Neural Networks 166 (2023) 174–187

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Dual memory model for experience-once task-incremental lifelong learning☆

Gehua Ma [a], Runhao Jiang [a], Lang Wang [b], Huajin Tang [a,*]

[a] College of Computer Science and Technology, Zhejiang University, Hangzhou, China
[b] Department of Neurology of the First Affiliated Hospital, Interdisciplinary Institute of Neuroscience and Technology, Zhejiang University School of Medicine, Hangzhou, China

ARTICLE INFO

Article history:
Received 1 July 2022
Received in revised form 21 June 2023
Accepted 7 July 2023
Available online 13 July 2023

Keywords:
Experience-once lifelong learning
Task-incremental lifelong learning
Continual learning
Catastrophic forgetting
Dual-store memory model

ABSTRACT

Experience replay (ER) is a widely-adopted neuroscience-inspired method to perform lifelong learning. Nonetheless, existing ER-based approaches consider very coarse memory modules with simple memory and rehearsal mechanisms that cannot fully exploit the potential of memory replay. Evidence from neuroscience has provided fine-grained memory and rehearsal mechanisms, such as the dual-store memory system consisting of PFC-HC circuits. However, the computational abstraction of these processes is still very challenging. To address these problems, we introduce the Dual-Memory (Dual-MEM) model emulating the memorization, consolidation, and rehearsal process in the PFC-HC dual-store memory circuit. Dual-MEM maintains an incrementally updated short-term memory to benefit current-task learning. At the end of the current task, short-term memories will be consolidated into long-term ones for future rehearsal to alleviate forgetting. For the Dual-MEM optimization, we propose two learning policies that emulate different memory retrieval strategies: Direct Retrieval Learning and Mixup Retrieval Learning. Extensive evaluations on eight benchmarks demonstrate that Dual-MEM delivers compelling performance while maintaining high learning and memory utilization efficiencies under the challenging experience-once setting.

© 2023 Elsevier Ltd. All rights reserved.

1. Introduction

Despite significant advances in machine learning, modern artificial neural networks still fail to perform lifelong learning due to catastrophic forgetting, which means the model forgets much of its previously learned knowledge when learning for a new task (French, 1999; Hassabis, Kumaran, Summerfield, & Botvinick, 2017; McCloskey & Cohen, 1989; Parisi, Kemker, Part, Kanan, & Wermter, 2019). As a result, anytime a new task becomes available, the entire neural network must go through the offline training process again, which is neither cost-effective nor practicable in the real world. Experience replay, which maintains a randomly updated memory buffer of samples from previous tasks, is a widely-used technique to reduce forgetting. However, existing experience replay methods only consider simplistic memory buffers and replay processes. More sophisticated models may be necessary to fully understand and exploit the potential of memory replay.

Neuroscience research on memory mechanisms has provided some ideas for refining the memory module. In the brain, short-term memory (STM) and long-term memory (LTM) serve distinct functions, and the idea that different systems support them has also become a key premise of modern cognitive psychology (James, Burkhardt, Bowers, & Skrupskelis, 1890). According to the framework of systems consolidation, long-term memories are initially bound by a fast-learning system in the hippocampus and followed by a stabilization of these memory traces (*i.e.*, consolidation) in long-term stores (Chen, Niknazar, Alaynick, Whitehurst, & Mednick, 2021). Neuroscientists suggest that the brain may facilitate memory consolidation by increasing communication between the neocortex–hippocampus circuits via complex microscopic neural activity (Gais et al., 2007; Mednick et al., 2013; Ngo, Fell, & Staresina, 2020; Rasch & Born, 2013). Research on STM mechanisms, on the other hand, suggests that the brain actively preserves the information held in STM through the activation of prefrontal neurons (Funahashi, Bruce, & Goldman-Rakic, 1989). The PFC-HC dual memory system (Fig. 1) implements a stratified memory paradigm in which STM, a repository for task-relevant information, is linked to LTM by consolidation and retrieval. Existing research has revealed a subtle bi-directional interaction between the updating of STM and memory consolidation. However, how these processes are carried out and how these memories are involved in cognition and learning is still unclear.

☆ This work was supported by the National Key Research and Development Program of China under Grant No. 2020AAA0105900 and the National Natural Science Foundation of China under Grant No. 62236007.

* Corresponding author.
E-mail address: htang@zju.edu.cn (H. Tang).

https://doi.org/10.1016/j.neunet.2023.07.009
0893-6080/© 2023 Elsevier Ltd. All rights reserved.

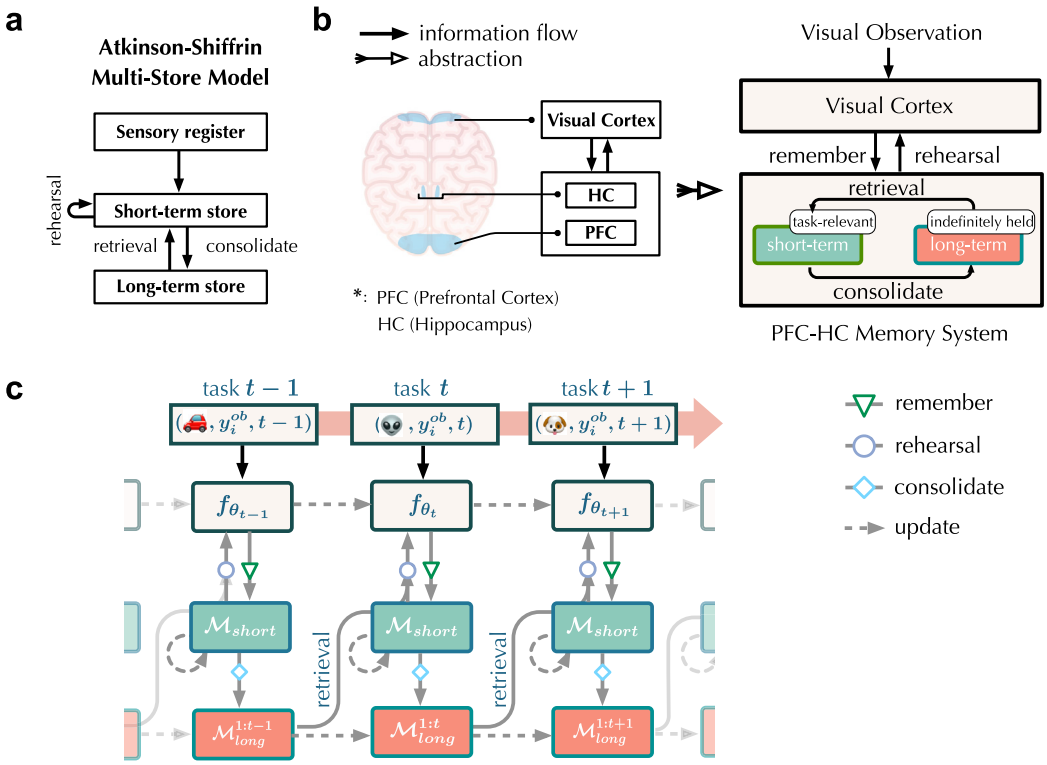

Fig. 1. Schematic of how Dual-MEM is mapped onto the brain cognitive system. **a**: Graphical illustration of Atkinson–Shiffrin Multi-store memory model (Atkinson & Shiffrin, 1968). **b**: A simplified structure of the brain's visual cognitive system, which views the vision recognition prefrontal cortex-hippocampus (PFC-HC) circuits as a system that consists of short-term (current-task related) and long-term memories. This abstraction can also be considered a visual case of the Atkinson–Shiffrin multi-store model. **c**: A brief illustration of the lifelong learning procedure of the proposed Dual-MEM model on a task sequence. The Dual-MEM model is optimized via a hybrid objective function (refer to Section 4.2.3) composed of the main objective and an LTM-based dark knowledge regularizer.

This work introduces an efficient rehearsal-based task-incremental lifelong learning approach, dubbed Dual-MEM (Dual Memory model), illuminated by the PFC-HC memory system. Specifically, we propose a concrete computational implementation of the abstract learning process involving a dual memory system mimicking the PFC-HC system. We offer the computational account, which includes online updating of STM, memory consolidation, and visual recognition learning with the STM and LTM system. The proposed Dual-MEM model (Fig. 1) employs an incrementally updated fixed-size STM to aid the current-task learning and consolidate STMs to long-term ones for future rehearsal to alleviate catastrophic forgetting. We propose two learning methodologies for Dual-MEM model optimization. In addition, to further increase the model's stability on previously learned tasks, we introduce a regularization term using response dark knowledge from the LTM. Based on thorough experiments, we show that the Dual-MEM model enables more efficient learning while maintaining low resource consumption compared to existing approaches.

2. Related works

2.1. Lifelong learning approaches

Lifelong learning, also termed continual learning, considers learning through a sequence of tasks. In lifelong learning, the learner has to continually accumulate general knowledge while retaining task-specific information about past tasks (Konidaris & Barto, 2006; Taylor & Stone, 2009). This definition led to implementations in both reinforcement learning problems (Abel, Jinnai, Guo, Konidaris, & Littman, 2018; Song, Sun, Song, & Stojanovic, 2022; Thrun, 1995; Xin et al., 2022; Xu, Li, & Stojanovic, 2021) and recognition problems (Lavda, Ramapuram, Gregorova, & Kalousis, 2018; Lopez-Paz & Ranzato, 2017; Rebuffi, Kolesnikov, Sperl, & Lampert, 2017). In general, the setting is the task-incremental setting, in which task descriptors are fed to the learner and observations (Wilson, Fern, Ray, & Tadepalli, 2007). In particular, in this work, we focus the lifelong learning for recognition problems in a realistic situation where the observations are experienced only once, memory is limited, and the learner is also given task indicators.

Regularization-based approaches. Some research, like Refs. Jung, Ju, Jung, and Kim (2016), Li and Hoiem (2017), Rannen, Aljundi, Blaschko, and Tuytelaars (2017), proposed to alleviate forgetting by reducing the changes of activation in intermediate feature space or final output space. However, these methods require either snapshotting the old model parameters or continually adding parameterized modules for new tasks. Synaptic Intelligence (SI) (Zenke, Poole, & Ganguli, 2017) and Memory Aware Synapses (MAS) (Aljundi, Babiloni, Elhoseiny, Rohrbach, & Tuytelaars, 2018) estimate the weight importance while EWC (Kirkpatrick et al., 2017) and GEM (Chaudhry, Marc'Aurelio, Rohrbach, & Elhoseiny, 2019; Lopez-Paz & Ranzato, 2017) derive angular constraints over the gradient steps to regularize the model optimization. In Ref. Zeng, Chen, Cui, and Yu (2019), an orthogonal prior was introduced for regularization, and a brain-inspired context-dependent module was designed to improve model flexibility. Since a stratified probabilistic model can formulate the learning and inference in the brain Friston (2003), some works introduce variational inference to avoid catastrophic forgetting (Ahn, Cha, Lee, & Moon, 2019; Nguyen, Li, Bui, & Turner, 2018; Zhao, Wang, Masoomi, & Dy, 2022). Nevertheless, these methods inevitably induce additional heavy computations to the training process.

Parameter isolation-based approaches. Parameter-isolation methods apply different subsets of model parameters for each task to prevent catastrophic forgetting. This is realized by activating task-related parameter mask, unit mask, or gates (Aljundi, Chakravarty, & Tuytelaars, 2017; Mallya & Lazebnik, 2018; Masse, Grant, & Freedman, 2018; Serra, Suris, Miron, & Karatzoglou, 2018). Unfortunately, these parameter isolation-based methods ineluctably increase the optimization difficulty and hinder general knowledge accumulation across tasks.

Rehearsal-based approaches. The reactivation of neuronal activity patterns indicating old-task memories in the brain is believed to be essential for preserving the ability to solve experienced tasks (Ji & Wilson, 2007; Qin, McNaughton, Skaggs, & Barnes, 1997; Rasch & Born, 2007; Wilson & McNaughton, 1994). With this motivation, researchers developed rehearsal-based approaches that periodically repeat the knowledge from earlier experiences to prevent catastrophic forgetting. Pseudo rehearsal methods (Kemker & Kanan, 2018; Lavda et al., 2018; Shin, Lee, Kim, & Kim, 2017; van de Ven, Siegelmann, & Tolias, 2020) use generative models with statistical information acquired from previous tasks and retrain on the synthetic samples and the new ones. However, these methods introduce extra model parameters and inevitably increase the complexity of the training. Exact rehearsal approaches (Chaudhry, Gordo, Dokania, Torr, & Lopez-Paz, 2021; Chaudhry, Rohrbach, et al., 2019; Rebuffi et al., 2017) store samples from previous tasks in a memory buffer and directly retrain the model on memory while learning a new task. However, the limitation of the aforementioned exact rehearsal-based methods is twofold. First, their memory allocation policy typically assumes that all classes within a task are equal and assigns the same memory budget to all categories. Second, they use memory updating strategies such as feature space moving average, which imposes heavy extra computational overhead.

Conventional lifelong learning methods have been effectively expanded in recent literature. Ref. Pham, Liu, Sahoo, and HOI (2021) proposed a contextual transformation-based approach to handle lifelong learning in an online paradigm. Similarly, in Ref. Lee, Ha, Zhang, and Kim (2020), the authors proposed a novel Dirichlet process mixture model for continual learning in a task-free setting, extending the work of Shin et al. (2017). This task-free setting is combined with zero-shot learning in Ref. Gautam, Parameswaran, Mishra, and Sundaram (2022). Ref. Sun, Cong, Wang, Zhong, and Fu (2022) introduced an efficient, flexible feature learning mechanism. Inspired by the trending self-attention research, Wang, Zhang, Ebrahimi, et al. (2022), Wang, Zhang, Lee, et al. (2022) creatively introduce prompts into continual learning and achieve competitive performance. In Ref. Dong et al. (2022), the class-incremental learning setting is further extended to the federated learning scenario.

2.2. Neuroscience-inspired lifelong learning methods

Some existing works draw inspiration from neuroscience theories and mechanisms to design algorithms to perform lifelong learning. The Complementary Learning System (CLS) is commonly employed as a biological evocator. Shin et al. (2017) proposed a CLS-inspired method that trains scholar and generative models sequentially. Pham, Liu, and Hoi (2021) introduced a hybrid method that includes a slow learner optimized by unsupervised learning and a fast learner trained by supervised learning, inspired by the long- and short-term learning mechanism in CLS. The intermediate feature-level adaptation connects the fast and slow learners in their method. Elahe Arani (2022) developed an approach that preserves semantic memories by storing previous model weights in addition to basic exemplar memory rehearsal. Unlike these works, this work focuses on the memory module and improving the efficiency of replay through more fine-grained processing within the memory module.

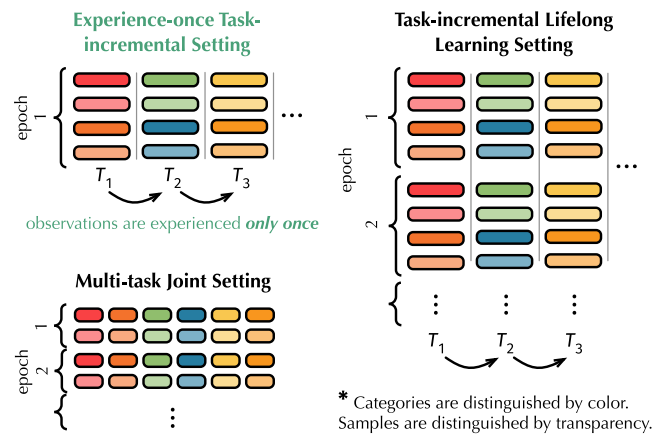

Fig. 2. A schematic illustration of the *experience-once* task-incremental lifelong learning setting (top-left) used in this work, unlike the normal task-incremental setting (right), observations are experienced once in the adopted setting, which is more realistic.

In particular, this work is closely related to previous lifelong learning methods that adopt dual memory designs. Ref. Kamra, Gupta, and Liu (2018) proposed a neuroscience-inspired model consisting of a set of generative models as STMs, interacting with LTM through deep generative replay to reduce forgetting, another research (Rostami, Kolouri, & Pilly, 2019) also adopted a similar strategy. In IL2M (Belouadah & Popescu, 2019), LTM is defined as task statistics, which are used to mitigate forgetting by test-time refinement of the model's predictive scores. In Ref. Sun, Cong, Dong, et al. (2022), the authors proposed a novel approach that considers orthogonal basis memories and embedding memories to achieve lifelong spectral clustering. Ref. Fahad Sarfraz (2023) considered the model parameters from the previous training stage as the LTM and integrated these parameters into the new model through moving averages to reduce forgetting. It can be seen that the dual-memory designs considered in the above works are different from this work. The proposed dual memory approaches are a brain-inspired extension of the single memory experience replay approach without the need for additional training of generative models or storage of previous model weights during learning. This is achieved by explicitly dividing the memory buffer into STM related to the current task and LTM related to past tasks.

3. Preliminary

3.1. Experience-once task-incremental setup

The tasks that biological agents encounter during their learning process have complex relationships that may be linked or even overlap. Thus, creating a completely realistic setting for computational implementations is challenging. Here we adopt a widely-used simplified setup (van de Ven et al., 2020; Zeng et al., 2019) with explicit task boundaries (Kudithipudi et al., 2022; Saxena, Shobe, & McNaughton, 2022), as shown in Fig. 2. Specifically, we consider an experience-once task-incremental setting following prior literature (Hayes, Cahill, & Kanan, 2019; Lopez-Paz & Ranzato, 2017). The task sequence consists of data (x_i^{ob}, y_i^{ob}, t) which includes observation x_i^{ob}, one-hot target y_i^{ob} and task indicator $t \in \mathcal{T} = \{1, \ldots, T\}$. This is a natural scenario for many recognition and reinforcement learning problems (Kirkpatrick et al., 2017; Van de Ven & Tolias, 2019). For task t, each

data pair (x_i^{ob}, y_i^{ob}) is an identical and independently distributed example drawn from the data generating process distribution P_t of that task. The training set contains T subsets $\mathcal{D}_t \overset{i.i.d}{\sim} P_t$. All data becomes available for *single-pass* (one-epoch) offline training for a given task, while data from other tasks cannot be accessed. Under this setup, the goal of the learner is to estimate a parameterized predictor $f(x, t; \theta) : \mathcal{X} \times \mathcal{T} \to \mathcal{Y}$ by minimizing the objective over all training data,

$$\arg\min_{\theta} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_{\mathcal{D}_t} \left[\mathcal{O}(\hat{y}_i^{ob}, y_i^{ob}) \right], \tag{1}$$

where the model prediction $\hat{y}_i^{ob} = \text{softmax}(z_i)$, model response $z_i = f_\theta(x_i^{ob}, t)$ and objective function $\mathcal{O} : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$.

3.2. Basic definitions and notations

We use a fixed-size STM $\mathcal{M}_{short} = \{(x_i^s, y_i^s, z_i^s)\}$ to store exemplars from the current task, and an LTM $\mathcal{M}_{long} = \{(x_i^l, y_i^l, z_i^l, t)\}$ to store samples from past tasks. We denote mini-batch sampled from the input data stream (observation), STM, and LTM as $\mathcal{B}_{ob}$, $\mathcal{B}_{short}$ and $\mathcal{B}_{long}$, respectively. We use ℓ to denote cross entropy function, and $|\cdot|$ to represent the cardinality, *e.g.* $|\mathcal{M}|$ denotes the memory size.

4. Method

We elucidate Dual-MEM as follows. Dual memory management rules in Section 4.1, and learning strategies in Sections 4.2.1 and 4.2.2 and additional regularization using model response in Section 4.2.3.

4.1. Dual memory management

4.1.1. Incremental surprise maximization updating

A critical difficulty for rehearsal-based lifelong learning algorithms is selecting the data to be saved in the memory buffer. Here, we propose a bio-inspired solution for updating short-term memory in Dual-MEM, coined **S**urprise **M**aximization **U**pdating (**SMU**). It is revealed that, for biological agents, the violation of a prior prediction from the internal cognitive distribution $\hat{P}$ leads to enhanced memory for the surprising item (Chen, Cook, & Wagner, 2015; Friston, 2010). This strategy is reminiscent of the online hard example mining policy (Felzenszwalb, McAllester, & Ramanan, 2008; Shrivastava, Gupta, & Girshick, 2016) frequently used in machine learning solutions. To mimic this procedure, we start by approximating the internal distribution $\hat{P}$ using P_θ, an artificial neural network that is parameterized by θ, whose deterministic mapping is denoted by f_θ. We adopt the *bayesian surprise* (*surprise* for short) (Itti & Baldi, 2009) to quantitatively measure the difference between the prior prediction and the observation as the Kullback–Leibler divergence of the observation distribution $P_{\mathcal{D}_t}$ and the internal distribution $\hat{P}$ (approximated by P_θ). For a given task, $\arg\max KL(P_{\mathcal{D}_t} \parallel P_\theta) = \arg\max H(P_{\mathcal{D}_t}, P_\theta)$ since $KL(P_{\mathcal{D}_t} \parallel P_\theta) = H(P_{\mathcal{D}_t}, P_\theta) - H(P_{\mathcal{D}_t})$ and $H(P_{\mathcal{D}_t})$ is constant during task t. Thus, we can re-formulate the objective of SMU and describe it by

$$\arg\max_{\mathcal{M}_{short}} \sum_{i=1}^{|\mathcal{M}_{short}|} \ell\left(\hat{y}_i^s, y_i^s\right), \text{ s.t., } (x_i^s, y_i^s, z_i^s) \in \mathcal{M}_{short} \tag{2}$$

where $\hat{y}_i^s = \text{softmax}(f_\theta(x_i^s, t))$ and ℓ is the cross entropy function.

Nevertheless, directly applying SMU described in Eq. (2) induces two problems: (a) It requires access to the entire distribution $P_{\mathcal{D}_t}$, which is not realizable under our lifelong learning setup; (b) It is expensive to estimate the expectation over whole $\mathcal{M}_{short}$ on-the-fly, especially when the size of STM $|\mathcal{M}_{short}|$ is large. To circumvent the above issues, we propose a variant of SMU, dubbed **ISMU** (**I**ncremental **S**urprise **M**aximization **U**pdating) to approximate Eq. (2). For every iteration, the ISMU updating is described by

$$\mathcal{M}_{short} \leftarrow \mathcal{M}_{short} \setminus \mathcal{B}_{short} \cup \mathcal{A}$$

where

$$\mathcal{A} \subsetneq (\mathcal{B}_{short} \cup \mathcal{B}_{ob}) \wedge |\mathcal{A}| = |\mathcal{B}_{short}| \wedge \ \mathcal{A} = \arg\max_{\mathcal{A}} \ell(\mathcal{A}), \tag{3}$$

the $\setminus, \cup$ denote set subtraction and set addition, respectively. In particular, we randomly sample $\mathcal{B}_{short}$ from $\mathcal{M}_{short}$, $\mathcal{B}_{ob}$ from $\mathcal{D}$ and compute the cross entropy *surprise* of all samples in $\mathcal{B}_{ob} \cup \mathcal{B}_{short}$. Exemplars in $\mathcal{B}_{short}$ with low *surprise* will be replaced with higher *surprise* samples from $\mathcal{B}_{ob}$, and when $\mathcal{B}_{short}$ is updated, the same updates will be made in $\mathcal{M}_{short}$ synchronously. As the number of samples traversed by the model f_θ increases, $\mathcal{M}_{short}$ will approach the ideal $\mathcal{M}_{short}^*$. We summarize the STM updating algorithm in Alg. 1.

Algorithm 1 ISMU: STM updating by incremental *surprise* maximization updating.

Input: t, $\mathcal{M}_{short}$
Input: ℓ_{ob}, ℓ_{short}
Input: $\mathcal{B}_{ob} = \{(x_i^{ob}, y_i^{ob})\}_{i=1}^{B}$
Input: $\mathcal{B}_{short} = \{(x_j^s, y_j^s, z_j^s)\}_{j=1}^{B}$
while $\min_j \ell(\hat{y}_j^s, y_j^s) < \max_i \ell(\hat{y}_i^{ob}, y_i^{ob})$ **do**
  $m \leftarrow \arg\min_j \ell(\hat{y}_j^s, y_j^s)$
  $n \leftarrow \arg\max_i \ell(\hat{y}_i^{ob}, y_i^{ob})$
  $\mathcal{B}_{short}[m] \leftarrow (x_n^{ob}, y_n^{ob}, f_\theta(x_n^{ob}, t))$
  ▷ update mini-batch slot
end while
update $\mathcal{B}_{short}$ to $\mathcal{M}_{short}$
▷ update memory slot
return $\mathcal{M}_{short}$

4.1.2. Memory consolidation

Consolidation is a fundamental mnemonic operation that relates STM to LTM (Paller, 2009; Preston & Eichenbaum, 2013; van de Ven, Trouche, McNamara, Allen, & Dupret, 2016). The neuroscience findings on memory consolidation serve as the foundation for our computational implementation of memory consolidation. During the memory consolidation, the hippocampal activation level declines progressively, while the ventral medial prefrontal region's activations are progressively stronger (Takashima et al., 2006). This implied functional region transfer during memory consolidation, which motivated us to use a simplified computational implementation to imitate this process. Specifically, we implement memory consolidation as data transfer of memory content from STM to LTM. All exemplars in $\mathcal{M}_{short}$ will be transferred to $\mathcal{M}_{long}$ after observations of the current task have been experienced once. Further, motivated by the fact that information stored in STM is cleared out during the consolidation process to make room for new memories (Rosenzweig, Barnes, & McNaughton, 2002), we simulate the clear-out process in our computational implementation. We abstract the memory consolidation as

$$\begin{aligned} \mathcal{M}_{long}^{1:t} &\leftarrow \mathcal{M}_{long}^{1:t-1} \cup \mathcal{M}_{short} \\ \mathcal{M}_{short} &\leftarrow \varnothing, \end{aligned} \tag{4}$$

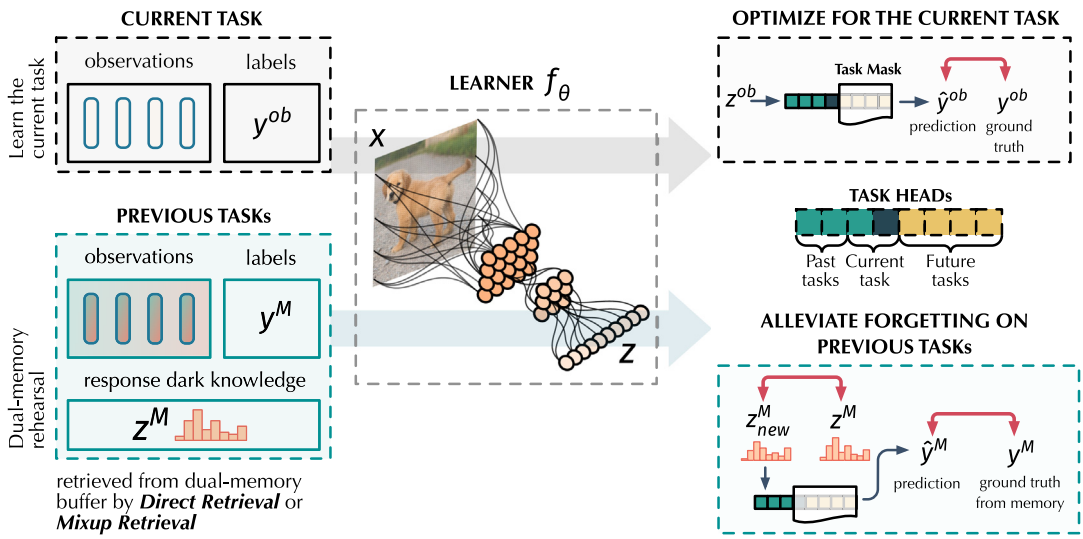

Fig. 3. A schematic illustration of the learning pipeline with the proposed dual memory. It includes optimization for the current task and forgetting alleviation by rehearsal using label and response dark knowledge information retrieved from the dual memory buffer.

where $\mathcal{M}_{long}^{1:t}$ denotes cumulated LTM at the end of task t. And the size of $\mathcal{M}_{long}$ is proportionate to the size of $\mathcal{M}_{short}$:

$$|\mathcal{M}_{long}^{1:T}| = T \times |\mathcal{M}_{short}|, \tag{5}$$

where T is the number of currently-experienced tasks, albeit an intuitive consolidation policy, it is easy to implement and does not introduce extra computation.

4.2. Learning with dual memory

Learning rules specify how the neural network interacts with the dual memory (DM) module for synaptic weights updating, especially how the memories in DM are retrieved and replayed (Fig. 3). The memory retrieval process defines how memories from STM and LTM are handled and involved in subsequent model parameter updates as a whole or as parts. Here, we introduce two types of learning: Direct Retrieval learning (DR) and Mixup Retrieval learning (MR). These two learning methods, partially referencing awake replay (Karlsson & Frank, 2009; Yamamoto & Tonegawa, 2017), perform new observation receiving and memory replay simultaneously. To penalize changes to model response for the same observations, we label the LTM with relevant response dark knowledge (Hinton, Vinyals, & Dean, 2014) in addition to integer labels. On that basis, we introduce a response dark knowledge loss term for further regularization.

4.2.1. Direct retrieval learning (DRL)

Direct retrieval. One practical way to learn with DM is to directly retrieve the long-term memories and joint training on the joint set of memories from long and short-term stores; we refer to this type of strategy as Direct Retrieval learning (DRL). DRL is a simple yet effective way to perform learning with dual memory. Aside from its intuitive simplicity, the direct retrieval and merging process does not necessitate additional computation costs.

In DRL, the neural network f_θ is optimized using the samples from the observation stream of current task t, samples from the STM, and samples of previous $t-1$ tasks from the LTM. In every iteration, we draw mini-batches (with batch size B) as $\mathcal{B}_{ob} \sim \mathcal{D}_t$, $\mathcal{B}_{long} \sim \mathcal{M}_{long}$ and $\mathcal{B}_{short} \sim \mathcal{M}_{short}$ for model parameter optimization. Cross entropies computed using $\mathcal{B}_{short}, \mathcal{B}_{ob}$ are also used for ISMU updating, which further reduces additional computational expenses. The STM rehearsal here amounts to bootstrapping (Efron & Tibshirani, 1994), which resamples the sample from $\mathcal{D}_t$. Under the single-pass task-incremental lifelong learning setting, the STM rehearsal can aid the learning of the current task and improve overall performance. We use a weighting factor λ_{short} to leverage the STM rehearsal during model optimization. Meanwhile, exemplars saved in $\mathcal{M}_{long}$ from previous tasks are also replayed to preserve the capability to alleviate catastrophic forgetting. We formulate the objective function of DRL as

$$\begin{aligned}\mathcal{L}_{DR} &= \mathcal{L}_{\text{ce}}(\mathcal{B}_{ob} \cup \mathcal{B}_{long}) + \lambda_{short} \cdot \mathcal{L}_{\text{ce}}(\mathcal{B}_{short}) \\ &= \mathbb{E}_{\mathcal{B}_{ob}\cup\mathcal{B}_{long}} \ell(\hat{y}, y) + \lambda_{short} \cdot \mathbb{E}_{\mathcal{B}_{short}} \ell(\hat{y}, y).\end{aligned} \tag{6}$$

Denoting gradient computed using $\mathcal{B}_{ob}$, $\mathcal{B}_{short}$ and $\mathcal{B}_{long}$ as g_{ob}, g_{short} and g_{long}, respectively, methods like GEM (Lopez-Paz & Ranzato, 2017) adopt a modified gradient $\tilde{g}_{ob}$, which is orthogonal to the gradient g_{long} computed on a sampled set $\mathcal{B}_{long}$ to update parameter θ. However, in DRL, we use a weighted average of g_{ob}, g_{short}, and g_{long} to compute the update step in the parameter space. The gradient g_{short} encourages the update step to the loss of the current task more radically, while g_{long} contributes to preventing θ from leaving the local minima found in the learning procedure of previous tasks.

4.2.2. Mixup retrieval learning (MRL)

Mixup retrieval. The MRL considers a more flexible retrieval instantiation by adopting the mixup (Zhang, Cisse, Dauphin, & Lopez-Paz, 2018) algorithm. Mixup provides a prior-knowledge-free way to perform data augmentation via linear interpolation (Chawla, Bowyer, Hall, & Kegelmeyer, 2002; DeVries & Taylor, 2017)

$$\tilde{x} = \gamma x_i + (1-\gamma)x_j, \ \tilde{y} = \gamma y_i + (1-\gamma y_j). \tag{7}$$

where $\gamma \sim \text{Beta}(\alpha, \alpha)$, $\alpha \in (0, \infty)$. The parameter α controls the strength of interpolation between the sample-label pairs. Following the linear interpolation operation, we propose to mix the short-term and long-term memories by mixup. For sample-label pairs from STM and LTM $(x_j^s, y_j^s) \sim \mathcal{M}_{short}$, $(x_i^l, y_i^l) \sim \mathcal{M}_{long}$, the mixup retrieval is realized by

$$\tilde{x} = \gamma x_i^l + (1-\gamma)x_j^s, \ \tilde{y} = \gamma y_i^l + (1-\gamma y_j^s). \tag{8}$$

The linear interpolation between samples from short-term and long-term memories can be understood as a form of joint training. Unlike DRL, MRL is not explicitly specified by weighting factors λ of separate loss terms. Specifically, when $\alpha \rightarrow 0$, the interpolation factor γ sampled from the Beta distribution with p.d.f $p_\gamma = \frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2} x^{\alpha-1}(1-x)^{\alpha-1}$ (Γ for Gamma distribution) is very likely to be either 1 or 0. And the mixup retrieval will degrade to

$$\begin{cases} \tilde{x} = x_j^s, \ \tilde{y} = y_j^s, & \text{if } \gamma = 0 \\ \tilde{x} = x_i^l, \ \tilde{y} = y_i^l, & \text{if } \gamma = 1 \end{cases} \tag{9}$$

equivalent to pure STM rehearsal ($\gamma = 0$) and LTM rehearsal ($\gamma = 1$), respectively. Intuitively, small α approximately corresponds to a learning policy of alternative rehearsal of short-term and long-term memories. Generally, the coefficient α sampled from a Beta distribution determines the specific implementation of the memory retrieval and thus influences the rehearsal process. As a result, MRL selectively recombines and replays memories in a stochastic way instead of replaying them all as in DRL.

In MRL, we optimize f_θ using the observation data and mixup retrieved memories. In every iteration, we draw mini-batches (with batch size B) as $\mathcal{B}_{ob} \sim \mathcal{D}_t$, $\mathcal{B}_{long} \sim \mathcal{M}_{long}$ and $\mathcal{B}_{short} \sim \mathcal{M}_{short}$. Every sample (x_i^l, y_i^l) from LTM batch is merged with random-selected samples (x_j^s, y_j^s) from STM using Eq. (8), we denote the merged batch as $\tilde{\mathcal{B}}_{mem}$, the objective function of MRL is formulated as

$$\mathcal{L}_{MR} = \mathcal{L}_{\text{ce}}(\mathcal{B}_{ob} \cup \tilde{\mathcal{B}}_{mem}) = \mathbb{E}_{\mathcal{B}_{ob} \cup \tilde{\mathcal{B}}_{mem}} \ell(\hat{y}, y). \tag{10}$$

4.2.3. Response dark knowledge regularization

For both DRL and MRL, instead of only maintaining the predicted labels consistent, we also keep the model response stable. We introduce a response dark knowledge (Hinton et al., 2014) based regularizer $\mathcal{L}_{dkr}$ around the LTM to imitate the brain's replay mechanism that strives for hidden state stability (van de Ven et al., 2020). This regularizer term encourages the learner to match the model response dark knowledge z of the old-time models $\{f_{\theta_k}\}_{k=1}^{t-1}$ at the learning procedure of task t to increase model stability and further alleviate forgetting. In particular, we define the dark knowledge regularizer as the KL divergence between the model response z^l derived from LTM $\mathcal{M}_{long}$ and current response value $f_{\theta_t}(x, t)$

$$\mathcal{L}_{dkr} \triangleq \mathbb{E}_{\mathcal{B}_{long}} KL\left(z_i^l \parallel f_{\theta_t}(x_i^l, t)\right), \tag{11}$$

where $\mathcal{B}_{long} \overset{B}{\sim} \mathcal{M}_{long}$ and B denotes batch size. This term is also associated with soft targets (Hinton, Vinyals, & Dean, 2015), in which the labels are replaced by the soft targets generated by a trained model, and self-distillation (Zhang et al., 2019), in which a group of identical networks is used to perform inter-model distillations. However, the previous response value z_i^l is not used to surrogate the one-hot target vector y_i like soft target does, but to provide additional regularization. Also, although akin to self-distillation, we calculate the $\mathcal{L}_{dkr}$ value given current learner state f_θ using response z_i^l retrieve from f_θ's past selves rather than from external teachers. The **overall** objective for optimizing DUAL-MEM model with the dark knowledge regularizer $\mathcal{L}_{dkr}$ is to minimize the weighted sum of learning with dual memory $\mathcal{L}_{dual} \in \{\mathcal{L}_{DRL}, \mathcal{L}_{MRL}\}$ and $\mathcal{L}_{dkr}$:

$$\theta^* \leftarrow \arg\min_\theta \mathcal{L}_{dual} + \lambda_{dkr} \cdot \mathcal{L}_{dkr} \tag{12}$$

where λ_{dkr} is a hyper-parameter balancing the main objective and the dark knowledge regularizer, we summarize the whole optimization process of DUAL-MEM in Alg. 2.

Algorithm 2 Learning procedure of the DUAL-MEM model.

Input: θ, $\mathcal{D}_t$, $\mathcal{M}_{short}$, $\mathcal{M}_{long}$
Input: B, K ▷ batch size, iteration number
Input: η ▷ learning rate
Input: λ_{short}, λ_{dkr} ▷ weighting factors
for $t = 1$ to T **do**
  for $k = 1$ to K **do**
    $\mathcal{B}_{ob} \overset{B}{\sim} \mathcal{D}_t$, $\mathcal{B}_{short} \overset{B}{\sim} \mathcal{M}_{short}$, $\mathcal{B}_{long} \overset{B}{\sim} \mathcal{M}_{long}$
    $\ell_{ob} \leftarrow \ell(\mathcal{B}_{ob})$
    $\ell_{short} \leftarrow \ell(\mathcal{B}_{short})$
    $\mathcal{M}_{short} \leftarrow \texttt{ISMU}(\mathcal{M}_{short})$
    if using DRL **then**
      $\ell_{long} \leftarrow \ell(\mathcal{B}_{long})$
      $\mathcal{L}_{dual} \leftarrow \ell_{ob} + \ell_{long} + \lambda_{short} \cdot \ell_{short}$
    else if using MRL **then**
      $\tilde{\mathcal{B}}_{mem} \leftarrow \text{mixup}(\mathcal{B}_{short}, \mathcal{B}_{long})$
      $\mathcal{L}_{dual} \leftarrow \ell_{ob} + \ell(\tilde{\mathcal{B}}_{mem})$
    end if
    $\theta \leftarrow \theta - \eta \nabla_\theta(\mathcal{L}_{dual} + \lambda_{dkr} \cdot \mathcal{L}_{dkr})$
  end for
  consolidate $\mathcal{M}_{short}$ to $\mathcal{M}_{long}$
  re-initialize $\mathcal{M}_{short}$
end for
return θ

Note that, besides ground truth target vectors $y \in \mathcal{Y}$, we also save the response vectors $z \in \mathbb{R}^{1 \times |\mathcal{Y}|}$:

$$\begin{aligned} \mathcal{M}_{short} &= \{(x_i^s, y_i^s, z_i^s)\}, \\ \mathcal{M}_{long} &= \{(x_j^l, y_j^l, z_j^l, t)\}. \end{aligned} \tag{13}$$

This increases the memory overhead slightly by adding $|\mathcal{Y}| \times |\mathcal{M}_{long}|$ floating numbers, which is trivial when compared with input memories x from high dimensional observation space. Also, computing this regularizer does not incur significant additional computational and storage overhead as the model responses are already saved or calculated during the $\mathcal{B}_{long}$ rehearsal and the previous learner snapshots of million-param-net $\{f_{\theta_k}\}_{k=1}^{t-1}$ are not required.

5. Experiments

5.1. Evaluation benchmarks

We consider benchmarks of various complexities, and task lengths to evaluate the proposed lifelong learning methods. These benchmarks are constructed using multiple datasets. Details of the evaluation benchmarks are summarized in Table 1.

First, we use the **CIFAR-100** (Rebuffi et al., 2017) dataset that consists of 6×10^4 color images in 32×32 pixels, 100 classes. Each class contains 600 samples subdivided into two disjoint subsets (500 and 100 samples) for training and evaluation, respectively. To construct a balanced dataset, we assign 5/10/20 randomly chosen classes to each task in a task sequence (refer to Table 1). This setting allows us to evaluate different lifelong learning methods on a task sequence, ensuring all tasks are similar. The second dataset is **Tiny ImageNet** (Stanford, 2017). It contains 1×10^5 color images of 200 classes from the ILSVRC (Russakovsky et al., 2015); each class has 500 training and 50 evaluation samples. All images are down-sampled to 64×64. To construct benchmarks of different task lengths, we divide the original dataset into 20/10/5 disjoint tasks, and each task has 10/20/40 randomly-selected classes. Besides, we draw a 100-class subset from the original Tiny ImageNet dataset to construct

Table 1
Details about the CIFAR-100, Tiny ImageNet, and Mini ImageNet benchmarks considered in experiments.

Benchmark name	Data source	Dimension	#Tasks	Task length	Samples	
					Train	*Test*
CIFAR-100	CIFAR-100	$3 \times 32 \times 32$	20	5 classes	5×10^4	1×10^4
CIFAR-100 T10	CIFAR-100	$3 \times 32 \times 32$	10	10 classes	5×10^4	1×10^4
CIFAR-100 T5	CIFAR-100	$3 \times 32 \times 32$	5	20 classes	5×10^4	1×10^4
Mini ImageNet-100	Mini ImageNet	$3 \times 84 \times 84$	20	5 classes	5×10^4	1×10^4
Tiny ImageNet-100	Tiny ImageNet	$3 \times 64 \times 64$	20	5 classes	5×10^4	5×10^3
Tiny ImageNet-200	Tiny ImageNet	$3 \times 64 \times 64$	20	10 classes	1×10^5	1×10^4
Tiny ImageNet-200 T10	Tiny ImageNet	$3 \times 64 \times 64$	10	20 classes	1×10^5	1×10^4
Tiny ImageNet-200 T5	Tiny ImageNet	$3 \times 64 \times 64$	5	40 classes	1×10^5	1×10^4

a simpler 20-task benchmark (Tiny ImageNet-100). Thirdly, we adopt **Mini ImageNet** dataset (Vinyals, Blundell, Lillicrap, Wierstra, et al., 2016), which is a down-sampled subset from the original ImageNet with a total of 100 classes and 600 images per class. We use Mini ImageNet by splitting it into 20 disjoint tasks.

5.2. Experimental settings

We adopt identical slim ResNet-18 architectures (Lopez-Paz & Ranzato, 2017) for all methods to be evaluated. By default, all models are optimized by error back-propagation (Rumelhart, Hinton, & Williams, 1986) from scratch under the experience-once task-incremental lifelong learning setup (refer to Section 3.1) using training sets and evaluated on corresponding test sets. We find optimal hyper-parameters for all methods by grid search and report their best scores.

We denote Dual-MEM with Direct Retrieval Learning, Mixup Retrieval Learning as Dual-MEM-DR, Dual-MEM-MR, respectively. In experiments on Tiny-ImageNet and Mini-ImageNet, we utilize a batch size of 10 for Dual-MEM-DR and Dual-MEM-MR methods. When conducting CIFAR-100 experiments, we decrease the batch size to 5 to ensure sufficient update iterations, since each sample is only observed once. We employ fixed learning rates for all Dual-MEM experiments. Specifically, learning rate $\eta = 0.01$ for CIFAR-100, Tiny ImageNet-100 and Tiny ImageNet-200 benchmarks and $\eta = 0.005$ for Mini ImageNet-100. For Dual-MEM-DR method, we adopt $\lambda_{short} = 5 \times 10^{-4}$ for all experiments and $\lambda_{dkr} = 0.3$ on CIFAR-100 and Tiny ImageNet-100, $\lambda_{dkr} = 0.7$ for Tiny ImageNet-200 and Mini ImageNet-100 benchmarks. For Dual-MEM-MR, we adopt $\lambda_{dkr} = 1$ for all benchmarks.

5.3. Evaluation metrics

We measure the lifelong learning performance using *average accuracy*, *backward/forward transfer* (Lopez-Paz & Ranzato, 2017). Acc measures the average accuracy across all tasks while the Bwt quantifies the average relative performance drop (< 0, *i.e.*, forgetting) or gain (> 0). The value of Bwt is related to the performance of the model tested just after a task is learned and depends on the accuracy metric. Forward transfer (Fwt) quantifies the influence that learning a task t has on the performance of a future task; in particular, positive Fwt is possible when the model can exploit the task structure or general knowledge across tasks.

Let $A_{T,t}$ be the mean accuracy of task t after the learning of T tasks, the *average accuracy*, *backward transfer* and *forward transfer* are defined as

$$\begin{aligned} \text{Acc} &= \frac{1}{T}\sum_{t=1}^{T} A_{T,t}, \\ \text{Bwt} &= \frac{1}{T-1}\sum_{t=1}^{T-1} A_{T,t} - A_{t,t}, \ \text{Fwt} = \frac{1}{T-1}\sum_{t-2}^{T} A_{t-1,t} - b_t, \end{aligned} \tag{14}$$

where b_t denotes test accuracy at random initialization. The larger these metrics, the better the lifelong learning method is. If two approaches can achieve similar Acc, the one with larger Bwt and Fwt is preferable.

5.4. Baselines

We compare Dual-MEM with following alternatives:

- Naïve Fine-tune, continually fine-tune the model on the task sequence. This method starts from the previous task parameters to optimize for the current task.
- Gem (Lopez-Paz & Ranzato, 2017) is a memory-using regularization-based lifelong learning method. It achieves new task updates with little interference to previous tasks by projecting the estimated gradient direction on a feasible region outlined by previous task gradients.
- iCaRL (Rebuffi et al., 2017) uses a nearest-mean-of-exemplar (NME) strategy for classification and avoids catastrophic forgetting by regularizing in the feature space via knowledge distillation loss.
- PODNet-NME (Douillard, Cord, Ollion, Robert, & Valle, 2020) applies temporal model multi-dimensional distillation by snapshotting the previous model to alleviate forgetting.
- ER-Reservoir (Chaudhry, Rohrbach, et al., 2019) is an exact-rehearsal method with reservoir sampling memory updating.
- Bir (van de Ven et al., 2020) implements conditional internal generative replay through a VAE generator, and for every task to be learned, a different subset of neurons in each layer is inhibited during the generative backward pass.
- DER, DER++ (Boschini, Bonicelli, Buzzega, Porrello, & Calderara, 2022; Buzzega, Boschini, Porrello, Abati, & Calderara, 2020) are dark-experience replay-based approaches that combine direct rehearsal and knowledge distillation.
- HAL (Chaudhry et al., 2021) complement regular experience replay with a new anchoring objective. The learner employs bi-level optimization to update new information while maintaining predictions on anchor points of past tasks.
- RPC (Pernici, Bruni, Baecchi, Turchini, & Del Bimbo, 2021) builds a classifier with several pre-allocated output nodes that are subject to the classification loss right from the beginning. The nodes are calculated using an N-simplex.

The proposed approaches to be evaluated are:

- Dual-MEM-DR uses the Dual-MEM model architecture, consisting of the main network and a dual memory system. The learning strategy of Dual-MEM-DR is Direct Retrieval learning.
- Dual-MEM-MR has the same architecture setting as Dual-MEM-DR and is optimized by the Mixup Retrieval learning.

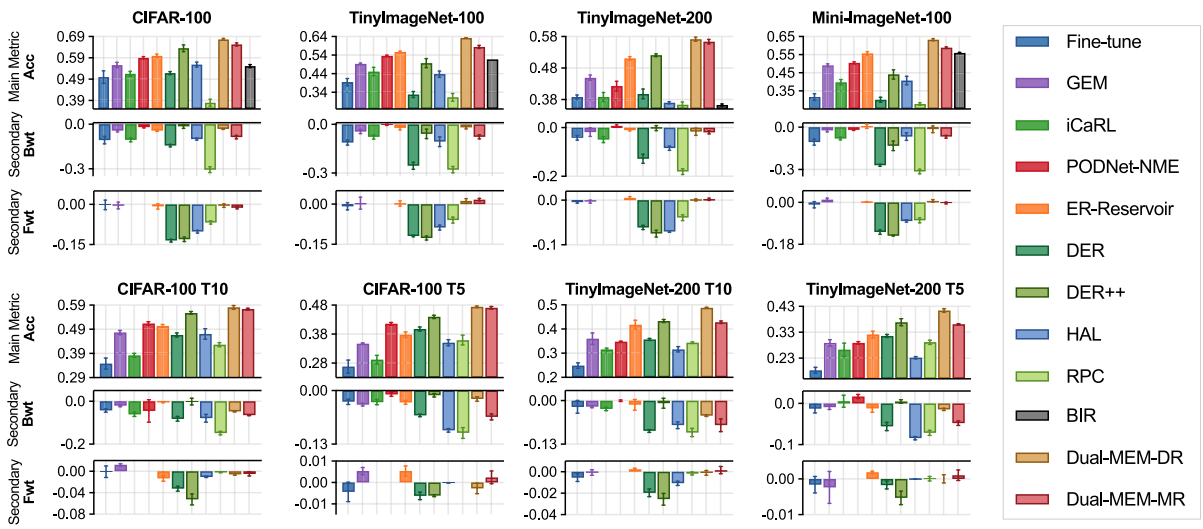

Fig. 4. Acc, Bwt and Fwt results of comparison experiments on eight benchmarks. The major metric average accuracy Acc measures the final performance while the minor metric Bwt/Fwt quantifies the relative knowledge transfer during learning. Methods that assign independent parameter groups for different tasks report zero-Bwt here. The results (mean and SD) are measured across three independent trials, and we adopt $|\mathcal{M}|_{\text{per cls}} = 5$ as default for all memory-using methods. All methods except Bir are trained using the default one-epoch setting.

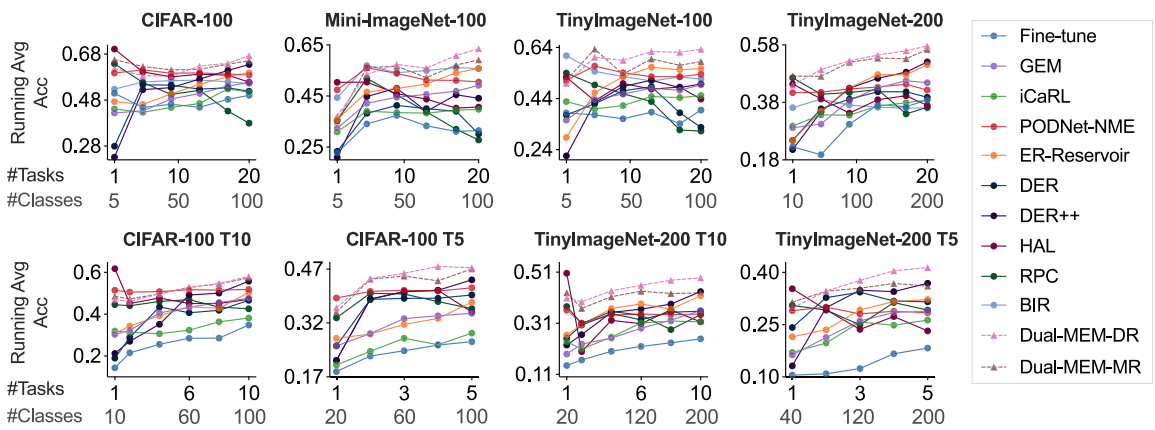

Fig. 5. Acc changes during lifelong learning. This figure presents running average Acc dynamics of twelve approaches on eight benchmarks under the default one-epoch task-incremental setting, Accs reported here are averaged Acc on test data based on all experienced tasks so far. We adopt $|\mathcal{M}|_{\text{per cls}} = 5$ for memory-using methods.

5.5. Comparison with baselines

We first compare the performance of the proposed Dual-MEM methods with alternatives. For this, we report the performance results on various benchmarks; all methods are trained under the one-epoch (experience-once) task-incremental lifelong learning setup except Bir. For the Bir method, which includes a generator network that needs more iterations to produce ideal hidden representations, we adopt the training hyper-parameters following the original implementation (van de Ven et al., 2020). For memory-using methods, the memory budgets are normalized to $|\mathcal{M}|_{\text{per cls}}$ uniformly, which is defined as

$$|\mathcal{M}|_{\text{per cls}} = \frac{|\mathcal{M}|_{\text{per task}}}{N_{\text{class per task}}}$$

for a fair comparison. Besides final metric results, we show the running average Acc results during the learning process in Fig. 5 to compare the performance dynamics of different methods more intuitively.

Overall results. In the task-incremental learning scenario, most lifelong learning methods are successful. Fig. 4 presents the evaluation results for the comparison experiments. Especially on all benchmarks, Dual-MEM approaches achieve much improved Acc while reporting low negative Bwt (minimal forgetting). We evaluate the performance at different task sizes using benchmarks from different divisions of the same data (*e.g.*, CIFAR-100, and CIFAR-100 T10/T5). According to results in Fig. 4, the Fwts of most approaches improve when the number of classes in a single task increases, demonstrating that the learner accumulates more general knowledge in single-task learning. At the same time, increasing the number of single tasks exacerbates forgetting, and the Dual-MEM approaches still achieve high performance on these challenging benchmarks, demonstrating consistent superiority over baselines.

Comparison with other replay-based approaches. Although the exact replay method ER succeeds in mitigating forgetting, its implementation of memory and replay is relatively coarse and cannot fully exploit the power of memory rehearsal. This problem might be solved by introducing a generative model (Wu et al., 2018) so that the more sophisticated conditional replay can be used. However, the concern here is that training such generative models is also a challenging problem. In particular, the BIR (Brain-inspired Replay) method, although trained with more iterations, is still lagging behind some competitors. The results indicated

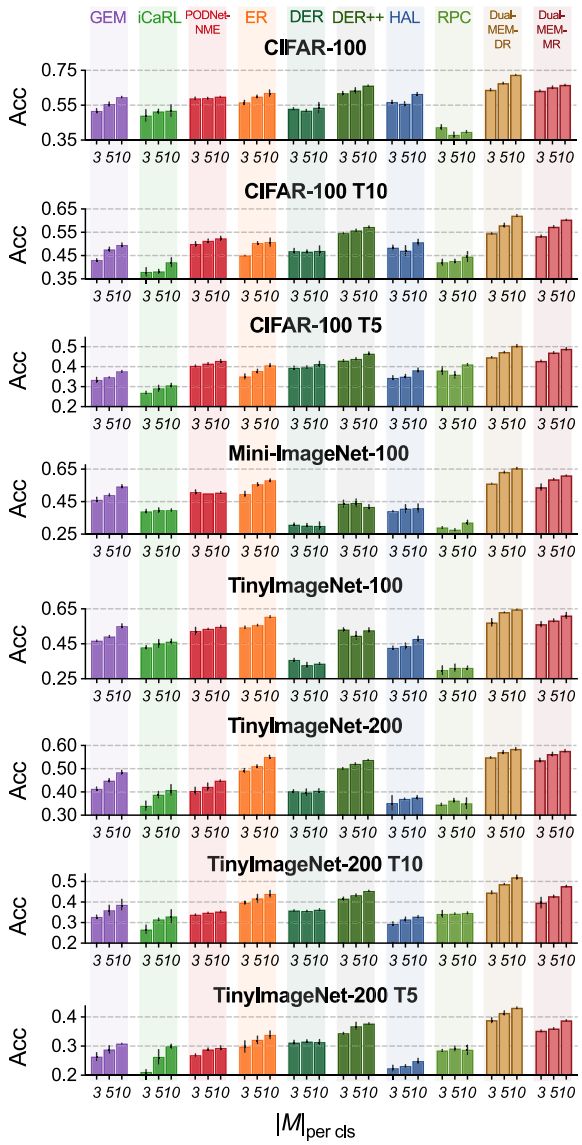

Fig. 6. Acc results under different memory budgets. We presents the Acc performances (mean and SD) of memory-using methods under $|\mathcal{M}|_{\text{per cls}} \in \{3, 5, 10\}$ across three independent trials.

that the Dual-MEM-DR's performance is more consistent and always superior to that of the competing estimators. Dual-MEM-DR obtains 11.0%, 12.8%, 11.7%, 8.4% higher Acc relative to the best counterparts on CIFAR-100, Tiny ImageNet-100/200 and Mini ImageNet-100, respectively. Furthermore, Dual-MEM-DR reduces the relative forgetting (indicated by negative Bwt) by 70.2%, 82.9%, 84.0%, 89.8% compared with Fine-tune on CIFAR-100, Tiny ImageNet-100/200 and Mini ImageNet-100, respectively. Dual-MEM model combined with MRL has a relatively high BWT despite its strong performance in the accuracy metric. The MRL uses a sampling-based linear interpolation to implement the retrieval procedure and therefore does not guarantee the exact rehearsal of samples from previous tasks as the DRL strategy does. Thus, the mixup retrieval strategy enhances the learning efficiency of the current task but deteriorates the previous tasks' performance. As a piece of evidence, it is shown in Fig. 5 that in the first few tasks of learning, Dual-MEM-MR achieves higher average Accs compared to Dual-MEM-DR method. Our experimental results show that the Dual-MEM method has significant advantages over other replay-based methods.

Dual-MEM *methods exhibits high learning efficiency.* For small-scale benchmarks (such as those based on CIFAR-100), the learner can achieve high performance after only a few tasks in the training set. On the other hand, for more complicated benchmarks, such as the Tiny ImageNet-based ones, the learner requires ongoing general knowledge accumulation, causing the moving average Acc to show a continually rising trend. As a result, the complex learning process involves combatting against forgetting and gathering general information. PODNet and Bir demonstrate stable performance during the task-incremental learning process, mainly attributed to their parameter-isolation strategies. This also leads to good Bwt performance as shown in Fig. 4. However, the task-specific parameter strategy hinders the inter-task general knowledge accumulation, resulting in trivial Acc improvements after learning many tasks (the slopes of these curves are small). Furthermore, we note that Dual-MEM-DR/MR can achieve good performance after learning a few tasks, and the performance improves as new tasks are added continuously, demonstrating Dual-MEM methods yield high learning efficiency.

5.6. Performance under different memory budgets

We further evaluate the performance of memory-using approaches across various memory sizes. Although replaying memorized data alleviates forgetting, the concern is that its exact effect will be closely related to the budget of the memory, which should be tiny in practical scenarios. Therefore, we expect lifelong learning methods to achieve competitive performance using tiny memory budgets. Apparently, a naive implementation of the exact rehearsal method requires a higher memory budget to achieve similar performance. Since the methods to be evaluated employ different memory-allocating strategies (store-by-class or store-by-task), we uniformly denote the memory size by category for a fair comparison.

Fig. 6 presents the Acc performance under different memory budgets. We note that the proposed Dual-MEM-DR/MR method substantially improves memory using efficiency over its competitors under all memory settings on evaluation benchmarks. We also find that the performance gains of increasing memory size for Dual-MEM-MR are less significant than Dual-MEM-DR. We conjecture it is the consequence of the additional stochasticity introduced by the sampling operation, which might hinder the model from directly benefiting from a larger memory buffer. In the Dual-MEM model, we use brain-inspired modifications to improve the straightforward use of memories in naive exact replay methods, which should enable more efficient memory utilization. Further control experiments confirmed this intuition, in which we trained all memory-using methods under several memory budget settings. In particular, we observe that Dual-MEM-DR with a small memory budget ($|\mathcal{M}|_{\text{per cls}} = 3$) outperforms most competitors with larger memory budgets ($|\mathcal{M}|_{\text{per cls}} = 5, 10$).

Specifically, the memories obtained utilizing the proposed ISMU method might better depict the task knowledge structure. We compare an experience replay baseline with and without the ISMU updating. The results listed in Table 2 show that the one that uses ISMU to update the memory demonstrates consistent superiority on all evaluation benchmarks. We present t-SNE (Van der Maaten & Hinton, 2008) plots in Fig. 7 to visualize the samples and memories in the learned feature space, revealing that the proposed updating strategy can better preserve the task knowledge structure, which benefits subsequent forgetting alleviation through rehearsal.

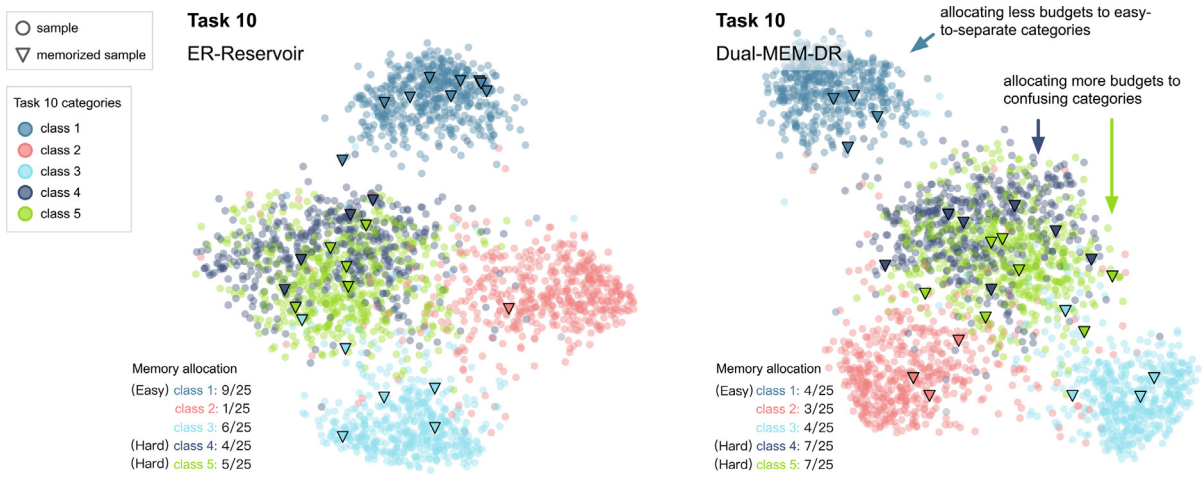

Fig. 7. t-SNE visualization of sample points in the feature space of a task from the CIFAR-100 benchmark, obtained using ER-Reservoir, Dual-MEM-DR approaches with $|\mathcal{M}|_{\text{per cls}} = 5$. The plots show that, in the feature space, the ISMU updating strategy used in Dual-MEM methods tends to memorize samples close to decision boundaries (the edge of the cluster of a class). In comparison, the distribution of the samples memorized by reservoir sampling looks random. Also, the ISMU tends to allocate more memory budgets for challenging classes to preserve this task's knowledge structure better.

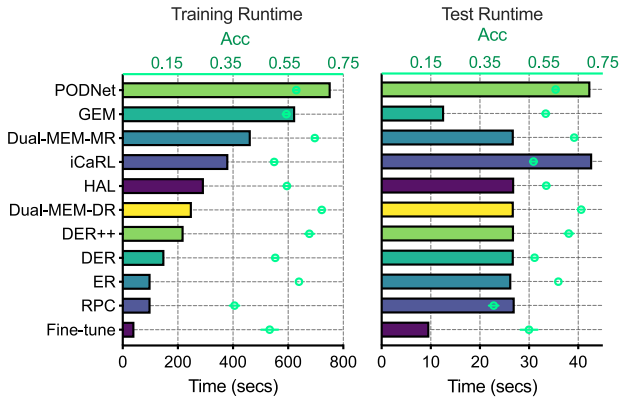

Fig. 8. Results of running cost evaluations on the CIFAR-100 benchmark under the default experience-once task-incremental setting.

Table 2
The proposed ISMU update strategy presents superior performance against reservoir updating. We conduct experiments to examine the effectiveness of the ISMU updating strategy by using an experience-replay baseline ($|\mathcal{M}|_{\text{per cls}} = 5$), results reported are Acc mean±SD obtained across three trials.

Benchmark	w/ ISMU	w/o ISMU
CIFAR-100	**0.638±0.006**	0.599±0.008
Tiny ImageNet-100	**0.584±0.007**	0.558±0.004
Tiny ImageNet-200	**0.516±0.006**	0.511±0.006
Mini ImageNet-100	**0.577±0.006**	0.561±0.010
CIFAR-100 T10	**0.545±0.005**	0.514±0.006
CIFAR-100 T5	**0.440±0.008**	0.388±0.011
Tiny ImageNet-200 T10	**0.427±0.003**	0.418±0.019
Tiny ImageNet-200 T5	**0.364±0.007**	0.322±0.013

5.7. Model complexity and computational cost

We further evaluate the lifelong learning methods in Section 5.6 (CIFAR-100, $|\mathcal{M}|_{\text{per cls}} = 5$) regarding memory cost and time complexities. The training and testing runtimes are recorded for each method. All results are reported using a platform with an Nvidia RTX 3090 GPU, an Intel 10400 CPU, and 128 GB RAM. The results in Fig. 8 demonstrate that while some methods are computationally daunting, Dual-MEM methods have low running costs while achieving high performance. Our evaluation results suggest that the rehearsal-based methods can be enhanced by improving memory management strategies and rehearsal mechanisms. In general, Dual-MEM approaches are in line with other rehearsal-based methods such as DER (DER++) and ER and are much less demanding than methods using additional past models (HAL, PODNet). Furthermore, we find that Dual-MEM methods significantly improve performance while only marginally increasing computation costs (both time and memory) when compared to the exact-replay baseline ER-Reservoir. On the contrary, although the PODNet approach yields competitive performance according to evaluation, it requires far more computational and memory consumption. Table 4 shows the additional space costs. Typically, the additional space cost of saving labels and model responses is minor since their dimensionality is significantly lower than that of observations. In contrast, the memory overhead introduced by keeping model snapshots is usually significant, even if we choose a simple backbone for the learner. However, methods that leverage additional model snapshots do not yield the expected performance in the test, suggesting that the strategies for utilizing learners' historical parameters may need to be improved further.

5.8. Ablation study

We study the effectiveness of different modules to further reveal the nature of Dual-MEM approaches on CIFAR-100, Tiny ImageNet-100, Tiny ImageNet-200, and Mini ImageNet-100 benchmarks. Specifically, for Dual-MEM-DR, we consider (a) the proposed STM updating strategy, Incremental Surprise Maximization Updating (ISMU), (b) STM rehearsal in the Direct Retrieval learning strategy, and (c) additional regularization using response dark knowledge, implemented by the term $\mathcal{L}_{dkr}$. The $\mathcal{M}_{short}$ is updated via reservoir policy to make a comparison with the ISMU. Similarly, for Dual-MEM-MR, we consider (a) ISMU and (c) additional regularization by the response information. We report results in Table 3. According to the results, the proposed ISMU STM updating rule helps maintain an effective task memory to alleviate forgetting during learning and contributes to the overall performance on all benchmarks. In particular, the role of ISMU in Dual-MEM-DR is more pronounced than that of Dual-MEM-MR as the mixup retrieval strategy introduces additional randomness in the rehearsal process via linear interpolation, which weakens the effect of the ISMU strategy. The resampling of high-surprise samples through $\mathcal{M}_{short}$ replay could help the learner in Dual-MEM-DR to fit the class boundaries in the feature space, leading

Table 3
Effectiveness study of (a) the proposed STM updating strategy, Incremental Surprise Maximization Updating (ISMU), (b) STM rehearsal in the Direct Retrieval learning strategy, and (c) additional regularization using response dark knowledge, implemented by the term $\mathcal{L}_{dkr}$. We report results by Acc (Bwt) across three independent trials, and denote the default Dual-MEM approaches as Full here and note the best ones with **bold**.

	Dual-MEM-DR			
	Full	w/o (a)	w/o (b)	w/o (c)
CIFAR-100	**0.6789 (−0.0318)**	0.6563 (−0.0400)	0.6689 (−0.0576)	0.6642 (−0.0738)
Tiny ImageNet-100	**0.6322 (−0.0154)**	0.6208 (−0.0207)	0.6283 (−0.0175)	0.6177 (−0.0236)
Tiny ImageNet-200	**0.5741 (−0.0084)**	0.5640 (−0.0452)	0.5614 (−0.0442)	0.5341 (−0.0604)
Mini ImageNet-100	**0.6389 (−0.0139)**	0.5993 (−0.0186)	0.6345 (0.0142)	0.5329 (−0.0765)
	Dual-MEM-MR			
CIFAR-100	**0.6532 (−0.0859)**	0.6144 (−0.1144)	–	0.6442 (−0.0862)
Tiny ImageNet-100	**0.5847 (−0.0764)**	0.5763 (−0.0648)	–	0.5412 (−0.1232)
Tiny ImageNet-200	**0.5638 (−0.0192)**	0.5562 (−0.0282)	–	0.5313 (−0.0530)
Mini ImageNet-100	**0.5888 (−0.0667)**	0.5873 (−0.0695)	–	0.5412 (−0.1183)

Table 4
Additional memory requirements (other than the main model f_θ) of lifelong learning approaches. We use $|a|$ to denote the memory consumption for saving variable a and denote the memory size (number of memories) by N.

Method	x, y	z	θ_{past}	Total
Finetune				-
GEM	✓			$N(\|x\| + \|y\|)$
iCaRL	✓			$N(\|x\| + \|y\|)$
PODNet	✓		✓	$N(\|x\| + \|y\|) + \|\theta\|$
ER	✓			$N(\|x\| + \|y\|)$
DER	✓			$N(\|x\| + \|y\|)$
DER++	✓	✓		$N(\|x\| + \|y\| + \|z\|)$
HAL	✓		✓	$N(\|x\| + \|y\|) + \|\theta\|$
RPC	✓			$N(\|x\| + \|y\|)$
DualMEM	✓	✓		$N(\|x\| + \|y\| + \|z\|)$

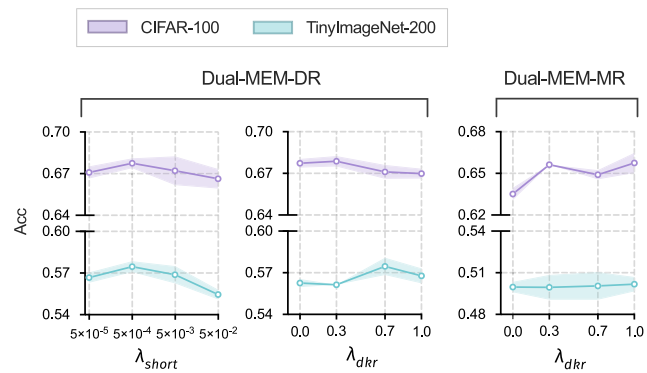

Fig. 9. Sensitivity analysis of Dual-MEM approaches to parameter λ_{short}, λ_{dkr} on CIFAR-100 and Tiny ImageNet-200 benchmarks.

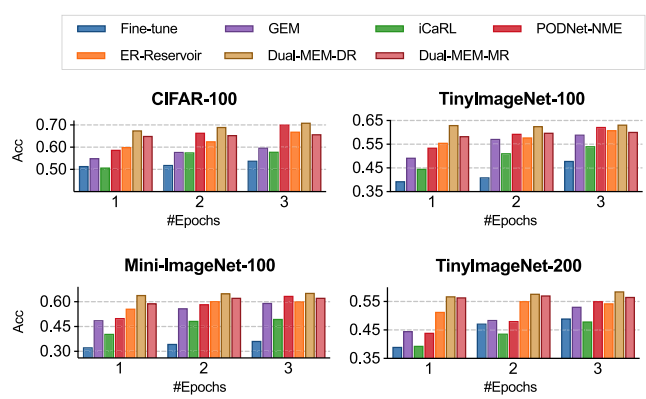

Fig. 10. Acc under multi-epoch learning settings. We present the Acc performances under one/two/three-epoch training settings. Dual-MEM trained under the one-epoch setting still has performance advantages over competitors trained with multi-epoch settings.

to better performance on all benchmarks. We also note that the additional response regularization $\mathcal{L}_{dkr}$ enhances both Dual-MEM methods on all benchmarks. The introduced $\mathcal{L}_{dkr}$ term provides weak supervision to the learner for maintaining the ability to perform old tasks and results in significantly lower relative forgetting (measured by Bwt).

5.9. Sensitivity analysis of hyper-parameters

In Fig. 9, we investigate how the choice of different values of balancing factors influences the final performance on CIFAR-100 and TinyImageNet-200 benchmarks. The results highlight that the Dual-MEM-MR approach is quite robust to λ_{dkr}. For the Dual-MEM-DR method, λ_{dkr} is stable enough but benefits from higher values for complex benchmarks and lower values for simple benchmarks. Also, λ_{short} (the weighting factor of short memory rehearsal) must be taken relatively low enough not to over-constrain the model when learning a new task.

5.10. Learning efficiency study under multi-epoch settings

To further assess the performance dynamics of different lifelong learning approaches across extended learning periods, we relax the default experience-once setup and expand the number of epochs on each task. We consider all approaches except Bir, which is trained until convergence by default. We expect the model to be competitive with fewer traversals for the same benchmark. Generally, the one epoch setting is insufficient to guarantee convergence, so the increase in the number of epochs will bring some performance improvement. As shown in Fig. 10, as the number of epochs on each task increases, the performance of all learners improves to varying degrees, attributing to more optimization iterations. However, we note that the proposed Dual-MEM-DR under the one-epoch setting still outperforms all competitors trained with multi-epoch settings, suggesting that Dual-MEM has substantially superior learning efficiency.

5.11. Comparison with representative dual-memory-design approaches

We also conducted comparison experiments with representative high-performance methods (IL2M (Belouadah & Popescu, 2019), SCoMMER (Fahad Sarfraz, 2023)) that adopt a dual memory design using the CIFAR-100 T10 and T5 benchmarks. As mentioned in the related works, the proposed dual-memory approach has obvious differences from these methods. Compared to previous methods, this work focuses more on borrowing brain memory mechanisms to improve the conventional exact replay methods to achieve competitive performance with a minor additional

Table 5
Comparisons with representative lifelong learning approaches that adopt "dual memory" design. Results reported are Acc mean±SD obtained across five trials.

	CIFAR-100 T10		
$\mathcal{M}_{\text{per cls}}$	3	5	10
Dual-MEM-DR	0.547±0.004	0.581±0.008	0.622±0.003
Dual-MEM-MR	0.533±0.005	0.572±0.004	0.603±0.002
IL2M	0.419±0.004	0.423±0.006	0.440±0.012
SCoMMER	0.475±0.011	0.514±0.013	0.558±0.008
	CIFAR-100 T5		
Dual-MEM-DR	0.448±0.002	0.474±0.004	0.504±0.006
Dual-MEM-MR	0.430±0.003	0.471±0.005	0.489±0.006
IL2M	0.383±0.011	0.394±0.007	0.418±0.005
SCoMMER	0.334±0.007	0.413±0.008	0.445±0.010

computational burden. Table 5 shows the accuracy results of comparisons. It can be seen that the two proposed methods (Dual-MEM-DR/MR) consistently demonstrate better performance than competitors under different memory buffer budgets.

6. Discussion

Aside from leveraging insights from neuroscience to improve memory-based task-incremental learning, another aim of this work was to explore new perspectives and hypotheses about the computational role and potential implementations of memory in lifelong learning in the brain. Regarding memory's role in lifelong learning, this study considered a more complex memory model, the dual-store memory model (Atkinson & Shiffrin, 1968), as a foundation. Although the extensively-studied hippocampal–prefrontal and hippocampal–neocortex interactions provide insights that the reactivation of memorized patterns could prevent forgetting and strength memory, it has remained an open question whether mechanisms can be transplanted to solve real recognition problems, which has been partially answered in this work. Regarding memory's function, our study highlights critical computational roles for more intricate memory using strategies in incrementally learning new tasks. Specifically, we use neuroscience-inspired modifications to refine and improve the exact memory rehearsal approach and experimentally demonstrate that these modifications can significantly boost learning efficiency and improve overall performance. Finally, we should note that some important aspects are missing from our study. One of them is the regularization in the cellular parameter space, in which the challenge is to assign credit to the synaptic weights of the network correctly and effectively. While the replay mechanisms restrict the activation over the function level, the cellular level constraints might provide beneficial complements and further protect the memories.

More recently, the continual learning community has increasingly focused on more general settings, among which the task-free setting (Aljundi, Kelchtermans, & Tuytelaars, 2019; Shin et al., 2017) is considered to be the ultimate form of lifelong learning that can adaptively adapt to non-stationary continuous data streams and is a promising for building artificial agents. The challenge of this setting is that the model no longer has a clear task boundary or contextual information during the inference process, which brings new challenges to the design of more general continuous learning methods (van de Ven, Tuytelaars, & Tolias, 2022). Some previous works (Gautam et al., 2022; Lee et al., 2020) have proposed novel solutions to this problem. However, it should be noted that although it is ideal for generalizing the task-incremental setting to a task-free setting, it is not feasible in all cases (Tadros, Krishnan, Ramyaa, & Bazhenov, 2022; van de Ven et al., 2020). In some tasks, we need to use contextual information to explore memory dynamics and forward and backward transfer during the entire learning process (van de Ven et al., 2022). This positive forward or backward transfer between contexts or tasks is necessary, and achieving such a positive transfer is not trivial (Lopez-Paz & Ranzato, 2017; van de Ven et al., 2022). In the next research stage, we will conduct research based on this work and strive to build a more flexible memory utilization paradigm and move further towards a task-free form.

At the same time, with the advancement of machine learning technology, lifelong learning may further expand in scale in the future. Due to its small additional computational cost, the method proposed in this work will have a great opportunity to play an important role in larger-scale tasks and provide a reference for the design of replay-based methods for larger tasks.

7. Conclusion

This work proposes the Dual-MEM model by emulating the brain PFC-HC memory system. Dual-MEM uses ISMU to update short-term memory and consolidates short-term memories to long-term ones at the end of the current task for future rehearsal. We introduce DRL and MRL learning rules and the dark knowledge regularizer around the dual memory to update the model parameters balancing stability and plasticity. Comprehensive experimental results demonstrate that Dual-MEM yields competitive performance while exhibiting high learning and running efficiencies.

Declaration of competing interest

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.

Data availability

Data will be made available on request.

References

Abel, D., Jinnai, Y., Guo, S. Y., Konidaris, G., & Littman, M. (2018). Policy and value transfer in lifelong reinforcement learning. In *Proceedings of international conference on machine learning* (pp. 20–29). PMLR.

Ahn, H., Cha, S., Lee, D., & Moon, T. (2019). Uncertainty-based continual learning with adaptive regularization. In *Advances in neural information processing systems*.

Aljundi, R., Babiloni, F., Elhoseiny, M., Rohrbach, M., & Tuytelaars, T. (2018). Memory aware synapses: Learning what (not) to forget. In *Proceedings of European conference on computer vision*.

Aljundi, R., Chakravarty, P., & Tuytelaars, T. (2017). Expert gate: Lifelong learning with a network of experts. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Aljundi, R., Kelchtermans, K., & Tuytelaars, T. (2019). Task-free continual learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Atkinson, R. C., & Shiffrin, R. M. (1968). *Human memory: A proposed system and its control processes*. Elsevier.

Belouadah, E., & Popescu, A. (2019). IL2m: Class incremental learning with dual memory. In *Proceedings of international conference on computer vision* (pp. 583–592).

Boschini, M., Bonicelli, L., Buzzega, P., Porrello, A., & Calderara, S. (2022). Class-incremental continual learning into the extended DER-verse. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Buzzega, P., Boschini, M., Porrello, A., Abati, D., & Calderara, S. (2020). Dark experience for general continual learning: a strong, simple baseline. In *Advances in neural information processing systems, vol. 33* (pp. 15920–15930).

Chaudhry, A., Gordo, A., Dokania, P. K., Torr, P., & Lopez-Paz, D. (2021). Using hindsight to anchor past knowledge in continual learning. In *Proceedings of AAAI conference on artificial intelligence*.

Chaudhry, A., Marc'Aurelio, R., Rohrbach, M., & Elhoseiny, M. (2019). Efficient lifelong learning with A-GEM. In *Proceedings of international conference on learning representations*.

Chaudhry, A., Rohrbach, M., Elhoseiny, M., Ajanthan, T., Dokania, P. K., Torr, P. H. S., et al. (2019). Continual learning with tiny episodic memories. In *Proceedings of international conference on machine learning*.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357.

Chen, J., Cook, P. A., & Wagner, A. D. (2015). Prediction strength modulates responses in human area CA1 to sequence violations. *Journal of Neurophysiology*, *114*(2), 1227–1238.

Chen, P.-C., Niknazar, H., Alaynick, W. A., Whitehurst, L. N., & Mednick, S. C. (2021). Competitive dynamics underlie cognitive improvements during sleep. *Proceedings of the National Academy of Sciences*, *118*(51).

DeVries, T., & Taylor, G. W. (2017). Dataset augmentation in feature space. arXiv.

Dong, J., Wang, L., Fang, Z., Sun, G., Xu, S., Wang, X., et al. (2022). Federated class-incremental learning. In *IEEE/CVF conference on computer vision and pattern recognition*.

Douillard, A., Cord, M., Ollion, C., Robert, T., & Valle, E. (2020). Podnet: Pooled outputs distillation for small-tasks incremental learning. In *Proceedings of European conference on computer vision*.

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC Press.

Elahe Arani, B. Z. (2022). Learning fast, learning slow: A general continual learning method based on complementary learning system. In *Proceedings of international conference on learning representations*.

Fahad Sarfraz, B. Z. (2023). Sparse coding in a dual memory system for lifelong learning. In *Proceedings of AAAI conference on artificial intelligence*.

Felzenszwalb, P., McAllester, D., & Ramanan, D. (2008). A discriminatively trained, multiscale, deformable part model. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1–8). Ieee.

French, R. M. (1999). Catastrophic forgetting in connectionist networks. *Trends in Cognitive Sciences*, *3*(4), 128–135.

Friston, K. (2003). Learning and inference in the brain. *Neural Networks*, *16*(9), 1325–1352.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, *61*(2), 331–349.

Gais, S., Albouy, G., Boly, M., Dang-Vu, T. T., Darsaud, A., Desseilles, M., et al. (2007). Sleep transforms the cerebral trace of declarative memories. *Proceedings of the National Academy of Sciences*, *104*(47), 18778–18783.

Gautam, C., Parameswaran, S., Mishra, A., & Sundaram, S. (2022). Tf-GCZSL: Task-free generalized continual zero-shot learning. *Neural Networks*, *155*, 487–497.

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. *Neuron*, *95*(2), 245–258.

Hayes, T. L., Cahill, N. D., & Kanan, C. (2019). Memory efficient experience replay for streaming learning. In *International conference on robotics and automation* (pp. 9769–9776). IEEE.

Hinton, G., Vinyals, O., & Dean, J. (2014). Dark knowledge 2(2). Presented as the Keynote in BayLearn.

Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. arXiv.

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295–1306.

James, W., Burkhardt, F., Bowers, F., & Skrupskelis, I. K. (1890). *The principles of psychology, vol. 1, no. 2*. Macmillan London.

Ji, D., & Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience*, *10*(1), 100–107.

Jung, H., Ju, J., Jung, M., & Kim, J. (2016). Less-forgetting learning in deep neural networks. arXiv.

Kamra, N., Gupta, U., & Liu, Y. (2018). Deep generative dual memory network for continual learning. ArXiv.

Karlsson, M. P., & Frank, L. M. (2009). Awake replay of remote experiences in the hippocampus. *Nature Neuroscience*, *12*(7), 913–918.

Kemker, R., & Kanan, C. (2018). FearNet: Brain-inspired model for incremental learning. In *Proceedings of international conference on learning representations*.

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., et al. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, *114*(13), 3521–3526.

Konidaris, G., & Barto, A. (2006). Autonomous shaping: Knowledge transfer in reinforcement learning. In *Proceedings of international conference on machine learning* (pp. 489–496).

Kudithipudi, D., Aguilar-Simon, M., Babb, J., Bazhenov, M., Blackiston, D., Bongard, J., et al. (2022). Biological underpinnings for lifelong learning machines. *Nature Machine Intelligence*, *4*(3), 196–210.

Lavda, F., Ramapuram, J., Gregorova, M., & Kalousis, A. (2018). Continual classification learning using generative model. In *Advances in neural information processing systems*.

Lee, S., Ha, J., Zhang, D., & Kim, G. (2020). A neural Dirichlet process mixture model for task-free continual learning. In *International conference on learning representations*.

Li, Z., & Hoiem, D. (2017). Learning without forgetting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(12), 2935–2947.

Lopez-Paz, D., & Ranzato, M. (2017). Gradient episodic memory for continual learning. In *Advances in neural information processing systems*.

Mallya, A., & Lazebnik, S. (2018). Packnet: Adding multiple tasks to a single network by iterative pruning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Masse, N. Y., Grant, G. D., & Freedman, D. J. (2018). Alleviating catastrophic forgetting using context-dependent gating and synaptic stabilization. *Proceedings of the National Academy of Sciences*, *115*(44), E10467–E10475.

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of learning and motivation*. Elsevier.

Mednick, S. C., McDevitt, E. A., Walsh, J. K., Wamsley, E., Paulus, M., Kanady, J. C., et al. (2013). The critical role of sleep spindles in hippocampal-dependent memory: a pharmacology study. *Journal of Neuroscience*, *33*(10), 4494–4504.

Ngo, H.-V., Fell, J., & Staresina, B. (2020). Sleep spindles mediate hippocampal-neocortical coupling during long-duration ripples. *ELife*, *9*, Article e57011.

Nguyen, C. V., Li, Y., Bui, T. D., & Turner, R. E. (2018). Variational continual learning. In *Proceedings of international conference on learning representations*.

Paller, K. A. (2009). Memory consolidation: systems. *Encyclopedia of Neuroscience*, *1*, 741–749.

Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. *Neural Networks*, *113*, 54–71.

Pernici, F., Bruni, M., Baecchi, C., Turchini, F., & Del Bimbo, A. (2021). Class-incremental learning with pre-allocated fixed classifiers. In *2020 25th international conference on pattern recognition* (pp. 6259–6266). IEEE.

Pham, Q., Liu, C., & Hoi, S. (2021). Dualnet: Continual learning, fast and slow. In *Advances in neural information processing systems, vol. 34* (pp. 16131–16144).

Pham, Q., Liu, C., Sahoo, D., & HOI, S. (2021). Contextual transformation networks for online continual learning. In *International conference on learning representations*.

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, *23*(17), R764–R773.

Qin, Y.-L., McNaughton, B. L., Skaggs, W. E., & Barnes, C. A. (1997). Memory reprocessing in corticocortical and hippocampocortical neuronal ensembles. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *352*(1360), 1525–1533.

Rannen, A., Aljundi, R., Blaschko, M. B., & Tuytelaars, T. (2017). Encoder based lifelong learning. In *Proceedings of international conference on computer vision*.

Rasch, B., & Born, J. (2007). Maintaining memories by reactivation. *Current Opinion in Neurobiology*, *17*(6), 698–703.

Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews*.

Rebuffi, S.-A., Kolesnikov, A., Sperl, G., & Lampert, C. H. (2017). Icarl: Incremental classifier and representation learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Rosenzweig, E. S., Barnes, C. A., & McNaughton, B. L. (2002). Making room for new memories. *Nature Neuroscience*, *5*(1), 6–8.

Rostami, M., Kolouri, S., & Pilly, P. K. (2019). Complementary learning for overcoming catastrophic forgetting using experience replay. In *IJCAI* (pp. 3339–3345).

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). ImageNet large scale visual recognition challenge. *International Journal of Computer Vision (IJCV)*, *115*(3), 211–252. http://dx.doi.org/10.1007/s11263-015-0816-y.

Saxena, R., Shobe, J. L., & McNaughton, B. L. (2022). Learning in deep neural networks and brains with similarity-weighted interleaved learning. *Proceedings of the National Academy of Sciences*, *119*(27), Article e2115229119.

Serra, J., Suris, D., Miron, M., & Karatzoglou, A. (2018). Overcoming catastrophic forgetting with hard attention to the task. In *Proceedings of international conference on machine learning*.

Shin, H., Lee, J. K., Kim, J., & Kim, J. (2017). Continual learning with deep generative replay. In *Advances in neural information processing systems*.

Shrivastava, A., Gupta, A., & Girshick, R. (2016). Training region-based object detectors with online hard example mining. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 761–769).

Song, X., Sun, P., Song, S., & Stojanovic, V. (2022). Event-driven NN adaptive fixed-time control for nonlinear systems with guaranteed performance. *Journal of the Franklin Institute*.

Stanford (2017). Tiny ImageNet challenge, CS231N course. https://tiny-imagenet.herokuapp.com/.

Sun, G., Cong, Y., Dong, J., Liu, Y., Ding, Z., & Yu, H. (2022). What and how: Generalized lifelong spectral clustering via dual memory. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(7), 3895–3908.

Sun, G., Cong, Y., Wang, Q., Zhong, B., & Fu, Y. (2022). Representative task self-selection for flexible clustered lifelong learning. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(4), 1467–1481.

Tadros, T., Krishnan, G. P., Ramyaa, R., & Bazhenov, M. (2022). Sleep-like unsupervised replay reduces catastrophic forgetting in artificial neural networks. *Nature Communications*, *13*(1), 7742.

Takashima, A., Petersson, K. M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M., et al. (2006). Declarative memory consolidation in humans: a prospective functional magnetic resonance imaging study. *Proceedings of the National Academy of Sciences*, *103*(3), 756–761.

Taylor, M. E., & Stone, P. (2009). Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*, *10*(7).

Thrun, S. (1995). A lifelong learning perspective for mobile robot control. In *Intelligent robots and systems* (pp. 201–214). Elsevier.

van de Ven, G. M., Siegelmann, H. T., & Tolias, A. S. (2020). Brain-inspired replay for continual learning with artificial neural networks. *Nature Communications*, *11*(1), 1–14.

Van de Ven, G. M., & Tolias, A. S. (2019). Three scenarios for continual learning. arXiv.

van de Ven, G. M., Trouche, S., McNamara, C. G., Allen, K., & Dupret, D. (2016). Hippocampal offline reactivation consolidates recently formed cell assembly patterns during sharp wave-ripples. *Neuron*, *92*(5), 968–974.

van de Ven, G. M., Tuytelaars, T., & Tolias, A. S. (2022). Three types of incremental learning. *Nature Machine Intelligence*, 1–13.

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*(11).

Vinyals, O., Blundell, C., Lillicrap, T., Wierstra, D., et al. (2016). Matching networks for one shot learning. In *Advances in neural information processing systems, vol. 29*.

Wang, Z., Zhang, Z., Ebrahimi, S., Sun, R., Zhang, H., Lee, C.-Y., et al. (2022). DualPrompt: Complementary prompting for rehearsal-free continual learning. In *Proceedings of European conference on computer vision* (pp. 631–648).

Wang, Z., Zhang, Z., Lee, C.-Y., Zhang, H., Sun, R., Ren, X., et al. (2022). Learning to prompt for continual learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 139–149).

Wilson, A., Fern, A., Ray, S., & Tadepalli, P. (2007). Multi-task reinforcement learning: a hierarchical bayesian approach. In *Proceedings of international conference on machine learning* (pp. 1015–1022).

Wilson, M. A., & McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, *265*(5172), 676–679.

Wu, C., Herranz, L., Liu, X., van de Weijer, J., Raducanu, B., et al. (2018). Memory replay gans: Learning to generate new categories without forgetting. In *Advances in neural information processing systems, vol. 31*.

Xin, X., Tu, Y., Stojanovic, V., Wang, H., Shi, K., He, S., et al. (2022). Online reinforcement learning multiplayer non-zero sum games of continuous-time Markov jump linear systems. *Applied Mathematics and Computation*, *412*, Article 126537.

Xu, Z., Li, X., & Stojanovic, V. (2021). Exponential stability of nonlinear state-dependent delayed impulsive systems with applications. *Nonlinear Analysis. Hybrid Systems*, *42*, Article 101088.

Yamamoto, J., & Tonegawa, S. (2017). Direct medial entorhinal cortex input to hippocampal CA1 is crucial for extended quiet awake replay. *Neuron*, *96*(1), 217–227.

Zeng, G., Chen, Y., Cui, B., & Yu, S. (2019). Continual learning of context-dependent processing in neural networks. *Nature Machine Intelligence*, *1*(8), 364–372.

Zenke, F., Poole, B., & Ganguli, S. (2017). Continual learning through synaptic intelligence. In *Proceedings of international conference on machine learning*.

Zhang, H., Cisse, M., Dauphin, Y. N., & Lopez-Paz, D. (2018). Mixup: Beyond empirical risk minimization. In *Proceedings of international conference on learning representations*.

Zhang, L., Song, J., Gao, A., Chen, J., Bao, C., & Ma, K. (2019). Be your own teacher: Improve the performance of convolutional neural networks via self distillation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Zhao, T., Wang, Z., Masoomi, A., & Dy, J. (2022). Deep Bayesian unsupervised lifelong learning. *Neural Networks*, *149*, 95–106.