MetaFSCIL: A Meta-Learning Approach for Few-Shot Class Incremental Learning
Supplementary document

1 Overview

In this supplementary document, we provide the following details to support the paper:

- Detailed comparison of network architecture between ours and ANML.
- Illustration of the BGM positions.
- Graphic comparison with state-of-the-art.

2 Architecture compared with ANML[1]

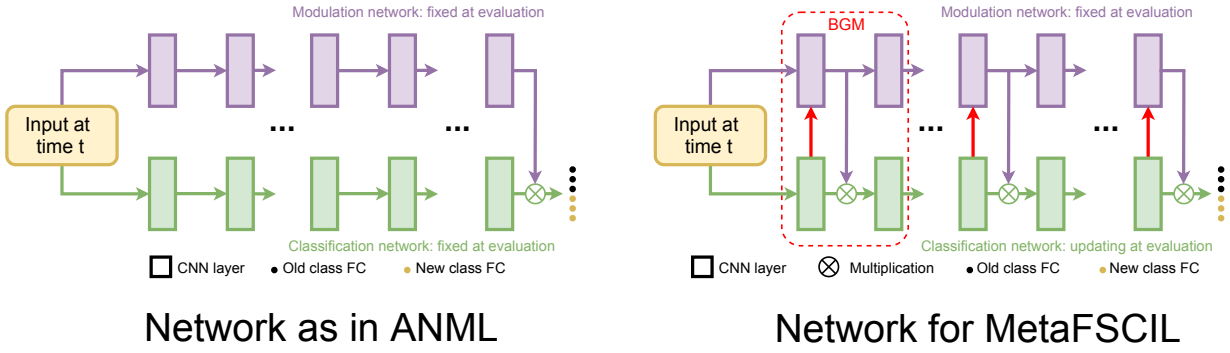

Figure 1: Network architecture comparison between ANML and MetaFSCIL (ours).

ANML also employs the modulation network for *selective activation*. The key differences are:

- ANML only modulates the last activation map, which scales poorly for deeper networks (ANML only conducted experiments with shallow networks). In comparison, our architecture provides more modulation throughout the network, which has shown the effectiveness in our ablation studies for the deeper network.
- In ANML, the modulation and classification network is decoupled. The modulation module is only conditioned on the input images while discarding the current status of the classification module. For example, for the same image of one class, the modulation module always produces the same output at different time steps. But the classification module may have learned different knowledge, which needs different modulation. In contrast, our BGM couples two modules for more effective modulation. It is conditioned on both new input images and the current status of the classification module.

- The trained classification module in ANML is fixed during incremental learning, which greatly limits the generalization for complex datasets (e.g., MiniImageNet [2]). Our network allows the classification module to continuously update with the new images to improve the adaptation.

3 Illustration of the BGM positions

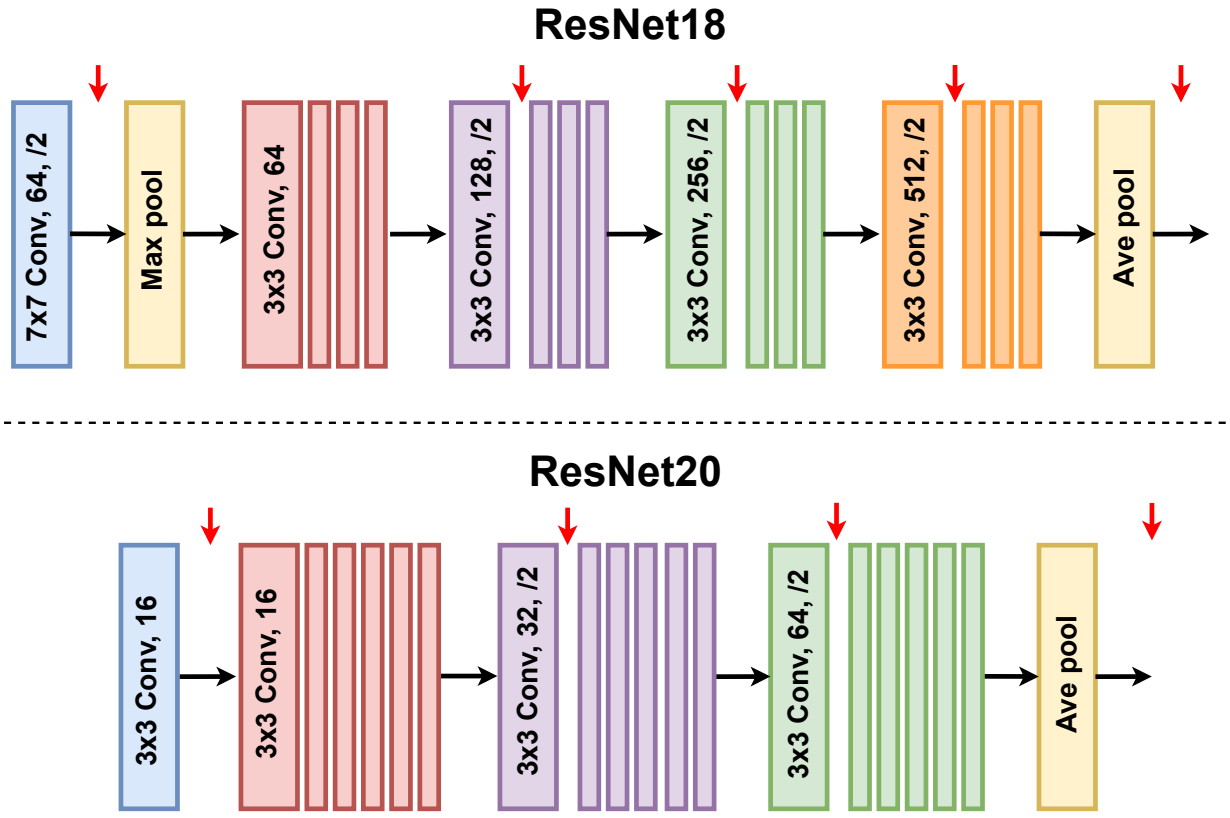

Figure 2: Illustration of the BGM positions for ResNet18 and ResNet20. The red arrows show the locations within the networks where we place the BGM module. We place the BGM module after the layers with downsampling operation. We also modulate the activation after the average pooing. For ResNet20, we also modulate the first activation map. There are total of 5 and 4 BGM modules for ResNet18 and ResNet20, respectively.

4 Graphic comparison with the state-of-the-art. (Table 1 in paper)

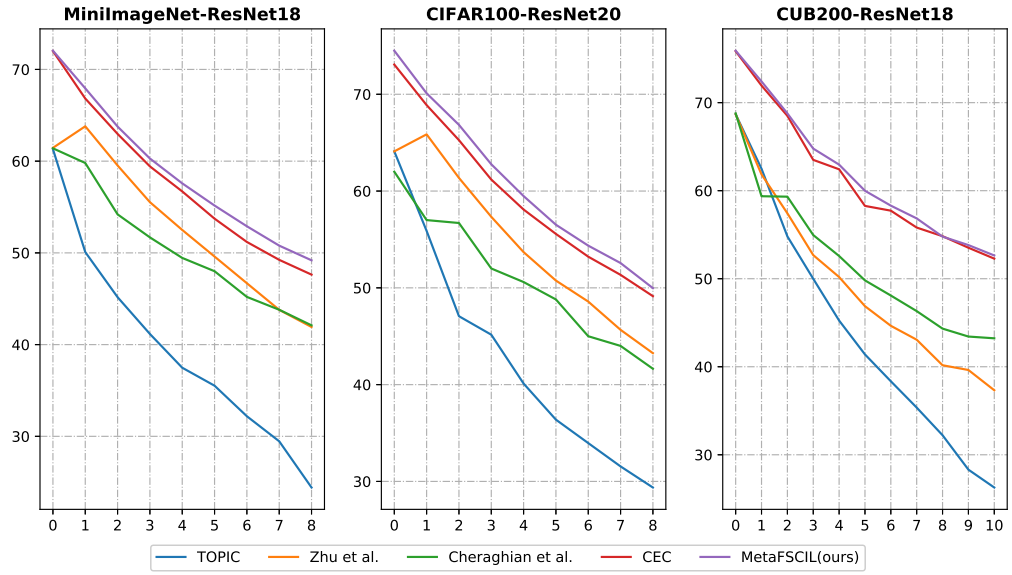

Figure 3: Plots of Table 1 in paper for comparison with TOPIC [3], Zhu *et al.* [4], Cheraghian *et al.* [5] and CEC [6]. The proposed method (MetaFSCIL) outperforms all other methods. Our method causes less forgetting and is more stable for all incremental sessions.

[1] Beaulieu, Shawn and Frati, Lapo and Miconi, Thomas and Lehman, Joel and Stanley, Kenneth O and Clune, Jeff and Cheney, Nick. Learning to continually learn. In *European Conference on Artificial Intelligence*, 2020.

[2] Russakovsky, Olga and Deng, Jia and Su, Hao and Krause, Jonathan and Satheesh, Sanjeev and Ma, Sean and Huang, Zhiheng and Karpathy, Andrej and Khosla, Aditya and Bernstein, Michael and others. Imagenet large scale visual recognition challenge. In *International Journal of Computer Vision*, 2015.

[3] Tao, Xiaoyu and Hong, Xiaopeng and Chang, Xinyuan and Dong, Songlin and Wei, Xing and Gong, Yihong. Few-shot class-incremental learning. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020.

[4] Zhu, Kai and Cao, Yang and Zhai, Wei and Cheng, Jie and Zha, Zheng-Jun. Self-Promoted Prototype Refinement for Few-Shot Class-Incremental Learning. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2021.

[5] Cheraghian, Ali and Rahman, Shafin and Ramasinghe, Sameera and Fang, Pengfei and Simon, Christian and Petersson, Lars and Harandi, Mehrtash. Synthesized Feature based Few-Shot Class-Incremental Learning on a Mixture of Subspaces. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2021.

[6] Zhang, Chi and Song, Nan and Lin, Guosheng and Zheng, Yun and Pan, Pan and Xu, Yinghui. Few-shot incremental learning with continually evolved classifiers. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2021.