

# Long-Tailed Class Incremental Learning

## – Supplementary Material

Xialei Liu<sup>1</sup>, Yu-Song Hu<sup>1</sup>, Xu-Sheng Cao<sup>1</sup>, Andrew D. Bagdanov<sup>2</sup>, Ke Li<sup>3</sup>, and Ming-Ming Cheng<sup>1</sup>

TMCC, CS, Nankai University, China<sup>1</sup>; MICC, University of Florence, Florence, Italy<sup>2</sup>  
Tencent YouTu Lab<sup>3</sup>

### 1 More results on ImageNet-Subset

**Different imbalance ratios on ImageNet-Subset** In this section, we analyze three different imbalance ratios  $\rho = 0.01$ ,  $\rho = 0.05$  and  $\rho = 0.1$  on ImageNet-Subset, the smaller the ratio the more skewed the distribution. Compared to CIFAR100, ImageNet-Subset contains more samples, which results in a more skewed distribution on different continual training steps. We report three baselines, i.e. EEIL, LUCIR and PODNET, and our two-stage approach applied to them, denoted as EEIL+, LUCIR+ and PODNET+. As we can see from Figure 1 (a-c), with more samples, for the ordered LT-CIL scenario, PODNET surpasses other approaches consistently with a large margin, obtaining the best performance in all scenarios. We consider that PODNET can learn much more information when the data is sufficient. Overall our two-stage method can consistently boost accuracy for most methods, especially for LUCIR+ with a significant gain. For shuffled LT-CIL scenario from Figure 1 (d-f), PODNET+ and LUCIR+ are very competitive in all three imbalance ratio  $\rho$ . The proposed two-stage method further improves the performance, especially for EEIL and LUCIR. Interestingly, we can see that compared to conventional settings, long-tailed scenarios with a large imbalance ratio can achieve competitive performance with less samples, which may due to the imbalance effect of training data.

**Long sequence on ImageNet-Subset** We evaluate on long sequence of 25 steps for all three scenarios with three state-of-the-art methods on ImageNet-Subset, and collect the results in Table 1. As we can see, our method also improves over different baselines in this more challenging setting like on CIFAR100 except

	EEIL	EEIL+	LUCIR	LUCIR+	PODNet	PODNet+
Conventional	44.14	55.92	53.60	56.96	60.25	<b>64.75</b>
Ordered	47.18	46.43	47.05	49.77	<b>58.30</b>	57.76
Shuffled	40.75	43.95	40.52	48.35	48.93	<b>51.52</b>

Table 1: The average accuracy on long sequence of 25 steps for three different scenarios on Imagenet-Subset. Methods with + sign indicate our two-stage method applied to the corresponding baseline.

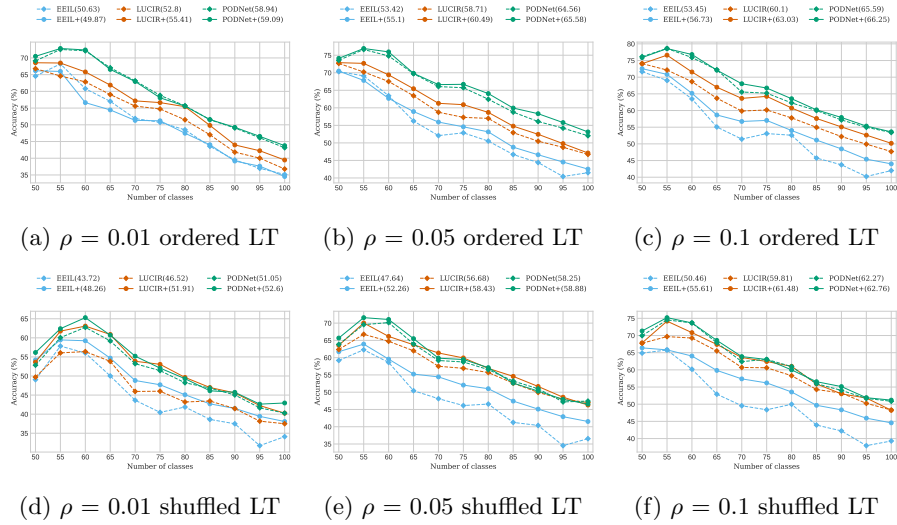


Fig. 1: Average accuracy on different imbalance ratios on Imagenet-Subset, the top row is on ordered LT-CIL and the bottom row is on shuffled LT-CIL. Methods with + sign indicate our two-stage method applied to the corresponding baseline.

for PODNET on ordered LT-CIL scenario. Further more, we can see that for 25-step scenario, two-stage methods can get much larger gain than in 5-step and 10-step scenarios in most cases. It shows that the two-stage methods are more robust for longer sequences.