1 SUPPLEMENTARY MATERIAL

To demonstrate the benefits of our proposed method, we also compared the DF-IL on CIFAR-100 with prominent works: iCaRL (Rebuffi et al., 2017), DER (Buzzega et al., 2020), Mnemonics (Liu et al., 2020), oEWC (Schwarz et al., 2018), BiC(Wu et al., 2019) and a-GEM (Chaudhry et al., 2018) as shown in Table 1.

Table 1: Classification results (average accuracy) for IL benchmarks, averaged across 10 runs. Best performance is marked in bold in each section. CPT stands for number of Classes Per Task.

	Class-IL		**Task-IL**	
CPT	**5**	**10**	**5**	**10**
oEWC	15.32	28.29	19.76	29.65
iCaRL	22.32	26.12	48.46	59.12
Mnemonics	35.41	44.93	62.73	**75.20**
a-GEM	23.54	37.62	55.23	62.21
BIC	28.12	38.36	61.76	68.47
DER	39.71	49.87	**62.51**	74.01
DF-IL (Ours)	**42.39**	**51.87**	61.01	73.31

REFERENCES

Pietro Buzzega, Matteo Boschini, Angelo Porrello, Davide Abati, and Simone Calderara. Dark experience for general continual learning: a strong, simple baseline. *arXiv preprint arXiv:2004.07211*, 2020.

Arslan Chaudhry, Marc'Aurelio Ranzato, Marcus Rohrbach, and Mohamed Elhoseiny. Efficient lifelong learning with a-gem. *arXiv preprint arXiv:1812.00420*, 2018.

Yaoyao Liu, Yuting Su, An-An Liu, Bernt Schiele, and Qianru Sun. Mnemonics training: Multi-class incremental learning without forgetting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12245–12254, 2020.

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*, pp. 2001–2010, 2017.

Jonathan Schwarz, Wojciech Czarnecki, Jelena Luketina, Agnieszka Grabska-Barwinska, Yee Whye Teh, Razvan Pascanu, and Raia Hadsell. Progress & compress: A scalable framework for continual learning. In *International Conference on Machine Learning*, pp. 4528–4537. PMLR, 2018.

Yue Wu, Yinpeng Chen, Lijuan Wang, Yuancheng Ye, Zicheng Liu, Yandong Guo, and Yun Fu. Large scale incremental learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 374–382, 2019.