Mnemonics Training: Multi-Class Incremental Learning without Forgetting

CVPR 2020 Oral

Webpage: https://mnemonics.yyliu.net/

Code: https://github.com/yaoyao-liu/mnemonics



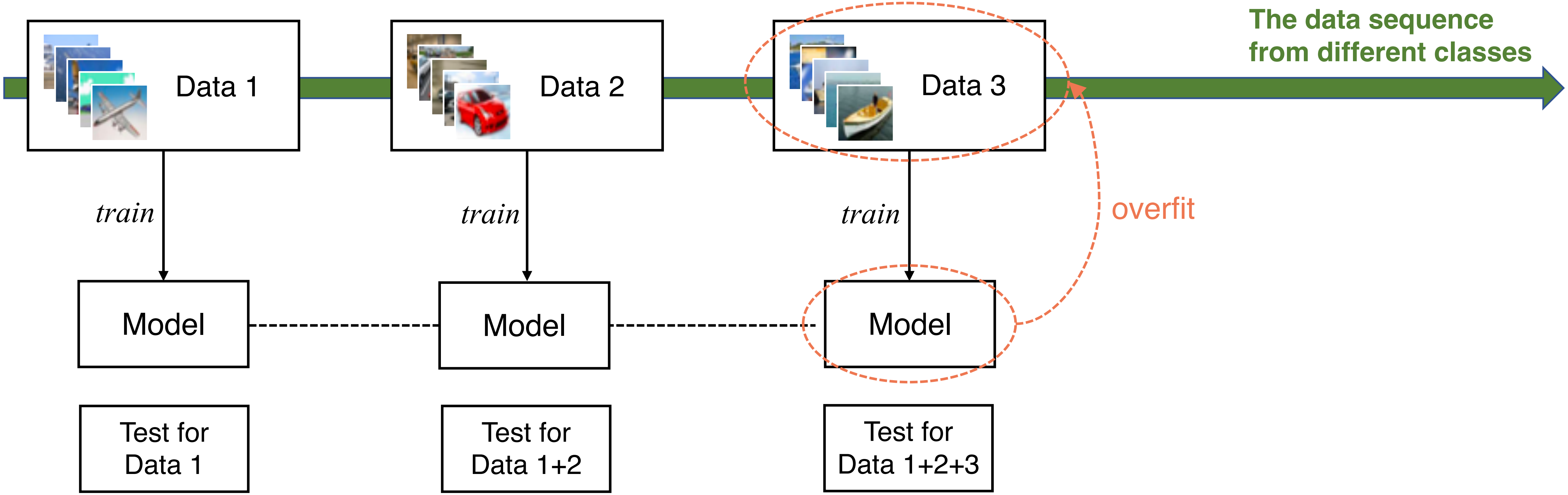


Yaoyao Liu Yuting Su An-An Liu Bernt Schiele Qianru Sun

max planck institut informatik
天津大学 Tianjin University
SMU SINGAPORE MANAGEMENT UNIVERSITY

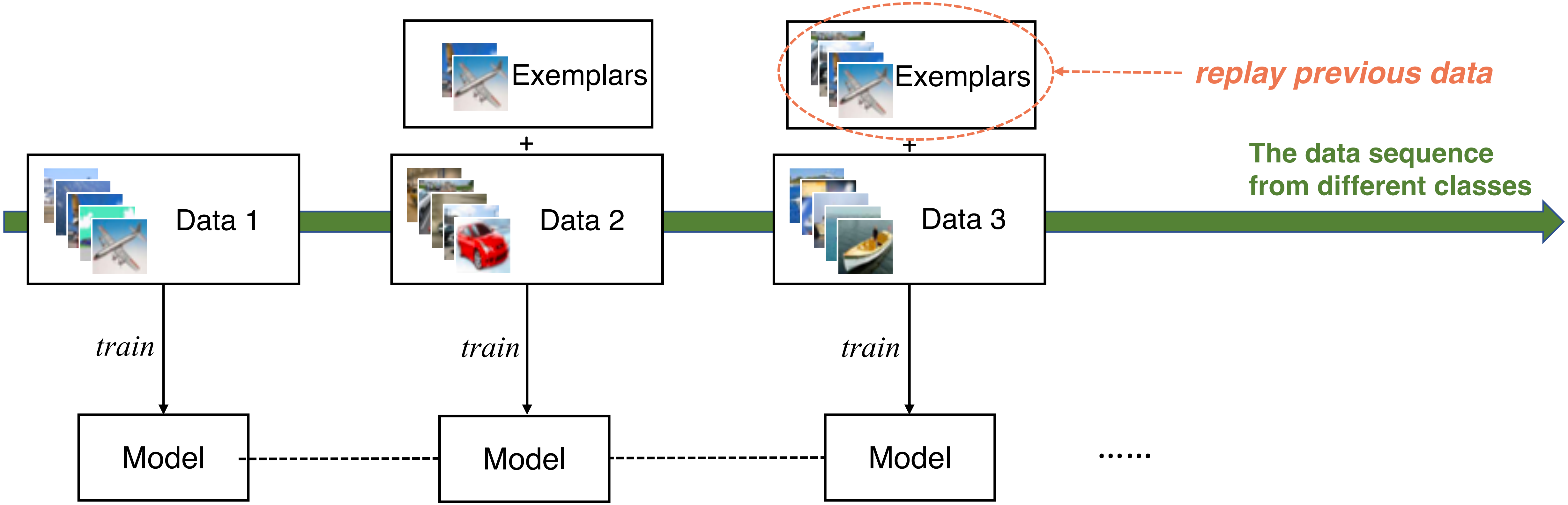
Multi-class incremental learning

Challenge: catastrophic forgetting



(Images from CIFAR-100 dataset)

Replay previous data within memory limitations



Question: how to extract the exemplars?

(Images from CIFAR-100 dataset)

Question: how to extract the exemplars?

Existing methods

- *Herding (nearest neighbor)* [1][2]
- *Random sampling* [2]

Limitations:
- *Heuristic selection, not performance-based*
- *Select from finite sets (real images)*

Our method*: mnemonics training*

Key idea*: bilevel optimaztion*

Benefits:
+ *Optimal selection by end-to-end training*
+ *Select from continuous (infinite) synthetic data*

Reference
[1] Rebuffi, Sylvestre-Alvise, et al. "iCaRL: Incremental Classifier and Representation Learning." CVPR 2017;
[2] Wu, Yue, et al. "Large scale incremental learning." CVPR 2019.

Mnemonics training: select exemplars to best represent the history data

In the i-th incremental phase,

Data for the current phase

We have: Exemplars for previous phases $\boxed{\mathcal{E}_{0:i-1}} \cup \boxed{D_i} \xrightarrow{\textit{train}} \Theta_i^*$ *(model)*

$\approx$

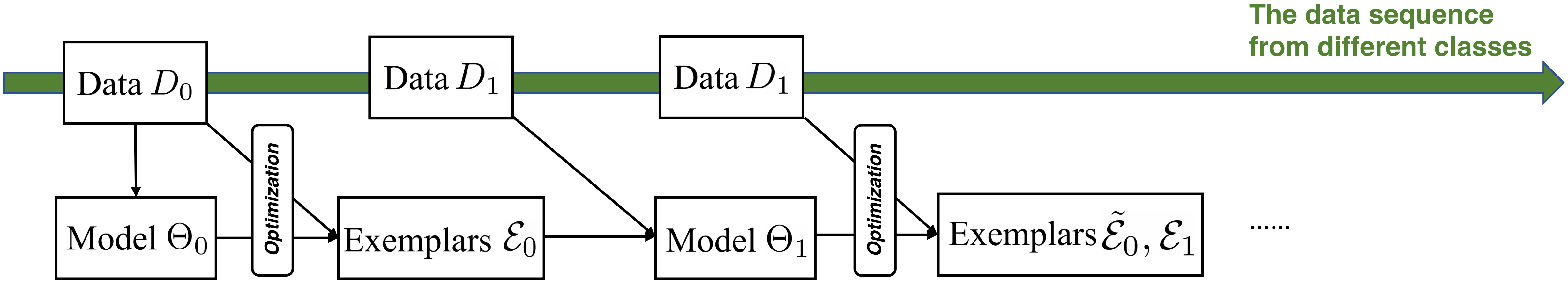
We aim to get: Exemplars for the current phase $\boxed{\mathcal{E}_{0:i}} \xrightarrow{\textit{train}} \Theta_i^{\mathcal{E}}$

Bilevel optimization formulation:

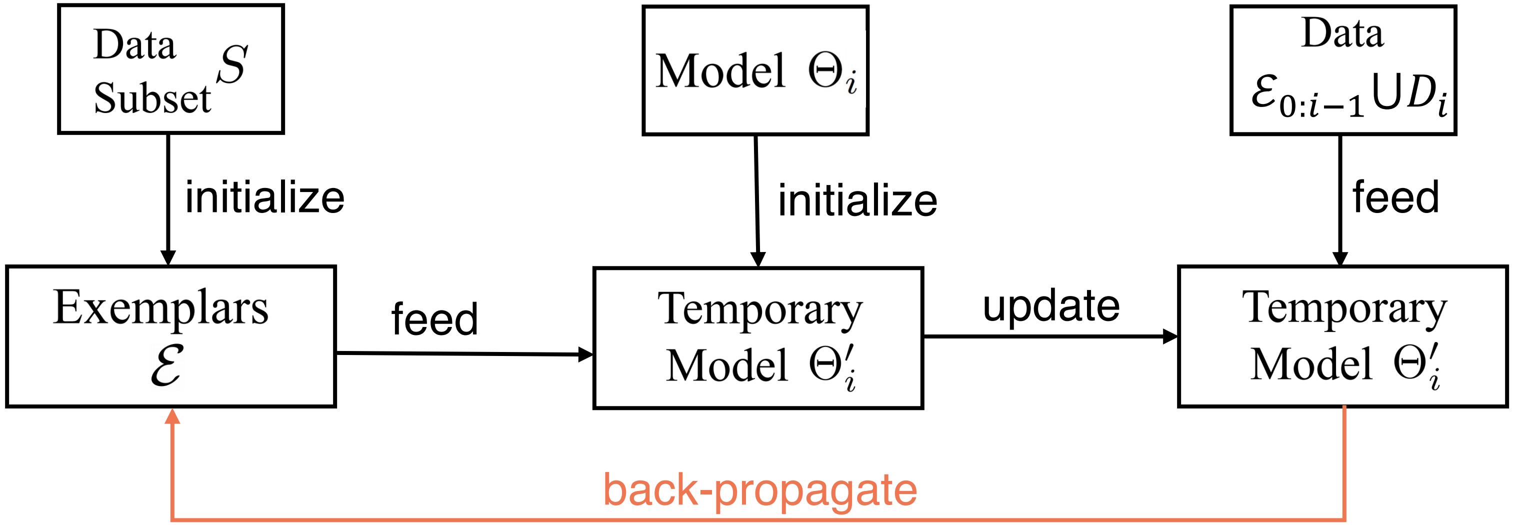
$$\min_{\mathcal{E}_{0:i}} \mathcal{L}\left(\Theta_i^{\mathcal{E}}; \mathcal{E}_{0:i-1} \cup D_i\right)$$

$$\text{s.t. } \Theta_i^{\mathcal{E}} = \min_{\mathcal{E}_{0:i}} \mathcal{L}\left(\Theta_i; \mathcal{E}_{0:i}\right)$$

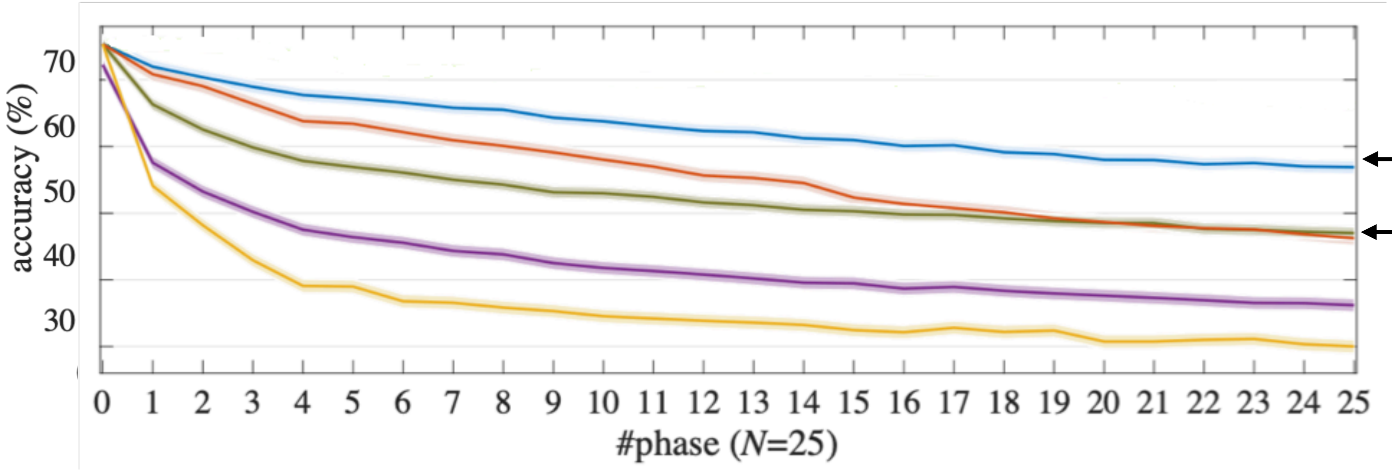
The global computing flow:



In the i-th incremental phase:



Our method boosts the performance



Ours

LUCIR [3]

Boost SOTA for 10%

Metric	Method	*CIFAR-100*			*ImageNet*		
		N=5	10	25	5	10	25
	LwF$^\diamond$ (2016)	49.59	46.98	45.51	44.35	38.90	36.87
	LwF *w/* ours	54.43	52.67	51.75	52.70	50.37	50.79
	iCaRL (2017)	57.12	52.66	48.22	51.50	46.89	43.14
Average acc. (%) ↑	iCaRL *w/* ours	59.88	57.53	54.30	60.61	58.62	53.46
$\bar{\mathcal{A}} = \frac{1}{N+1}\sum_{i=0}^{N}\mathcal{A}_i$	BiC (2019)	59.36	54.20	50.00	62.65	58.72	53.47
	BiC *w/* ours	60.67	58.11	55.51	64.63	62.71	60.20
	LUCIR (2019)	63.17	60.14	57.54	64.45	61.57	56.56
	LUCIR *w/* ours	**64.95**	**63.25**	**63.70**	**66.15**	**63.12**	**63.08**

- *Generic*
- *Boost the performance for **FOUR** different baselines*

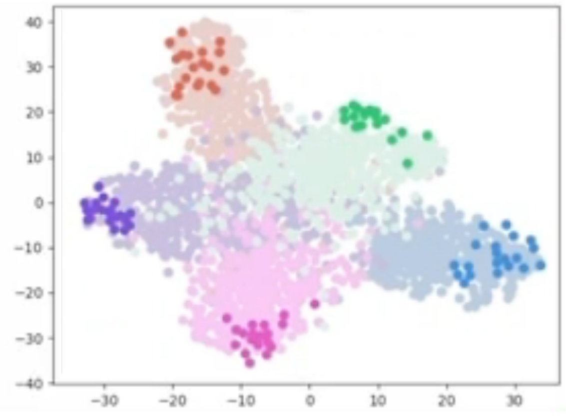
Reference

[3] Hou, Saihui, et al. "Learning a unified classifier incrementally via rebalancing." CVPR 2019.

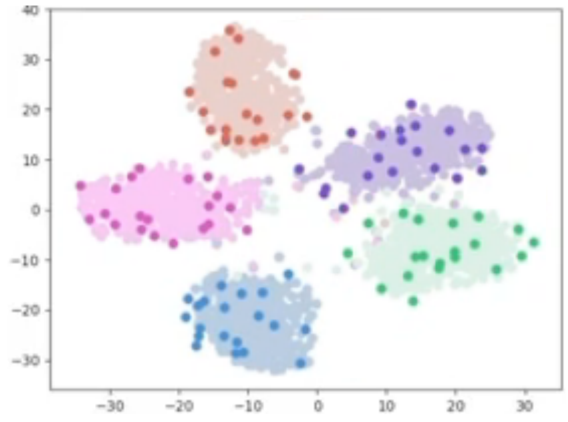
t-SNE results: clearer separation in the data

Phase 25

One region for one class
Light color: original data
Deep color: exemplars



Herding [1] [2]



Mnemonics (ours)

Our method:

- *Clearer separation in data*
- *Exemplars locate on the class boundaries*

Reference
[1] Rebuffi, Sylvestre-Alvise, et al. “iCaRL: Incremental Classifier and Representation Learning.” CVPR 2017;
[2] Wu, Yue, et al. “Large scale incremental learning.” CVPR 2019.

arXiv:2002.10211v3 [cs.CV] 7 May 2020

Mnemonics Training: Multi-Class Incremental Learning without Forgetting

Yaoyao Liu[1,2*] Yuting Su[1†] An-An Liu[1†] Bernt Schiele[2] Qianru Sun[3]

[1]School of Electrical and Information Engineering, Tianjin University
[2]Max Planck Institute for Informatics, Saarland Informatics Campus
[3]School of Information Systems, Singapore Management University
{yaoyao.liu, schiele, qsun}@mpi-inf.mpg.de
{liuyaoyao, ytsu, liuanan}@tju.edu.cn qianrusun@smu.edu.sg

Abstract

Multi-Class Incremental Learning (MCIL) aims to learn new concepts by incrementally updating a model trained on previous concepts. However, there is an inherent trade-off to effectively learning new concepts without catastrophic forgetting of previous ones. To alleviate this issue, it has been proposed to keep around a few examples of the previous concepts but the effectiveness of this approach heavily depends on the representativeness of these examples. This paper proposes a novel and automatic framework we call mnemonics, *where we parameterize exemplars and make them optimizable in an end-to-end manner. We train the framework through bilevel optimizations, i.e., model-level and exemplar-level. We conduct extensive experiments on three MCIL benchmarks, CIFAR-100, ImageNet-Subset and ImageNet, and show that using* mnemonics *exemplars can surpass the state-of-the-art by a large margin. Interestingly and quite intriguingly, the* mnemonics *exemplars tend to be on the boundaries between different classes*[1].

1. Introduction

Natural learning systems such as humans inherently work in an incremental manner as the number of concepts increases over time. They naturally learn new concepts while not forgetting previous ones. In contrast, current machine learning systems, when continuously updated using novel incoming data, suffer from catastrophic forgetting (or catastrophic interference), as the updates can override knowledge acquired from previous data [12, 20, 21, 24, 28]. This is especially true for multi-class incremental learning (MCIL) where one cannot replay all previous inputs. Catastrophic forgetting, therefore, becomes a major problem for MCIL systems.

Motivated by this, a number of works have recently emerged [2, 9, 16, 17, 25, 36]. Rebuffi et al. [25] firstly defined a protocol for evaluating MCIL methods, i.e., to tackle the image classification task where the training data for different classes comes in sequential training phases. As it is neither desirable nor scaleable to retain all data from previous concepts, in their protocol, they restrict the number of exemplars that can be kept around per class, e.g., only 20 exemplars per class can be stored and passed to the subsequent training phases. These "20 exemplars" are important to MCIL as they are the key resource for the model

Early phase (50 classes used, 5 classes visualized in color):

Late phase (100 classes used, 5 classes visualized in color):

random (baseline) *herding* (related) *mnemonics* (ours)

Figure 1. The t-SNE [18] results of three exemplar methods in two phases. The original data of 5 colored classes occur in the early phase. In each colored class, deep-color points are exemplars, and light-color ones show the original data as reference of the real data distribution. Gray crosses represent other participating classes, and each cross for one class. We have two main observations. (1) Our approach results in much clearer separation in the data, than *random* (where exemplars are randomly sampled in the early phase) and *herding* (where exemplars are nearest neighbors of the mean sample in the early phase) [2,9,25,36]. (2) Our learned exemplars mostly locate on the boundaries between classes.

*This work was done during Yaoyao's internship supervised by Qianru.
†Corresponding authors.
[1]Code: https://github.com/yaoyao-liu/mnemonics

Mnemonics Training: Multi-Class Incremental Learning without Forgetting

Webpage: https://mnemonics.yyliu.net/

Code: https://github.com/yaoyao-liu/mnemonics

Thank you!

Yaoyao Liu

Webpage: https://mpi-inf.mpg.de/~yaliu/

Code: yaoyao.liu@mpi-inf.mpg.de