LoRA Subtraction for Drift-Resistant Space in Exemplar-Free Continual Learning

Supplementary Material

I. Experiments Across Diverse Datasets

We apply our method to DomainNet and CUB datasets, following the experimental setup of InfLoRA [2] and EASE [5], respectively. As shown in Table A, while our method does not achieve the highest accuracy on DomainNet, it performs comparably to the SOTA methods. DomainNet consists of five short tasks, where our method's strengths are less evident. However, on CUB, which involves longer tasks, our method excels with an $\overline{ACC}_{20}$ of 92.78%, outperforming both InfLoRA and EASE.

Table A. Comparisons in DomainNet and CUB.

Method	DomainNet		CUB	
	ACC_5	$\overline{ACC}_5$	ACC_{20}	$\overline{ACC}_{20}$
InfLoRA	69.68	**76.93**	62.68	76.57
EASE	66.39	72.21	86.13	91.68
Ours	**70.37**	76.65	**87.74**	**92.78**

II. Variants for Computing Drift-Resistance Space

We conduct experiments to evaluate the effectiveness of the LoRA subtraction method. Specifically, we employ the initial pre-trained weights W_0 to design DRS. Tab. B presents the results of our method alongside its variant. From the results, we observe that the variant does not perform as effectively as our method, thus LoRA$^-$ is necessary.

Table B. DRS Computation with W_0 v.s. LoRA$^-$ on CIFAR-100.

	ACC_{10}	$\overline{ACC}_{10}$	ACC_{50}	$\overline{ACC}_{50}$
$W_0 \rightarrow$ DRS	63.04	90.33	52.88	78.32
$LoRA^- \rightarrow$ DRS	**89.40**	**92.78**	**86.82**	**91.29**

III. Further Performance Analysis

We provide a detailed analysis of the performance of the old and new classes compared to existing methods [1–5], as shown in Tab. C. Specifically, the results demonstrate a 5.5% improvement in A_{old} over EASE, and a 9.5% improvement in A_{new} compared to EASE. Our method is consistently better than all methods on both tasks and thus more stable and plastic.

Table C. The old class accuracy (A_{old}) and new class accuracy (A_{new}) at different stages of CIFAR-100 50 tasks.

	Stage-10		Stage-20		Stage-40		Stage-50	
	A_{old}	A_{new}	A_{old}	A_{new}	A_{old}	A_{new}	A_{old}	A_{new}
LAE	90.17	92.5	83.13	83.5	79.22	80.5	73.54	85.5
L2P	88.72	91.5	79.95	94.5	76.33	90.0	76.72	68.5
InfLoRA	83.83	79.0	70.76	94.0	61.29	67.0	60.84	78.5
Adam-NSCL	78.67	61.5	65.24	79.0	56.83	59.5	53.19	68.5
EASE	92.67	95.0	87.53	94.0	81.77	89.5	81.34	76.5
Ours	**95.61**	**95.0**	**91.34**	**94.5**	**87.53**	**92.5**	**86.84**	**86.0**

IV. Memory Usage and Storage Efficiency

Existing methods typically store statistical information from previous tasks to reduce the impact of feature drift. We compare the memory usage of prior methods that explore related ideas, including InfLoRA and Adam-NSCL. As shown in Tab. D, our method demonstrates the lowest memory requirement with ViT-B/16-IN21K for CIFAR100 50 tasks, as it does not retain statistics from previous tasks. Our LoRA$^-$ approach enables efficient storage and computation while effectively handling feature drifts without explicit feature modeling.

Table D. Memory usage of storing statistics on CIFAR100 50-task.

Method	Adam-NSCL	InfLoRA	Ours
Memory (KB)	720.9	2861.3	**0.0**

References

[1] Qiankun Gao, Chen Zhao, Yifan Sun, Teng Xi, Gang Zhang, Bernard Ghanem, and Jian Zhang. A unified continual learning framework with general parameter-efficient tuning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 11483–11493, 2023.

[2] Yan-Shuo Liang and Wu-Jun Li. Inflora: Interference-free low-rank adaptation for continual learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 23638–23647, 2024.

[3] Shipeng Wang, Xiaorong Li, Jian Sun, and Zongben Xu. Training networks in null space of feature covariance for continual learning. In *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*, pages 184–193, 2021.

[4] Zifeng Wang, Zizhao Zhang, Chen-Yu Lee, Han Zhang, Ruoxi Sun, Xiaoqi Ren, Guolong Su, Vincent Perot, Jennifer Dy, and Tomas Pfister. Learning to prompt for continual learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 139–149, 2022.

[5] Da-Wei Zhou, Hai-Long Sun, Han-Jia Ye, and De-Chuan Zhan. Expandable subspace ensemble for pre-trained model-based class-incremental learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 23554–23564, 2024.