 10$)	77.05 ± 2.17	66.63 ± 1.57	86.77 ± 1.05	86.68 ± 1.49	79.28 ± 0.98
VPT-DEEP ($p = 2, d = 2$)	78.01 ± 1.40	67.89 ± 1.25	86.89 ± 1.88	87.52 ± 0.46	80.08 ± 0.35
VPT-DEEP ($p = 10, d = 9$)	77.11 ± 1.09	65.71 ± 1.16	85.79 ± 2.36	87.56 ± 1.15	79.05 ± 0.37
FedAKT (CLIP)	0.77 ± 0.08	0.93 ± 0.10	0.85 ± 0.09	0.93 ± 0.09	0.87 ± 0.09
FedGH (CLIP)	1.81 ± 0.12	1.99 ± 0.14	1.99 ± 0.13	1.85 ± 0.13	1.91 ± 0.13
FedLT	78.49 ± 1.19	67.05 ± 0.50	88.09 ± 0.73	88.67 ± 0.27	80.57 ± 0.31
FedAdapter	79.45 ± 0.12	$\mathbf{68.71 \pm 1.06}$	88.26 ± 0.63	88.90 ± 0.47	81.33 ± 0.26
FedOT (All Global)	79.38 ± 2.09	68.54 ± 1.72	87.90 ± 1.20	88.44 ± 1.06	81.07 ± 0.85
FedOT	79.11 ± 1.27	66.93 ± 1.92	$\mathbf{88.61 \pm 1.05}$	87.71 ± 0.90	80.59 ± 0.47
FedOT+(B)	79.11 ± 1.27	66.93 ± 1.92	$\mathbf{88.61 \pm 1.05}$	87.71 ± 0.90	80.59 ± 0.47

Table 6. The generalization accuracy of Office-Home.

Method	C	L	S	V	AVG
Zero Shot	$\mathbf{100.00 \pm 0.00}$	$\mathbf{65.22 \pm 2.72}$	70.96 ± 0.33	84.79 ± 1.07	80.59 ± 0.96
FedCLIP	99.29 ± 0.71	63.40 ± 0.54	73.48 ± 1.40	85.63 ± 1.71	80.45 ± 0.86
PromptFL ($M = 4$)	99.06 ± 0.54	62.96 ± 1.52	78.96 ± 1.30	86.17 ± 1.56	81.79 ± 0.15
PromptFL ($M = 16$)	99.18 ± 0.54	60.20 ± 1.21	77.75 ± 2.51	83.90 ± 2.20	80.26 ± 0.77
CoCoOp (All Global)	99.88 ± 0.20	62.65 ± 4.18	74.34 ± 4.05	84.05 ± 4.20	80.23 ± 1.96
CoCoOp (Meta-Net Random Init)	67.02 ± 9.08	49.46 ± 6.01	52.75 ± 9.02	55.41 ± 7.63	56.16 ± 6.06
CoCoOp (Server CoOp)	89.28 ± 18.57	53.99 ± 12.15	60.36 ± 3.58	74.77 ± 0.17	69.60 ± 7.67
VPT-SHALLOW ($p = 2$)	99.65 ± 0.36	63.03 ± 1.96	75.15 ± 4.36	85.73 ± 1.07	80.89 ± 1.02
VPT-SHALLOW ($p = 10$)	99.65 ± 0.36	64.78 ± 1.47	76.63 ± 1.41	85.33 ± 0.68	81.60 ± 0.16
VPT-DEEP ($p = 2, d = 2$)	99.41 ± 0.21	61.83 ± 3.58	76.58 ± 5.21	85.93 ± 0.74	80.94 ± 1.70
VPT-DEEP ($p = 10, d = 9$)	99.17 ± 0.20	61.40 ± 0.82	77.75 ± 1.88	86.42 ± 0.31	81.18 ± 0.19
FedAKT (CLIP)	11.21 ± 1.25	12.24 ± 1.35	11.05 ± 1.30	11.66 ± 1.30	11.54 ± 1.30
FedGH (CLIP)	47.12 ± 2.00	48.12 ± 2.05	46.26 ± 2.02	48.22 ± 2.01	47.43 ± 2.02
FedLT	99.53 ± 0.12	64.28 ± 0.17	80.08 ± 0.94	84.94 ± 0.87	82.21 ± 0.44
FedAdapter	96.88 ± 0.16	62.71 ± 0.97	77.52 ± 1.53	82.20 ± 0.70	79.83 ± 0.81
FedOT (All Global)	99.65 ± 0.36	61.77 ± 2.74	76.73 ± 3.98	$\mathbf{87.16 \pm 2.14}$	81.33 ± 1.19
FedOT	99.65 ± 0.20	65.03 ± 0.17	$\mathbf{80.33 \pm 0.99}$	86.22 ± 1.11	$\mathbf{82.81 \pm 0.47}$
FedOT+(B)	99.65 ± 0.20	65.03 ± 0.17	$\mathbf{80.33 \pm 0.99}$	86.22 ± 1.11	$\mathbf{82.81 \pm 0.47}$

Table 7. The generalization accuracy of VLCS.

Method	L38	L43	L46	L100	AVG
Zero Shot	13.35 ± 0.76	29.90 ± 1.95	19.59 ± 1.31	13.47 ± 0.52	19.08 ± 0.26
FedCLIP	33.28 ± 4.81	30.98 ± 3.84	30.67 ± 1.80	$\mathbf{43.56 \pm 9.00}$	34.62 ± 1.63
PromptFL ($M = 4$)	38.54 ± 2.98	33.02 ± 2.20	36.99 ± 2.26	41.47 ± 4.42	37.51 ± 1.92
PromptFL ($M = 16$)	37.31 ± 7.34	34.43 ± 3.48	35.28 ± 3.40	36.96 ± 1.72	35.99 ± 2.50
CoCoOp (All Global)	38.33 ± 7.94	29.82 ± 2.76	38.84 ± 1.38	37.61 ± 1.36	36.15 ± 1.44
CoCoOp (Meta-Net Random Init)	44.65 ± 2.62	26.09 ± 5.39	25.52 ± 4.81	23.18 ± 2.98	29.86 ± 1.21
CoCoOp (Server CoOp)	34.15 ± 18.23	26.34 ± 5.58	24.16 ± 5.63	20.03 ± 3.40	26.17 ± 5.38
VPT-SHALLOW ($p = 2$)	40.52 ± 5.47	$\mathbf{39.86 \pm 2.56}$	32.49 ± 6.35	26.19 ± 1.36	34.76 ± 0.45
VPT-SHALLOW ($p = 10$)	$\mathbf{45.79 \pm 2.19}$	34.01 ± 6.58	35.06 ± 4.36	19.42 ± 3.51	33.57 ± 3.40
VPT-DEEP ($p = 2, d = 2$)	40.26 ± 0.92	32.85 ± 2.75	36.04 ± 3.21	28.14 ± 2.29	34.32 ± 0.71
VPT-DEEP ($p = 10, d = 9$)	15.31 ± 9.09	32.52 ± 2.51	27.54 ± 2.96	5.30 ± 3.08	20.17 ± 1.92
FedAKT (CLIP)	10.21 ± 5.70	14.78 ± 5.80	13.01 ± 5.85	11.00 ± 5.77	12.25 ± 5.78
FedGH (CLIP)	7.01 ± 2.75	7.49 ± 2.85	8.12 ± 2.80	6.50 ± 2.80	7.28 ± 2.80
FedLT	29.27 ± 2.30	39.31 ± 2.41	33.82 ± 3.65	17.09 ± 6.07	29.87 ± 2.29
FedAdapter	34.92 ± 3.21	37.09 ± 1.58	$\mathbf{41.14 \pm 0.92}$	40.03 ± 0.33	$\mathbf{38.30 \pm 0.89}$
FedOT (All Global)	37.75 ± 4.34	32.48 ± 1.48	34.54 ± 2.82	26.02 ± 1.74	32.70 ± 1.02
FedOT	30.65 ± 12.20	25.43 ± 4.12	35.17 ± 6.28	27.08 ± 4.26	29.58 ± 4.52
FedOT+(B)	38.59 ± 3.99	34.33 ± 1.32	41.10 ± 0.91	37.61 ± 0.21	37.91 ± 0.90

Table 8. The generalization accuracy of TerraIncognita.

F.2. Personalization

Method	Target	0	25	26	AVG
Zero Shot	33	48.40 ± 1.47	45.19 ± 2.50	48.24 ± 2.73	47.28 ± 0.73
FedCLIP		70.03 ± 8.40	72.44 ± 11.70	69.23 ± 9.38	70.57 ± 9.77
PromptFL ($M = 4$)		96.96 ± 0.74	95.83 ± 1.94	96.96 ± 2.00	96.58 ± 1.49
PromptFL ($M = 16$)		**97.12 ± 0.97**	97.60 ± 0.48	**98.24 ± 1.47**	**97.65 ± 0.72**
CoCoOp (All Global)		88.46 ± 7.26	89.43 ± 6.68	87.50 ± 9.36	88.46 ± 7.63
CoCoOp (Meta-Net Random Init)		87.18 ± 6.13	87.34 ± 7.28	85.42 ± 10.59	86.64 ± 8.00
CoCoOp (Server CoOp)		94.39 ± 1.94	94.71 ± 1.44	96.32 ± 1.39	95.14 ± 1.09
VPT-SHALLOW ($p = 2$)		87.98 ± 2.09	90.06 ± 2.42	89.75 ± 2.37	89.26 ± 2.24
VPT-SHALLOW ($p = 10$)		94.07 ± 3.74	94.87 ± 2.22	95.03 ± 3.61	94.66 ± 3.13
VPT-DEEP ($p = 2, d = 2$)		90.54 ± 5.46	91.35 ± 5.91	91.51 ± 6.09	91.13 ± 5.80
VPT-DEEP ($p = 10, d = 9$)		86.54 ± 1.92	88.78 ± 1.95	89.42 ± 2.89	88.25 ± 2.25
FedAKT (CLIP)		49.82 ± 9.52	59.78 ± 9.88	54.80 ± 9.73	54.80 ± 9.71
FedGH (CLIP)		37.57 ± 1.00	39.88 ± 1.10	39.55 ± 1.02	39.00 ± 1.04
FedLT		95.99 ± 0.58	97.12 ± 0.56	96.95 ± 0.80	96.69 ± 0.65
FedAdapter		96.07 ± 0.66	96.96 ± 0.32	96.55 ± 0.21	96.53 ± 0.40
FedOT (All Global)		94.71 ± 5.46	94.87 ± 4.31	95.20 ± 3.75	94.93 ± 4.51
FedOT (All Local)		92.63 ± 1.95	89.10 ± 3.73	82.85 ± 7.67	88.19 ± 1.20
FedOT		95.99 ± 0.73	96.47 ± 0.28	96.80 ± 1.95	96.42 ± 0.99
FedOT+(B)		96.33 ± 0.48	**97.62 ± 0.45**	97.50 ± 0.47	97.15 ± 0.38
		33	25	26	AVG
Zero Shot	0	50.88 ± 3.42	48.40 ± 1.47	48.24 ± 2.73	48.10 ± 1.33
FedCLIP		67.75 ± 9.90	73.56 ± 10.64	70.03 ± 9.20	70.45 ± 9.81
PromptFL ($M = 4$)		95.10 ± 1.47	96.31 ± 0.74	96.47 ± 0.56	95.96 ± 0.63
PromptFL ($M = 16$)		95.64 ± 1.03	96.64 ± 1.28	**98.24 ± 0.73**	96.84 ± 0.74
CoCoOp (All Global)		85.71 ± 10.49	89.58 ± 7.55	86.86 ± 7.93	87.39 ± 8.65
CoCoOp (Meta-Net Random Init)		84.90 ± 11.88	87.82 ± 8.07	84.94 ± 10.82	85.89 ± 10.06
CoCoOp (Server CoOp)		95.92 ± 1.23	95.51 ± 1.54	95.35 ± 1.00	95.59 ± 0.93
VPT-SHALLOW ($p = 2$)		90.07 ± 5.42	92.79 ± 3.82	93.27 ± 3.37	92.04 ± 4.19
VPT-SHALLOW ($p = 10$)		87.75 ± 2.16	89.10 ± 3.13	89.10 ± 3.54	88.65 ± 1.49
VPT-DEEP ($p = 2, d = 2$)		91.56 ± 5.17	91.19 ± 5.74	91.51 ± 6.28	91.42 ± 5.51
VPT-DEEP ($p = 10, d = 9$)		88.03 ± 2.25	88.62 ± 2.65	89.58 ± 2.90	88.74 ± 1.29
FedAKT (CLIP)		58.21 ± 9.60	51.19 ± 10.00	55.14 ± 9.53	55.18 ± 9.71
FedGH (CLIP)		40.80 ± 1.10	38.89 ± 0.98	40.31 ± 1.04	40.00 ± 1.04
FedLT		96.19 ± 0.14	96.47 ± 0.58	96.96 ± 0.43	96.54 ± 0.38
FedAdapter		95.99 ± 0.45	96.56 ± 0.56	96.79 ± 0.43	96.45 ± 0.48
FedOT (All Global)		90.07 ± 5.10	92.95 ± 3.61	94.07 ± 4.36	92.36 ± 4.31
FedOT (All Local)		86.39 ± 0.47	89.10 ± 3.73	82.85 ± 7.67	86.12 ± 1.19
FedOT		96.19 ± 1.84	95.83 ± 1.00	97.92 ± 0.55	96.65 ± 1.13
FedOT+(B)		**98.19 ± 0.32**	**97.31 ± 0.48**	97.45 ± 0.53	**97.65 ± 0.52**

Table 9. The personalization accuracy of FEMNIST (Target 33, 0).

Method	Target	33	0	26	AVG
Zero Shot	25	50.88 ± 3.42	48.40 ± 1.47	48.24 ± 2.73	49.17 ± 1.71
FedCLIP		67.75 ± 8.19	71.96 ± 7.93	70.35 ± 10.67	70.02 ± 8.93
PromptFL ($M = 4$)		95.92 ± 0.70	96.63 ± 0.49	97.28 ± 0.74	96.61 ± 0.12
PromptFL ($M = 16$)		96.33 ± 1.08	96.80 ± 0.56	97.92 ± 0.55	97.01 ± 0.44
CoCoOp (All Global)		84.63 ± 12.12	85.58 ± 10.10	85.90 ± 10.01	85.37 ± 10.72
CoCoOp (Meta-Net Random Init)		85.17 ± 11.83	83.33 ± 9.86	84.62 ± 11.34	84.37 ± 10.95
CoCoOp (Server CoOp)		96.33 ± 1.23	94.55 ± 1.00	95.67 ± 1.66	95.52 ± 0.62
VPT-SHALLOW ($p = 2$)		88.03 ± 1.93	88.62 ± 1.21	89.91 ± 2.10	88.85 ± 0.43
VPT-SHALLOW ($p = 10$)		92.65 ± 5.02	93.27 ± 2.68	94.71 ± 3.63	93.54 ± 3.70
VPT-DEEP ($p = 2, d = 2$)		87.89 ± 2.01	87.34 ± 2.16	90.23 ± 2.47	88.49 ± 0.99
VPT-DEEP ($p = 10, d = 9$)		88.03 ± 2.62	87.18 ± 2.16	89.90 ± 3.15	88.37 ± 0.97
FedAKT (CLIP)		45.17 ± 9.80	65.63 ± 9.50	54.65 ± 9.83	55.15 ± 9.71
FedGH (CLIP)		39.13 ± 1.08	40.67 ± 1.00	39.60 ± 1.04	39.80 ± 1.04
FedLT		96.05 ± 0.36	91.19 ± 0.43	**98.08 ± 0.28**	95.11 ± 0.36
FedAdapter		95.58 ± 0.76	96.55 ± 0.40	97.36 ± 0.37	96.50 ± 0.51
FedOT (All Global)		90.21 ± 4.95	92.15 ± 4.31	93.11 ± 4.01	91.82 ± 4.40
FedOT (All Local)		86.39 ± 0.47	92.63 ± 1.95	89.10 ± 3.73	89.37 ± 1.90
FedOT		89.80 ± 4.60	91.50 ± 4.09	92.79 ± 4.88	91.36 ± 4.48
FedOT+(B)		**97.50 ± 0.31**	**98.35 ± 0.57**	98.03 ± 0.22	**97.96 ± 0.42**
		33	0	25	AVG
Zero Shot	26	50.88 ± 3.42	48.40 ± 1.47	45.19 ± 2.50	48.16 ± 1.98
FedCLIP		68.71 ± 9.43	72.28 ± 6.74	71.79 ± 13.74	70.93 ± 9.81
PromptFL ($M = 4$)		95.92 ± 1.63	96.96 ± 0.28	96.95 ± 1.00	96.61 ± 0.66
PromptFL ($M = 16$)		95.64 ± 1.03	97.44 ± 0.74	96.96 ± 0.28	96.68 ± 0.41
CoCoOp (All Global)		86.40 ± 10.02	89.10 ± 5.93	88.94 ± 6.36	88.15 ± 7.37
CoCoOp (Meta-Net Random Init)		84.62 ± 9.06	85.74 ± 9.40	86.86 ± 8.47	85.74 ± 8.69
CoCoOp (Server CoOp)		95.92 ± 1.47	94.71 ± 0.83	95.19 ± 1.44	95.27 ± 0.50
VPT-SHALLOW ($p = 2$)		95.24 ± 3.66	94.07 ± 5.70	93.75 ± 6.25	94.35 ± 5.20
VPT-SHALLOW ($p = 10$)		91.84 ± 4.60	94.07 ± 3.64	94.87 ± 2.73	93.59 ± 3.63
VPT-DEEP ($p = 2, d = 2$)		93.33 ± 5.95	93.75 ± 5.46	94.71 ± 4.17	93.93 ± 5.18
VPT-DEEP ($p = 10, d = 9$)		88.03 ± 2.59	86.86 ± 1.54	88.78 ± 2.27	87.89 ± 0.66
FedAKT (CLIP)		52.91 ± 10.10	56.91 ± 9.60	54.97 ± 9.43	54.93 ± 9.71
FedGH (CLIP)		40.64 ± 1.10	38.05 ± 1.00	39.51 ± 1.02	39.40 ± 1.04
FedLT		95.78 ± 0.36	96.15 ± 0.83	88.78 ± 0.42	93.57 ± 0.54
FedAdapter		95.44 ± 0.60	96.79 ± 0.08	96.87 ± 0.24	96.37 ± 0.31
FedOT (All Global)		92.52 ± 4.55	94.39 ± 3.92	95.52 ± 3.20	94.14 ± 3.87
FedOT (All Local)		86.39 ± 0.47	92.63 ± 1.95	89.10 ± 3.73	89.37 ± 1.90
FedOT		96.63 ± 0.49	96.31 ± 1.47	95.51 ± 0.28	96.15 ± 0.75
FedOT+(B)		**98.30 ± 0.66**	**97.60 ± 0.72**	**97.50 ± 0.52**	**97.80 ± 0.62**

Table 10. The personalization accuracy of FEMNIST (Target 25, 26).

Method	Target	C	P	S	AVG
Zero Shot		96.94 ± 0.81	99.60 ± 0.46	85.56 ± 0.99	94.03 ± 0.38
FedCLIP		97.01 ± 1.67	99.70 ± 0.30	89.94 ± 0.46	95.55 ± 0.62
PromptFL ($M = 4$)		98.36 ± 0.81	**99.80 ± 0.35**	92.78 ± 1.02	96.98 ± 0.56
PromptFL ($M = 16$)		98.15 ± 0.50	99.70 ± 0.30	92.74 ± 0.44	96.86 ± 0.23
CoCoOp (All Global)		98.08 ± 0.98	99.30 ± 0.46	91.04 ± 1.91	96.14 ± 1.01
CoCoOp (Meta-Net Random Init)		98.15 ± 1.18	99.70 ± 0.30	92.02 ± 1.54	96.62 ± 0.66
CoCoOp (Server CoOp)		97.86 ± 0.86	99.70 ± 0.30	91.42 ± 0.51	96.33 ± 0.09
VPT-SHALLOW ($p = 2$)		98.58 ± 0.66	99.70 ± 0.52	91.85 ± 0.77	96.71 ± 0.38
VPT-SHALLOW ($p = 10$)		98.58 ± 0.54	**99.80 ± 0.35**	91.25 ± 0.92	96.54 ± 0.39
VPT-DEEP ($p = 2, d = 2$)	A	98.57 ± 0.89	99.70 ± 0.30	92.14 ± 0.81	96.81 ± 0.49
VPT-DEEP ($p = 10, d = 9$)		97.79 ± 0.69	99.50 ± 0.62	91.25 ± 0.96	96.18 ± 0.45
FedAKT (CLIP)		71.21 ± 3.50	75.00 ± 3.80	72.82 ± 3.74	73.00 ± 3.68
FedGH (CLIP)		61.86 ± 1.50	64.90 ± 1.70	63.44 ± 1.57	63.40 ± 1.59
FedLT		**99.22 ± 0.26**	99.60 ± 0.26	93.25 ± 0.22	97.36 ± 0.25
FedAdapter		98.68 ± 0.36	99.70 ± 0.30	92.67 ± 0.39	97.02 ± 0.35
FedOT (All Global)		98.65 ± 0.45	99.70 ± 0.52	92.10 ± 0.67	96.82 ± 0.24
FedOT (All Local)		98.93 ± 0.57	**99.80 ± 0.35**	92.53 ± 0.82	97.09 ± 0.34
FedOT		99.15 ± 0.25	99.70 ± 0.30	**93.50 ± 0.29**	**97.45 ± 0.28**
FedOT+(B)		99.15 ± 0.25	99.70 ± 0.30	**93.50 ± 0.29**	**97.45 ± 0.28**
		A	P	S	AVG
Zero Shot		95.35 ± 0.98	99.60 ± 0.46	85.56 ± 0.99	93.51 ± 0.37
FedCLIP		95.52 ± 1.49	99.70 ± 0.30	90.66 ± 0.63	95.29 ± 0.68
PromptFL ($M = 4$)		97.07 ± 0.49	99.70 ± 0.30	92.95 ± 0.63	96.57 ± 0.21
PromptFL ($M = 16$)		96.66 ± 0.38	99.70 ± 0.30	92.57 ± 0.19	96.31 ± 0.22
CoCoOp (All Global)		96.58 ± 0.98	99.30 ± 0.46	91.46 ± 1.96	95.78 ± 0.79
CoCoOp (Meta-Net Random Init)		96.33 ± 0.88	99.60 ± 0.35	91.93 ± 1.79	95.96 ± 0.46
CoCoOp (Server CoOp)		96.09 ± 1.69	99.70 ± 0.30	91.51 ± 0.27	95.76 ± 0.54
VPT-SHALLOW ($p = 2$)		96.01 ± 0.51	99.70 ± 0.52	91.97 ± 1.11	95.89 ± 0.33
VPT-SHALLOW ($p = 10$)		96.01 ± 0.99	99.60 ± 0.69	91.30 ± 1.20	95.63 ± 0.39
VPT-DEEP ($p = 2, d = 2$)	C	96.01 ± 0.62	99.70 ± 0.30	92.02 ± 0.49	95.91 ± 0.18
VPT-DEEP ($p = 10, d = 9$)		96.25 ± 0.79	99.50 ± 0.35	91.63 ± 0.97	95.79 ± 0.16
FedAKT (CLIP)		74.05 ± 3.60	71.10 ± 3.90	74.30 ± 3.54	73.15 ± 3.68
FedGH (CLIP)		64.13 ± 1.60	62.10 ± 1.55	64.87 ± 1.62	63.70 ± 1.59
FedLT		**97.15 ± 0.35**	99.60 ± 0.26	93.38 ± 0.29	96.71 ± 0.30
FedAdapter		97.03 ± 0.50	99.60 ± 0.26	92.92 ± 0.29	96.52 ± 0.35
FedOT (All Global)		96.82 ± 0.85	99.70 ± 0.30	92.57 ± 0.29	96.36 ± 0.25
FedOT (All Local)		96.90 ± 1.49	**99.80 ± 0.35**	92.53 ± 0.82	96.41 ± 0.46
FedOT		97.07 ± 0.38	99.60 ± 0.26	**93.59 ± 0.30**	**96.75 ± 0.31**
FedOT+(B)		97.07 ± 0.38	99.60 ± 0.26	**93.59 ± 0.30**	**96.75 ± 0.31**

Table 11. The personalization accuracy of PACS (Target A, C).

Method	Target	A	C	S	AVG
Zero Shot	P	95.35 ± 0.98	96.94 ± 0.81	85.56 ± 0.99	92.62 ± 0.78
FedCLIP		95.11 ± 1.36	98.00 ± 1.73	89.94 ± 0.26	94.35 ± 0.98
PromptFL ($M = 4$)		96.82 ± 0.65	98.29 ± 0.74	92.27 ± 1.28	95.80 ± 0.50
PromptFL ($M = 16$)		96.41 ± 1.21	98.43 ± 0.66	92.65 ± 0.32	95.83 ± 0.71
CoCoOp (All Global)		96.58 ± 0.89	98.08 ± 0.77	91.04 ± 1.80	95.23 ± 0.78
CoCoOp (Meta-Net Random Init)		96.50 ± 0.93	98.29 ± 0.98	91.85 ± 1.60	95.54 ± 0.92
CoCoOp (Server CoOp)		96.01 ± 1.77	97.86 ± 0.86	91.46 ± 0.44	95.11 ± 0.76
VPT-SHALLOW ($p = 2$)		96.01 ± 1.10	98.58 ± 0.86	91.72 ± 0.92	95.43 ± 0.78
VPT-SHALLOW ($p = 10$)		96.17 ± 1.49	98.43 ± 0.44	91.04 ± 0.91	95.21 ± 0.68
VPT-DEEP ($p = 2, d = 2$)		96.09 ± 0.25	98.43 ± 0.86	91.97 ± 0.78	95.50 ± 0.56
VPT-DEEP ($p = 10, d = 9$)		96.09 ± 1.36	97.65 ± 0.93	91.25 ± 0.97	95.00 ± 0.49
FedAKT (CLIP)		71.75 ± 3.70	75.65 ± 3.50	72.62 ± 3.84	73.34 ± 3.68
FedGH (CLIP)		63.75 ± 1.70	65.22 ± 1.50	64.11 ± 1.57	64.36 ± 1.59
FedLT		96.90 ± 0.45	**99.60 ± 0.26**	93.42 ± 0.23	96.64 ± 0.31
FedAdapter		96.96 ± 0.52	**99.60 ± 0.26**	92.83 ± 0.24	96.46 ± 0.34
FedOT (All Global)		96.57 ± 1.27	98.65 ± 0.45	92.31 ± 0.63	95.84 ± 0.72
FedOT (All Local)		96.90 ± 1.49	98.93 ± 0.57	92.53 ± 0.82	96.12 ± 0.76
FedOT		**97.15 ± 0.71**	99.15 ± 0.12	**93.71 ± 0.35**	**96.67 ± 0.39**
FedOT+(B)		**97.15 ± 0.71**	99.15 ± 0.12	**93.71 ± 0.35**	**96.67 ± 0.39**
		A	C	P	AVG
Zero Shot	S	95.35 ± 0.98	96.94 ± 0.81	99.60 ± 0.46	97.30 ± 0.47
FedCLIP		96.58 ± 1.76	97.87 ± 1.40	99.80 ± 0.17	98.08 ± 1.07
PromptFL ($M = 4$)		96.74 ± 1.59	98.79 ± 0.69	**99.90 ± 0.17**	98.48 ± 0.69
PromptFL ($M = 16$)		96.82 ± 1.12	98.57 ± 0.89	**99.90 ± 0.17**	98.43 ± 0.61
CoCoOp (All Global)		96.49 ± 1.34	98.36 ± 1.08	99.70 ± 0.52	98.18 ± 0.59
CoCoOp (Meta-Net Random Init)		96.25 ± 0.93	98.15 ± 1.08	99.60 ± 0.35	98.00 ± 0.54
CoCoOp (Server CoOp)		96.25 ± 1.55	98.15 ± 0.81	99.80 ± 0.17	98.06 ± 0.71
VPT-SHALLOW ($p = 2$)		96.58 ± 1.60	98.79 ± 0.65	99.70 ± 0.30	98.36 ± 0.58
VPT-SHALLOW ($p = 10$)		95.76 ± 1.10	98.65 ± 0.62	99.70 ± 0.30	98.04 ± 0.50
VPT-DEEP ($p = 2, d = 2$)		95.85 ± 1.12	98.51 ± 0.57	99.60 ± 0.17	97.98 ± 0.49
VPT-DEEP ($p = 10, d = 9$)		96.17 ± 1.23	97.65 ± 0.77	99.60 ± 0.17	97.81 ± 0.46
FedAKT (CLIP)		73.25 ± 3.60	72.25 ± 3.90	72.73 ± 3.54	72.75 ± 3.68
FedGH (CLIP)		64.33 ± 1.50	61.05 ± 1.70	61.52 ± 1.58	62.30 ± 1.59
FedLT		96.82 ± 0.85	**99.60 ± 0.20**	85.48 ± 0.70	93.97 ± 0.58
FedAdapter		96.66 ± 0.71	**99.60 ± 0.26**	84.95 ± 0.76	93.74 ± 0.58
FedOT (All Global)		96.66 ± 1.25	98.79 ± 0.54	99.70 ± 0.30	98.38 ± 0.49
FedOT (All Local)		96.90 ± 1.49	98.93 ± 0.57	99.80 ± 0.35	**98.54 ± 0.55**
FedOT		**97.15 ± 0.92**	98.72 ± 0.44	99.60 ± 0.20	98.49 ± 0.52
FedOT+(B)		**97.15 ± 0.92**	98.72 ± 0.44	99.60 ± 0.20	98.49 ± 0.52

Table 12. The personalization accuracy of PACS (Target P, S).

Method	Target	C	P	R	AVG
Zero Shot		65.29 ± 0.89	86.89 ± 1.61	86.83 ± 0.74	79.67 ± 0.32
FedCLIP		69.11 ± 0.70	88.32 ± 1.43	87.60 ± 0.99	81.67 ± 0.89
PromptFL ($M = 4$)		74.88 ± 2.43	90.91 ± 1.09	89.40 ± 0.81	85.06 ± 0.88
PromptFL ($M = 16$)		74.91 ± 1.38	91.02 ± 0.72	89.17 ± 0.66	85.03 ± 0.42
CoCoOp (All Global)		69.11 ± 3.52	88.95 ± 0.69	88.37 ± 0.96	82.14 ± 1.05
CoCoOp (Meta-Net Random Init)		71.25 ± 2.25	89.10 ± 1.02	88.37 ± 1.16	82.91 ± 0.58
CoCoOp (Server CoOp)		74.88 ± 2.43	90.91 ± 1.09	89.40 ± 0.81	85.06 ± 0.88
VPT-SHALLOW ($p = 2$)		70.10 ± 1.60	89.59 ± 0.86	88.37 ± 0.87	82.69 ± 0.32
VPT-SHALLOW ($p = 10$)		65.64 ± 3.74	88.01 ± 0.35	86.91 ± 2.04	80.19 ± 1.26
VPT-DEEP ($p = 2, d = 2$)	A	69.11 ± 1.40	89.59 ± 1.05	88.06 ± 0.83	82.25 ± 0.10
VPT-DEEP ($p = 10, d = 9$)		65.67 ± 1.86	87.41 ± 1.24	85.65 ± 1.79	79.58 ± 0.65
FedAKT (CLIP)		40.23 ± 0.42	65.42 ± 0.55	60.25 ± 0.25	55.30 ± 0.14
FedGH (CLIP)		6.91 ± 2.52	7.12 ± 2.88	6.97 ± 1.82	7.00 ± 1.21
FedLT		80.18 ± 0.83	93.16 ± 0.74	90.36 ± 0.50	87.90 ± 0.24
FedAdapter		78.39 ± 1.60	92.14 ± 0.47	89.80 ± 0.60	86.78 ± 0.34
FedOT (All Global)		73.81 ± 1.67	91.10 ± 0.81	88.98 ± 1.41	84.63 ± 0.54
FedOT (All Local)		75.30 ± 1.79	92.22 ± 0.90	89.44 ± 0.90	85.65 ± 0.43
FedOT		**80.68 ± 0.79**	**93.24 ± 0.70**	**90.58 ± 0.34**	**88.17 ± 0.21**
FedOT+(B)		**80.68 ± 0.79**	**93.24 ± 0.70**	**90.58 ± 0.34**	**88.17 ± 0.21**
		A	P	R	AVG
Zero Shot		77.59 ± 1.33	86.89 ± 1.61	86.83 ± 0.74	83.77 ± 0.32
FedCLIP		78.91 ± 0.84	87.90 ± 0.95	87.64 ± 0.13	84.82 ± 0.57
PromptFL ($M = 4$)		82.34 ± 1.75	90.49 ± 0.97	89.70 ± 0.81	87.51 ± 0.59
PromptFL ($M = 16$)		81.72 ± 0.43	90.60 ± 1.30	90.01 ± 0.91	87.44 ± 0.25
CoCoOp (All Global)		78.62 ± 2.07	88.54 ± 0.83	88.10 ± 0.92	85.09 ± 0.52
CoCoOp (Meta-Net Random Init)		79.04 ± 1.38	88.88 ± 0.85	88.33 ± 0.92	85.41 ± 0.38
CoCoOp (Server CoOp)		82.34 ± 1.75	90.49 ± 0.97	89.70 ± 0.81	87.51 ± 0.59
VPT-SHALLOW ($p = 2$)		79.93 ± 0.78	89.25 ± 0.95	88.25 ± 1.26	85.81 ± 0.34
VPT-SHALLOW ($p = 10$)		77.32 ± 1.15	88.54 ± 0.53	87.49 ± 1.73	84.45 ± 0.35
VPT-DEEP ($p = 2, d = 2$)	C	80.00 ± 1.09	89.33 ± 0.83	88.14 ± 0.78	85.82 ± 0.21
VPT-DEEP ($p = 10, d = 9$)		76.63 ± 1.73	87.71 ± 1.63	85.88 ± 1.84	83.41 ± 0.55
FedAKT (CLIP)		45.52 ± 0.35	63.32 ± 0.51	57.11 ± 0.32	55.32 ± 0.16
FedGH (CLIP)		7.11 ± 2.01	6.95 ± 2.80	7.04 ± 0.83	7.03 ± 0.72
FedLT		84.19 ± 0.60	92.82 ± 0.60	**90.74 ± 0.54**	89.25 ± 0.20
FedAdapter		83.40 ± 0.52	92.22 ± 0.57	90.17 ± 0.36	88.60 ± 0.16
FedOT (All Global)		82.40 ± 1.75	90.49 ± 1.16	88.90 ± 1.24	87.27 ± 0.53
FedOT (All Local)		81.86 ± 0.54	92.22 ± 0.90	89.44 ± 0.90	87.84 ± 0.26
FedOT		**84.40 ± 0.96**	**92.97 ± 0.84**	90.62 ± 0.27	**89.33 ± 0.25**
FedOT+(B)		**84.40 ± 0.96**	**92.97 ± 0.84**	90.62 ± 0.27	**89.33 ± 0.25**

Table 13. The personalization accuracy of Office-Home (Target A, C).

Method	Target	A	C	R	AVG
Zero Shot	P	77.59 ± 1.33	65.29 ± 0.89	86.83 ± 0.74	76.57 ± 0.64
FedCLIP		78.08 ± 1.37	68.00 ± 0.82	87.49 ± 0.41	77.86 ± 0.09
PromptFL ($M = 4$)		82.34 ± 1.73	73.88 ± 1.69	89.44 ± 0.70	81.89 ± 0.90
PromptFL ($M = 16$)		83.09 ± 2.03	74.42 ± 1.98	89.82 ± 0.75	82.44 ± 1.41
CoCoOp (All Global)		78.08 ± 1.45	69.15 ± 2.78	88.40 ± 1.10	78.54 ± 1.52
CoCoOp (Meta-Net Random Init)		78.49 ± 1.34	71.02 ± 2.40	88.37 ± 1.09	79.29 ± 1.32
CoCoOp (Server CoOp)		82.34 ± 1.73	73.88 ± 1.69	89.44 ± 0.70	81.89 ± 0.90
VPT-SHALLOW ($p = 2$)		79.52 ± 0.52	70.14 ± 1.59	88.06 ± 1.32	79.24 ± 0.67
VPT-SHALLOW ($p = 10$)		77.46 ± 1.13	64.22 ± 2.92	86.37 ± 1.95	76.02 ± 1.34
VPT-DEEP ($p = 2, d = 2$)		79.66 ± 1.13	68.16 ± 0.99	87.52 ± 0.84	78.45 ± 0.56
VPT-DEEP ($p = 10, d = 9$)		76.77 ± 1.17	64.53 ± 2.31	85.11 ± 1.39	75.47 ± 1.31
FedAKT (CLIP)		62.48 ± 0.38	49.27 ± 0.40	56.45 ± 0.35	55.40 ± 0.19
FedGH (CLIP)		7.18 ± 2.02	6.82 ± 2.53	7.10 ± 1.75	7.03 ± 0.73
FedLT		**86.15 ± 1.02**	79.69 ± 0.91	90.51 ± 0.42	**85.45 ± 0.28**
FedAdapter		83.54 ± 0.60	78.39 ± 1.18	89.71 ± 0.70	83.88 ± 0.29
FedOT (All Global)		81.51 ± 1.26	73.12 ± 1.61	88.75 ± 1.32	81.13 ± 1.25
FedOT (All Local)		81.86 ± 0.54	75.30 ± 1.79	89.44 ± 0.90	82.20 ± 0.74
FedOT		84.67 ± 1.19	**80.26 ± 0.84**	**90.66 ± 0.40**	85.20 ± 0.29
FedOT+(B)		84.67 ± 1.19	**80.26 ± 0.84**	**90.66 ± 0.40**	85.20 ± 0.29
		A	C	P	AVG
Zero Shot	R	77.59 ± 1.33	65.29 ± 0.89	86.89 ± 1.61	76.59 ± 0.76
FedCLIP		77.25 ± 0.83	68.15 ± 0.64	87.67 ± 0.36	77.69 ± 0.02
PromptFL ($M = 4$)		82.41 ± 2.69	74.68 ± 2.07	90.75 ± 1.27	82.61 ± 1.27
PromptFL ($M = 16$)		82.27 ± 2.18	74.15 ± 2.16	90.68 ± 1.27	82.37 ± 1.39
CoCoOp (All Global)		78.49 ± 1.86	69.68 ± 3.09	88.99 ± 0.64	79.05 ± 1.50
CoCoOp (Meta-Net Random Init)		78.70 ± 1.86	71.29 ± 2.37	89.03 ± 0.91	79.67 ± 1.25
CoCoOp (Server CoOp)		82.41 ± 2.69	74.68 ± 2.07	90.75 ± 1.27	82.61 ± 1.27
VPT-SHALLOW ($p = 2$)		79.59 ± 0.35	70.52 ± 2.25	89.52 ± 0.97	79.88 ± 0.78
VPT-SHALLOW ($p = 10$)		76.43 ± 0.43	65.25 ± 3.64	87.82 ± 0.23	76.50 ± 1.33
VPT-DEEP ($p = 2, d = 2$)		80.14 ± 0.66	68.88 ± 0.93	89.37 ± 0.98	79.46 ± 0.53
VPT-DEEP ($p = 10, d = 9$)		76.70 ± 1.25	65.83 ± 0.96	87.33 ± 1.38	76.62 ± 0.40
FedAKT (CLIP)		60.32 ± 0.32	44.28 ± 0.38	61.36 ± 0.52	55.32 ± 0.14
FedGH (CLIP)		7.99 ± 1.33	6.24 ± 1.50	6.62 ± 0.89	6.95 ± 0.94
FedLT		**83.99 ± 1.04**	80.18 ± 1.08	93.20 ± 0.84	85.79 ± 0.33
FedAdapter		82.95 ± 0.69	78.45 ± 1.59	92.35 ± 0.67	84.58 ± 0.36
FedOT (All Global)		82.13 ± 1.86	73.58 ± 1.63	90.83 ± 1.09	82.18 ± 1.12
FedOT (All Local)		81.86 ± 0.54	75.30 ± 1.79	92.22 ± 0.90	83.12 ± 0.50
FedOT		83.85 ± 0.89	**80.53 ± 1.21**	**93.57 ± 0.45**	**85.98 ± 0.30**
FedOT+(B)		83.85 ± 0.89	**80.53 ± 1.21**	**93.57 ± 0.45**	**85.98 ± 0.30**

Table 14. The personalization accuracy of Office-Home (Target P, R).

Method	Target	L	S	V	AVG
Zero Shot		65.22 ± 2.72	70.96 ± 0.33	84.79 ± 1.07	74.06 ± 1.24
FedCLIP		70.37 ± 2.68	81.35 ± 1.64	86.96 ± 1.31	79.56 ± 1.33
PromptFL ($M = 4$)		72.44 ± 0.71	84.75 ± 2.97	88.99 ± 1.44	82.06 ± 1.27
PromptFL ($M = 16$)		73.07 ± 0.98	83.69 ± 1.58	88.39 ± 1.38	81.72 ± 0.66
CoCoOp (All Global)		72.69 ± 2.29	82.37 ± 1.96	87.90 ± 3.33	80.99 ± 0.96
CoCoOp (Meta-Net Random Init)		78.66 ± 1.44	84.70 ± 1.01	88.10 ± 1.13	83.82 ± 1.13
CoCoOp (Server CoOp)		79.35 ± 0.66	83.23 ± 0.93	88.39 ± 1.57	83.66 ± 0.58
VPT-SHALLOW ($p = 2$)		75.64 ± 1.46	84.70 ± 1.88	88.94 ± 1.20	83.09 ± 1.49
VPT-SHALLOW ($p = 10$)		75.83 ± 0.71	84.25 ± 1.56	88.35 ± 1.26	82.81 ± 0.96
VPT-DEEP ($p = 2, d = 2$)	C	76.71 ± 1.90	84.10 ± 2.76	88.34 ± 1.43	83.05 ± 1.90
VPT-DEEP ($p = 10, d = 9$)		76.33 ± 3.07	84.50 ± 0.72	89.38 ± 1.48	83.41 ± 1.63
FedAKT (CLIP)		70.50 ± 0.55	64.19 ± 0.66	72.11 ± 0.23	68.93 ± 0.16
FedGH (CLIP)		46.50 ± 1.54	62.23 ± 1.65	63.77 ± 1.31	57.50 ± 0.52
FedLT		**80.67 ± 0.80**	**86.84 ± 0.28**	88.94 ± 0.65	**85.48 ± 0.21**
FedAdapter		77.43 ± 0.54	85.60 ± 0.12	88.32 ± 0.72	83.78 ± 0.17
FedOT (All Global)		73.38 ± 0.39	85.42 ± 0.97	88.79 ± 0.89	82.53 ± 0.72
FedOT (All Local)		80.03 ± 1.30	86.59 ± 1.00	89.29 ± 1.71	85.30 ± 1.10
FedOT		80.29 ± 0.27	**86.84 ± 0.18**	**89.09 ± 0.66**	85.41 ± 0.14
FedOT+(B)		80.29 ± 0.27	**86.84 ± 0.18**	**89.09 ± 0.66**	85.41 ± 0.14
		C	S	V	AVG
Zero Shot		**100.00 ± 0.00**	70.96 ± 0.33	84.79 ± 1.07	87.67 ± 4.41
FedCLIP		**100.00 ± 0.00**	81.05 ± 1.83	87.26 ± 1.12	89.44 ± 0.54
PromptFL ($M = 4$)		99.88 ± 0.20	86.99 ± 1.12	89.38 ± 1.51	**92.09 ± 0.90**
PromptFL ($M = 16$)		99.88 ± 0.20	86.13 ± 0.67	89.09 ± 1.40	91.70 ± 0.27
CoCoOp (All Global)		**100.00 ± 0.00**	83.84 ± 0.55	87.90 ± 2.16	90.58 ± 0.90
CoCoOp (Meta-Net Random Init)		**100.00 ± 0.00**	84.65 ± 1.14	88.25 ± 1.43	90.97 ± 0.85
CoCoOp (Server CoOp)		99.65 ± 0.36	82.27 ± 1.96	88.30 ± 1.51	90.07 ± 0.78
VPT-SHALLOW ($p = 2$)		99.65 ± 0.36	85.72 ± 1.24	**89.78 ± 0.25**	91.72 ± 0.40
VPT-SHALLOW ($p = 10$)		99.29 ± 0.36	85.32 ± 0.79	89.24 ± 1.15	91.28 ± 0.44
VPT-DEEP ($p = 2, d = 2$)	L	99.76 ± 0.41	85.36 ± 1.35	89.04 ± 1.03	91.39 ± 0.71
VPT-DEEP ($p = 10, d = 9$)		99.88 ± 0.20	85.72 ± 0.84	89.58 ± 1.05	91.73 ± 0.63
FedAKT (CLIP)		79.95 ± 0.42	59.27 ± 0.63	68.98 ± 0.37	69.40 ± 0.21
FedGH (CLIP)		58.86 ± 1.24	58.16 ± 1.61	56.50 ± 1.39	58.20 ± 0.52
FedLT		99.77 ± 0.12	86.89 ± 0.23	89.19 ± 0.60	91.95 ± 0.13
FedAdapter		99.65 ± 0.20	**87.02 ± 0.11**	89.06 ± 0.28	91.91 ± 0.07
FedOT (All Global)		99.77 ± 0.20	86.43 ± 0.61	89.58 ± 1.12	91.93 ± 0.44
FedOT (All Local)		**100.00 ± 0.00**	86.59 ± 1.00	89.29 ± 1.71	91.96 ± 0.70
FedOT		99.88 ± 0.12	86.58 ± 0.49	89.29 ± 0.65	91.92 ± 0.16
FedOT+(B)		99.88 ± 0.12	86.58 ± 0.49	89.29 ± 0.65	91.92 ± 0.16

Table 15. The personalization accuracy of VLCS (Target C, L).

Method	Target	C	L	V	AVG
Zero Shot	S	**100.00 ± 0.00**	65.22 ± 2.72	84.79 ± 1.07	83.34 ± 1.11
FedCLIP		100.00 ± 0.00	74.45 ± 1.83	87.06 ± 0.73	87.17 ± 0.72
PromptFL ($M = 4$)		99.65 ± 0.61	76.46 ± 0.98	88.74 ± 1.17	88.28 ± 0.37
PromptFL ($M = 16$)		99.77 ± 0.20	76.96 ± 1.28	88.00 ± 0.74	88.24 ± 0.56
CoCoOp (All Global)		**100.00 ± 0.00**	77.09 ± 0.47	87.06 ± 2.64	88.05 ± 0.73
CoCoOp (Meta-Net Random Init)		**100.00 ± 0.00**	78.21 ± 1.23	88.05 ± 1.30	88.75 ± 0.78
CoCoOp (Server CoOp)		99.77 ± 0.20	79.53 ± 0.85	88.20 ± 1.78	89.17 ± 0.87
VPT-SHALLOW ($p = 2$)		99.76 ± 0.41	78.60 ± 2.14	87.95 ± 2.27	88.77 ± 1.44
VPT-SHALLOW ($p = 10$)		99.53 ± 0.41	78.41 ± 1.58	87.76 ± 0.94	88.56 ± 0.85
VPT-DEEP ($p = 2, d = 2$)		99.88 ± 0.20	78.78 ± 1.83	87.46 ± 2.10	88.71 ± 1.29
VPT-DEEP ($p = 10, d = 9$)		99.77 ± 0.20	77.46 ± 1.89	88.15 ± 1.34	88.46 ± 1.02
FedAKT (CLIP)		78.90 ± 0.45	59.52 ± 0.55	68.98 ± 0.33	69.15 ± 0.22
FedGH (CLIP)		68.73 ± 1.25	52.54 ± 1.59	51.73 ± 1.32	57.67 ± 0.58
FedLT		99.77 ± 0.12	79.91 ± 0.54	88.49 ± 0.58	89.39 ± 0.15
FedAdapter		99.71 ± 0.16	78.63 ± 0.54	87.53 ± 0.80	88.62 ± 0.19
FedOT (All Global)		99.88 ± 0.20	77.40 ± 1.05	88.00 ± 1.80	88.43 ± 0.74
FedOT (All Local)		**100.00 ± 0.00**	80.03 ± 1.30	89.29 ± 1.71	89.77 ± 1.00
FedOT		99.77 ± 0.12	**80.85 ± 0.44**	**88.89 ± 0.67**	**89.84 ± 0.16**
FedOT+(B)		99.77 ± 0.12	**80.85 ± 0.44**	**88.89 ± 0.67**	**89.84 ± 0.16**
		C	L	S	AVG
Zero Shot	V	**100.00 ± 0.00**	65.22 ± 2.72	70.96 ± 0.33	79.93 ± 1.85
FedCLIP		**100.00 ± 0.00**	72.69 ± 2.72	77.79 ± 2.71	83.50 ± 0.95
PromptFL ($M = 4$)		99.41 ± 0.41	73.89 ± 0.71	84.50 ± 2.12	85.93 ± 0.69
PromptFL ($M = 16$)		75.14 ± 0.19	83.38 ± 1.91	83.90 ± 2.20	86.10 ± 0.71
CoCoOp (All Global)		99.77 ± 0.20	73.89 ± 1.70	82.72 ± 2.03	85.46 ± 1.17
CoCoOp (Meta-Net Random Init)		**100.00 ± 0.00**	78.47 ± 1.07	84.81 ± 1.88	87.76 ± 0.98
CoCoOp (Server CoOp)		99.88 ± 0.20	79.41 ± 0.87	81.91 ± 1.14	87.07 ± 0.60
VPT-SHALLOW ($p = 2$)		99.65 ± 0.36	77.03 ± 2.55	85.06 ± 1.76	87.24 ± 1.35
VPT-SHALLOW ($p = 10$)		99.65 ± 0.00	77.59 ± 1.99	85.11 ± 1.62	87.45 ± 1.17
VPT-DEEP ($p = 2, d = 2$)		99.77 ± 0.20	78.34 ± 2.72	83.94 ± 2.03	87.35 ± 1.49
VPT-DEEP ($p = 10, d = 9$)		99.77 ± 0.20	77.59 ± 1.86	84.70 ± 0.32	87.35 ± 0.55
FedAKT (CLIP)		79.36 ± 0.41	63.11 ± 0.52	66.15 ± 0.55	69.20 ± 0.17
FedGH (CLIP)		78.75 ± 1.20	40.52 ± 0.98	54.43 ± 2.01	57.90 ± 1.06
FedLT		99.88 ± 0.12	**80.92 ± 0.38**	**86.33 ± 0.31**	**89.04 ± 0.10**
FedAdapter		99.71 ± 0.06	78.88 ± 0.52	85.04 ± 0.51	87.88 ± 0.14
FedOT (All Global)		99.77 ± 0.20	74.70 ± 1.07	84.45 ± 1.53	86.31 ± 0.47
FedOT (All Local)		**100.00 ± 0.00**	80.03 ± 1.30	86.59 ± 1.00	88.87 ± 0.58
FedOT		99.88 ± 0.12	80.67 ± 0.23	86.28 ± 0.32	88.94 ± 0.08
FedOT+(B)		99.88 ± 0.12	80.67 ± 0.23	86.28 ± 0.32	88.94 ± 0.08

Table 16. The personalization accuracy of VLCS (Target S, V).

Method	Target	L43	L46	L100	AVG
Zero Shot	L38	29.90 ± 1.95	19.59 ± 1.31	13.47 ± 0.52	20.99 ± 0.43
FedCLIP		44.17 ± 9.98	50.91 ± 7.83	62.60 ± 19.22	52.56 ± 12.32
PromptFL ($M = 4$)		54.71 ± 1.85	58.16 ± 0.37	67.97 ± 2.87	60.28 ± 1.66
PromptFL ($M = 16$)		53.09 ± 1.39	58.81 ± 0.28	67.69 ± 2.54	59.87 ± 0.56
CoCoOp (All Global)		47.29 ± 0.66	54.02 ± 0.32	65.34 ± 5.52	55.55 ± 1.75
CoCoOp (Meta-Net Random Init)		58.48 ± 4.88	59.36 ± 0.53	78.63 ± 1.40	65.49 ± 1.52
CoCoOp (Server CoOp)		59.31 ± 0.77	58.00 ± 5.81	76.07 ± 3.18	64.46 ± 1.69
VPT-SHALLOW ($p = 2$)		51.85 ± 7.22	52.85 ± 5.23	68.34 ± 3.12	57.68 ± 5.04
VPT-SHALLOW ($p = 10$)		51.97 ± 4.21	55.00 ± 2.51	69.10 ± 0.36	58.69 ± 2.02
VPT-DEEP ($p = 2, d = 2$)		51.43 ± 7.31	53.66 ± 6.38	70.84 ± 2.93	58.64 ± 5.25
VPT-DEEP ($p = 10, d = 9$)		47.83 ± 6.43	51.13 ± 6.33	68.68 ± 3.58	55.88 ± 5.30
FedAKT (CLIP)		39.64 ± 0.95	43.82 ± 1.01	36.04 ± 0.78	39.83 ± 0.30
FedGH (CLIP)		12.77 ± 6.51	9.41 ± 6.85	8.32 ± 5.52	10.17 ± 2.13
FedLT		**66.71 ± 1.47**	**65.96 ± 1.14**	**82.67 ± 0.78**	**71.78 ± 0.39**
FedAdapter		59.62 ± 1.60	62.69 ± 0.35	77.38 ± 0.32	66.56 ± 0.32
FedOT (All Global)		50.44 ± 4.64	54.10 ± 6.45	70.29 ± 1.18	58.28 ± 3.35
FedOT (All Local)		51.49 ± 2.28	57.60 ± 3.77	74.60 ± 0.21	61.23 ± 0.83
FedOT		56.37 ± 4.64	56.93 ± 6.03	75.28 ± 1.06	62.86 ± 1.48
FedOT+(B)		63.30 ± 1.41	65.11 ± 0.78	81.37 ± 0.44	69.93 ± 0.32
		L38	L46	L100	AVG
Zero Shot	L43	13.35 ± 0.76	19.59 ± 1.31	13.47 ± 0.52	15.47 ± 0.31
FedCLIP		60.52 ± 14.05	43.96 ± 7.29	60.92 ± 19.02	55.14 ± 13.40
PromptFL ($M = 4$)		64.77 ± 3.55	52.68 ± 2.04	67.21 ± 1.90	61.56 ± 1.67
PromptFL ($M = 16$)		63.82 ± 1.25	52.98 ± 3.66	69.40 ± 3.79	62.07 ± 2.23
CoCoOp (All Global)		64.70 ± 1.14	49.55 ± 0.95	63.62 ± 2.48	59.29 ± 0.51
CoCoOp (Meta-Net Random Init)		74.84 ± 0.88	59.52 ± 0.46	78.53 ± 1.40	70.96 ± 0.36
CoCoOp (Server CoOp)		73.38 ± 2.10	57.61 ± 4.90	77.09 ± 2.93	69.36 ± 2.08
VPT-SHALLOW ($p = 2$)		58.25 ± 4.68	53.01 ± 6.86	72.92 ± 1.69	61.40 ± 4.23
VPT-SHALLOW ($p = 10$)		55.81 ± 7.30	49.58 ± 1.71	71.08 ± 4.01	58.82 ± 4.07
VPT-DEEP ($p = 2, d = 2$)		62.45 ± 1.93	51.05 ± 3.91	73.81 ± 1.56	62.44 ± 2.46
VPT-DEEP ($p = 10, d = 9$)		56.13 ± 6.72	51.16 ± 6.85	71.18 ± 3.77	59.49 ± 4.92
FedAKT (CLIP)		36.84 ± 0.94	43.66 ± 1.05	36.20 ± 2.01	38.90 ± 1.32
FedGH (CLIP)		10.02 ± 5.02	10.44 ± 6.89	9.24 ± 5.54	9.90 ± 2.03
FedLT		**78.27 ± 1.00**	**67.16 ± 1.43**	**82.84 ± 0.54**	**76.09 ± 0.35**
FedAdapter		70.71 ± 0.50	60.40 ± 0.99	73.44 ± 1.05	68.18 ± 0.29
FedOT (All Global)		64.24 ± 4.23	51.27 ± 2.37	69.02 ± 0.98	61.51 ± 1.64
FedOT (All Local)		73.36 ± 0.70	57.60 ± 3.77	74.60 ± 0.21	68.52 ± 1.19
FedOT		74.13 ± 2.51	55.46 ± 1.73	77.09 ± 3.33	68.89 ± 0.87
FedOT+(B)		76.63 ± 0.43	65.25 ± 0.71	81.30 ± 0.19	74.39 ± 0.16

Table 17. The personalization accuracy of TerraIncognita (Target L38, L43).

Method	Target	L38	L43	L100	AVG
Zero Shot	L46	13.35 ± 0.76	29.90 ± 1.95	13.47 ± 0.52	18.91 ± 0.68
FedCLIP		61.53 ± 13.73	37.16 ± 6.02	66.77 ± 18.25	55.15 ± 12.60
PromptFL ($M = 4$)		64.43 ± 1.79	48.98 ± 2.73	69.44 ± 2.96	60.95 ± 0.59
PromptFL ($M = 16$)		63.22 ± 2.08	48.65 ± 2.09	69.92 ± 3.23	60.60 ± 0.21
CoCoOp (All Global)		64.67 ± 1.51	38.91 ± 5.01	67.49 ± 2.42	57.02 ± 1.30
CoCoOp (Meta-Net Random Init)		74.14 ± 0.32	58.56 ± 4.42	78.33 ± 1.44	70.34 ± 1.21
CoCoOp (Server CoOp)		73.29 ± 2.06	59.14 ± 1.30	77.20 ± 3.10	69.88 ± 1.50
VPT-SHALLOW ($p = 2$)		60.27 ± 7.97	49.23 ± 2.89	74.77 ± 2.31	61.42 ± 4.37
VPT-SHALLOW ($p = 10$)		55.73 ± 7.45	50.77 ± 5.31	72.45 ± 2.84	59.65 ± 4.11
VPT-DEEP ($p = 2, d = 2$)		60.40 ± 3.97	51.23 ± 6.04	74.70 ± 2.20	62.11 ± 3.97
VPT-DEEP ($p = 10, d = 9$)		56.34 ± 6.90	48.03 ± 5.07	63.15 ± 3.18	55.84 ± 4.91
FedAKT (CLIP)		33.20 ± 0.92	43.04 ± 1.21	36.36 ± 0.99	37.53 ± 0.50
FedGH (CLIP)		10.07 ± 5.03	10.82 ± 6.54	9.21 ± 5.51	10.03 ± 1.92
FedLT		**77.91 ± 0.71**	**65.47 ± 1.07**	**83.25 ± 0.32**	**75.54 ± 0.25**
FedAdapter		71.39 ± 0.33	60.71 ± 1.42	76.07 ± 0.61	69.39 ± 0.30
FedOT (All Global)		64.04 ± 5.64	49.36 ± 4.33	73.06 ± 1.46	62.15 ± 3.37
FedOT (All Local)		73.36 ± 0.70	51.49 ± 2.28	74.60 ± 0.21	66.48 ± 1.00
FedOT		73.72 ± 1.97	55.58 ± 5.16	75.56 ± 2.72	68.29 ± 1.18
FedOT+(B)		76.79 ± 0.38	63.93 ± 1.49	80.82 ± 0.46	73.85 ± 0.31
		L38	L43	L46	AVG
Zero Shot	L100	13.35 ± 0.76	29.90 ± 1.95	19.59 ± 1.31	20.95 ± 0.44
FedCLIP		60.32 ± 10.31	49.44 ± 11.44	50.26 ± 7.98	53.34 ± 9.86
PromptFL ($M = 4$)		63.56 ± 1.78	53.17 ± 1.95	55.76 ± 2.09	57.50 ± 0.95
PromptFL ($M = 16$)		61.97 ± 2.09	53.84 ± 3.11	57.31 ± 0.12	57.71 ± 0.50
CoCoOp (All Global)		63.54 ± 1.87	47.16 ± 4.81	52.98 ± 0.27	54.56 ± 2.15
CoCoOp (Meta-Net Random Init)		74.30 ± 0.26	59.56 ± 5.27	59.52 ± 0.85	64.46 ± 1.72
CoCoOp (Server CoOp)		73.10 ± 2.12	59.48 ± 2.17	57.97 ± 5.11	63.52 ± 2.26
VPT-SHALLOW ($p = 2$)		56.82 ± 4.74	48.86 ± 5.91	54.21 ± 7.36	53.30 ± 5.90
VPT-SHALLOW ($p = 10$)		55.74 ± 8.39	49.44 ± 4.82	51.81 ± 2.99	52.33 ± 3.84
VPT-DEEP ($p = 2, d = 2$)		58.27 ± 4.21	50.94 ± 6.67	54.48 ± 7.61	54.56 ± 5.98
VPT-DEEP ($p = 10, d = 9$)		48.54 ± 4.03	45.38 ± 2.47	48.98 ± 4.72	47.63 ± 3.31
FedAKT (CLIP)		37.02 ± 1.11	39.54 ± 0.73	43.64 ± 0.95	40.07 ± 0.30
FedGH (CLIP)		8.03 ± 3.04	13.83 ± 1.55	9.44 ± 2.88	10.43 ± 2.11
FedLT		**78.20 ± 0.89**	**66.21 ± 1.05**	**66.48 ± 1.19**	**70.30 ± 0.35**
FedAdapter		70.87 ± 0.39	59.94 ± 1.17	61.43 ± 1.01	64.08 ± 0.31
FedOT (All Global)		62.01 ± 6.81	49.02 ± 3.89	53.56 ± 4.00	54.86 ± 4.56
FedOT (All Local)		73.36 ± 0.70	51.49 ± 2.28	57.60 ± 3.77	60.82 ± 0.81
FedOT		73.43 ± 1.71	52.97 ± 4.45	58.65 ± 5.36	61.68 ± 1.38
FedOT+(B)		76.68 ± 0.52	63.80 ± 1.48	65.11 ± 1.12	68.53 ± 0.37

Table 18. The personalization accuracy of TerraIncognita (Target L46, L100).

F.3. Comprehensive

Method	33	0	25	26	AVG
Zero Shot	50.88 ± 3.42	48.40 ± 1.47	45.19 ± 2.50	48.24 ± 2.73	48.18 ± 1.31
FedCLIP	70.07 ± 9.35	70.58 ± 8.74	70.55 ± 9.69	70.82 ± 9.57	70.51 ± 9.30
PromptFL ($M = 4$)	96.01 ± 1.56	96.13 ± 0.49	96.66 ± 0.30	96.82 ± 0.67	96.40 ± 0.74
PromptFL ($M = 16$)	$\mathbf{97.12 \pm 0.64}$	96.87 ± 0.53	96.92 ± 0.52	96.95 ± 0.27	96.96 ± 0.48
CoCoOp (All Global)	87.91 ± 8.43	87.29 ± 8.86	85.90 ± 9.89	87.94 ± 7.46	87.26 ± 8.65
CoCoOp (Meta-Net Random Init)	70.29 ± 8.78	69.58 ± 10.48	69.17 ± 10.76	70.31 ± 9.64	69.84 ± 9.87
CoCoOp (Server CoOp)	75.13 ± 1.06	76.02 ± 0.82	76.04 ± 2.84	73.86 ± 1.03	75.26 ± 0.82
VPT-SHALLOW ($p = 2$)	84.80 ± 1.66	83.14 ± 0.98	83.02 ± 1.33	83.02 ± 1.33	84.78 ± 2.39
VPT-SHALLOW ($p = 10$)	83.34 ± 2.06	78.75 ± 2.21	83.10 ± 1.24	83.34 ± 0.91	82.13 ± 0.50
VPT-DEEP ($p = 2, d = 2$)	84.23 ± 2.80	84.67 ± 2.50	78.70 ± 0.94	83.99 ± 3.48	82.90 ± 1.12
VPT-DEEP ($p = 10, d = 9$)	82.48 ± 2.21	79.38 ± 3.04	80.78 ± 4.37	79.70 ± 5.02	80.58 ± 3.52
FedAKT (CLIP)	44.85 ± 5.53	45.10 ± 6.52	45.08 ± 6.84	45.01 ± 5.90	45.01 ± 6.17
FedGH (CLIP)	38.69 ± 0.51	39.77 ± 0.73	39.67 ± 0.56	39.38 ± 0.52	39.38 ± 0.58
FedLT	93.23 ± 0.51	95.44 ± 0.48	95.49 ± 0.54	94.42 ± 0.45	94.65 ± 0.30
FedAdapter	96.00 ± 0.61	96.47 ± 0.58	96.41 ± 0.49	96.35 ± 0.40	96.31 ± 0.33
FedOT (All Global)	94.39 ± 4.75	92.35 ± 4.18	91.98 ± 4.43	94.53 ± 3.62	93.31 ± 3.17
FedOT	94.58 ± 0.61	95.11 ± 0.58	95.16 ± 0.49	94.91 ± 0.42	94.94 ± 1.01
FedOT+(B)	$\mathbf{97.12 \pm 0.59}$	$\mathbf{97.63 \pm 0.49}$	$\mathbf{97.94 \pm 0.21}$	$\mathbf{97.78 \pm 1.01}$	$\mathbf{97.61 \pm 0.44}$

Table 19. The comprehensive accuracy of FEMNIST.

Method	A	C	P	S	AVG
Zero Shot	95.35 ± 0.98	$\mathbf{96.94 \pm 0.81}$	$\mathbf{99.60 \pm 0.46}$	85.56 ± 0.99	94.36 ± 0.48
FedCLIP	95.50 ± 0.75	95.61 ± 0.62	95.59 ± 0.62	94.99 ± 1.08	95.42 ± 0.76
PromptFL ($M = 4$)	96.56 ± 0.42	96.74 ± 0.30	96.72 ± 0.34	94.92 ± 0.44	96.23 ± 0.37
PromptFL ($M = 16$)	96.16 ± 0.36	96.52 ± 0.22	96.73 ± 0.54	95.31 ± 0.31	96.18 ± 0.33
CoCoOp (All Global)	95.92 ± 1.05	96.11 ± 0.81	96.15 ± 0.74	95.06 ± 0.65	95.81 ± 0.80
CoCoOp (Meta-Net Random Init)	87.99 ± 3.92	91.30 ± 2.33	87.35 ± 4.49	91.64 ± 2.48	89.57 ± 3.27
CoCoOp (Server CoOp)	93.68 ± 0.22	90.47 ± 6.63	93.69 ± 1.32	93.60 ± 3.39	92.86 ± 2.22
VPT-SHALLOW ($p = 2$)	96.49 ± 0.50	96.28 ± 0.49	96.50 ± 0.46	95.02 ± 0.46	96.07 ± 0.47
VPT-SHALLOW ($p = 10$)	96.21 ± 0.44	96.12 ± 0.43	96.36 ± 0.43	95.30 ± 0.52	96.00 ± 0.45
VPT-DEEP ($p = 2, d = 2$)	96.42 ± 0.66	96.33 ± 0.39	96.52 ± 0.31	94.72 ± 0.79	96.00 ± 0.52
VPT-DEEP ($p = 10, d = 9$)	95.91 ± 0.48	96.13 ± 0.42	96.17 ± 0.29	94.51 ± 0.73	95.68 ± 0.41
FedAKT (CLIP)	60.46 ± 1.02	60.53 ± 1.23	60.64 ± 1.34	60.08 ± 1.12	60.43 ± 1.16
FedGH (CLIP)	61.63 ± 1.32	61.91 ± 1.63	62.32 ± 1.74	60.11 ± 1.32	61.49 ± 1.48
FedLT	$\mathbf{96.68 \pm 0.33}$	97.18 ± 0.29	98.98 ± 0.32	94.19 ± 0.44	96.50 ± 0.24
FedAdapter	96.66 ± 0.34	96.73 ± 0.32	97.04 ± 0.32	94.97 ± 0.58	96.35 ± 0.23
FedOT (All Global)	96.63 ± 0.36	96.61 ± 0.34	96.78 ± 0.43	95.07 ± 0.61	96.28 ± 0.43
FedOT	96.40 ± 0.35	96.30 ± 0.34	96.15 ± 0.40	$\mathbf{97.70 \pm 0.47}$	$\mathbf{96.64 \pm 0.25}$
FedOT+(B)	96.40 ± 0.35	96.30 ± 0.34	96.15 ± 0.40	$\mathbf{97.70 \pm 0.47}$	$\mathbf{96.64 \pm 0.25}$

Table 20. The comprehensive accuracy of PACS.

Method	A	C	P	R	AVG
Zero Shot	77.59 ± 1.33	65.29 ± 0.89	**86.89 ± 1.61**	**86.83 ± 0.74**	79.15 ± 0.40
FedCLIP	80.59 ± 0.41	80.06 ± 0.09	79.88 ± 0.42	79.94 ± 0.49	80.12 ± 0.17
PromptFL ($M = 4$)	83.14 ± 1.12	82.66 ± 0.69	83.38 ± 0.62	84.04 ± 1.02	83.31 ± 0.86
PromptFL ($M = 16$)	83.74 ± 0.42	82.71 ± 0.57	83.95 ± 0.95	84.06 ± 0.96	83.61 ± 0.69
CoCoOp (All Global)	81.04 ± 1.26	80.21 ± 1.01	81.00 ± 0.94	81.36 ± 1.15	80.90 ± 1.07
CoCoOp (Meta-Net Random Init)	77.42 ± 1.92	76.59 ± 0.92	76.35 ± 1.70	77.00 ± 1.05	76.84 ± 1.30
CoCoOp (Server CoOp)	83.14 ± 1.12	82.66 ± 0.69	83.38 ± 0.62	84.04 ± 1.02	83.31 ± 0.86
VPT-SHALLOW ($p = 2$)	81.62 ± 0.73	81.23 ± 0.24	81.42 ± 0.51	81.80 ± 0.53	81.52 ± 0.45
VPT-SHALLOW ($p = 10$)	79.40 ± 1.47	79.99 ± 0.62	78.71 ± 0.82	79.05 ± 1.17	79.29 ± 0.99
VPT-DEEP ($p = 2, d = 2$)	81.19 ± 0.33	81.34 ± 0.46	80.56 ± 0.50	81.48 ± 0.29	81.14 ± 0.33
VPT-DEEP ($p = 10, d = 9$)	78.96 ± 0.76	78.98 ± 0.67	78.05 ± 1.13	79.36 ± 0.17	78.84 ± 0.55
FedAKT (CLIP)	55.34 ± 0.35	55.31 ± 0.42	55.37 ± 0.45	55.37 ± 0.28	55.34 ± 0.37
FedGH (CLIP)	7.00 ± 2.05	6.94 ± 2.52	7.06 ± 2.74	7.06 ± 1.79	7.00 ± 2.27
FedLT	85.55 ± 0.43	83.70 ± 0.28	85.86 ± 0.40	86.51 ± 0.43	85.40 ± 0.32
FedAdapter	84.95 ± 0.44	83.63 ± 0.33	84.98 ± 0.41	85.66 ± 0.48	84.80 ± 0.45
FedOT (All Global)	83.32 ± 0.92	82.58 ± 0.79	82.82 ± 0.75	83.75 ± 0.84	83.12 ± 0.81
FedOT	**85.91 ± 0.42**	**83.73 ± 0.58**	86.05 ± 0.46	86.43 ± 0.45	**85.53 ± 0.32**
FedOT+(B)	**85.91 ± 0.42**	**83.73 ± 0.58**	86.05 ± 0.46	86.43 ± 0.45	**85.53 ± 0.32**

Table 21. The comprehensive accuracy of Office-Home.

Method	C	L	S	V	AVG
Zero Shot	**100.00 ± 0.00**	65.22 ± 2.72	70.96 ± 0.33	84.79 ± 1.07	80.59 ± 0.96
FedCLIP	84.49 ± 1.04	82.93 ± 0.54	83.75 ± 0.59	84.03 ± 0.40	83.80 ± 0.63
PromptFL ($M = 4$)	86.31 ± 0.88	84.81 ± 0.30	85.95 ± 0.51	85.99 ± 0.72	85.77 ± 0.57
PromptFL ($M = 16$)	86.08 ± 0.39	83.82 ± 0.18	85.62 ± 1.05	85.55 ± 0.90	85.27 ± 0.54
CoCoOp (All Global)	85.71 ± 0.67	83.60 ± 0.93	84.62 ± 1.47	85.11 ± 1.85	84.76 ± 1.01
CoCoOp (Meta-Net Random Init)	83.82 ± 1.13	**90.97 ± 0.85**	**88.75 ± 0.78**	87.76 ± 0.98	79.91 ± 1.06
CoCoOp (Server CoOp)	85.06 ± 4.23	81.05 ± 2.72	81.97 ± 0.78	83.99 ± 0.49	83.02 ± 1.46
VPT-SHALLOW ($p = 2$)	87.23 ± 1.07	84.54 ± 0.36	85.37 ± 2.05	86.87 ± 1.24	86.00 ± 1.08
VPT-SHALLOW ($p = 10$)	87.02 ± 0.65	84.66 ± 0.14	85.58 ± 0.74	86.92 ± 1.01	86.04 ± 0.58
VPT-DEEP ($p = 2, d = 2$)	87.14 ± 1.47	84.00 ± 1.02	85.67 ± 2.15	87.00 ± 1.30	85.95 ± 1.27
VPT-DEEP ($p = 10, d = 9$)	87.35 ± 1.23	84.15 ± 0.34	85.78 ± 1.14	87.12 ± 0.45	86.10 ± 0.74
FedAKT (CLIP)	69.83 ± 0.42	68.60 ± 0.53	69.19 ± 0.58	69.12 ± 0.27	69.18 ± 0.44
FedGH (CLIP)	58.73 ± 1.24	56.60 ± 1.57	58.18 ± 1.51	57.80 ± 1.32	56.27 ± 1.38
FedLT	89.00 ± 0.27	85.03 ± 0.17	87.06 ± 0.31	88.02 ± 0.25	87.28 ± 0.34
FedAdapter	87.06 ± 0.23	84.61 ± 0.26	85.85 ± 0.45	86.45 ± 0.25	85.99 ± 0.34
FedOT (All Global)	86.81 ± 0.48	84.39 ± 0.73	85.50 ± 1.55	86.52 ± 0.75	85.80 ± 0.73
FedOT	88.97 ± 0.19	85.20 ± 0.21	87.46 ± 0.32	**88.26 ± 0.29**	**87.47 ± 0.26**
FedOT+(B)	88.97 ± 0.19	85.20 ± 0.21	87.46 ± 0.32	**88.26 ± 0.29**	**87.47 ± 0.26**

Table 22. The comprehensive accuracy of VLCS.

Method	L38	L43	L46	L100	AVG
Zero Shot	13.35 ± 0.76	29.90 ± 1.95	19.59 ± 1.31	13.47 ± 0.52	19.08 ± 0.26
FedCLIP	47.74 ± 8.38	49.10 ± 9.72	49.03 ± 9.81	50.89 ± 8.74	49.19 ± 9.11
PromptFL ($M = 4$)	54.84 ± 0.50	54.42 ± 1.80	54.96 ± 0.91	53.49 ± 1.57	54.43 ± 0.75
PromptFL ($M = 16$)	54.23 ± 2.03	55.16 ± 1.98	54.27 ± 0.98	52.52 ± 0.58	54.04 ± 1.07
CoCoOp (All Global)	51.24 ± 3.28	51.93 ± 0.56	52.48 ± 0.73	50.32 ± 1.83	51.49 ± 0.16
CoCoOp (Meta-Net Random Init)	60.28 ± 1.33	59.75 ± 1.47	59.14 ± 1.49	54.14 ± 1.61	58.33 ± 0.86
CoCoOp (Server CoOp)	56.88 ± 5.74	58.61 ± 1.00	58.45 ± 0.79	52.65 ± 1.74	56.65 ± 2.22
VPT-SHALLOW ($p = 2$)	53.39 ± 5.01	56.01 ± 3.54	54.19 ± 1.70	46.52 ± 4.63	52.53 ± 3.71
VPT-SHALLOW ($p = 10$)	55.47 ± 1.00	52.62 ± 1.42	53.50 ± 2.69	44.11 ± 2.01	51.42 ± 1.68
VPT-DEEP ($p = 2, d = 2$)	54.05 ± 3.84	55.04 ± 1.43	55.59 ± 3.58	47.96 ± 4.80	53.16 ± 3.40
VPT-DEEP ($p = 10, d = 9$)	45.74 ± 5.38	52.75 ± 3.83	48.76 ± 4.06	37.05 ± 1.84	46.07 ± 3.55
FedAKT (CLIP)	36.92 ± 0.95	39.40 ± 0.82	43.62 ± 1.01	36.33 ± 0.84	39.44 ± 0.91
FedGH (CLIP)	10.12 ± 5.04	10.67 ± 6.59	10.31 ± 6.72	9.39 ± 5.49	10.12 ± 5.98
FedLT	61.15 ± 0.76	$\mathbf{66.90 \pm 0.76}$	65.11 ± 0.97	57.00 ± 1.58	62.54 ± 0.09
FedAdapter	58.65 ± 0.90	60.41 ± 0.55	62.33 ± 0.46	58.07 ± 0.41	59.86 ± 0.50
FedOT (All Global)	53.15 ± 3.23	54.25 ± 1.30	55.25 ± 1.91	47.65 ± 3.00	52.57 ± 3.17
FedOT	54.81 ± 5.81	58.03 ± 2.27	60.01 ± 3.85	53.03 ± 1.88	56.47 ± 3.40
FedOT+(B)	$\mathbf{62.09 \pm 1.08}$	64.38 ± 0.39	$\mathbf{65.66 \pm 0.46}$	$\mathbf{60.80 \pm 0.48}$	$\mathbf{63.23 \pm 0.44}$

Table 23. The comprehensive accuracy of TerraIncognita.

References

[1] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, and Shyamal Anadkat. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023. 1

[2] Sara Beery, Grant Van Horn, and Pietro Perona. Recognition in terra incognita. In *ECCV*, pages 456–473, 2018. 2

[3] Rishi Bommasani, Drew A. Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S. Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, Erik Brynjolfsson, S. Buch, Dallas Card, Rodrigo Castellon, Niladri S. Chatterji, Annie S. Chen, Kathleen A. Creel, Jared Davis, Dora Demszky, Chris Donahue, Moussa Doumbouya, Esin Durmus, Stefano Ermon, John Etchemendy, Kawin Ethayarajh, Li Fei-Fei, Chelsea Finn, Trevor Gale, Lauren E. Gillespie, Karan Goel, Noah D. Goodman, Shelby Grossman, Neel Guha, Tatsunori Hashimoto, Peter Henderson, John Hewitt, Daniel E. Ho, Jenny Hong, Kyle Hsu, Jing Huang, Thomas F. Icard, Saahil Jain, Dan Jurafsky, Pratyusha Kalluri, Siddharth Karamcheti, Geoff Keeling, Fereshte Khani, O. Khattab, Pang Wei Koh, Mark S. Krass, Ranjay Krishna, Rohith Kuditipudi, Ananya Kumar, Faisal Ladhak, Mina Lee, Tony Lee, Jure Leskovec, Isabelle Levent, Xiang Lisa Li, Xuechen Li, Tengyu Ma, Ali Malik, Christopher D. Manning, Suvir P. Mirchandani, Eric Mitchell, Zanele Munyikwa, Suraj Nair, Avanika Narayan, Deepak Narayanan, Benjamin Newman, Allen Nie, Juan Carlos Niebles, Hamed Nilforoshan, J. F. Nyarko, Giray Ogut, Laurel Orr, Isabel Papadimitriou, Joon Sung Park, Chris Piech, Eva Portelance, Christopher Potts, Aditi Raghunathan, Robert Reich, Hongyu Ren, Frieda Rong, Yusuf H. Roohani, Camilo Ruiz, Jack Ryan, Christopher R'e, Dorsa Sadigh, Shiori Sagawa, Keshav Santhanam, Andy Shih, Krishna Parasuram Srinivasan, Alex Tamkin, Rohan Taori, Armin W. Thomas, Florian Tramèr, Rose E. Wang, William Wang, Bohan Wu, Jiajun Wu, Yuhuai Wu, Sang Michael Xie, Michihiro Yasunaga, Jiaxuan You, Matei A. Zaharia, Michael Zhang, Tianyi Zhang, Xikun Zhang, Yuhui Zhang, Lucia Zheng, Kaitlyn Zhou, and Percy Liang. On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*, 2021. 1

[4] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, and Amanda Askell. Language models are few-shot learners. In *NeurIPS*, pages 1877–1901, 2020. 1

[5] Sebastian Caldas, Sai Meher Karthik Duddu, Peter Wu, Tian Li, Jakub Konečný, H Brendan McMahan, Virginia Smith, and Ameet Talwalkar. Leaf: A benchmark for federated settings. *arXiv preprint arXiv:1812.01097*, 2018. 2

[6] Arthur Cayley. Sur quelques propriétés des déterminants gauches. 1846. 2

[7] Chen Fang, Ye Xu, and Daniel N Rockmore. Unbiased metric learning: On the utilization of multiple datasets and web images for softening bias. In *ICCV*, pages 1657–1664, 2013. 2

[8] Dashan Gao, Xin Yao, and Qiang Yang. A survey on heterogeneous federated learning. *arXiv preprint arXiv:2210.04505*, 2022. 1

[9] Tao Guo, Song Guo, Junxiao Wang, Xueyang Tang, and Wenchao Xu. Promptfl: Let federated participants cooperatively learn prompts instead of models–federated learning in age of foundation model. *IEEE Transactions on Mobile Computing*, 23(5):5179–5194, 2023. 1

[10] Chao Jia, Yinfei Yang, Ye Xia, Yi-Ting Chen, Zarana Parekh, Hieu Pham, Quoc Le, Yun-Hsuan Sung, Zhen Li, and Tom Duerig. Scaling up visual and vision-language representation learning with noisy text supervision. In *Proc. Int. Conf. Mach. Learn. (ICML)*, pages 4904–4916, 2021. 1

[11] Da Li, Yongxin Yang, Yi-Zhe Song, and Timothy M Hospedales. Deeper, broader and artier domain generalization. In *ICCV*, pages 5542–5550, 2017. 2

[12] Wang Lu, Xixu Hu, Jindong Wang, and Xing Xie. Fedclip: Fast generalization and personalization for clip in federated learning. *arXiv preprint arXiv:2302.13485*, 2023. 1

[13] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Proc. Int. Conf. Artif. Intell. Stat. (AISTATS)*, pages 1273–1282, 2017. 1

[14] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, and Luca Antiga. Pytorch: An imperative style, high-performance deep learning library. In *NeurIPS*, pages 8026–8037, 2019. 2

[15] Krishna Pillutla, Kshitiz Malik, Abdel-Rahman Mohamed, Mike Rabbat, Maziar Sanjabi, and Lin Xiao. Federated learning with partial model personalization. In *Proc. Int. Conf. Mach. Learn. (ICML)*, pages 17716–17758, 2022. 1

[16] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. OpenAI blog, 2019. 1

[17] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *Proc. Int. Conf. Mach. Learn. (ICML)*, pages 8748–8763, 2021. 1, 2

[18] Karan Singhal, Hakim Sidahmed, Zachary Garrett, Shanshan Wu, John Rush, and Sushant Prakash. Federated reconstruction: Partially local federated learning. In *NeurIPS*, pages 11220–11232, 2021. 1

[19] Hemanth Venkateswara, Jose Eusebio, Shayok Chakraborty, and Sethuraman Panchanathan. Deep hashing network for unsupervised domain adaptation. In *CVPR*, pages 5018–5027, 2017. 2