Supplementary Materials for FCS: Feature Calibration and Separation for Non-Exemplar Class Incremental Learning

Qiwei Li[1, 2], Yuxin Peng[1], Jiahuan Zhou[1*]
[1] Wangxuan Institute of Computer Technology, Peking University, Beijing 100871, China
[2] School of Electronics Engineering and Computer Science, Peking University
{lqw, pengyuxin, jiahuanzhou}@pku.edu.cn

In supplementary materials, we introduce the evaluation metric in detail and provide more visualization results including accuracy curves and confusion matrices in different settings to further demonstrate the effectiveness of our proposed method.

1. Evaluation Metric

We follow [1, 2] to choose *Accuracy* and *Average Forgetting* for evaluation. Let T denote the number of incremental learning tasks. *Accuracy* is the average accuracy of all the classes after training the model from 1 to T. *Average forgetting* represents the average performance degradation of different tasks. Specifically, the forgetting of j-th incremental stage after training on t-th stage can be computed as :

$$f_j^t = \max_{l \in \{1,\ldots,t-1\}} a_{l,j} - a_{t,j}, \tag{1}$$

where $a_{t,j}$ represents the accuracy of task j, after training the model from 1 to t. Then the average forgetting at stage t can be computed as:

$$F_t = \frac{1}{t-1} \sum_{j=1}^{t-1} f_j^t. \tag{2}$$

2. More Visualization Results

2.1. Accuracy Curve.

Similar to Fig.4 in the main paper, in Fig. 1, we present the complete classification accuracy of each stage on CIFAR-100, TinyImageNet, and ImageNet-Subset. The results show that our method is superior to other non-exemplar and exemplar-based CIL methods demonstrating that our method achieves a better balance between knowledge forgetting and acquisition.

*Corresponding author

2.2. Confusion Matrix.

Furthermore, we present the confusion matrices of iCaRL-NME, PASS, and our method on different datasets in Fig. 2. It can be observed that, without retaining any exemplars, our method achieves better performance than the exemplar-based CIL method, iCaRL. This phenomenon can be attributed to our carefully designed FCN and PIC can fully utilize the information represented by prototypes, thus mitigating knowledge forgetting. In comparison to the NECIL method, PASS, our methods can predict classes correctly. This is because the PIC alleviates the feature overlap of different classes. And the FCN transfers features of old classes to the feature space of the new model, thereby retaining more knowledge.

References

[1] Zhi Gao, Chen Xu, Feng Li, Yunde Jia, Mehrtash Harandi, and Yuwei Wu. Exploring data geometry for continual learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023. 1

[2] Fei Zhu, Xu-Yao Zhang, Chuang Wang, Fei Yin, and Cheng-Lin Liu. Prototype augmentation and self-supervision for incremental learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021. 1

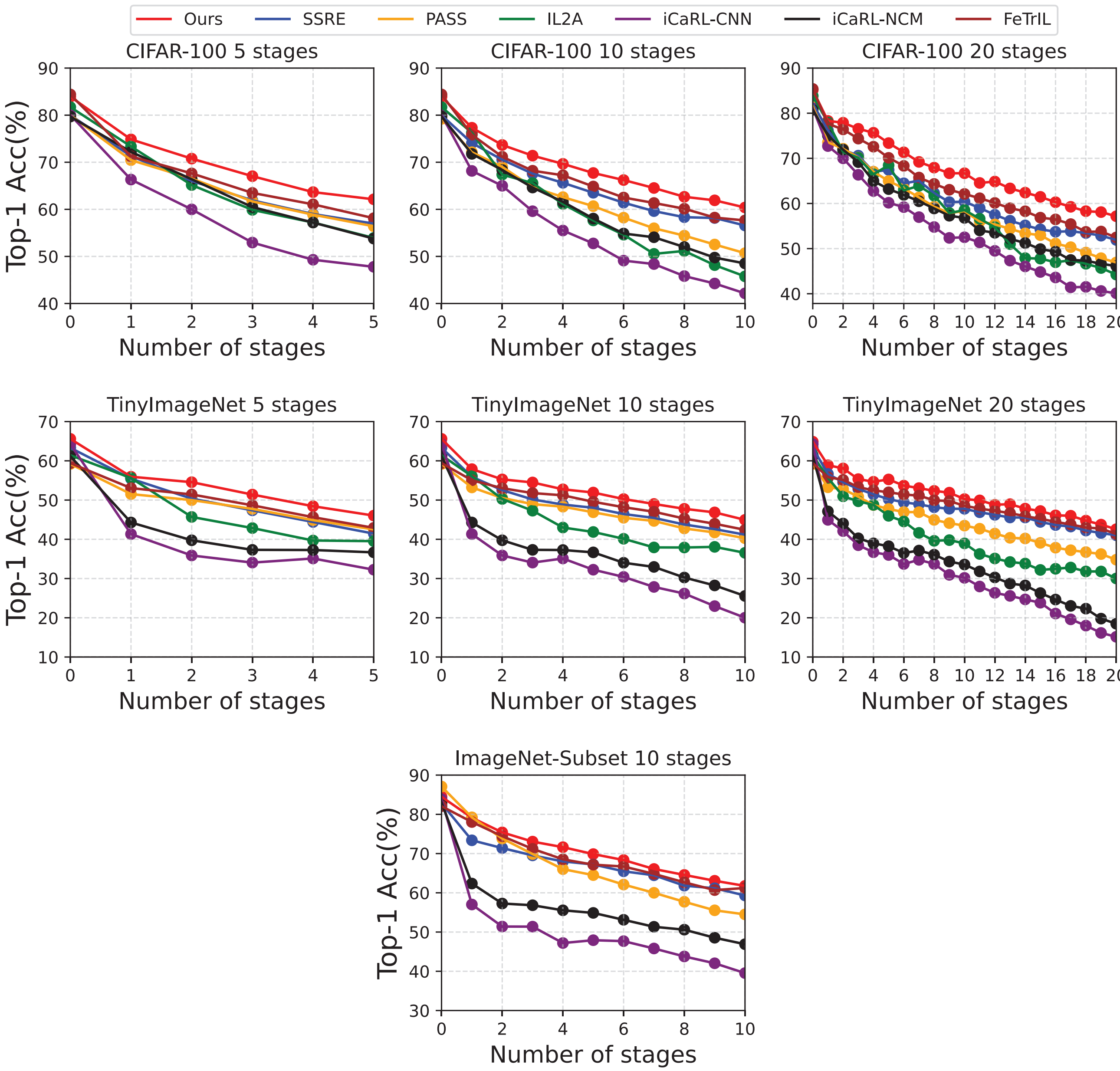

Figure 1. Complete classification accuracy of each stage on CIFAR-100, TinyImageNet, and ImageNet-Subset.

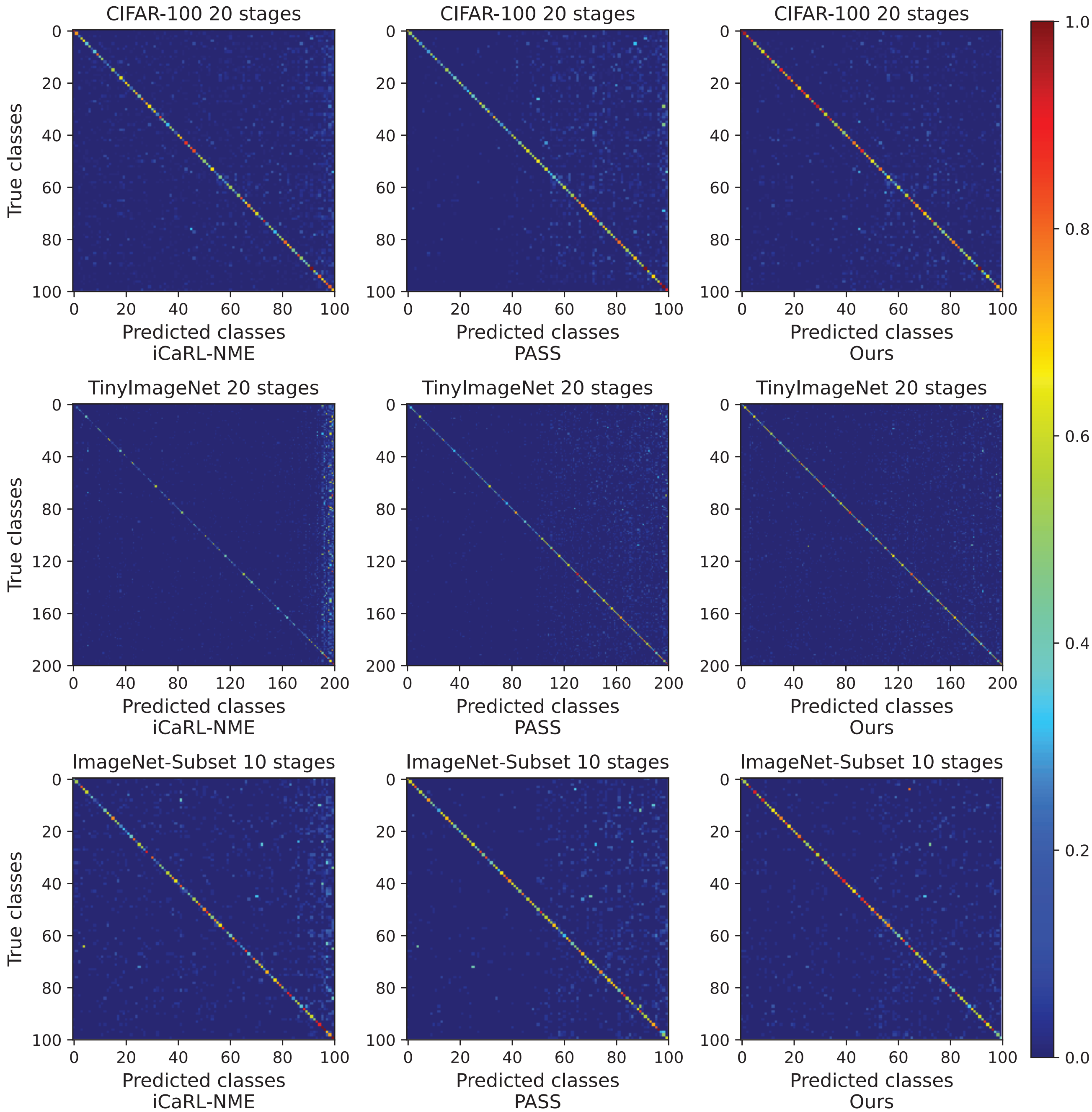

Figure 2. More confusion matrices of iCaRL-NME, PASS, and our method on CIFAR-100, TinyImageNet, and ImageNet-Subset.