GKEAL: Gaussian Kernel Embedded Analytic Learning for Few-shot Class Incremental Task

Huiping Zhuang[1]*, Zhenyu Weng [2], Run He[1], Zhiping Lin [2] and Ziqian Zeng[1]

[1]Shien-Ming Wu School of Intelligent Engineering, South China University of Technology, China,
[2]School of Electrical and Electronic Engineering, Nanyang Technological University.

Introduction & Motivation

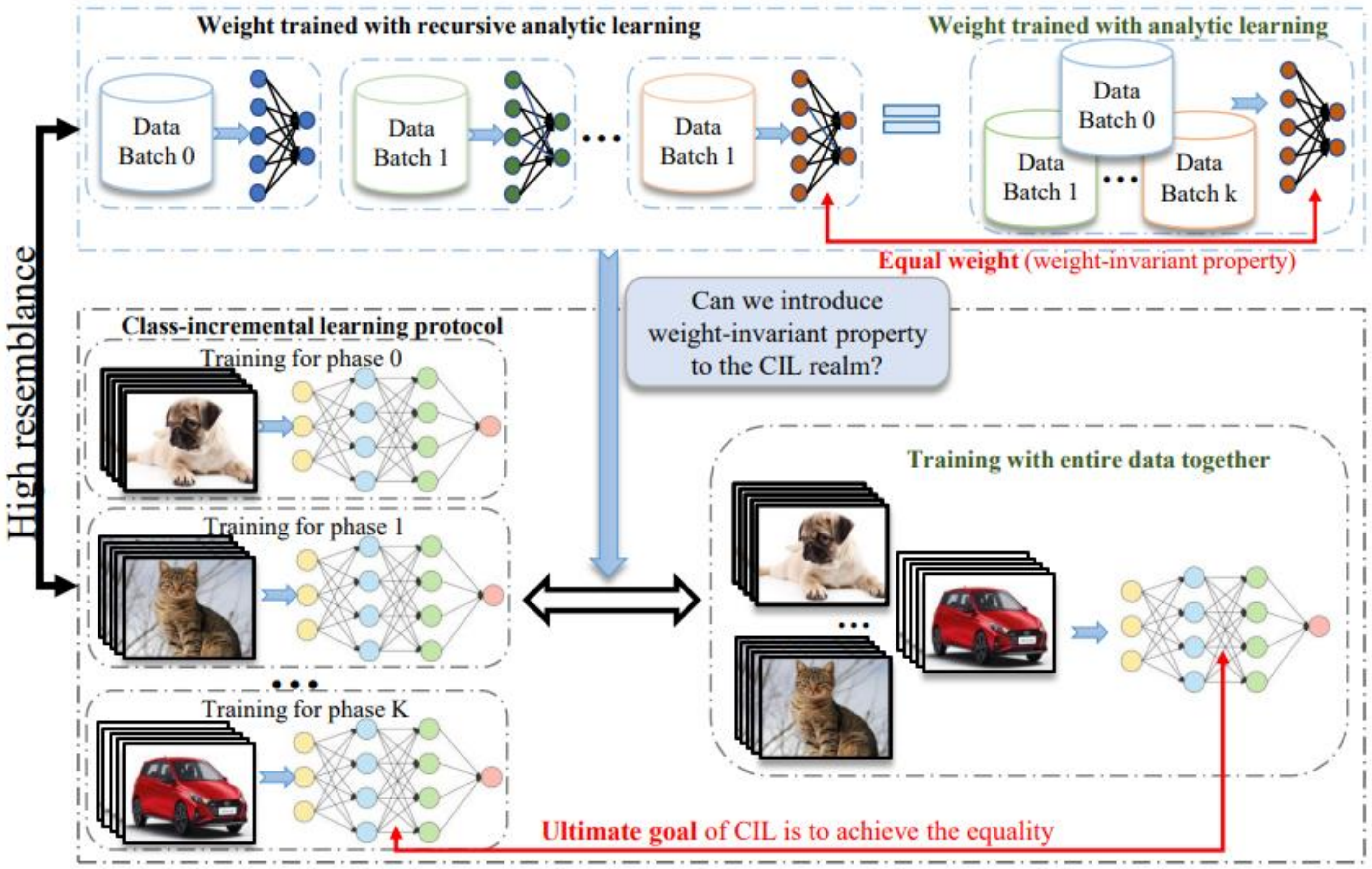
- Class-incremental learning (CIL) can continuously absorbs new category knowledge phase-by-phase while faces the challenge of catastrophic forgetting that renders the networks losing grasp of the learned knowledge when accepting new tasks.

- The few-shot setting in FSCIL further impose the data insufficiency constraint on data avalability that each category/task is given only a few training samples.

- Analytic learning allows the training to be implemented in a recursive manner where training data are scattered into multiple batches and the weights trained recursively are identical to those trained jointly with entire dataset.

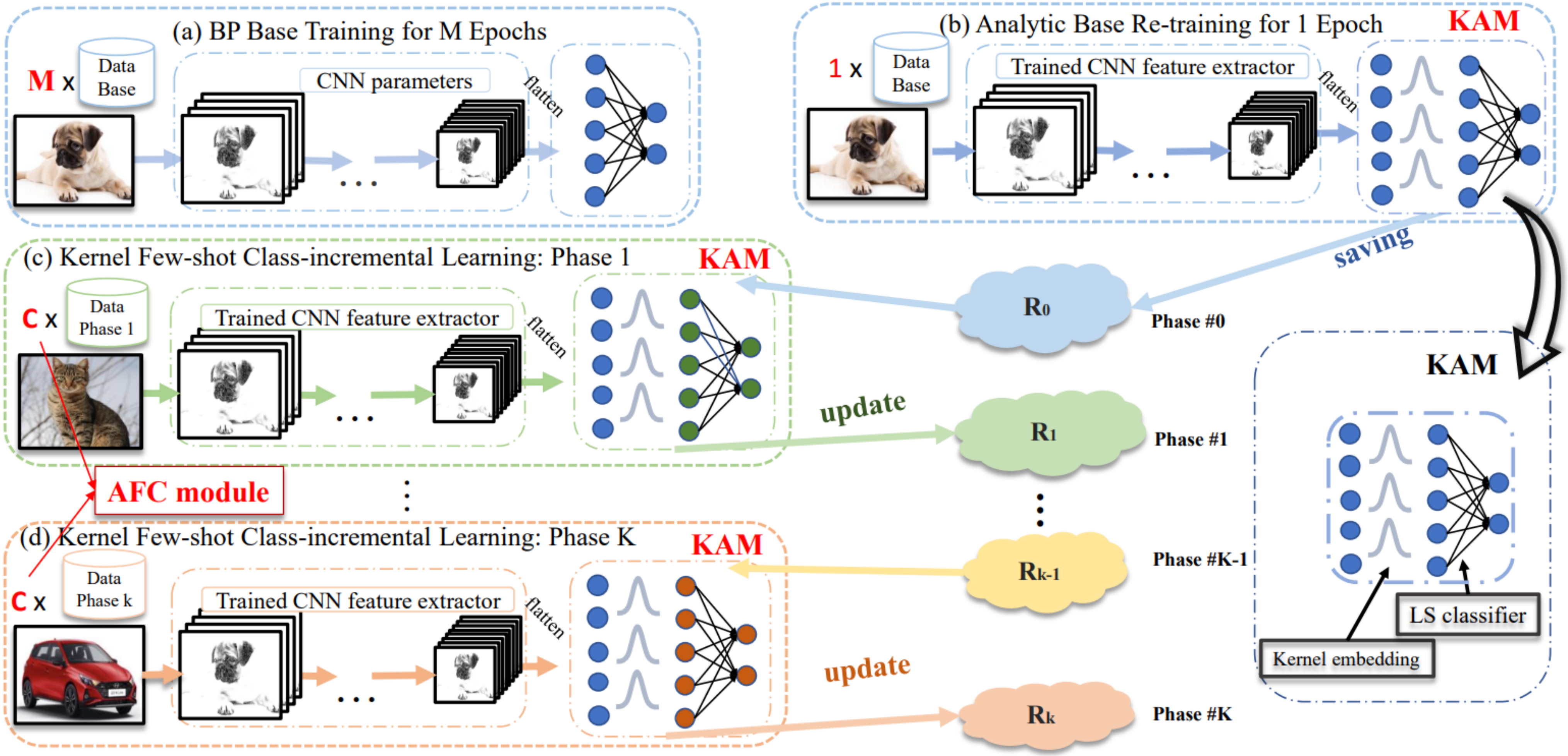
Introduction & Motivation

❑This weight-invariant property in analytic learning highly resemble the incremental learning paradigm and its objective of avoiding forgetting. Can we implement the resemblance?



The Proposed Method

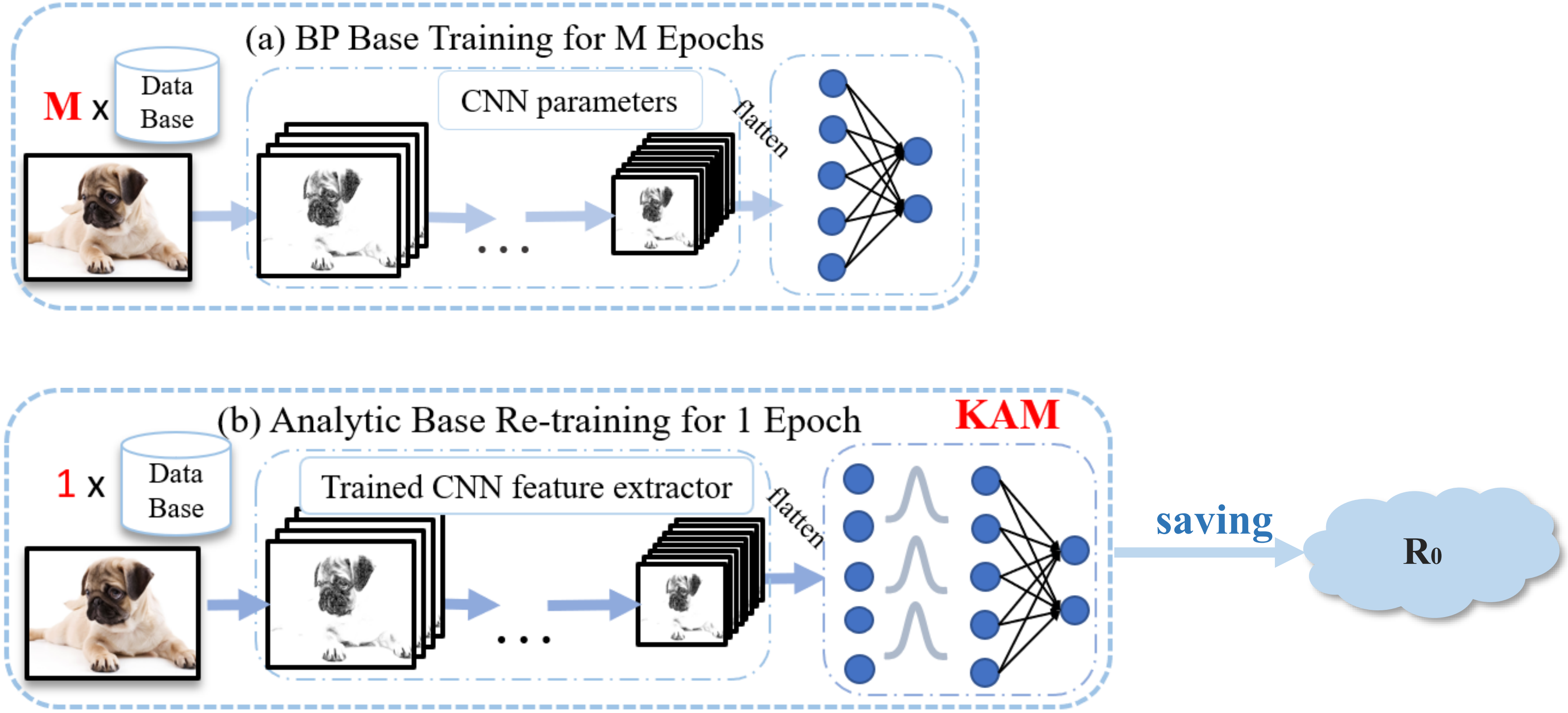
- Two phases: **Base training** and **few-shot class incremental learning**.

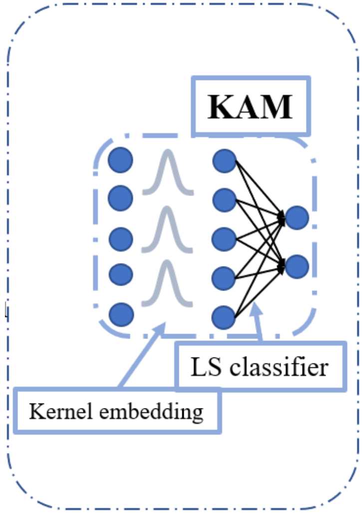


The Proposed Method

- ❑BP Base training for training the backbone. Then the backbone is frozen.

- ❑Analytic base retraining for initialize the LS classifier in KAM.

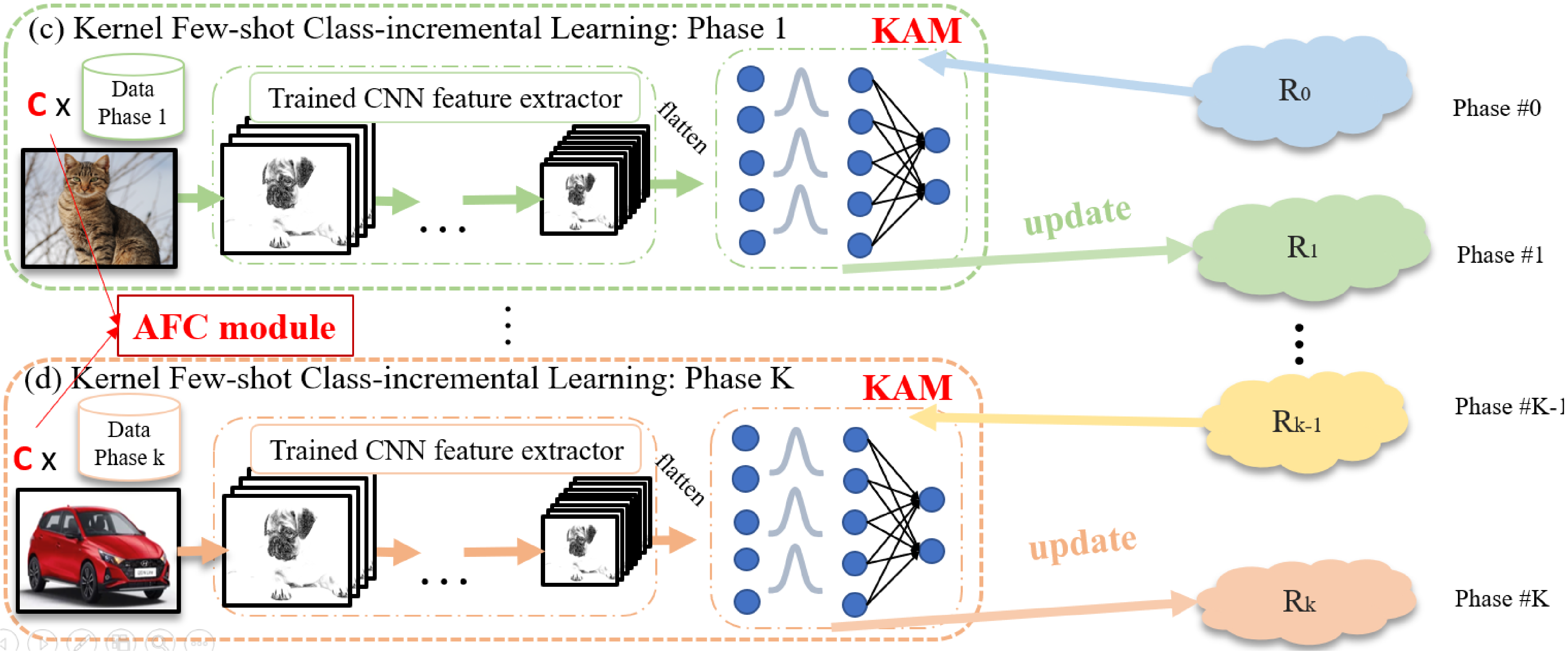




The Proposed Method

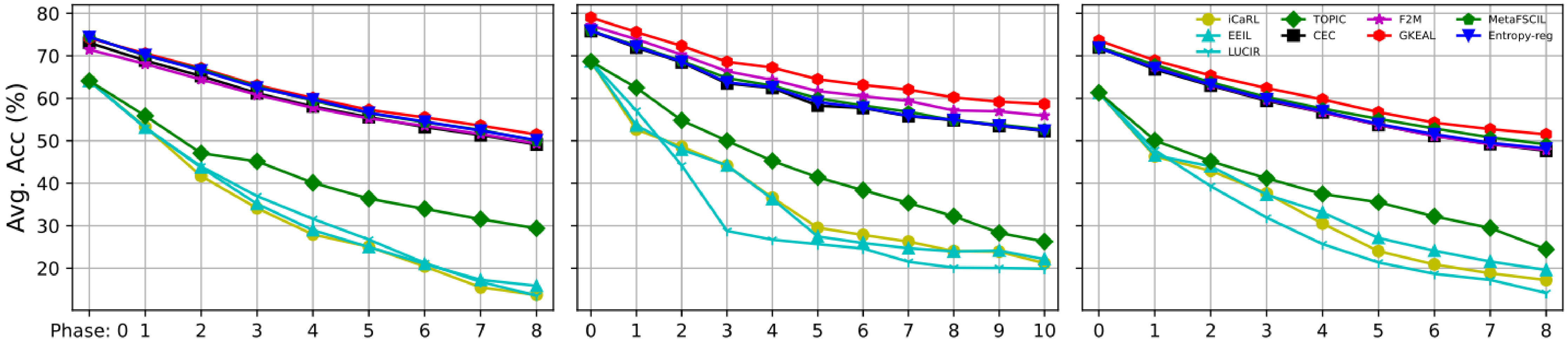
❑Information stored in $\mathbf{R_i}$ is used to update the LS classifiers.

❑AFC module augment features to balance the old-new preference.



Experiments

- ❑For validation we conduct FSCIL tasks of classification on the *CIFAR-100, CUB-200* and *mini-ImageNet* datasets.
- ❑The setting in CIFAR-100/mini-ImageNet is 5-way 5-shot (total 8 phases) and 10-way 5-shot (total 10 phases) in CUB-200.
- ❑In the comparison with State-of-the-arts, we can find that our method outperform the other methods.



(a) CIFAR-100 (a) CUB200-2011 (c) mini-ImageNet

Ablation Study

❑Ablation study on GKE and AFC shows that:

1. Lacking GKE results in catastrophic forgetting;

2. AFC with GKE can improve the performance.

Table 4. Ablation study of the GKE (w: $I = 5k$, w/o: removed) and AFC (w: $C = 200$, w/o: $C = 1$) modules.

GKE	AFC	Phase: 0	1	2	3	4	5	6	7	8
×	×	13.20	11.99	11.29	10.01	9.39	9.22	8.81	8.10	7.99
×	✓	12.56	10.80	10.29	9.81	9.36	8.60	8.00	7.89	7.22
✓	×	**74.80**	68.98	64.11	59.35	55.78	52.28	49.08	47.02	43.79
✓	✓	74.35	**70.32**	**66.21**	**62.37**	**60.01**	**56.98**	**55.12**	**53.39**	**51.21**

Hyperparameters Analysis

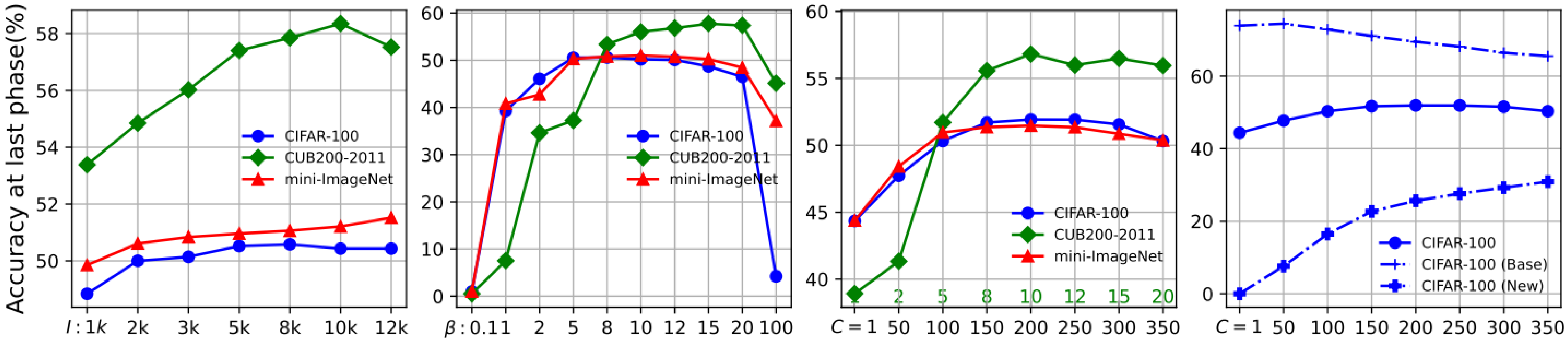
JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

❑Analyze the parameters including GKE parameter, width parameter, AFC parameter:

1. GKEAL hungers for a larger GKE parameter;

2. Exceeding bound of width parameter will cause performance drop;

3. AFC balance the preference of base and new data.



(a) Effect of GKE parameter (b) Effect of width parameter (c) Effect of AFC parameter (d) Performance trade-off

Conclusion

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

- ❑ GKEAL handles the FSCIL problem with the kernel embedded module.
- ❑ AFC is another contribution to balance the base-new knowledge.
- ❑ GKEAL shows outstanding performance compared with SOTA in various experiments.