MetaNet: Rehearsal-based Meta Continual Learning

Young D. Kwon*, Jagmohan Chauhan†, and Cecilia Mascolo*
*University of Cambridge, United Kingdom
†University of Southampton, United Kingdom
Email: {ydk21, cm542}@cam.ac.uk, J.Chauhan@soton.ac.uk

Introduction. With the rise of mobile, wearable devices, and the Internet of Things (IoT), the proliferation of sensory type data has fostered the adoption of deep neural networks (DNN) in the modeling of a variety of mobile sensing applications [1]; Then, a crucial characteristic common to the mobile applications on the edge is the need for a trained model to adapt to accommodate new classes/tasks and to a dynamically changing environment. In these settings, the ability to perform *continual learning (CL)* becomes essential, i.e., to learn consecutive tasks without forgetting how to perform previously learned tasks [2].

Researchers have proposed many CL methods. However, those methods typically require many training samples to learn new tasks/classes [2]. This limits its applicability to real-world applications where the labeled user data is not abundant. Hence, Meta CL methods are proposed to solve the limitation by reducing the amount of required training data to a few samples [3], [4]. However, Meta CL methods also have limitations. Meta CL's performance degrades when a large number of classes are added. New Meta CL method (OML-AIM and ANML-AIM) [4] is proposed to improve upon prior works, however, it uses 9× more parameters (81 Million) compared to ANML (9 Million) [3].

This Work. We propose **MetaNet** to resolve the drawbacks and challenges of prior works. First, to solve the performance degradation problem of ANML, MetaNet extends the current Meta CL method (ANML) to be rehearsal-based Meta CL to ensure high accuracy without making the model too large such as (OML-AIM and ANML-AIM). Second, to examine the storage and memory overheads incurred from saving rehearsal samples, we analyze the two rehearsal strategies ((1) raw data rehearsal, (2) latent representations rehearsal).

In detail, we first present the overview of MetaNet that extends the existing Meta CL method (i.e., ANML) to be a rehearsal-based method. Given a stream of tasks, $T_1, T_2, ..., T_t, ..., T_N$, the aim of CL is to learn new tasks one after another instead of learning all the tasks at once. For each task, only a few samples (e.g., 30) are given for a model to learn. Note that in conventional training, many samples (e.g., 500) are used. In prior Meta CL methods, the given samples are discarded once used for training. Whereas, inspired by rehearsal techniques in the CL literature [2], MetaNet stores the given samples to prevent forgetting when it learns a new class by replaying the saved samples of learned classes.

Further, MetaNet adopts two schemes as a rehearsal strategy. The first scheme is to save the raw data samples as exemplars. Then, the second scheme is to store the activations of a chosen layer as exemplars (denoted as latent representations). As Meta CL freezes the feature extraction part during the meta-testing phase, MetaNet chooses the last layer of the feature extraction network as the latent replay layer and stores the latent layer's activations as exemplars.

Performance: We first examine the performance of MetaNet in comparison to the prior Meta CL methods (ANML, OML-AIM, ANML-AIM) for the CIFAR-100 dataset. For our proposed method, we use a fixed budget to store up to 900 exemplars to study the effectiveness of exemplars. We find that all the Meta CL methods largely prevent forgetting previously learned classes and get to learn new classes. However, their performance degrades as they learn many classes. Besides, while AIM variants (OML-AIM and ANML-AIM) outperform the ANML, our method (MetaNet) achieves the highest accuracy improving the accuracy by 10.3% over the current state-of-the-art (AIM variants).

Storage: We measure the storage overhead of the Meta CL methods to investigate the trade-offs between accuracy and resource overheads. The size of the prior Meta CL methods is the model parameter size (M), as they do not require storing exemplars. MetaNet consists of the model parameter (M) and budget size for exemplars (B). We find that ANML requires a minor storage requirement of 35 MB. AIM variants show the most considerable storage overhead (209 MB for OML-AIM and 310 MB for ANML-AIM) as their AIM module is very expensive. Then, MetaNet shows a moderate storage size of 46-49 MB with 900 exemplars, incurring a modest overhead of 11-14 MB. MetaNet requires 6.3-6.8× less storage than AIM variants do.

Conclusion. We proposed the rehearsal-based Meta CL method, MetaNet, that improves performance over the state-of-the-art Meta CL method (OML-AIM, ANML-AIM) using less storage requirement. As future work, we want to study different rehearsal strategies and propose more efficient Meta CL using various quantization techniques.

References

[1] Y. D. Kwon, J. Chauhan, and C. Mascolo, "Yono: Modeling multiple heterogeneous neural networks on microcontrollers," in *Proc. IPSN '22*.
[2] Y. D. Kwon, J. Chauhan, A. Kumar, P. Hui, and C. Mascolo, "Exploring system performance of continual learning for mobile and embedded sensing applications," in *Proc. SEC '21*.
[3] S. Beaulieu, L. Frati, T. Miconi, J. Lehman, K. O. Stanley, J. Clune, and N. Cheney, "Learning to continually learn," in *ECAI 2020*, pp. 992–1001.
[4] E. Lee, C.-H. Huang, and C.-Y. Lee, "Few-shot and continual learning with attentive independent mechanisms," in *Proc. ICCV '21*.