

## A. Limitations and Scope

**Dependence on a Source-Task Model Library.** BOLT assumes access to a collection of fine-tuned source-task models from which the shared spectral bases are constructed. In scenarios where such a model library is unavailable, the proposed approach may not be directly applicable. Moreover, constructing the library itself—i.e., training multiple source-task models—can incur non-trivial computational cost.

However, we emphasize that BOLT does not require access to source-task data, but only to the resulting task vectors. In many modern model-development workflows, fine-tuned checkpoints naturally accumulate over time (e.g., in large-scale model repositories or continual deployment settings). In such cases, BOLT leverages existing artifacts rather than requiring additional meta-training or joint optimization across tasks.

Importantly, the spectral basis construction (layer-wise SVD and orthogonalization) is performed once offline and reused across all downstream tasks. This cost is amortized and independent of target-task training iterations. Our contribution is therefore not to reduce the cost of building the source library, but to demonstrate a principled and parameter-efficient way to reuse it for new-task adaptation.

**Expressiveness of Diagonal Spectral Updates.** By restricting adaptation to diagonal coefficients in the shared orthogonal basis, BOLT limits updates to a task-informed low-dimensional subspace. While this constraint improves stability and parameter efficiency, it may reduce expressiveness for tasks that lie far outside the span of the source-task manifold. Empirically, we observe strong performance across diverse few-shot and domain-shift benchmarks, suggesting that the constructed spectral bases capture broadly transferable directions, but extreme domain gaps may require expanding the rank or enriching the source library.

## B. Implementation Details

This section provides additional information about (i) the compute environment used in all experiments, and (ii) training and hyperparameter settings for few-shot, OOD, and test-time adaptation, (iii) high-level pseudo-code for the BOLT pipeline.

### B.1. Compute Environment

All experiments were conducted on the same compute environment:

- CPU: INTEL(R) XEON(R) PLATINUM 8570
- GPU: NVIDIA B200
- OS: Ubuntu 22.04.5 LTS
- Python: 3.11.12
- PyTorch: 2.7.0

Few-shot, OOD, and TTA experiments all use this software stack.

### B.2. Backbones and Trainable Parameters

We use CLIP vision encoders ViT-B/32, ViT-B/16, and ViT-L/14 as frozen backbones. The CLIP text encoder is kept fixed and is only used to construct zero-shot classifiers. For BOLT, the only trainable parameters are the layer-wise spectral diagonal coefficients in the shared orthogonal basis; all CLIP weights and bases remain frozen. For baselines (Linear Probe, LoRA, TIP, LP++, aTLAS) we follow their standard parameterizations while keeping the backbone frozen and sharing the same data splits.

### B.3. Initialization and Alpha Grid Search

Before any sigma-only adaptation (few-shot, OOD, or TTA), we perform a common initialization step:

- **Basis construction:** from a set of fine-tuned models on source tasks, we derive task vectors and compute SVD-based orthogonal bases  $U_{\text{orth}}, V_{\text{orth}}$  for each weight matrix, together with an initial diagonal coefficient vector  $\sigma$  per module.
- **Sigma parametrization:** each matrix update is represented as  $\Delta(\sigma) = U_{\text{orth}} \text{diag}(\sigma) V_{\text{orth}}$ , and the collection of all  $\sigma$  forms the only trainable parameters.
- **Global scaling:** we run a short grid search over a global scale  $\alpha \in \{1, 3, 5, 7, 10\}$ , evaluate the merged encoder  $\Theta(\alpha) = \Theta_0 + \Delta(\alpha \cdot \sigma)$  on the train data loader, and select the best  $\hat{\alpha}$ , which is then fixed for the rest of training.

This procedure is shared across all adaptation scenarios and provides a strong, data-informed initialization for the sigma parameters.

### B.4. Few-shot Adaptation Settings

In the few-shot setting, each dataset in turn is treated as a held-out target task, while the remaining datasets are used to construct the spectral basis. For each target dataset and  $k \in \{1, 2, 4, 8, 16\}$ , we sample class-balanced  $k$ -shot support sets from the training split and reuse the same indices across all methods.

Training on the target dataset uses the backbone’s standard train-time augmentations (random resized crops and horizontal flips followed by CLIP-style normalization), while evaluation uses the standard validation preprocessing (resize, center crop, normalization).

Unless otherwise stated, sigma-only fine-tuning for BOLT uses:

- Optimizer: AdamW on all sigma parameters.
- Learning rate:  $1 \times 10^{-3}$ .
- Weight decay: 0.
- Epochs: 20.
- Batch size: 32.

- Schedule: cosine learning-rate decay with a warmup of 2 epochs.

During few-shot training, all encoder weights remain frozen and only the sigma coefficients in the spectral basis are updated using a standard cross-entropy loss on the  $k$ -shot labeled examples. For each target dataset, all remaining datasets from the same domain are used to construct the spectral basis unless otherwise noted.

### B.5. OOD Training Settings

We use the same optimizer, learning rate, weight decay, epoch count, and batch size as in the few-shot setting. All methods share the same 16-shot support sets and identical training schedules.

### B.6. Test-Time Adaptation Settings

For test-time adaptation (TTA) we use a fully label-free protocol on a held-out target dataset. The model receives only unlabeled images from the target split.

**Trusted sample mining.** Using the sigma-initialized model with the chosen  $\hat{\alpha}$ , we run a forward pass over the entire target split and collect softmax predictions  $p(y | x)$  for each image. Let  $C$  be the number of classes and  $N$  the number of target samples. We select a fixed number of high-confidence “trusted” samples per class: for each class  $c$  we sort examples by  $p(y = c | x)$  and keep the top

$$k_{\text{trusted}} = \min\left(\frac{N/C}{10}, 100\right)$$

indices for class  $c$ . The union of these indices forms the trusted set  $\mathcal{D}_{\text{trusted}}$ , and the complement forms the unlabeled set  $\mathcal{D}_{\text{unlabeled}}$ . We also store one-hot targets for all samples based on the argmax predictions of this initial model; these targets are used only for trusted indices.

**Two-stream batch construction.** During TTA training we use a two-stream batch sampler. Each mini-batch of size  $B$  is constructed by sampling  $B/2$  indices from  $\mathcal{D}_{\text{unlabeled}}$  and  $B/2$  indices from  $\mathcal{D}_{\text{trusted}}$ , with independent random permutations for the two streams.

**Weak/strong augmentations.** We use an asymmetric transform to generate weak and strong views:

- **Weak view**  $x^{\text{weak}}$ : the standard validation preprocessing of the encoder (resize, center crop, normalization).
- **Strong view**  $x^{\text{strong}}$ : a composition of RandomResizedCrop to size 224 with scale (0.5,1.0) and bicubic interpolation, followed by RandomHorizontalFlip with probability 0.5, and finally the last normalization transforms from the validation pipeline.

**UFM loss with trusted samples.** Let  $\ell_{\text{UFM}}$  denote the loss. Given logits from the weak and strong views,  $\ell_{\text{UFM}}$  proceeds as follows:

1. Compute soft predictions from the weak view  $q(x) = \text{softmax}(\text{logits}^{\text{weak}})$  and apply a simple sharpening by scaling and renormalization:  $\tilde{q}(x) \propto 0.5 \cdot q(x)$ .
2. For trusted indices, overwrite  $\tilde{q}(x)$  with the one-hot targets obtained from the initial sigma-initialized model (trusted pseudo-labels).
3. Define a confidence score  $w(x) = \max_c \tilde{q}_c(x)$  and a binary mask  $m(x) = \mathbf{1}[w(x) > \tau]$  with a fixed threshold  $\tau = 0.99$ .
4. Compute the per-sample cross-entropy between the strong-view logits and the *soft* pseudo-labels, and weight it by  $m(x)$ :

$$\ell_{\text{UFM}} = \frac{1}{\sum_x m(x)} \sum_x m(x) \text{CE}(\text{logits}^{\text{strong}}(x), \tilde{q}(x)).$$

Thus, TTA optimizes a single UFM-style loss that uses high-confidence pseudo-labels from the weak view, with trusted examples anchored to fixed one-hot targets and low-confidence examples masked out.

### B.7. Pseudo-code for BOLT

Below we summarize the BOLT pipeline in two stages: (i) offline construction of a shared spectral basis and pooled diagonals, and (ii) online adaptation to a new task in spectral coordinates. The mathematical details are given in the main paper; here we focus on implementation flow.

#### Offline: shared spectral basis and pooled diagonals.

1. For each source task  $i$ , obtain a fine-tuned model  $\Theta_i$  and define the task vector  $\Delta_i = \Theta_i - \Theta_0$  with respect to the pre-trained CLIP weights  $\Theta_0$ .
2. For each layer  $\ell$ , extract the layer-wise update  $M_i^{(\ell)}$  from  $\Delta_i$  (reshaped as a matrix).
3. For each  $M_i^{(\ell)}$ , compute a thin SVD and keep the top singular directions per task and layer.
4. Stack all singular directions across tasks and apply whitening-based orthogonalization to obtain a shared spectral basis  $U_{\text{orth}}^{(\ell)}, V_{\text{orth}}^{(\ell)}$  for each layer.
5. Project each task update  $M_i^{(\ell)}$  into the shared basis and extract the diagonal coefficients  $s_i^{(\ell)}$ .
6. Average the layer-wise diagonals across tasks to obtain a pooled initializer  $s_{\text{pool}}^{(\ell)}$  for each layer.

#### Online: new task adaptation in spectral coordinates.

1. On a small held-out subset, perform a short sweep over a set of scalar scales  $\mathcal{A}$  and select  $\hat{\alpha}$  that maximizes train accuracy.
2. Initialize the trainable spectral coefficients as  $s_0^{(\ell)} = \hat{\alpha} s_{\text{pool}}^{(\ell)}$  for all layers.

Table 1. **General-domain datasets.** We report the number of classes and fully fine-tuned validation accuracy for three CLIP backbones.

Dataset	Classes	ViT-B/32	ViT-B/16	ViT-L/14
DTD	47	78.55	82.09	85.11
GTSRB	43	99.92	99.92	99.96
MNIST	10	99.56	99.50	99.70
SVHN	10	96.38	96.76	97.24
STL10	10	98.40	99.60	99.40
Oxford-IIIT Pet	37	92.39	94.84	95.92
Flowers102	102	95.10	97.06	99.02
CIFAR100	100	89.52	91.08	93.68
PCAM	2	97.36	97.86	98.04
CIFAR10	10	97.88	98.42	99.06
Food101	101	84.62	89.56	93.06
Fashion-MNIST	10	95.52	95.28	95.66
RenderedSST2	2	71.39	77.31	82.51
EMNIST	47	99.82	99.78	99.78
FGVCAircraft	100	40.65	47.28	68.11
CUB200	200	73.56	77.37	86.35
Country211	211	21.99	27.64	38.06

- For a new task, define the objective (few-shot cross-entropy or TTA loss) in terms of the parameters  $\{s^{(\ell)}\}$  only.
- For each training step:
  - Reconstruct the weight update  $\Delta(s)$  from the current diagonals  $\{s^{(\ell)}\}$  and form  $\Theta(s) = \Theta_0 + \Delta(s)$ .
  - Run a forward pass on a mini-batch and compute the loss (few-shot or TTA).
  - Backpropagate gradients into  $\{s^{(\ell)}\}$  and update them using AdamW.
- After a fixed number of epochs, evaluate the final model  $\Theta(s)$  on the target test split.

## C. Dataset Details

### C.1. General-Domain Benchmarks

The general-domain task pool includes the following 17 datasets: DTD, GTSRB, MNIST, SVHN, STL10, Oxford-IIIT Pet, Flowers102, CIFAR100, PCAM, CIFAR10, Food101, Fashion-MNIST, RenderedSST2, EMNIST, FGVCAircraft, CUB200, and Country211. We use the official train/validation/test splits whenever available, or widely adopted splits from the CLIP and prompt-tuning literature. All methods share the same splits and  $k$ -shot support sets. Details for the general-domain datasets are shown in Table 1.

### C.2. Remote-Sensing Benchmarks

The remote-sensing task pool includes the following 15 datasets: AID, CLRS, EuroSAT RGB, MLRSNet, NWPU-RESISC45, Optimal-31, PatternNet, RS C11, RSD46-WHU, RSI-CB128, RSSCN7, SAT-4, SIRI-WHU, UC

Table 2. **Remote-sensing datasets.** We report the number of classes and fully fine-tuned validation accuracy for three CLIP backbones.

Dataset	Classes	ViT-B/32	ViT-B/16	ViT-L/14
AID	10	98.50	99.00	99.17
CLRS	25	89.43	90.60	91.10
EuroSAT RGB	10	98.19	98.11	99.06
MLRSNet	30	96.28	96.39	97.12
NWPU-RESISC45	45	93.97	95.44	98.03
Optimal-31	31	95.16	95.94	96.51
PatternNet	38	99.72	99.77	99.81
RS C11	11	96.76	97.57	96.82
RSD46-WHU	46	89.61	90.35	91.82
RSI-CB128	45	99.14	99.17	99.40
RSSCN7	7	95.71	97.68	96.79
SAT-4	4	96.19	99.82	99.67
SIRI-WHU	12	97.29	98.54	98.58
UC Merced	21	98.81	99.52	99.60
WHU-RS19	19	98.51	99.50	99.58

Merced, and WHU-RS19. All remote-sensing datasets are partitioned into training, validation, and test sets using an 8:1:1 ratio. Details for the remote-sensing datasets are shown in Table 2.

### C.3. Qualitative Examples

Figure 1 provides qualitative examples from both domains. The left column shows a collage of general-domain images (textures, digits, traffic signs, objects, and animals). The right column shows satellite and aerial scenes from remote-sensing datasets (urban, agricultural, coastal, and natural regions).

## D. Extended Few-shot Results

### D.1. Detailed Few-shot Results on General-Domain Datasets

Table 5 report per-dataset few-shot accuracy for all 17 general-domain benchmarks across three CLIP backbones (ViT-B/32, ViT-B/16, and ViT-L/14). Although a few datasets contain individual settings where certain baselines (such as TIP or aTLAS) obtain competitive scores, BOLT achieves the best overall performance in the vast majority of cases. When averaging across all datasets for each backbone, BOLT consistently outperforms all competing methods at every  $k \in \{1, 2, 4, 8, 16\}$ , demonstrating strong robustness across domains, dataset sizes, and visual characteristics.

A notable trend is that the performance gap becomes larger as  $k$  decreases: with extremely small support sets (1- or 2-shot), BOLT shows substantial gains over prior parameter-efficient and training-based approaches. This confirms that constructing a shared spectral basis and learning only task-specific diagonal coefficients is particularly

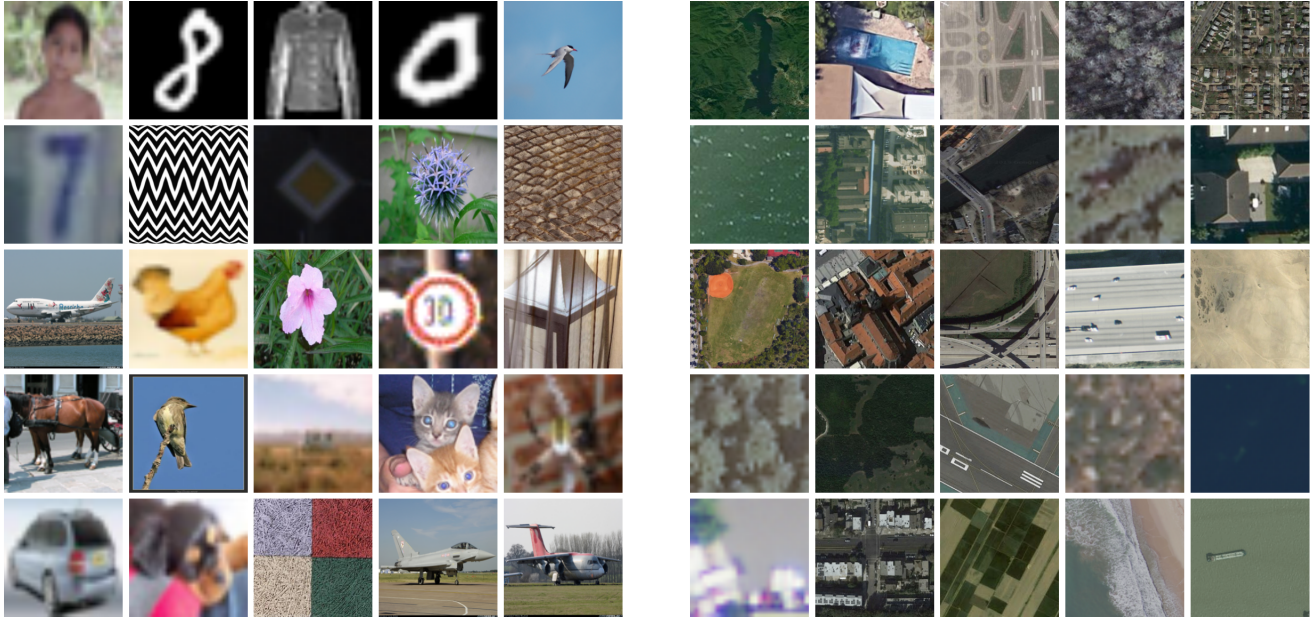


Figure 1. **Example images from the two domains.** General-domain benchmarks (left) and remote-sensing benchmarks (right).

effective in the low-data regime.

## D.2. Detailed Few-shot Results on Remote-Sensing Datasets

Complete few-shot results for all 15 remote-sensing datasets are shown in Table 6. As in the general-domain setting, BOLT achieves the strongest average accuracy for every backbone and every  $k$ -shot configuration. While a small number of datasets may exhibit close competition with alternative methods, BOLT remains the top-performing approach overall, providing stable improvements across diverse Earth-observation tasks.

The advantage of BOLT is again most pronounced in the smallest  $k$ -shot settings. When only one or two labeled samples per class are available, BOLT achieves clear margins over all baselines—highlighting that the proposed spectral subspace parameterization captures transferable task structure especially well when supervision is extremely limited.

## E. Extended OOD and TTA Results

### E.1. OOD Results

Table 3 summarizes the additional 16-shot OOD evaluation results on ImageNet variants for ViT-B/16 and ViT-L/14. Across both backbones, BOLT consistently delivers the highest top-1 accuracy on all four OOD datasets—ImageNet-A, ImageNet-R, ImageNet-S, and ImageNet-V2.

For ViT-B/16, BOLT achieves the highest accuracy on all four OOD datasets—ImageNet-A, ImageNet-R, ImageNet-

S, and ImageNet-V2—outperforming the next-best method (typically aTLAS) by noticeable margins. In particular, the gains over both linear-probing and adapter-based methods (LoRA, TIP, LP++) are substantial on the more challenging benchmarks such as ImageNet-A and ImageNet-S.

For ViT-L/14, the pattern becomes even stronger. While several baselines achieve competitive second-best numbers on individual datasets, BOLT consistently ranks first across all four OOD datasets, delivering the strongest overall robustness among all evaluated methods. The improvements are especially pronounced on ImageNet-A and ImageNet-R, where the larger backbone benefits significantly from BOLT’s diagonal spectral adaptation.

Overall, these extended results confirm that BOLT’s OOD generalization improvements are not limited to ViT-B/32 but generalize robustly across backbones of different scales. The performance margin tends to enlarge as the model capacity increases, suggesting that BOLT effectively leverages the additional representational power of larger models for robust transfer.

### E.2. Per-dataset TTA Results (General-domain)

Table 4 reports the full per-dataset test-time adaptation accuracy on all 17 general-domain datasets for the three CLIP backbones ViT-B/32, ViT-B/16, and ViT-L/14. Each entry corresponds to top-1 accuracy after label-free adaptation with the UFM-style objective and provides a more fine-grained analysis beyond the backbone-averaged results.

Across ViT-B/32 and ViT-B/16, BOLT achieves the highest average accuracy and delivers strong improvements

Table 3. **Additional OOD results on ImageNet variants (16-shot setting)**. All entries are top-1 accuracy (%). Best and second-best per dataset are shown in **bold** and underlined, respectively.

Method	ViT-B/16				ViT-L/14			
	ImageNet-A	ImageNet-R	ImageNet-S	ImageNet-V2	ImageNet-A	ImageNet-R	ImageNet-S	ImageNet-V2
Zero-shot	<u>27.67</u>	57.53	44.22	57.88	46.36	70.30	55.38	66.52
LinearProbe	27.44	58.31	44.91	59.83	46.47	70.69	55.78	68.02
LoRA	17.39	44.65	39.61	52.59	40.32	58.39	51.73	62.06
TIP	27.09	56.99	43.79	57.26	45.84	69.63	55.04	66.11
LP++	22.76	50.31	41.40	50.54	<u>47.43</u>	65.88	50.15	68.97
aTLAS	26.75	<u>59.21</u>	<u>45.62</u>	<u>60.90</u>	47.23	<u>73.02</u>	<u>57.62</u>	<u>69.11</u>
<b>BOLT</b>	<b>28.41</b>	<b>60.91</b>	<b>47.35</b>	<b>61.76</b>	<b>49.63</b>	<b>74.68</b>	<b>58.69</b>	<b>70.12</b>

over Zero-shot CLIP, LayerNorm, and aTLAS baselines. This trend is most evident on challenging fine-grained datasets such as FGVC Aircraft and CUB200, where the structured spectral-diagonal updates appear particularly effective. These results suggest that for smaller CLIP backbones, BOLT is able to capture meaningful task-specific variations even when pseudo-labels are imperfect.

For ViT-L/14, LayerNorm becomes a more competitive baseline—especially on difficult datasets—sometimes outperforming BOLT. This behavior is reasonable: larger backbones have substantially more parameters, and updating even a small subset via spectral coefficients can be less stable during TTA, whereas normalization-only adjustments (as in LayerNorm) remain lightweight and robust. Nevertheless, BOLT still achieves the best overall average accuracy among all methods on ViT-L/14, indicating that the spectral basis remains effective even at large scale.

Overall, the detailed results in Table 4 show that BOLT provides consistent gains across a wide range of datasets and model sizes. The method is especially effective for smaller backbones and for harder tasks, while remaining competitive with stronger normalization-based baselines on larger models. These findings complement the supervised few-shot experiments and demonstrate that BOLT is a robust approach for label-free test-time adaptation without additional supervised fine-tuning.

Table 4. **Per-dataset TTA accuracy (%) on general-domain datasets with ViT-B/32, ViT-B/16, ViT-L/14** Each entry reports top-1 accuracy after label-free test-time adaptation using the UFM-style objective.

Backbone	Method	DTD	GTSRB	MINIST	SVHN	STL10	OxfordIIITPet	Flowers102	CIFAR100	PCAM	CIFAR10	Food101	FashionMNIST	RenderedSST2	EMNIST	FGVCAircraft	CUB200	Country211	Average
ViT-B-32	Zero-shot CLIP	44.41	32.56	48.25	31.61	97.12	87.44	66.32	64.20	<u>60.62</u>	89.83	82.72	63.02	58.59	50.24	19.56	53.00	17.21	56.87
	LayerNorm	<u>46.22</u>	<u>41.94</u>	<u>60.22</u>	<u>59.01</u>	<u>97.17</u>	<u>89.10</u>	<u>66.74</u>	<b>73.29</b>	<b>64.93</b>	<u>93.57</u>	<u>83.73</u>	<u>75.86</u>	<u>61.34</u>	<u>61.46</u>	<u>20.04</u>	<u>53.52</u>	<b>17.42</b>	<u>62.68</u>
	aTLAS	44.41	32.56	48.25	31.61	97.12	87.44	66.32	64.20	<u>60.62</u>	89.83	82.72	63.02	58.59	50.24	19.56	53.00	17.21	56.87
	BOLT (ours)	<b>48.62</b>	<b>51.21</b>	<b>83.22</b>	<b>75.27</b>	<b>97.67</b>	<b>90.71</b>	<b>69.82</b>	<u>72.77</u>		<b>58.56</b>	<b>94.13</b>	<b>84.41</b>	<b>80.41</b>	<b>64.20</b>	<b>89.22</b>	<b>20.64</b>	<b>53.57</b>	<u>17.24</u>
ViT-B/16	Zero-shot CLIP	44.68	43.34	51.79	51.98	98.25	89.04	71.15	66.91	54.02	90.80	87.68	67.32	60.52	66.43	24.30	55.37	22.84	61.55
	LayerNorm	<u>45.74</u>	<u>49.15</u>	<u>75.59</u>	<u>70.49</u>	<u>98.40</u>	<u>91.39</u>	<u>71.98</u>	<u>75.21</u>	<b>64.37</b>	<u>94.40</u>	<u>88.39</u>	<b>78.27</b>	<u>61.07</u>	<u>79.53</u>	<u>25.11</u>	<u>57.73</u>	<u>22.93</u>	<u>67.63</u>
	aTLAS	44.68	43.34	51.79	51.98	98.25	89.04	71.15	66.91	54.02	90.80	87.68	67.32	60.52	66.57	24.30	55.37	22.84	61.56
	BOLT (ours)	<b>46.17</b>	<b>50.73</b>	<b>86.83</b>	<b>70.82</b>	<b>98.52</b>	<b>91.52</b>	<b>72.95</b>	<b>75.53</b>	<u>63.17</u>	<b>95.43</b>	<b>88.46</b>	<u>77.29</u>	<b>62.00</b>	<b>97.84</b>	<b>25.32</b>	<b>57.75</b>	<b>23.12</b>	<b>69.62</b>
ViT-L/14	Zero-shot CLIP	55.37	50.55	76.36	58.45	99.36	93.43	79.17	75.82	51.21	95.57	92.33	66.94	68.92	65.04	31.77	62.19	31.86	67.90
	LayerNorm	<u>56.81</u>	<b>58.91</b>	<b>92.19</b>	<u>71.29</u>	<b>99.45</b>	<b>94.24</b>	<u>79.26</u>	<b>83.96</b>	<u>59.11</u>	<u>97.52</u>	<b>93.02</b>	<b>79.13</b>	<b>69.36</b>	<u>93.95</u>	<b>34.20</b>	<b>64.31</b>	<b>32.60</b>	<u>74.08</u>
	aTLAS	55.37	50.55	76.36	58.45	99.36	93.43	79.17	75.82	51.21	95.57	92.33	66.94	68.92	65.34	31.77	62.19	31.86	67.92
	BOLT (ours)	<b>56.86</b>	<u>57.74</u>	<u>90.50</u>	<b>71.44</b>	<u>99.38</u>	<u>94.44</u>	<b>79.46</b>	<u>83.54</u>	<b>64.01</b>	<b>97.68</b>	<u>92.73</u>	<u>77.88</u>	64.96	<b>98.76</b>	<u>34.14</u>	<u>64.08</u>	<u>32.35</u>	<b>74.11</b>

Table 5. Few-shot general domain accuracy (%) for each dataset and backbone. Results are reported for  $k \in \{1, 2, 4, 8, 16\}$ .

Backbone	$k$	Method	DTD	GTSRB	MNIST	SVHN	STL10	OxfordIIITPet	Flowers102	CIFAR100	PCAM	CIFAR10	Food101	FashionMNIST	RenderedSST2	EMNIST	FGVCAircraft	CUB200	Country211	Average
ViT-B/32	1	Linear Probe	41.97	32.87	53.95	28.79	96.21	86.94	64.32	63.00	62.80	86.80	80.55	62.87	57.72	50.40	18.39	52.57	16.36	56.26
		LoRA	43.09	41.27	60.91	41.71	93.16	85.53	47.96	63.04	56.02	83.19	76.33	61.35	59.20	24.92	11.52	32.57	13.85	52.68
		TIP	42.93	33.05	55.98	29.14	96.17	87.19	64.69	63.25	62.70	87.08	80.53	62.78	57.06	51.92	18.72	53.04	16.41	56.63
		LP++	31.81	32.34	42.25	13.47	68.01	51.89	63.12	31.07	59.44	60.98	41.35	52.24	53.54	39.80	13.77	21.57	3.83	40.03
		aTLAS	45.64	48.21	85.12	63.64	95.29	83.65	63.59	70.29	62.22	89.78	78.53	70.38	58.15	93.19	19.53	51.05	13.47	64.22
	BOLT (ours)	<b>48.24</b>	<b>58.38</b>	<b>86.14</b>	<b>69.70</b>	<b>96.78</b>	<b>88.53</b>	<b>68.86</b>	<b>71.27</b>	<b>68.24</b>	<b>91.50</b>	<b>82.11</b>	<b>73.28</b>	<b>60.41</b>	<b>97.88</b>	<b>21.72</b>	<b>53.52</b>	<b>16.51</b>	<b>67.83</b>	
	2	Linear Probe	41.97	32.88	54.04	28.79	96.21	86.92	64.32	63.00	62.81	86.83	80.55	62.87	57.61	50.40	18.42	52.69	16.37	56.28
		LoRA	45.59	53.92	66.96	40.63	96.19	85.94	55.15	67.45	63.17	87.01	77.70	62.37	53.65	17.09	8.28	43.06	13.68	55.17
		TIP	43.72	34.35	56.90	25.75	96.14	87.24	65.62	63.42	62.46	87.28	80.53	64.22	55.96	51.04	19.29	53.09	16.54	56.68
		LP++	45.21	46.83	69.24	18.87	86.59	63.37	71.02	40.75	64.58	74.34	44.45	66.68	49.37	63.83	17.73	32.97	4.78	50.62
		aTLAS	46.38	49.96	85.68	65.07	96.54	85.20	66.40	71.70	62.31	88.66	80.78	71.12	55.41	92.23	20.94	53.23	15.59	65.13
	BOLT (ours)	<b>51.91</b>	<b>64.14</b>	<b>86.23</b>	<b>70.24</b>	<b>97.12</b>	<b>89.32</b>	<b>70.45</b>	<b>73.23</b>	<b>68.01</b>	<b>91.18</b>	<b>82.71</b>	<b>74.76</b>	<b>59.80</b>	<b>96.70</b>	<b>24.21</b>	<b>55.06</b>	<b>16.61</b>	<b>68.92</b>	
	4	Linear Probe	42.02	32.88	54.00	28.80	96.21	86.92	64.32	63.03	62.81	86.83	80.56	62.87	57.61	50.45	18.36	52.68	16.38	56.28
		LoRA	48.67	59.05	74.77	41.92	95.62	88.58	66.06	67.37	74.33	85.37	79.02	64.69	55.57	22.71	4.38	48.76	14.15	58.30
		TIP	45.32	36.14	53.90	28.65	96.25	87.46	66.38	63.88	63.28	87.57	80.66	65.88	57.28	53.27	20.04	53.37	16.46	57.40
		LP++	50.05	59.13	77.08	17.92	91.31	61.95	82.21	46.56	74.10	77.79	57.21	66.36	53.10	75.05	24.03	47.13	7.20	56.95
		aTLAS	46.49	58.08	87.89	72.50	95.58	86.89	69.10	72.89	72.52	89.06	81.43	74.31	54.09	95.70	22.68	54.12	16.41	67.63
	BOLT (ours)	<b>55.64</b>	<b>71.77</b>	<b>90.35</b>	<b>75.07</b>	<b>97.59</b>	<b>89.59</b>	<b>72.53</b>	<b>74.54</b>	<b>75.74</b>	<b>92.30</b>	<b>83.30</b>	<b>77.73</b>	<b>60.74</b>	<b>97.56</b>	<b>26.43</b>	<b>57.04</b>	<b>16.98</b>	<b>71.46</b>	
	8	Linear Probe	42.07	32.94	54.08	28.86	96.20	87.00	64.42	63.10	62.80	86.86	80.59	62.91	57.61	50.50	18.45	52.73	16.36	56.32
		LoRA	55.64	71.09	79.00	54.61	96.36	88.88	82.47	72.76	66.88	89.15	80.04	38.51	56.56	17.37	6.54	55.47	14.92	60.37
TIP		47.23	37.59	57.28	24.78	96.30	87.41	68.22	64.50	60.65	88.49	80.61	66.46	56.23	52.57	20.16	53.81	16.55	57.58	
LP++		58.67	68.77	85.59	25.56	95.53	71.65	91.53	54.79	71.14	82.14	67.73	72.73	54.64	81.83	31.74	60.85	9.75	63.76	
aTLAS		51.17	65.04	88.22	72.03	96.89	88.96	71.02	74.99	73.24	93.06	83.71	78.82	58.54	97.88	23.73	55.02	17.19	69.97	
BOLT (ours)	<b>60.00</b>	<b>79.90</b>	<b>92.43</b>	<b>79.71</b>	<b>97.46</b>	<b>90.41</b>	<b>75.62</b>	<b>76.26</b>	<b>75.52</b>	<b>92.89</b>	<b>83.71</b>	<b>80.58</b>	<b>60.52</b>	<b>97.97</b>	<b>29.94</b>	<b>60.03</b>	<b>17.70</b>	<b>73.57</b>		
16	Linear Probe	42.23	33.06	54.19	28.91	96.21	87.03	64.56	63.11	62.79	86.88	80.63	62.98	57.61	50.52	18.45	52.93	16.40	56.38	
	LoRA	61.76	80.46	84.37	61.01	96.16	88.61	85.09	76.30	71.60	89.83	81.75	36.00	60.13	20.87	23.22	63.77	16.35	64.55	
	TIP	49.95	44.83	67.94	23.40	96.28	87.82	68.53	64.69	60.68	88.77	80.94	68.62	55.96	59.27	20.55	50.04	15.64	59.17	
	LP++	64.47	74.34	90.62	33.37	96.54	79.34	91.75	61.66	71.74	85.90	74.29	76.98	54.97	87.62	36.45	67.85	12.50	68.26	
	aTLAS	53.35	68.95	90.77	80.14	96.96	90.11	71.56	75.89	73.25	93.25	84.13	82.13	53.54	97.78	23.38	55.73	17.38	71.08	
BOLT (ours)	<b>63.72</b>	<b>82.76</b>	<b>94.87</b>	<b>81.48</b>	<b>97.81</b>	<b>90.65</b>	<b>78.76</b>	<b>77.37</b>	<b>76.80</b>	<b>94.07</b>	<b>84.26</b>	<b>83.06</b>	<b>61.34</b>	<b>98.08</b>	<b>33.09</b>	<b>63.00</b>	<b>18.33</b>	<b>75.26</b>		
ViT-B/16	1	Linear Probe	42.29	42.83	62.05	50.48	97.69	87.54	67.80	62.38	58.97	88.35	85.84	64.32	58.26	69.20	23.73	53.33	21.01	60.95
		LoRA	49.15	54.80	70.67	59.94	96.95	81.47	78.35	68.93	63.50	90.25	84.50	72.08	54.75	81.15	26.82	57.78	19.30	65.32
		TIP	42.87	42.91	63.29	51.61	97.62	87.76	68.43	62.62	59.40	88.44	85.83	64.40	58.76	70.68	24.00	53.75	21.06	61.38
		LP++	31.70	33.64	50.08	18.42	65.94	53.12	67.38	33.06	60.69	67.44	52.18	43.52	52.11	51.17	17.88	28.88	5.13	43.14
		aTLAS	45.11	56.46	80.08	70.81	96.66	88.61	69.82	73.45	61.65	92.54	85.75	70.49	58.92	97.07	24.90	53.97	18.92	67.37
	BOLT (ours)	<b>47.93</b>	<b>62.30</b>	<b>88.13</b>	<b>77.42</b>	<b>98.10</b>	<b>91.44</b>	<b>73.62</b>	<b>72.89</b>	<b>68.51</b>	<b>92.93</b>	<b>87.65</b>	<b>73.63</b>	<b>61.18</b>	<b>98.92</b>	<b>27.99</b>	<b>57.40</b>	<b>21.44</b>	<b>70.67</b>	
	2	Linear Probe	42.29	42.87	62.13	50.48	97.69	87.52	67.85	62.43	58.96	88.34	85.85	64.32	58.15	69.20	23.79	53.28	21.00	60.95
		LoRA	54.73	63.15	78.40	58.01	97.52	88.20	84.00	71.84	60.57	90.60	86.28	76.36	60.68	77.16	30.93	60.13	19.43	68.12
		TIP	43.94	43.81	66.61	51.86	97.64	87.79	69.12	62.63	60.16	88.64	85.89	65.13	57.83	71.56	24.60	54.09	21.25	61.89
		LP++	45.16	49.25	69.84	22.30	87.90	71.25	75.23	43.91	62.27	75.08	53.82	62.85	51.89	70.22	23.73	40.71	6.02	53.61
		aTLAS	47.82	61.77	79.77	70.33	97.25	89.23	71.90	74.62	61.98	92.82	86.47	74.29	59.53	95.25	27.00	56.02	20.61	68.63
	BOLT (ours)	<b>52.87</b>	<b>68.76</b>	<b>87.37</b>	<b>76.87</b>	<b>98.16</b>	<b>91.31</b>	<b>77.48</b>	<b>74.28</b>	<b>69.07</b>	<b>92.83</b>	<b>88.05</b>	<b>75.42</b>	<b>61.72</b>	<b>98.78</b>	<b>30.75</b>	<b>58.77</b>	<b>21.56</b>	<b>72.00</b>	
	4	Linear Probe	42.39	42.88	62.13	50.48	97.69	87.54	67.88	62.47	58.97	88.36	85.85	64.31	58.15	69.22	23.91	53.42	21.05	60.98
		LoRA	58.51	76.06	76.68	70.31	98.12	91.82	88.57	74.70	73.74	90.48	86.49	74.09	56.84	80.14	34.14	64.08	19.97	71.46
		TIP	44.89	43.81	70.24	51.53	97.69	87.98	70.21	62.97	61.36	88.84	86.10	66.07	58.05	72.15	25.80	55.09	21.50	62.60
		LP++	51.06	59.03	76.65	25.85	94.00	64.79	86.73	51.06	75.42	78.34	64.78	63.21	57.00	82.77	30.90	55.38	8.98	60.35
		aTLAS	48.94	63.01	90.48	76.68	97.75	90.38	74.22	74.43	64.99	92.20	87.66	76.88	58.65	97.77	27.66	57.92	21.56	70.66
	BOLT (ours)	<b>57.34</b>	<b>77.86</b>	<b>90.41</b>	<b>80.11</b>	<b>98.32</b>	<b>92.42</b>	<b>81.88</b>	<b>75.31</b>	<b>75.22</b>	<b>93.51</b>	<b>88.34</b>	<b>78.30</b>	<b>62.38</b>	<b>98.32</b>	<b>33.24</b>	<b>61.30</b>	<b>22.18</b>	<b>74.50</b>	
	8	Linear Probe	42.39	42.94	62.32	50.46	97.69	87.63	68.04	62.52	58.97	88.35	85.88	64.35	58.21	69.22	23.91	53.71	21.06	61.04
		LoRA	66.60	83.46	82.54	69.71	97.66	92.34	94.44	77.73	58.33	90.80	87.59	80.05	59.03	81.39	42.09	69.73	20.25	73.75
TIP		46.91	46.70	68.83	47.71	97.72	88.69	71.70	63.62	63.06	88.80	86.19	68.27	62.27	73.81	27.00	55.16	21.56	63.41	
LP++		59.10	70.00	86.83	34.41	97.25	79.34	92.37	58.10	73.27	87.46	75.48	71.35	60.08	87.30	38.31	67.24	12.28	67.66	
aTLAS		51.60	71.16	92.64	82.03	97.70	91.85	76.17	76.54	74.20	94.20	88.28	79.68	63.21	97.93	29.88	55.25	22.82	73.43	
BOLT (ours)	<b>62.98</b>	<b>82.14</b>	<b>92.13</b>	<b>83.30</b>	<b>98.50</b>															

Table 6. Few-shot remote-sensing accuracy (%) for each dataset and backbone. Results are reported for  $k \in \{1, 2, 4, 8, 16\}$ .

Backbone	$k$	Method	AID	CLRS	EuroSAT_RGB	MLRSNet	NWPU-RESISC45	Optimal-31	PatternNet	RSD46-WHU	RS1-CB128	RSSCN7	RS-C11	SAT-4	SIRI-WHU	UC_Merced	WHU-RS19	Average
ViT-B/32	1	Linear Probe	71.67	55.33	50.37	55.71	58.59	69.62	60.90	29.34	28.74	54.11	51.42	44.95	45.83	62.38	83.58	54.84
		LoRA	<u>81.50</u>	<u>62.97</u>	<u>60.06</u>	<u>65.51</u>	<u>66.27</u>	<u>78.76</u>	<u>79.41</u>	<u>26.57</u>	<u>44.42</u>	<u>57.68</u>	<u>78.14</u>	<u>70.45</u>	<u>62.29</u>	<u>76.90</u>	<u>86.57</u>	<u>66.83</u>
		TIP	72.00	55.70	53.65	57.10	59.57	70.43	63.60	31.14	29.60	54.11	53.44	46.87	47.50	64.05	84.58	56.22
		LP++	66.67	33.40	53.11	44.12	41.62	54.30	62.86	27.14	<u>51.78</u>	43.21	57.09	41.16	51.46	51.67	61.69	49.42
		aTLAS	77.83	<u>71.77</u>	58.65	<u>72.85</u>	<u>78.95</u>	<u>92.20</u>	78.31	39.67	44.91	<u>61.07</u>	74.49	<u>74.02</u>	<u>67.29</u>	<u>87.86</u>	<u>92.54</u>	<u>71.49</u>
		BOLT (ours)	<b>92.83</b>	<b>76.73</b>	<b>79.30</b>	<b>78.68</b>	<b>81.79</b>	<b>92.74</b>	<b>81.97</b>	<b>46.83</b>	<b>54.96</b>	<b>77.50</b>	<b>78.95</b>	<b>83.89</b>	<b>76.46</b>	<b>89.76</b>	<b>96.52</b>	<b>79.26</b>
	2	Linear Probe	71.67	55.43	50.37	55.83	58.67	69.62	61.28	29.48	28.82	54.11	51.82	44.89	45.83	62.38	83.58	54.92
		LoRA	84.83	<u>67.50</u>	<u>69.93</u>	70.30	69.86	81.72	<u>85.87</u>	38.81	59.48	62.50	<u>82.19</u>	65.00	65.62	85.95	89.05	71.91
		TIP	74.00	56.73	55.54	58.01	60.76	71.24	65.64	34.50	31.24	54.29	56.28	48.20	49.79	66.43	86.07	57.91
		LP++	79.33	51.07	69.65	57.53	62.51	69.35	79.52	<u>46.26</u>	<b>69.50</b>	59.82	71.26	55.22	60.42	68.81	79.60	65.32
		aTLAS	<u>88.67</u>	<u>73.47</u>	63.87	<u>75.95</u>	<u>80.06</u>	<b>94.62</b>	78.42	41.27	47.04	<u>64.82</u>	76.52	<u>67.56</u>	<u>70.62</u>	<u>90.00</u>	<u>98.01</u>	<u>74.06</u>
		BOLT (ours)	<b>96.00</b>	<b>78.97</b>	<b>84.24</b>	<b>79.13</b>	<b>82.14</b>	<b>93.82</b>	<b>87.40</b>	<b>49.91</b>	<b>60.90</b>	<b>80.00</b>	<b>87.85</b>	<b>85.67</b>	<b>82.08</b>	<b>93.57</b>	<b>98.51</b>	<b>82.68</b>
	4	Linear Probe	71.67	55.50	50.48	56.04	59.05	70.16	61.45	30.31	29.13	54.11	52.23	45.20	45.83	62.38	83.58	55.14
		LoRA	<u>92.17</u>	71.30	<u>72.87</u>	77.40	78.33	86.56	<b>90.56</b>	51.83	59.23	73.75	84.21	<u>83.48</u>	<u>80.21</u>	89.52	96.02	<u>79.50</u>
		TIP	75.33	58.07	58.63	60.62	63.05	73.92	69.31	37.56	34.62	54.64	62.35	53.99	53.54	69.29	88.06	60.86
		LP++	87.50	61.40	69.43	64.75	64.83	73.12	83.75	<u>53.60</u>	<b>78.17</b>	<u>75.89</u>	<u>86.23</u>	79.66	73.75	81.19	94.03	75.15
		aTLAS	91.83	<u>76.60</u>	74.46	<u>78.94</u>	<u>83.24</u>	<b>95.70</b>	81.58	45.80	49.10	63.04	84.21	79.59	77.08	91.19	<b>98.51</b>	78.06
		BOLT (ours)	<b>96.33</b>	<b>81.67</b>	<b>89.30</b>	<b>81.83</b>	<b>84.63</b>	<b>94.89</b>	<u>90.12</u>	<b>54.20</b>	<u>70.10</u>	<b>85.18</b>	<b>91.50</b>	<b>89.42</b>	<b>87.92</b>	<b>93.57</b>	<b>98.01</b>	<b>85.91</b>
8	Linear Probe	71.67	56.07	50.94	56.49	59.75	70.70	62.19	30.99	29.68	54.29	53.04	45.14	45.83	62.86	84.58	55.61	
	LoRA	95.17	75.77	<u>86.30</u>	<u>82.18</u>	81.62	90.32	<b>94.56</b>	57.71	84.19	80.36	90.69	<u>89.45</u>	<u>85.21</u>	<u>94.05</u>	<u>95.02</u>	<u>85.51</u>	
	TIP	78.67	61.37	63.41	64.85	65.94	77.42	75.33	42.87	38.57	52.50	70.45	56.19	58.13	69.76	90.55	64.40	
	LP++	<u>95.67</u>	71.80	78.80	74.97	75.25	80.38	90.95	<b>63.90</b>	<b>86.56</b>	<u>82.14</u>	90.28	83.41	81.04	83.33	93.53	82.13	
	aTLAS	92.67	<u>78.87</u>	77.98	81.09	<u>84.48</u>	<u>96.77</u>	84.44	46.78	54.52	70.18	85.43	76.57	83.54	92.86	<u>98.01</u>	<u>80.28</u>	
	BOLT (ours)	<b>97.17</b>	<b>82.20</b>	<b>92.07</b>	<b>84.12</b>	<b>85.71</b>	<b>97.31</b>	<u>93.34</u>	<u>58.45</u>	80.97	<b>90.00</b>	<b>93.12</b>	<b>91.00</b>	<b>91.25</b>	<b>98.81</b>	<b>98.51</b>	<b>88.73</b>	
16	Linear Probe	72.83	56.70	51.02	57.64	60.92	71.51	63.31	32.91	30.63	54.64	55.47	46.44	46.88	64.76	84.58	56.68	
	LoRA	97.17	78.93	92.00	85.92	85.70	92.47	<b>96.61</b>	<b>71.58</b>	<b>93.64</b>	86.43	94.33	92.51	91.67	96.19	98.51	90.24	
	TIP	83.17	64.53	67.22	70.10	69.02	80.65	81.20	47.12	46.49	62.32	72.87	60.34	68.54	75.24	92.04	69.39	
	LP++	96.17	75.33	83.46	81.98	80.14	87.90	<u>95.07</u>	<u>70.89</u>	<u>91.17</u>	85.54	92.31	88.24	89.38	92.86	97.01	87.16	
	aTLAS	96.17	79.43	88.17	82.70	84.78	<u>96.51</u>	85.87	48.49	56.13	85.36	91.50	88.72	86.46	94.76	98.01	84.20	
	BOLT (ours)	<b>97.83</b>	<b>83.97</b>	<b>93.52</b>	<b>86.76</b>	<b>87.60</b>	<b>96.04</b>	94.77	62.41	85.03	<b>91.96</b>	<b>94.74</b>	<b>93.08</b>	<b>94.17</b>	<b>96.90</b>	<b>99.50</b>	<b>90.62</b>	
ViT-B/16	1	Linear Probe	74.83	60.17	50.02	62.77	64.16	71.77	63.50	33.30	28.26	50.36	63.56	46.81	53.54	62.86	84.08	58.00
		LoRA	<u>88.33</u>	<u>66.63</u>	<u>74.26</u>	<u>72.75</u>	71.95	76.61	<u>84.46</u>	34.53	50.35	<u>67.50</u>	<u>82.19</u>	<u>65.75</u>	47.08	79.29	89.55	70.08
		TIP	75.50	60.83	53.41	63.81	65.19	73.39	65.84	35.36	30.36	50.89	65.18	51.90	55.42	63.81	86.07	59.80
		LP++	69.00	44.47	56.85	51.56	52.95	61.02	67.04	29.54	<u>55.20</u>	30.00	64.37	45.57	50.00	60.24	63.18	53.40
		aTLAS	82.83	51.03	59.83	<u>77.55</u>	<u>78.19</u>	<b>92.47</b>	78.29	<u>44.15</u>	46.27	65.00	79.76	63.22	68.54	85.00	<u>96.52</u>	<u>71.24</u>
		BOLT (ours)	<b>91.50</b>	<b>77.40</b>	<b>79.20</b>	<b>79.18</b>	<b>83.25</b>	<u>91.40</u>	<b>86.50</b>	<b>51.86</b>	<b>56.97</b>	<b>73.04</b>	<b>91.09</b>	<b>77.86</b>	<b>76.88</b>	<b>89.29</b>	<b>97.01</b>	<b>80.16</b>
	2	Linear Probe	74.83	60.17	50.04	62.79	64.21	71.77	63.67	33.45	28.28	50.36	63.56	46.47	53.54	62.86	85.07	58.07
		LoRA	<u>89.17</u>	56.33	<u>82.83</u>	74.48	76.65	83.60	<u>90.12</u>	<u>49.26</u>	58.23	59.82	49.39	57.89	<u>72.71</u>	83.57	71.14	70.35
		TIP	76.33	61.30	58.26	64.66	66.32	75.81	68.32	38.01	32.69	52.86	64.78	54.19	56.25	64.29	87.06	61.41
		LP++	70.50	58.10	69.98	61.27	66.97	71.77	81.07	47.20	<b>69.19</b>	61.25	<u>80.16</u>	55.89	62.92	61.10	85.57	67.33
		aTLAS	80.33	<u>74.90</u>	51.69	<u>77.89</u>	<u>80.73</u>	<b>94.89</b>	81.17	45.55	50.64	<u>73.39</u>	76.52	<u>64.13</u>	<u>70.62</u>	<u>88.81</u>	<u>96.52</u>	<u>73.85</u>
		BOLT (ours)	<b>93.50</b>	<b>80.37</b>	<b>85.50</b>	<b>80.27</b>	<b>85.06</b>	<u>92.74</u>	<b>90.38</b>	<b>54.59</b>	<b>64.90</b>	<b>82.14</b>	<b>89.47</b>	<b>73.20</b>	<b>84.38</b>	<b>93.81</b>	<b>98.01</b>	<b>83.22</b>
	4	Linear Probe	75.00	60.40	50.17	62.96	64.59	72.58	64.14	33.82	28.49	50.36	63.97	46.38	53.54	62.86	85.07	58.29
		LoRA	<u>94.50</u>	<u>73.83</u>	<u>88.46</u>	77.44	73.98	87.37	<u>91.35</u>	<u>57.28</u>	<u>75.93</u>	<u>75.54</u>	56.28	86.22	68.33	86.19	97.01	<u>79.31</u>
		TIP	78.83	63.00	62.00	66.18	68.16	76.17	72.19	41.24	38.15	56.07	68.83	58.55	58.33	67.14	89.05	64.29
		LP++	93.17	62.50	75.59	69.27	69.25	76.08	84.80	56.51	<b>75.97</b>	71.61	86.23	<u>86.50</u>	<u>79.38</u>	76.67	88.06	76.77
		aTLAS	80.16	60.13	63.48	<u>79.93</u>	<u>83.06</u>	<u>94.62</u>	83.98	47.92	55.54	74.29	<u>87.04</u>	45.12	75.62	90.24	<u>98.51</u>	74.93
		BOLT (ours)	<b>96.17</b>	<b>82.87</b>	<b>90.33</b>	<b>81.67</b>	<b>87.22</b>	<b>95.16</b>	<b>92.94</b>	<b>61.07</b>	74.72	<b>87.32</b>	<b>94.74</b>	<b>88.33</b>	<b>88.12</b>	<b>96.90</b>	<b>99.50</b>	<b>87.80</b>
8	Linear Probe	75.00	60.73	50.63	63.40	65.06	73.66	65.38	35.05	29.42	50.18	63.97	46.27	53.75	63.10	85.57	58.74	
	LoRA	95.00	79.77	<b>91.69</b>	<b>84.60</b>	80.75	90.05	<b>95.89</b>	<b>67.35</b>	80.92	86.79	<b>95.14</b>	84.78	88.96	89.05	98.51	87.28	
	TIP	82.50	64.57	66.30	69.44	71.10	79.57	77.65	45.83	45.36	59.46	72.06	63.20	61.46	71.90	90.55	68.06	
	LP++	94.67	72.00	80.41	78.35	79.44	83.33	92.58	<u>65.33</u>	<b>86.64</b>	84.82	92.31	84.61	82.71	85.71	89.05	83.46	
	aTLAS	87.50	70.63	83.11	81.25	<u>81.63</u>	<u>96.51</u>	86.94	50.91	61.10	78.39	71.26	83.06	82.92	94.76	97.01	80.47	
	BOLT (ours)	<b>96.83</b>	<b>83.90</b>	<u>91.67</u>	<u>84.27</u>	<b>89.21</b>	<b>97.04</b>	<u>95.23</u>	64.36	80.32	<b>89.64</b>	<u>94.33</u>	<b>91.52</b>	<b>92.08</b>	<b>97.86</b>	<b>99.00</b>	<b>89.82</b>	
16	Linear Probe	75.33	61.33	51.19	64.05	66.27	74.19	66.58	36.90	31.35	50.71	65.18	46.62	54.58	63.81	86.57	59.64	
	LoRA	95.33	<u>82.70</u>	<b>94.24</b>	<b>87.70</b>	86.06	93.82	<b>97.86</b>	<b>76.03</b>	<b>93.57</b>	<u>90.71</u>	<u>95.14</u>	88.31	<u>93.96</u>	96.43	<b>99.50</b>	<u>91.42</u>	
	TIP	87.17	69.63	70.54	72.81	75.10	82.80	83.77										