Class-Incremental Learning using Diffusion Model for Distillation and Replay
Supplementary materials

Quentin Jodelet[1,2], Xin Liu[2], Yin Jun Phua[1], Tsuyoshi Murata[1,2]
[1] Department of Computer Science, Tokyo Institute of Technology, Japan
[2] Artificial Intelligence Research Center, AIST, Japan
jodelet@net.c.titech.ac.jp, xin.liu@aist.go.jp, phua@c.titech.ac.jp, murata@c.titech.ac.jp

1. Alternative version of ImageNet-Subset

Additionally, we considered an alternative set of classes for ImageNet-Subset, denoted as "ImageNet-Subset Alt." in this work, used by the authors of FOSTER [4]. In Table 1, we report the performance of FOSTER combined with our method on this variation of ImageNet-Subset.

Compared to ImageNet-Subset, our proposed method SDDR achieves both higher average incremental accuracy and higher improvement over the baseline method on ImageNet-Subset Alt. .

Methods	ImageNet-Subset Alt.		
	T=5	10	25
FOSTER*	78.52	76.49	71.24
FOSTER w/ SDDR (ours)	80.35	79.21	77.09
Improvement in *p.p.*	+1.83	+2.72	+5.85

Table 1. Average incremental accuracy (Top-1) on ImageNet-Subset Alt. with a base step containing half of the classes followed by 5, 10, and 25 incremental steps. "ImageNet-Subset Alt." is a different version of ImageNet-Subset following the definition of the authors of FOSTER [4]. Results marked with "*" correspond to our own experiments. Results averaged over 3 random runs.

2. Dual branch FOSTER

In our main experiments, to fairly compare our proposed method with other methods, we only evaluated the accuracy of FOSTER after performing the feature compression. To better compare with expansion-based method such as DER [5] whose number of parameters is growing at each incremental steps, we report in Table 2 the accuracy of the dual branch FOSTER model before the compression, denoted FOSTER-B4.

Experiments show that, if the memory and computational budget of the system allow it, using the dual branch FOSTER for inference with our method further improve the accuracy. On CIFAR100, FOSTER combined with SDDR achieves higher performances than DER for both single and dual branch evaluation. Furthermore, on ImageNet-Subset with 10 incremental steps, FOSTER-B4 combined with SDDR achieves a Top-1 Average incremental accuracy 0.49*p.p.* lower and a Top-5 Average incremental accuracy 0.50*p.p.* higher than DER while requiring about five times less parameters.

3. CIFAR-100 with 5 incremental steps

Following Liu *et al.* [2], we report in Table 3 additional results for the different baseline methods on CIFAR100 with a base step containing half of the classes followed by 5 incremental steps depending on the number of exemplars saved in memory for each class.

When FOSTER [4] is used with a really small replay memory, we observed that the improvement resulting from using our proposed method SDDR is limited. We suppose that this is due to the logits alignment loss used by the authors that too strongly limits the plasticity of the model and restricts our proposed method. By tuning the hyperparameters of the logits aligment loss, it should be possible to further improve the performance of FOSTER when combined with our proposed method.

References

[1] Saihui Hou, Xinyu Pan, Chen Change Loy, Zilei Wang, and Dahua Lin. Learning a unified classifier incrementally via rebalancing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[2] Yaoyao Liu, Yingying Li, Bernt Schiele, and Qianru Sun. Online placebos for class-incremental learning, 2023.

[3] Sylvestre-Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl, and Christoph H. Lampert. icarl: Incremental classifier and representation learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[4] Fu-Yun Wang, Da-Wei Zhou, Han-Jia Ye, and De-Chuan Zhan. Foster: Feature boosting and compression for class-

Methods	CIFAR100			ImageNet-Subset		
	T=5	10	25	5	10	25
DER (w/o Pruning) [5]	68.52	67.09	-	-	**78.20**	-
FOSTER* [4]	71.17	68.89	65.07	76.09	75.05	70.96
FOSTER w/ SDDR (ours)	<u>72.18</u>	<u>70.88</u>	<u>68.06</u>	77.13	76.77	<u>75.50</u>
FOSTER-B4* [4]	72.08	69.32	65.41	<u>77.48</u>	76.11	71.40
FOSTER-B4 w/ SDDR (ours)	**73.28**	**71.43**	**68.36**	**78.48**	<u>77.71</u>	**75.97**

Table 2. Average incremental accuracy (Top-1) on CIFAR100, and ImageNet-Subset with a base step containing half of the classes followed by 5, 10, and 25 incremental steps, using a growing memory of 20 exemplars per class. Results for DER are reported from [5]. Results marked with "*" correspond to our own experiments. Results on CIFAR100 and ImageNet-Subset are averaged over 3 random runs. Best result is marked in bold and second best is underlined.

Methods	CIFAR100 with T=5							
	5 exemplars/class		10 exemplars/class		20 exemplars/class		50 exemplars/class	
	Average	Last	Average	Last	Average	Last	Average	Last
iCaRL* [3]	43.40	31.90	52.30	40.11	57.68	47.45	62.09	54.23
w/ PlaceboCIL [2]	51.55	39.35	59.11	46.42	61.24	51.47	-	-
w/ SDDR (ours)	55.36	43.13	58.53	48.33	62.01	52.91	65.53	58.46
LUCIR* [1]	53.14	41.52	60.77	49.65	63.37	53.91	65.63	57.58
w/ PlaceboCIL [2]	62.74	53.25	64.79	55.44	65.28	56.23	-	-
w/ SDDR (ours)	62.90	50.94	64.47	54.53	65.77	57.00	67.21	59.73
FOSTER* [4]	56.94	43.74	63.09	54.34	71.17	63.97	70.00	64.10
w/ PlaceboCIL [2]	62.78	50.72	65.12	54.81	71.97	64.43	-	-
w/ SDDR (ours)	57.76	44.32	63.26	54.32	72.18	64.78	71.66	66.04

Table 3. Performances on CIFAR100 with a base step containing half of the classes followed by 5 incremental steps depending on the number of exemplars saved in memory for each class. Results for PlaceboCIL reported from [2]. Results marked with "*" correspond to our own experiments. "Average" is Average Incremental Accuracy (Top-1) and "Last" is the final overall accuracy of the model after the last incremental step. Results averaged over 3 random runs.

incremental learning. In *European conference on computer vision*, pages 398–414. Springer, 2022.

[5] Shipeng Yan, Jiangwei Xie, and Xuming He. Der: Dynamically expandable representation for class incremental learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3014–3023, 2021.