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]

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.

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


