

# Supplementary Materials: Prior-free Balanced Replay: Uncertainty-guided Reservoir Sampling for Long-Tailed Continual Learning

Anonymous Authors

## 1 APPENDIX

### 1.1 Dataset Details

**Seq-CIFAR-10-LT and Seq-CIFAR-100-LT.** The original versions of CIFAR-10 and CIFAR-100 contain 50,000 training and 10,000 testing images, respectively. Moreover, the class numbers are 10 and 100. Following [1], the long-tailed versions of them are sampled from the original ones with different imbalance ratios, mainly including 0.01, 0.02, 0.05, and 0.1. The testing set is unchanged. For the CL setting, the dataset is split into 5 tasks, each of which has 2 or 20 classes.

**Seq-TinyImageNet-LT.** The original versions of TinyImageNet contain 100,000 training and 10,000 testing images, respectively. The class number is 200. Following [1], the long-tailed version is sampled from the original ones with different imbalance ratios (IR), mainly including 0.01, 0.02, 0.05, and 0.1. The testing set is unchanged. For the CL setting, the dataset is split into 10 tasks, each of which has 20 classes.

### 1.2 Varying Imbalance Ratios

**Different IRs on Seq-CIFAR10-LT and Seq-CIFAR100-LT under Ordered-LTCL.** In this section, we provide the detailed experimental results under the ordered-LTCL setting on Seq-CIFAR10-LT and Seq-CIFAR100-LT with various IRs and buffer sizes. As shown in Table 2 and 3, our method outperforms previous methods by a large margin for both Class-IL and Task-IL settings. Particularly, with the different buffer sizes, the overall accuracy on the total classes is improved compared with the previous strong baselines DER and DER++. Additionally, it is observed that the methods resorting to gradients (GEM, A-GEM, GSS) seem to be less effective under this scenario, since the gradient information is influenced by imbalance distribution. In particular, compared to rehearsal methods, our method exhibits strong performance in most settings.

**Different IRs on Seq-TinyImageNet-LT under Ordered-LTCL.** Table 4 reports the detailed experimental results under the ordered-LTCL setting on the Seq-TinyImageNet-LT dataset with various IRs and buffer sizes. The observations are consistent with those on Seq-CIFAR10-LT and Seq-CIFAR100-LT datasets. Our method could achieve SOTA performance in almost all settings on such a large-scale dataset. We notice that a larger buffer size generally performs higher accuracy results but has a higher probability of influencing the learning of new tasks. Thus, we could conclude that boundary supporting samples carry more valuable information to prevent forgetting. Besides, the similarity and prototype information can be preserved in the memory.

**Visualization of Selected Samples.** Visualization can intuitively reflect the role of uncertainty-guided sampling for the memory buffer. During task evolution on the Seq-CIFAR-10-LT, we report the 2D t-SNE visualization of selected samples in memory

---

#### Algorithm 1 Reservoir Sampling with Uncertainty

---

**Input:** buffer  $\mathcal{M}$ , parameters  $\theta$ , number of seen samples  $N$ , current data  $D_c$

```

 $\hat{D}_c[\cdot] \xleftarrow{MI} D_c, i = 0, N++$ 
for  $l = 0$  in  $\text{range}(\text{length}(D_c))$  do
  if  $|\mathcal{M}| > N$  then
     $\mathcal{M} \leftarrow \hat{D}_c[i++]$ 
  else
     $j = \text{randomInteger}(\text{min} = 0, \text{max} = N)$ 
    if  $j < |\mathcal{M}|$  then
       $\mathcal{M} \leftarrow \hat{D}_c[i++]$ 
return  $\mathcal{M}$ 

```

---



---

#### Algorithm 2 Balanced Experience Replay++

---

**Input:** dataset  $D$ , parameters  $\theta$ , learning rate  $\eta$ , scalar  $\tau_1$ ,  $\tau_2$ , and scalar  $\alpha$

```

 $\mathcal{M} \leftarrow \{\}$ 
while D not end do
  for  $(x, y)$  in  $\{D \cup \mathcal{M}\}$  do
     $\mathcal{L}_{kl} \leftarrow D_{KL}(f_{\theta^*}(x) || f_{\theta}(x))$ 
     $\mathcal{L}_{tc} \leftarrow \text{MCE}(f_{\theta}(x), y)$ 
     $\theta \leftarrow \theta + \eta \nabla_{\theta}(\mathcal{L}_{tc} + \alpha \mathcal{L}_{kl})$ 
return Updated  $\theta$ 

```

---

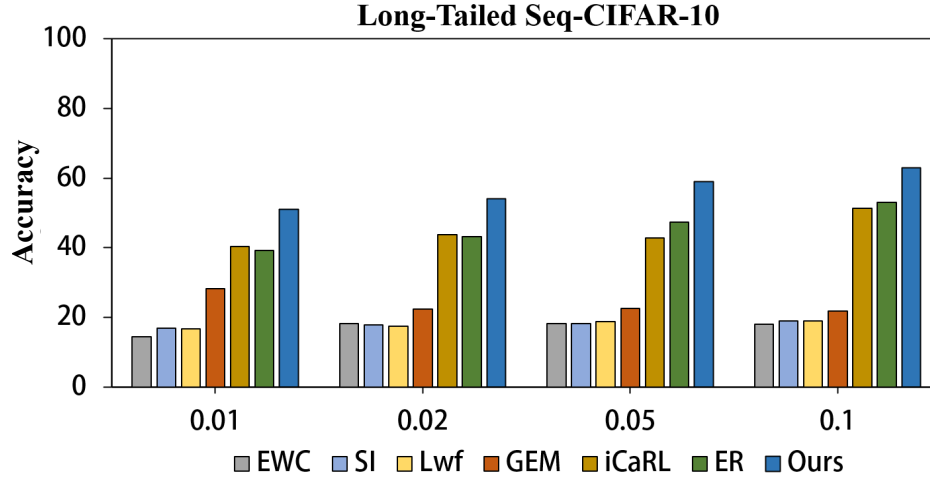
buffer for each task under the ordered-LTCL setting. As shown in Figure 3, different columns denotes the different tasks. Colorful points are associated with the samples that are selected into the memory buffer, while gray points are not. Guided by the uncertainty, most boundary supporting samples are stored in the memory buffer, which can reduce the influence of the forgetting problem. Due to class imbalance, new classes with fewer samples are influenced by both forgetting and majority classes, thus producing more uncertain samples for each task. Besides, with increasing buffer size, more boundary supporting samples of old classes can be stored, which can help to model task boundaries with new data.

## REFERENCES

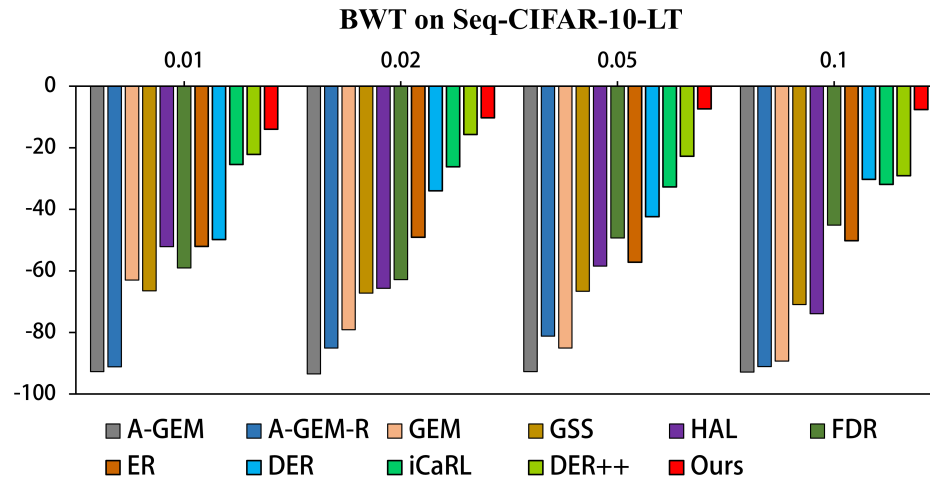
- [1] Bingyi Kang, Saining Xie, Marcus Rohrbach, Zhicheng Yan, Albert Gordo, Jiashi Feng, and Yannis Kalantidis. 2019. Decoupling Representation and Classifier for Long-Tailed Recognition. In *International Conference on Learning Representations*.
- [2] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, et al. 2017. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences* 114, 13 (2017), 3521–3526.
- [3] Zhizhong Li and Derek Hoiem. 2017. Learning without forgetting. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40, 12 (2017), 2935–2947.
- [4] David Lopez-Paz and Marc Aurelio Ranzato. 2017. Gradient episodic memory for continual learning. *Advances in Neural Information Processing Systems* 30 (2017),

**Table 1: Detailed Dataset Information.**

Dataset	#Class	#Training	#Test	#Task
Seq-CIFAR-10-LT	10	$\leq 20,431$	10,000	5
Seq-CIFAR-100-LT	100	$\leq 19,573$	10,000	5
Seq-TinyImageNet-LT	200	$\leq 21,549$	10,000	10



**Figure 1: Performance comparisons of several classical CL methods (EWC [2], SI [8], Lwf [3], GEM [4], iCaRL [6], and ER [5, 7]) and the proposed PBR method. Existing method mainly focus on weight regularization or experience relay to alleviate the catastrophic forgetting problem, which still suffer from the imbalance issue due to the deteriorated feature space. Our method can obtain higher recognition accuracy via learning an evolved feature space with prototype and boundary constraints.**



**Figure 2: BWT results on Seq-CIFAR-10-LT under the ordered-LTCL setting. Buffer size is 200. Our method can achieve best results compared previous solutions.**

6467–6476.

[5] Roger Ratcliff. 1990. Connectionist models of recognition memory: constraints imposed by learning and forgetting functions. *Psychological Review* 97, 2 (1990), 285.

[6] Sylvestre-Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl, and Christoph H Lampert. 2017. icarl: Incremental classifier and representation learning. In *Proceedings*

of the IEEE conference on Computer Vision and Pattern Recognition. 2001–2010.

[7] Matthew Riemer, Ignacio Cases, Robert Ajemian, Miao Liu, Irina Rish, Yuhai Tu, and Gerald Tesauro. 2018. Learning to Learn without Forgetting by Maximizing Transfer and Minimizing Interference. In *International Conference on Learning Representations*.

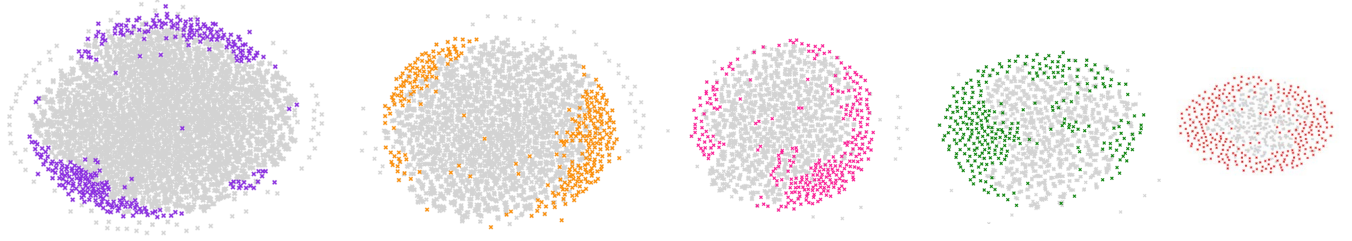


Figure 3: Selected samples during task evolution on Seq-CIFAR-10-LT under the ordered-LTCL setting. IR is 0.01. Buffer size is 50. The colorful points denote the selected samples into the memory buffer, while the gray points are not selected.

Table 2: Ordered-LTCL Accuracy on Seq-CIFAR-10-LT. The architecture is ResNet18.

Buffer	Method	Seq-CIFAR-10-LT							
		0.01		0.02		0.05		0.1	
		Class-IL	Task-IL	Class-IL	Task-IL	Class-IL	Task-IL	Class-IL	Task-IL
-	oEWC	17.53±0.47	62.26±4.52	18.20±0.45	61.87±4.31	18.24±0.26	62.61±3.85	18.01±0.15	58.27±4.21
	SI	16.95±0.33	62.48±4.51	17.79±0.30	63.72±4.27	18.27±0.18	64.11±4.32	19.08±0.18	64.11±4.53
	LwF	16.70±0.20	59.44±1.32	17.56±0.18	27.45±1.33	18.85±0.12	61.39±1.25	18.97±0.01	33.93±1.17
	PNN	-	84.72±0.88	-	87.61±0.72	-	90.39±0.71	-	90.00±0.68
200	ER	39.14±1.68	85.72±1.01	43.10±1.52	86.97±0.98	47.38±1.43	89.17±1.01	52.99±1.56	90.10±0.88
	GEM	29.20±0.97	82.83±0.87	22.40±0.89	86.60±0.97	22.52±1.12	86.22±1.24	21.91±0.87	89.35±0.98
	A-GEM	27.00±0.67	77.56±1.42	18.24±0.60	79.57±1.71	18.82±0.69	77.79±1.53	19.03±0.58	84.35±1.68
	A-GEM-R	17.86±0.87	71.44±1.64	17.24±0.71	76.80±1.32	18.14±1.02	75.49±1.68	18.13±0.61	73.29±1.59
	iCaRL	40.27±3.24	86.90±2.12	43.85±3.01	87.33±1.99	42.85±3.54	86.35±2.23	51.31±2.97	90.75±2.08
	FDR	28.54±3.17	72.52±0.99	31.13±2.36	85.67±1.02	43.06±1.98	90.17±0.88	40.78±2.68	88.53±0.97
	GSS	35.71±3.64	85.74±1.87	36.85±3.67	87.07±2.03	38.90±3.54	86.76±1.67	37.40±3.76	88.46±1.94
	HAL	30.91±2.97	75.03±2.12	28.59±2.53	73.59±2.62	33.27±2.43	76.32±2.45	27.98±2.54	80.21±2.77
	DER	28.00±1.81	74.75±1.44	43.23±1.74	85.17±1.22	51.20±1.79	87.57±1.38	61.52±1.63	90.54±1.13
	DER++	37.65±1.87	83.25±1.24	44.03±1.75	87.35±1.11	55.63±1.66	89.64±1.17	60.68±1.48	90.98±0.95
	<b>Ours</b>	<b>51.00±1.74</b>	<b>89.99±1.32</b>	<b>54.24±1.23</b>	<b>91.53±0.99</b>	<b>58.93±1.45</b>	<b>92.06±1.05</b>	<b>62.69±1.51</b>	<b>92.19±0.92</b>
500	ER	56.01±0.61	87.42±0.54	60.08±0.49	88.97±0.43	63.39±0.52	92.82±0.47	60.06±0.41	92.46±0.35
	GEM	28.46±1.77	84.16±0.75	22.41±1.45	89.05±0.66	20.57±1.32	90.83±0.67	27.65±1.13	90.01±0.56
	A-GEM	18.09±0.71	79.27±1.55	18.28±0.60	79.35±1.43	18.80±0.57	81.08±1.12	19.18±0.66	83.39±1.32
	A-GEM-R	11.32±0.67	57.97±1.54	17.54±0.63	78.14±1.54	18.05±0.62	74.95±1.52	20.75±0.67	79.06±1.32
	iCaRL	45.08±3.99	87.53±3.01	46.83±3.96	88.75±2.64	52.81±3.82	91.29±2.67	53.54±3.72	90.80±2.54
	FDR	30.76±3.58	84.11±1.01	49.54±3.32	87.94±0.96	37.59±3.57	90.14±1.11	46.13±3.31	90.42±0.63
	GSS	47.63±4.15	87.65±1.32	42.21±4.33	87.60±1.64	44.60±4.03	90.19±1.42	49.93±4.56	91.54±1.57
	HAL	39.18±4.44	82.50±2.36	37.16±4.05	76.37±1.98	36.28±4.20	78.85±2.32	49.32±4.12	85.01±2.01
	DER	42.25±1.94	86.69±0.62	52.05±1.88	89.11±0.56	58.47±1.84	89.64±0.53	66.01±1.77	92.09±0.40
	DER++	48.68±1.88	88.42±0.64	44.30±1.68	88.50±0.61	61.46±1.74	91.23±0.53	67.22±1.63	92.21±0.41
	<b>Ours</b>	<b>54.61±1.87</b>	<b>91.03±1.11</b>	<b>58.66±1.34</b>	<b>90.29±0.79</b>	<b>61.08±1.44</b>	<b>92.96±0.97</b>	<b>62.27±0.45</b>	<b>92.43±0.83</b>

[8] Friedemann Zenke, Ben Poole, and Surya Ganguli. 2017. Continual learning through synaptic intelligence. In *International Conference on Machine Learning*.

PMLR, 3987–3995.

Table 3: Ordered-LTCL Accuracy on Seq-CIFAR-100-LT. The architecture is ResNet18.

Buffer	Method	Seq-CIFAR-100-LT							
		0.01		0.02		0.05		0.1	
		Class-IL	Task-IL	Class-IL	Task-IL	Class-IL	Task-IL	Class-IL	Task-IL
-	oEWC	8.09±0.30	18.23±3.22	6.63±0.24	15.17±3.17	11.95±0.22	21.95±3.15	11.31±0.20	21.20±3.18
	SI	7.85±0.52	18.23±4.17	6.70±0.51	19.59±4.06	13.56±0.60	29.53±4.54	9.93±0.25	24.93±3.97
	LwF	8.47±0.20	14.80±2.13	8.29±0.21	15.22±2.20	13.52±0.17	19.51±2.10	13.66±0.10	20.37±2.04
	PNN	-	48.89±0.86	-	43.61±0.83	-	63.02±0.84	-	57.99±0.71
200	ER	14.72±2.31	42.32±1.17	16.54±2.11	46.81±1.14	19.21±2.34	53.03±1.27	19.57±1.94	55.73±1.07
	GEM	16.15±1.02	43.54±1.17	14.94±1.09	39.90±1.01	19.30±0.97	51.89±0.99	11.10±0.86	43.28±1.14
	A-GEM	11.91±0.45	30.57±1.57	7.88±0.57	26.48±1.62	13.15±0.38	37.08±1.49	8.12±0.40	29.81±1.49
	A-GEM-R	6.59±0.40	21.62±1.57	4.15±0.39	16.58±1.52	9.58±0.41	32.60±1.47	9.12±0.39	27.99±1.45
	iCaRL	22.85±3.54	49.62±2.21	20.40±3.43	46.17±2.27	30.33±3.56	54.83±2.39	26.67±3.47	54.55±2.21
	FDR	19.63±3.01	46.03±0.97	19.32±2.97	45.76±1.32	21.86±2.99	54.06±1.30	15.82±3.01	48.73±1.10
	GSS	11.78±4.12	40.56±1.57	14.15±4.52	43.45±1.68	15.30±3.97	45.61±1.27	16.34±3.87	47.20±1.09
	HAL	5.18±3.20	17.25±2.74	5.28±3.19	21.95±2.68	8.52±2.97	24.54±2.45	7.89±3.14	22.15±2.38
	DER	18.31±2.03	47.81±1.27	9.38±2.17	33.47±1.64	28.16±1.91	57.55±0.97	9.75±1.87	42.45±1.02
	DER++	19.62±1.95	46.22±1.14	7.75±1.99	31.81±1.43	26.45±1.74	56.70±1.03	11.28±1.79	46.23±0.93
	Ours	24.18±1.67	50.66±0.85	20.65±1.87	47.53±0.92	31.32±1.58	54.99±0.98	28.69±1.67	55.19±1.13
	ER	20.50±0.36	49.63±0.31	25.05±0.38	54.87±0.32	25.28±0.33	58.02±0.29	27.33±0.31	61.23±0.36
500	GEM	22.47±2.45	49.63±1.01	7.88±1.58	34.81±0.57	25.71±1.69	59.28±0.68	13.14±1.57	34.07±0.89
	A-GEM	8.68±2.57	49.63±2.23	7.77±1.78	28.22±0.73	12.96±1.77	34.82±0.93	13.78±1.86	34.00±1.01
	A-GEM-R	6.70±2.46	23.03±2.07	4.27±1.88	18.93±0.86	10.78±1.89	32.24±1.04	10.44±1.77	31.20±0.86
	iCaRL	24.51±2.03	50.74±1.08	24.3±1.81	50.40±0.97	33.61±1.73	60.94±0.92	30.80±1.60	58.97±0.89
	FDR	24.99±0.54	50.37±0.86	18.32±0.67	45.07±0.73	30.04±0.66	60.67±0.71	17.92±0.62	53.96±0.58
	GSS	14.74±3.97	44.82±0.95	16.22±3.86	45.29±1.00	18.55±3.74	52.66±0.91	19.09±3.60	55.26±0.88
	HAL	14.74±4.56	44.82±3.19	7.86±4.44	21.43±3.09	5.77±4.15	19.34±3.01	9.29±4.19	27.07±3.13
	DER	20.47±1.01	50.23±1.23	15.99±1.18	43.41±1.06	31.76±0.97	61.67±1.03	13.46±0.93	50.82±0.97
	DER++	20.90±0.98	51.88±1.14	15.80±1.05	43.55±0.94	30.69±1.06	62.66±1.11	17.43±0.93	53.89±1.02
	Ours	25.05±1.32	52.23±1.29	25.48±1.17	50.85±1.09	34.33±1.25	63.86±1.27	32.40±1.20	60.59±1.11

**Table 4: Ordered-LTCL Accuracy on Seq-TinyImageNet-LT. The model is ResNet18.**

Buffer	Method	Seq-TinyImageNet-LT							
		0.01		0.02		0.05		0.1	
		Class-IL	Task-IL	Class-IL	Task-IL	Class-IL	Task-IL	Class-IL	Task-IL
50	ER	3.89±0.37	23.38±2.32	4.73±0.34	25.27±1.97	6.04±0.37	27.73±2.10	6.8±0.34	26.22±2.13
	A-GEM	2.99±0.14	11.12±0.20	3.27±0.11	11.56±0.18	5.12±0.08	14.57±0.14	6.32±0.06	18.17±0.07
	A-GEM-R	2.38±0.09	9.93±0.17	3.39±0.14	12.85±0.12	4.23±0.10	12.92±0.07	4.98±0.09	14.48±0.06
	FDR	3.86±0.23	21.13±0.77	4.57±0.20	20.86±0.75	5.77±0.43	24.05±0.88	6.49±0.71	26.79±0.74
	DER	4.52±0.78	23.75±0.76	5.01±0.68	23.81±0.78	6.36±0.69	24.83±0.80	6.41±0.62	24.68±0.77
	DER++	3.18±1.41	23.19±1.53	4.88±1.26	24.83±1.44	5.77±1.57	25.32±1.31	7.92±1.32	26.58±1.48
	<b>Ours</b>	<b>4.83±1.67</b>	<b>23.80±1.55</b>	<b>5.49±1.54</b>	<b>25.07±1.63</b>	<b>6.78±1.34</b>	<b>26.08±1.23</b>	<b>8.32±1.45</b>	<b>29.98±1.52</b>
200	ER	5.58±0.57	35.31±0.67	5.78±0.22	34.72±1.78	6.54±0.55	37.63±1.33	7.21±0.23	38.54±1.47
	A-GEM	3.34±0.24	12.16±0.21	2.75±0.25	10.95±0.17	5.06±0.23	15.22±0.20	6.17±0.14	17.52±0.10
	A-GEM-R	2.60±0.27	11.02±0.29	3.22±0.26	11.06±0.20	3.54±0.16	13.31±0.17	4.62±0.14	14.51±0.11
	iCaRL	7.01±0.62	29.13±1.68	8.18±0.55	33.49±1.54	8.36±0.60	34.05±1.65	10.39±0.57	36.88±1.53
	FDR	5.98±0.34	31.15±0.77	6.79±0.40	31.73±0.79	7.40±0.27	34.84±0.85	7.85±0.23	37.22±0.77
	DER	6.56±0.82	35.29±0.78	6.61±0.67	36.57±0.81	9.69±0.58	39.54±0.79	11.77±0.59	40.00±0.67
	DER++	6.92±1.53	34.52±1.36	8.00±1.45	35.48±1.33	8.75±1.14	37.20±1.32	10.61±0.97	39.96±1.01
	<b>Ours</b>	<b>10.24±1.44</b>	<b>36.01±1.32</b>	<b>8.25±1.47</b>	<b>37.59±1.52</b>	<b>9.80±1.33</b>	<b>42.64±1.08</b>	<b>12.54±1.44</b>	<b>40.45±0.98</b>
500	ER	8.49±0.59	39.61±0.94	9.23±0.60	41.42±0.89	8.88±0.56	42.36±0.66	9.61±0.54	44.91±0.75
	A-GEM	3.09±0.44	10.35±0.77	4.17±0.34	13.94±0.69	4.96±0.47	15.09±0.76	6.0±0.32	15.83±0.50
	A-GEM-R	2.42±0.43	11.11±0.78	3.04±0.37	11.09±0.62	4.53±0.38	13.25±0.74	5.07±0.42	14.51±0.61
	iCaRL	10.51±1.76	11.87±3.01	11.37±1.54	38.67±3.19	12.74±1.44	40.39±3.20	13.13±1.33	11.37±3.22
	FDR	10.21±0.45	39.42±1.03	10.76±0.37	43.22±1.00	10.46±0.35	45.16±1.14	10.00±0.33	45.86±0.86
	DER	9.22±1.23	43.38±1.08	9.71±1.39	44.59±0.97	11.28±1.15	45.83±1.04	13.70±1.15	49.11±0.98
	DER++	10.42±1.47	41.94±1.23	12.39±1.40	45.17±1.08	14.34±1.22	47.98±1.17	15.66±0.97	47.88±1.28
	<b>Ours</b>	<b>10.78±1.25</b>	<b>44.13±1.36</b>	<b>13.12±1.34</b>	<b>46.54±1.17</b>	<b>16.33±1.18</b>	<b>49.24±1.17</b>	<b>16.34±1.07</b>	<b>50.43±1.29</b>
1000	ER	12.81±0.43	46.99±0.35	13.17±0.47	49.06±0.33	13.44±0.39	52.06±0.36	13.38±0.37	53.25±0.31
	A-GEM	3.45±0.33	12.58±0.74	3.90±0.38	13.05±0.69	5.12±0.38	14.41±0.72	6.14±0.30	16.94±0.71
	A-GEM-R	2.47±0.37	10.44±0.64	2.68±0.37	12.54±0.74	4.46±0.31	13.04±0.77	4.78±0.29	14.12±0.64
	iCaRL	14.01±2.12	42.99±3.01	15.38±1.87	45.15±3.14	16.79±2.14	47.77±3.36	17.71±2.11	49.91±3.37
	FDR	16.77±0.67	47.01±0.79	16.49±0.77	49.29±0.84	15.74±0.71	50.74±0.83	16.36±0.56	54.05±0.69
	DER	8.81±1.01	43.53±1.23	9.44±0.97	41.19±1.10	13.02±0.90	49.92±1.04	17.53±0.79	55.44±0.94
	DER++	12.54±1.12	47.02±1.23	15.76±1.07	50.00±1.19	16.65±1.11	52.89±1.27	17.98±1.09	55.76±1.16
	<b>Ours</b>	<b>12.97±1.32</b>	<b>50.13±1.17</b>	<b>16.92±1.22</b>	<b>51.00±1.21</b>	<b>17.33±1.17</b>	<b>55.44±1.20</b>	<b>18.45±1.10</b>	<b>58.55±1.16</b>

Table 5: BWT results under ordered-LTCL.

Buffer	Method	Seq-CIFAR-10-LT			
		0.01	0.02	0.05	0.1
-	oEWC	-58.50	-58.33	-67.50	-70.82
	SI	-59.36	-53.78	-71.52	-71.03
	LwF	-58.93	-60.23	-68.93	-72.32
	PNN	-	-	-	-
200	ER	-52.28	-57.05	-61.71	-67.56
	GEM	-47.25	-41.58	-64.06	-73.13
	A-GEM	-57.80	-45.75	-72.16	-74.02
	A-GEM-R	-48.01	-48.06	-60.95	-62.99
	iCaRL	-21.91	-22.61	-24.14	-24.77
	FDR	-42.80	-45.87	-61.81	-67.93
	GSS	-56.68	-58.93	-66.27	-68.97
	HAL	-33.32	-40.08	-39.13	-41.21
	DER	-28.03	-40.67	-44.68	-66.50
	DER++	-35.93	-40.01	-49.57	-68.47
	Ours	-14.00	-10.30	-7.40	-7.58
500	ER	-43.67	-46.81	-53.41	-56.65
	GEM	-36.18	-51.52	-56.08	-73.05
	A-GEM	-60.53	-58.37	-70.98	-74.61
	A-GEM-R	-49.63	-46.85	-66.17	-65.11
	iCaRL	-18.42	-20.91	-19.53	-21.15
	FDR	-31.70	-42.18	-49.21	-62.52
	GSS	-50.77	-54.40	-62.19	-66.00
	HAL	-30.54	-34.88	-37.13	-41.21
	DER	-25.78	-32.83	-33.41	-56.67
	DER++	-24.00	-36.70	-35.10	-54.88
	Ours	1.70	-2.53	4.00	13.30