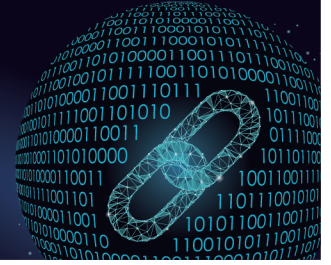




中国科技期刊卓越行动计划项目入选期刊

# 控制与决策

CONTROL AND DECISION



## 机器人增量学习研究综述

马旭森, 徐德

引用本文:

马旭森, 徐德. 机器人增量学习研究综述[J]. 控制与决策, 2024, 39(5): 1409–1423.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.0631>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### [移动机器人运动规划中的深度强化学习方法](#)

Deep reinforcement learning for motion planning of mobile robots

控制与决策. 2021, 36(6): 1281–1292 <https://doi.org/10.13195/j.kzyjc.2020.0470>

#### [基于接触状态感知发育的机器人柔性装配方法](#)

Flexible assembly method based on contact state perception development

控制与决策. 2021, 36(4): 876–884 <https://doi.org/10.13195/j.kzyjc.2019.1079>

#### [基于生物启发神经网络和DMPC的多机器人协同搜索算法](#)

Multi-robot cooperative search algorithm based on bio-inspired neural network and DMPC

控制与决策. 2021, 36(11): 2699–2706 <https://doi.org/10.13195/j.kzyjc.2020.0959>

#### [机器人抓取检测技术的研究现状](#)

Recent researches on robot autonomous grasp technology

控制与决策. 2020, 35(12): 2817–2828 <https://doi.org/10.13195/j.kzyjc.2019.1145>

#### [模仿学习示教轨迹自动分割方法的研究进展](#)

Recent advances on automatic segmentation method of teaching trajectory for imitation learning

控制与决策. 2019, 34(7): 1345–1354 <https://doi.org/10.13195/j.kzyjc.2018.0704>

# 机器人增量学习研究综述

马旭淼<sup>1</sup>, 徐德<sup>1,2†</sup>

(1. 中国科学院大学 人工智能学院, 北京 100049;

2. 中国科学院自动化研究所 中国科学院工业视觉智能装备技术工程实验室, 北京 100190)

**摘要:** 机器人的应用场景正在不断更新换代, 数据量也在日益增长. 传统的机器学习方法难以适应动态的环境, 而增量学习技术能够模拟人类的学习过程, 使机器人能利用旧知识来加快新任务的学习, 在不遗忘旧技能的前提下学习新的技能. 目前对于机器人增量学习的相关研究仍然较少, 对此, 主要介绍机器人增量学习研究进展. 首先, 对增量学习进行简介; 其次, 从参数和模型的角度出发, 将当前机器人增量学习主流方法分为变参数方法、变模型方法、混合方法 3 类, 分别对每一类进行论述, 并给出相应的增量学习技术在机器人领域中的应用实例; 然后, 对机器人增量学习中常用的数据集和评价指标进行介绍; 最后, 对增量学习未来的发展趋势进行展望.

**关键词:** 增量学习; 变参数方法; 变模型方法; 混合方法; 技能学习; 机器人

中图分类号: TP242

文献标志码: A

DOI: 10.13195/j.kzyjc.2023.0631

引用格式: 马旭淼, 徐德. 机器人增量学习研究综述[J]. 控制与决策, 2024, 39(5): 1409-1423.

## Incremental learning for robots: A survey

MA Xu-miao<sup>1</sup>, XU De<sup>1,2†</sup>

(1. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China; 2. CAS Engineering Laboratory for Intelligent Industrial Vision, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

**Abstract:** Nowadays, the application scenarios of robots are constantly updated, and the amount of data is also growing. Traditional machine learning methods are difficult to adapt to the dynamic environment. Incremental learning technology simulates the human learning process, enabling robots to use old knowledge to speed up the learning of new tasks and learn new skills without forgetting old skills. Currently, there is still relatively little research on robot incremental learning. This paper mainly introduces the research progress of robot incremental learning. Firstly, a brief introduction to incremental learning is given. Secondly, from the perspective of parameters and models, this paper classifies the current mainstream methods of robot incremental learning into three categories: variable parameter methods, variable model methods and hybrid methods, which are discussed in details, separately. Furthermore, the corresponding application examples of incremental learning technology in the field of robotics are provided. Thirdly, the data sets and evaluation metrics commonly used in incremental learning are introduced. Finally, the future development trends are prospected.

**Keywords:** incremental learning; variable parameter method; variable model method; hybrid method; skill learning; robot

## 0 引言

人工智能技术的发展正在推进新一轮工业革命以及产业结构变革的进程. 作为人工智能技术的重要载体和现代工业不可或缺的基础装备, 机器人在智能制造领域承担着加工、装配、搬运等一系列重要任务. 近年来, 随着科学技术的快速发展, 机器人行业取得了长足进步.

虽然机器人技术在不断取得新的进展, 但机器人仍然面临着一个颇具挑战性的问题, 即如何让机器人能够自我适应现代社会复杂且动态变化的环境, 从中学习新的技能并且不断成长, 这也是人类设计智能机器人最终想要达到的理想目标. 目前的大多数机器人, 其解决问题的实际手段还是通过预先设计好的算法进行计算, 而非自主学习. 对于封闭、静止环境中的

收稿日期: 2023-05-11; 录用日期: 2023-08-16.

基金项目: 国家自然科学基金项目(62273341).

责任编辑: 方勇纯.

†通讯作者. E-mail: de.xu@ia.ac.cn.

任务,机器人往往能有较好的表现,但对于动态环境中的任务,机器人的表现却难以达到预期的效果. 机器人静态工作环境是指机器人在工作过程中所处的是确定和无变化的环境. 这种环境下往往只需要解决单一或者固定的任务,例如送餐、运货、烹饪等. 机器人动态工作环境是指机器人在工作过程中面临不断变化和不确定性的环境. 这些环境通常涉及复杂的实时情况和多变的条件,需要机器人及时去适应,例如水下勘测、地面巡逻、搜救任务等<sup>[1-3]</sup>.

为了解决上述问题,增量学习 (incremental learning) 技术应运而生. 增量学习也被称为终身学习 (lifelong learning)<sup>[4]</sup>、连续学习 (continual learning)<sup>[5-9]</sup> 或者序列学习 (sequential learning)<sup>[10-11]</sup>. 在机器人领域中,增量学习指的是机器人能够学习一系列序列化的任务,在这个过程中,机器人不仅可以学习到新的知识,而且不会遗忘旧知识. 能够进行增量式学习的机器人可以被看作是强智能体,它能处理多种类型的任务和应对不断产生的新情况. 以轴孔装配领域为例,装配机器人在掌握单轴孔、双轴孔装配技能的基础上,可以通过增量学习进一步学会多轴孔、多形状轴孔的装配方法,并且在学会新技能之后仍然不会忘记旧技能. 如果装配过程受到阻碍,机器人则会根据实际情况重新规划装配路径、调整装配姿态或者停止装配.

本文首先对增量学习进行简介;其次,从参数以及模型的角度对增量学习中的主流方法进行论述,并给出应用实例;然后,介绍增量学习中常用的数据集以及评价指标;最后展望了增量学习的未来研究趋势.

## 1 增量学习简介

本节阐述增量学习的定义,并介绍稳定性-可塑性困境问题和相关理论及术语. 增量学习的早期定义是由 Thrun 等<sup>[12]</sup> 在 1995 年提出的. 随着增量学习的发展,Chen 等<sup>[4]</sup> 对 Thrun 的定义进行了拓展. 增量学习的过程是一个连续的过程. 假设增量学习模型已经学习了  $N$  个旧任务,这些任务形成一个连续的序列  $\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_N$ , 每个任务都对应着一个数据集  $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_N$ , 模型学习到的知识被保存在知识库中. 当开始学习第  $N+1$  个新任务  $\mathcal{T}_{N+1}$  时,模型可以利用知识库中的旧知识来帮助学习  $\mathcal{T}_{N+1}$ . 增量学习的目标通常是用从知识库中学习和积累的知识来优化新任务的性能. 在模型完成学习  $\mathcal{T}_{N+1}$  之后,可以用从  $\mathcal{T}_{N+1}$  中获得的知识来更新知识库. 在学习完成之后,模型不仅能够掌握新的知识,并且不会遗忘之前

的旧知识,即在旧任务上的性能表现仍然较好.

在增量学习中,稳定性-可塑性困境问题<sup>[13-14]</sup> 是一个较为关键的问题. 模型的稳定性是指模型对旧知识的保持能力,稳定性较差的模型会出现灾难性遗忘的问题,即学习新知识后会遗忘旧知识;模型的可塑性是指模型学习新知识的能力. 实际学习过程中,模型的稳定性和可塑性往往不能够兼得,二者之间存在着此消彼长的矛盾关系. 因此,如何平衡二者之间的关系是至关重要的.

互补学习系统 (complementary learning systems, CLS) 是增量学习中的一个重要理论,最早是由 McClelland 等<sup>[15]</sup> 于 1995 年提出的,后来, Kumaran 等<sup>[16]</sup> 又对其进行了补充和完善. CLS 是对人类的大脑记忆模式进行研究得到的理论体系. 人类大脑记忆和概括的互补任务在海马体和新皮质的相互作用下完成. 1987 年, Hinton 等<sup>[17]</sup> 提出了 CLS 早期计算模型,模型中每个神经元都有两个权重:一个权重变化率较低,用于存储长期知识,相当于新皮质;另一个权重变化率较高,用于学习临时知识,相当于海马体. 这种方法体现了互补学习系统的核心要点,可以缓解模型学习过程中的灾难性遗忘问题.

近几年,迁移学习 (transfer learning)<sup>[18-20]</sup>、多任务学习 (multi task learning)<sup>[21]</sup>、在线学习 (online learning)<sup>[22-23]</sup>、课程学习 (curriculum learning)<sup>[24-25]</sup>、元学习 (meta learning)<sup>[26-28]</sup> 等概念被相继提出,增量学习与这些术语之间有着相似之处,但也存在一定的区别.

迁移学习通常是单向的迁移过程,在迁移完成后通常不会再关注模型在旧任务上的性能;而增量学习还需要保证不会遗忘旧的知识. 多任务学习重点关注如何解决多个相关程度较高的任务;而增量学习主要关注如何在连续的数据流中不断积累知识并提升模型性能. 在线学习主要专注于单个任务的学习,其训练数据都来源于单个数据集,并且满足独立同分布的条件;而增量学习需要学习多个任务,并且多个任务的数据之间不满足独立同分布的条件. 课程学习与增量学习都需要处理序列形式的多个任务,但课程学习对于任务的顺序有着较为严格的要求;而增量学习对任务的顺序则没有特定的要求. 元学习也被称为“学会学习” (learning to learn), 主要目的是使用在许多其他非常类似的任务上训练过的模型,通过少量训练示例学习新任务,目前大多数的元学习策略都是假设旧任务与新任务的数据分布有着较高的一致性;而增量学习一般不作这种假设.

## 2 机器人增量学习方法

本节对目前主流的机器人增量学习方法进行分类论述,并给出在具体机器人平台上的应用实例。

目前的增量学习方法较多. de Lange等<sup>[5]</sup>将增量学习方法分为正则化方法、回放方法以及参数隔离方法. Mundt等<sup>[29]</sup>将增量学习方法分为正则化方法、回放方法、结构方法以及混合方法. Hadsell等<sup>[30]</sup>将增量学习的主流方法分为独立同分布学习方法、基于梯度的方法、基于模块化的方法、基于记忆的方法以及基于元学习的方法. Wang等<sup>[31]</sup>将增量学习方法分为基于正则化的方法、基于回放的方法、基于优化的方法、基于表示学习的方法、基于模型结构的方法。

近年来,国内学者对增量学习的关注度也在逐渐提高. 朱飞等<sup>[32]</sup>对类别增量学习进行了研究,将现有的类别增量学习方法分为基于参数正则化、基于知识蒸馏、基于数据回放、基于特征回放以及基于网络结构这5类方法. 韩亚楠等<sup>[33]</sup>对连续学习的定义、代表性方法进行了详细阐述,并将其总结为正则化策略、记忆回放以及动态框架策略3类主流方法。

机器人增量学习是增量学习的一个重要分支. 增量学习是一类机器学习方法,可用于处理分类、回归等问题. 而机器人增量学习是指机器人系统应用增量学习方法不断地积累和处理新数据或经验,进而提高自身智能水平的过程. 与传统的增量学习不同,在机器人增量学习模型实际训练的过程中,主要的难点如下:

1) 根据机器人平台的不同,需要注意相应的约束条件<sup>[34]</sup>. 例如,在轮式移动机器人的运动过程中通常存在非完整性约束问题,此时需要对增量学习网络的输出加以限制。

2) 由于机器人需要在实际环境中完成任务,对于计算速度的需求会更高. 在机器人上部署增量学习算法时,应满足实际工作的实时性需求,因此,往往需

要对机器人增量学习模型进行优化调整。

当前大部分工作主要关注增量学习的研究进展,而对于机器人增量学习的研究仍然较少. 根据现有方法的主要特点,本文从参数和模型两个宏观角度对现有的机器人增量学习方法进行分类,即变参数方法、变模型方法以及混合方法,如图1所示。

### 2.1 变参数方法

变参数是指在同一模型架构下,通过调整模型的参数来适应不同的任务或数据集. 这些参数可以是权重、偏置、学习率等超参数,它们的改变可以影响模型的性能和泛化能力. 在学习新任务时,为了解决灾难性遗忘问题,可以在损失函数中加入正则化项来限制模型的更新,这样的方法被称为正则化方法;也可以利用历史任务数据进行回放,使模型从以往的经验中学习,这样的方法被称为回放方法。

#### 2.1.1 正则化方法

正则化方法通过对神经网络权重的更新施加约束来减轻灾难性遗忘的问题. 主要可以分为基于数据约束的方法和基于参数约束的方法。

##### 1) 基于数据约束的方法.

基于数据约束的方法通常基于知识蒸馏策略来实现增量式学习. 在传统的深度学习方法中,如果来实现由旧任务到新任务的学习,可以采用迁移学习中的微调(fine-tuning)<sup>[35]</sup>、特征提取(feature extraction)<sup>[36]</sup>或联合训练(joint training)<sup>[37]</sup>等方法. 然而,微调方法会改变新旧网络之间的共享参数,有可能使得模型旧任务上的性能下降;特征提取方法由于需要保持共享参数不变,只能训练特定的几层网络,使得网络的拟合能力受到很大限制;联合训练方法需要优化所有的参数,其训练成本较高,训练难度也会很大。

为了解决上述问题,Furlanello等<sup>[38]</sup>提出了主动长期记忆网络模型(active long term memory network, A-LTM),基于知识蒸馏的思想,通过按序列学习多任务,可以在学习新知识的基础上保留旧知识. 尽管A-LTM在一定程度上解决了灾难性遗忘的问题,但其仅适用于在小数据集上的训练,如果数据集规模较大,则模型的性能仍会有较大程度的下降。

2018年,Li等<sup>[39]</sup>提出了由卷积神经网络作为模型主要架构的无遗忘学习(learning without forgetting, LwF)方法,将高度正则化的教师模型转化为更小的学生模型. 根据LwF算法,所有的任务会共享一组参数 $\theta_s$ ,新任务的参数为 $\theta_n$ ,旧任务的参数为 $\theta_o$ . 给定新任务的训练数据 $(X_n, Y_n)$ 、旧任务的输出 $Y_o$ 以及随

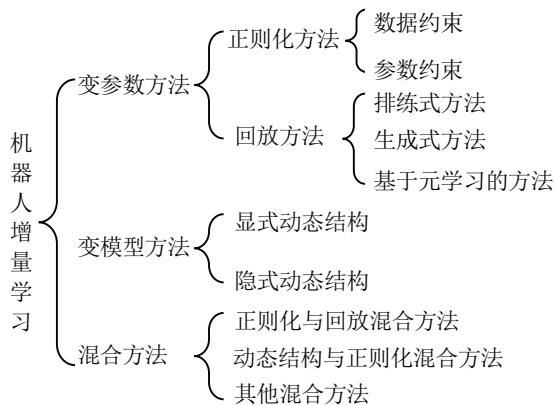


图1 机器人增量学习分类

机初始化的新参数 $\theta_n$ ,更新后的参数 $\theta_s^*$ 、 $\theta_o^*$ 、 $\theta_n^*$ 由下式给出:

$$\theta_s^*, \theta_o^*, \theta_n^* \leftarrow \arg \min_{\hat{\theta}_s, \hat{\theta}_o, \hat{\theta}_n} (\lambda_o \mathcal{L}_{\text{old}}(Y_o, \hat{Y}_o) + \mathcal{L}_{\text{new}}(Y_n, \hat{Y}_n) + \mathcal{R}(\hat{\theta}_s, \hat{\theta}_o, \hat{\theta}_n)). \quad (1)$$

其中: $\mathcal{L}_{\text{old}}$ 为蒸馏损失函数; $\mathcal{L}_{\text{new}}$ 为新任务上的损失函数; $\hat{Y}_n$ 、 $Y_n$ 分别为模型在新任务上的预测值和真实值; $\hat{Y}_o$ 、 $Y_o$ 分别为模型在旧任务上的预测值和真实值; $\lambda_o$ 为平衡系数,用于平衡新旧任务的性能; $\mathcal{R}(\hat{\theta}_s, \hat{\theta}_o, \hat{\theta}_n)$ 为正则化项。

LwF算法只需要新任务的数据就能掌握新的技能,同时较好地保留旧的知识.实验表明,LwF的实验效果优于A-LTM.然而,LwF算法的缺点是高度依赖于任务的相关性,如果任务之间的数据相关性太低,则容易出现任务混淆的现象(inter-task confusion).此外,LwF算法总体的训练时间随着学习任务的数量快速增加,当面临多任务情况时,训练模型可能会变得很困难.

由于LwF对新旧任务的相关性需求过高,Rannen等<sup>[40]</sup>对LwF方法进行了拓展,提出了基于编码器的增量学习方法(encoder based lifelong learning, EBLL).在LwF方法中,保留旧任务的特征是其核心.然而,如果完全保留旧特征,则会限制模型的学习能力.因此,EBLL保留了一部分重要低维特征,设计了自编码器来寻找其他特征,以增加模型的灵活性.EBLL将特征投影到一个低维流形上,在训练时向损失函数中添加正则化项,其作用是约束当前特征在低维流形上的投影不偏离先前任务上的最优投影,从而达到不遗忘旧知识的目的.在上述方法的基础上,Ayub等<sup>[41]</sup>提出了一种少样本的增量学习方法,用于逐步训练Baxter机器人识别不同的物体类别,并最终执行桌面整理任务.在Baxter机器人仅从人类提供的几个示例开始学习的情况下,逐渐扩展其对新类别物体的识别能力.Isele等<sup>[42]</sup>提出了一种针对移动机器人扰动抑制的方法,通过基于数据约束的增量学习技术,并采用线性回归算法来更新控制器参数,将其与样本数据作为训练集输入模型中进行学习.通过在Turtlebot机器人上部署该方法,机器人可以在不断遭受外部干扰时保持稳定性和高精度姿态控制,有效提高了机器人在实际工作中的鲁棒性.

## 2) 基于参数约束的方法.

除了基于数据约束的方法,另一类正则化方法是基于参数约束的方法.其中最为经典的是Kirkpatrick等<sup>[43]</sup>提出的弹性权重整合(elastic weight

consolidation, EWC)方法.

在EWC方法的思想中,神经网络发生灾难性遗忘的主要原因是参数权重更新的过程中,对于旧任务较为重要的一部分神经元学习率过高,导致其偏移程度过高,在旧任务上的性能就会下降.因此,EWC方法提出,应根据各个神经元对于旧任务的重要性排序,降低重要权重的学习率.通过为权重添加正则项,从而控制权重优化方向,最终达到增量学习的效果.假设一个学习场景具有两个独立的任务A和B,其数据集分别为 $\mathcal{D}_A$ 和 $\mathcal{D}_B$ ,EWC方法通过模型参数 $\theta$ 的后验概率分布 $p(\theta|\mathcal{D})$ 对任务训练数据集 $\mathcal{D}$ 的模型参数 $\theta$ 进行建模.根据贝叶斯法则给出的后验概率的对数值为

$$\log p(\theta|\mathcal{D}) = \log p(\mathcal{D}|\theta) + \log p(\theta) - \log p(\mathcal{D}). \quad (2)$$

由于任务A和B相互独立,式(2)可以进一步表示为

$$\log p(\theta|\mathcal{D}) = \log p(\mathcal{D}_B|\theta) + \log p(\theta|\mathcal{D}_A) - \log p(\mathcal{D}_B), \quad (3)$$

其中后验概率 $\log p(\theta|\mathcal{D}_A)$ 体现了关于先前任务A的所有信息.然而,由于该项难以处理,EWC方法通过拉普拉斯近似,将其转换为高斯分布

$$p(\theta|\mathcal{D}_A) \sim N\left(\theta_A^*, -\frac{1}{f''(\theta_A^*)}\right). \quad (4)$$

该高斯分布的平均值就是训练完任务A之后网络的最优参数 $\theta_A^*$ ,方差为任务A模型参数的Hessian矩阵 $f''(\theta_A^*)$ .但由于Hessian矩阵的计算开销太大,EWC方法使用Fisher信息矩阵 $\mathcal{F}$ (Fisher information matrix, FIM)来代替Hessian矩阵.因此,EWC的损失函数由下式给出:

$$\mathcal{L}(\theta) = \mathcal{L}_B(\theta) + \sum_i \frac{\lambda}{2} \mathcal{F}_i (\theta_i - \theta_{A,i}^*)^2. \quad (5)$$

其中: $\mathcal{L}_B$ 是任务B的损失函数; $\lambda$ 是一个权衡参数,表示旧任务与新任务的相关性; $i$ 表示参数的索引; $\theta_{A,i}^*$ 表示训练完任务A之后网络参数 $\theta_i$ 的最优值; $\mathcal{F}_i$ 是FIM的第*i*个对角线元素,实际上 $\mathcal{F}_i$ 就是参数 $\theta_i$ 对于旧任务的重要性程度.因此,该方法需要对学习任务的参数进行加权,该加权重与Fisher信息度量的对角线元素成正相关.

尽管EWC方法有着超越LwF方法的性能,但是Kemker等<sup>[44]</sup>的实验表明,EWC方法对于新分类的学习能力较差,并且在计算Fisher矩阵时的成本过于高昂,导致其只能训练规模较小的数据集,这是限制EWC方法应用的重要原因.

为了解决EWC方法的弊端,Zenke等<sup>[45]</sup>提出了SI(synaptic intelligence)算法.SI算法可以在线计算

参数权重重要性,即能够在训练时根据损失函数对参数的敏感性来实时地改变参数的权重大小,如果损失函数对参数 $\theta_i$ 越敏感,则说明该参数对任务学习的重要程度越高.与EWC方法类似,这种方法对相关神经元的更新添加了正则化项,从而可以在遗忘最少知识的情况下学习新任务.SI方法与EWC方法的主要区别在于,EWC方法通过FIM的对角元素值来估计参数的重要程度,而SI方法是通过损失函数对参数的敏感性来分析参数的重要性.为了在学习新任务时减少重要参数 $\theta_i$ 的大幅度变化,文献[45]使用了一个修正的成本函数 $\mathcal{L}_n^*$ 来替代所有先前任务的总损失函数 $\mathcal{L}_o^*$ ,即

$$\mathcal{L}_n^* = \mathcal{L}_n + c \sum_n \frac{\lambda}{2} \Omega_i^n (\theta_i^* - \theta_i)^2. \quad (6)$$

其中: $c$ 是平衡新任务与旧任务的加权参数, $\theta_i^*$ 是前一任务结束时的参数, $\Omega_i^n$ 是每个参数的调节强度参数.

SI方法在旧任务中表现良好,并且SI对于EWC方法的改进使Fisher矩阵的计算速度和难度得到了改善.

除了SI方法之外,许多学者也提出了EWC的改进方法.由于EWC方法假设FIM是对角的,这显然过于理想化,对此,Liu等<sup>[46]</sup>提出了基于旋转空间的EWC(EWC in the rotated space, R-EWC)方法,其主要思想是通过任务参数空间的因式旋转,使其坐标轴与FIM空间坐标轴对齐,这样就很好地解决了部分情况下EWC表现较差的问题.此外,Schwarz等<sup>[47]</sup>提出了Online-EWC方法,增加了EWC方法的拓展性;Chaudhry等<sup>[48]</sup>提出了RWalk方法,该方法将EWC和SI方法的正则化项进行结合,进一步提高了准确性.

受参数约束方法的启发,Valipour等<sup>[49]</sup>提出了一种基于带参数正则化的深度神经网络增量学习框架,包括对新类别物体的自动分配、逐步学习和模型更新等多个环节.训练过程中,在一个7自由度的Wam机械臂上搭载深度相机,通过与人类用户的交互,获取带有语义标签的物体图像样本,从而进行增量式学习和知识融合.Calinon等<sup>[50]</sup>提出了一种基于高斯混合模型及参数约束的增量式手势学习方法,用于训练HOAP-3人形机器人执行各种手势动作.该方法能够使机器人从人类示范中自适应地学习新的手势和运动方式,并且通过在线学习和更新,不断优化和改进动作表现.王磊等<sup>[51]</sup>提出了一种增量式正则化回声状态网络,得到的模型可以自适应设定储备池及相关参数,具有较好的泛化性.

## 2.1.2 回放方法

在现实场景中,旧任务的数据往往能够获得,如果可以利用旧任务的数据,则能改善知识库的知识积累过程.基于回放的增量学习方法是指在内存中开辟经验回放池作为存储器,以此保存过去任务的部分样本或者学习到的经验,在学习新任务的同时回顾经验池中的数据,以削弱灾难性遗忘的影响.基于回放的方法虽然需要额外的计算内存空间,但往往能为模型带来不错的性能提升.具体而言,增量学习中的回放方法可以分为排练式方法(rehearsal approaches)、生成式方法(generative approaches)、基于元学习的方法(meta-learning based approaches).

### 1) 排练式方法.

排练式方法是回放方法中的一类经典方法,当数据样本足够多时,将其放入经验池中供模型学习.数据样本的选择有多种方式:可以采用所有新旧数据样本,也可以挑选具有代表性的样本,或者是通过特征提取方法抽取出数据样本的特征作为核心集.理想情况下,将所有样本作为核心集的效果显然是最好的,但是考虑到实际计算资源以及内存的限制,采用代表性样本或者样本特征才是符合实际应用情况的方法.

2017年,Lopez-Paz等<sup>[52]</sup>提出了用于增量学习的梯度情景记忆方法(gradient episodic memory, GEM).情景记忆是指保存旧任务中效果好的数据,该方法为每个任务 $i$ 都创建了一个情景记忆模型 $\mathcal{M}_i$ ,其损失函数可以表示为

$$\ell(f_\theta, \mathcal{M}_i) = \frac{1}{|\mathcal{M}_i|} \sum_{(x_k, i, y_k) \in \mathcal{M}_i} \ell(f_\theta(x_k, i), y_k). \quad (7)$$

其中: $\theta$ 为模型参数, $f_\theta$ 为模型的预测值, $\ell(f_\theta, \mathcal{M}_i)$ 为第 $i$ 个任务的损失函数, $(x_k, y_k)$ 为样本数据.

因为直接最小化式(7)容易造成模型过拟合,所以GEM考虑了一个不等式约束,希望使得上述损失函数变小,即

$$\begin{aligned} \min_{\tilde{g}} \quad & \frac{1}{2} \|g - \tilde{g}\|^2; \\ \text{s.t.} \quad & \langle \tilde{g}, g_i \rangle \geq 0, \forall i \leq t. \end{aligned} \quad (8)$$

其中: $g$ 表示当前任务 $t$ 对应的梯度, $\tilde{g}$ 表示 $g$ 在 $L_2$ 范数条件下的投影.

通过式(8)中的不等式约束,可以对模型更新的梯度方向进行不断修正.由于该不等式约束的存在,式(7)中损失函数始终会朝着更小的方向更新其值,相比于EWC等方法,GEM能主动地实现知识的反向迁移.由于GEM的计算复杂度较高,Chaudhry等<sup>[53]</sup>

提出了平均梯度情景记忆方法 (averaged gradient episodic memory, A-GEM). 相比于 GEM, A-GEM 的主要改进之处是将减少先前任务记忆损失更改为减少先前任务的平均记忆损失, 通过计算均值, 可以有效减少计算的复杂度, 加快算法的收敛速度.

在互补学习系统 (complementary learning systems, CLS) 的思想启发下<sup>[15,16,54]</sup>, 采用双层记忆系统的回放方法也是一类较为常见的方法, 如 Geppert 等<sup>[55]</sup> 提出的 GeppNet+STM 方法. GeppNet 实际上是一个 3 层的神经网络, GeppNet+STM 方法使用短期记忆模块 (short-term memory, STM) 来拓展 SOM 方法. 当有新的知识输入时, 模型会将其存储在 STM 模块中, 与人类的记忆系统类似, STM 会将其中存储的知识不定时回放. GeppNet+STM 方法吸收了课程学习的思想, 使得模型可以逐步学习数据集中的特征分布. GeppNet+STM 方法的局限性在于 STM 的容量是有限的, 而且, 该方法需要在训练期间存储整个训练数据集, 这对模型的规模提出了较大的挑战. 在排练式方法的基础上, Liu 等<sup>[56]</sup> 提出了一种增量学习的移动机器人导航算法, 通过融合排练式方法以及深度强化学习算法, 在 Jackal 移动机器人上进行部署, 使得机器人能够在不同的环境中进行自适应和长期的导航任务.

排练式方法确保了记忆不会随着时间的推移而退化, 通常具有比正则化方法更高的性能. 但其缺点是模型需要在训练阶段始终保存所有旧任务的数据, 这是一种不尊重数据隐私的保存知识的方法, 并且容易出现多次学习同一数据的现象, 最重要的问题是往往会被计算内存所限制.

## 2) 生成式方法.

生成式方法也被称为伪排练式方法 (pseudo rehearsal approaches), 其与排练式方法的不同主要在于数据样本的获取方式不同. 排练式方法是在旧任务数据充足的前提下成立的, 但在任务数据不足的情况下, 排练式方法就不适用了. 生成对抗网络 (generative adversarial network, GAN)<sup>[57]</sup> 和自动编码器 (auto encoder, AE) 为这种情况提供了解决思路, 即可以通过生成网络或自动编码器来生成伪样本. Shin 等<sup>[58]</sup> 提出了用于增量学习的深度生成回放方法 (deep generative replay, DGR), 该方法受到 GAN 的启发, 其主要思想是将生成的旧任务伪样本与新任务的真实样本混合, 放入经验池中. 然而, DGR 的计算成本较高, 在模型训练上存在一定困难, 因此 DGR 在一般情况下用作初始框架, 用于拓展或者

对比. Wu 等<sup>[59]</sup> 提出了回放对抗生成网络 (memory replay GAN, MeRGAN), 加强了新旧生成模型产生的数据之间的一致性. Ostapenko 等<sup>[60]</sup> 提出了动态生成记忆方法 (dynamic generative memory, DGM), 通过引入注意力机制, 使得模型依靠单个生成网络就能进行训练, 简化了训练流程, 解决了 DGR 方法训练较为困难的问题.

利用生成网络可以获得较高质量的伪样本, 但是, Ostapenko 等<sup>[60]</sup> 以及 Ayub 等<sup>[61]</sup> 的研究指出, 这种方法存在数据标签不一致的问题, 因此可以用自动编码器代替生成网络. Kemker 等<sup>[62]</sup> 提出了 FearNet, 基于双层记忆系统, 利用概率神经网络快速学习存放短期记忆; 利用自动编码器存放长期记忆, 经过数个训练阶段, 概率神经网络中的知识会逐步转移至自动编码器中. Gopalakrishnan 等<sup>[63]</sup> 提出了知识捕获及回放的方法, 通过自动编码器将输入的数据转换为抽认卡 (flashcards) 形式, 使得模型可以通过抽认卡来回顾知识, 有效削弱了灾难性遗忘的问题. 此外, 扩展式重采集模块 (scalable recollection module, SRM)<sup>[64]</sup>、高斯混合回放 (Gaussian mixture replay, GMR)<sup>[65]</sup> 等方法也是基于自动编码器的经典生成式回放方法.

除了单独使用生成网络或者自动编码器作为骨干网络的方法, 也可将二者混合使用. Ye 等<sup>[66]</sup> 将变分自动编码器与生成对抗网络混合使用, 提出了 L-VAEGAN 方法, 能够生成高质量样本. 基于生成式方法, Soh 等<sup>[67]</sup> 提出了一种基于触觉信息的生成式增量学习方法, 旨在实现机器人对物体的识别和分类. 基于生成式方法的思想, 将采集到的触觉信息输入生成模型中以得到更高质量的数据, 进一步提高识别和分类的准确性. 通过在 iCub 人形机器人上应用生成式增量学习方法, 可以实现对新物体的快速自适应和识别.

生成式方式普遍存在的问题是生成网络的训练较为困难, 如果数据特征较为复杂, 则很难训练出效果较好的生成网络. 因此, 也可以采用特征提取的方法, 将数据的特征作为样本, 让生成网络学习生成特征, 这样便可简化网络的训练流程. 2020 年, Liu 等<sup>[68]</sup> 提出了训练条件生成对抗网络, 用以生成特征来进行回放; van de Ven 等<sup>[69]</sup> 提出了在自动变分编码器中加入语义调制反馈连接模块, 从而回放相关内在数据特征.

## 3) 基于元学习的方法.

基于元学习的回放方法是从元学习思想中受到启发而得. Riemer 等<sup>[70]</sup> 提出了一种元经验回放方法

(meta experience replay, MER), 该方法基于 Finn 等<sup>[71]</sup>提出的元学习优化方法, 能够使得模型在保持现有知识的基础上, 尽可能减少旧知识对学习新知识知识的干扰. Henning 等<sup>[72]</sup>提出了一种基于后验的元回放方法 (posterior meta-replay, PMR), 该方法采用贝叶斯推断来代替原有的任务推断方法, 能够从数据中连续推断任务条件参数分布. 实验表明, 该方法对灾难性遗忘的问题有较好的效果. PMR 方法的缺点是, 由于其需要对每个预测后验进行并行评估, 导致模型在推断不确定性较强的任务时计算成本较为高昂, 对计算内存的要求较高. Wang 等<sup>[73]</sup>针对少样本图像识别问题提出了一种增量元学习方法, 首先使用少量的目标任务数据来训练初始模型, 然后使用随机抽样和 SLDA 策略来选择相关历史经验数据, 并将其与目标任务数据组成一个临时性任务集合进行增量式元学习. 该方法能够实现快速学习和泛化, 有效解决了少样本图像识别问题, 对于基于元学习方法的回放方法在实际机器人上的部署有着较为重要的意义.

## 2.2 变模型方法

变模型是指在不同的任务或数据上使用不同的模型架构来解决问题. 这些不同的模型架构可以是从头开始设计的, 也可以通过对已有模型进行修改或组合得到. 变模型的优点是可以更好地适应不同的任务和数据集, 但需要相应地增加模型训练、调参和部署的成本. 变模型方法可以分为显式动态结构方法和隐式动态结构方法. 显式动态结构方法的主要思想是通过参数隔离, 将每个任务的神经元独立训练, 模型可以拓展网络结构, 在训练新任务时加入新的神经元进行训练; 隐式动态结构方法的主要思想是将任务网络参数稀疏化, 在训练不同任务时冻结不同部分权重或者改变其传播路径, 从而达到参数隔离的效果.

### 2.2.1 显式动态结构方法

由于显式动态结构方法需要持续拓展模型结构, 通常用于对模型规模限制要求不高的环境中, 其最终效果一般要优于隐式方法. Rusu 等<sup>[74]</sup>提出了渐进式神经网络方法 (progressive neural networks, PNN). PNN 的主要思想十分简洁: 对于要学习的每一项新任务, PNN 都会创建一个新网络, 新网络的每一层都与所有旧网络的相同层连接, 即旧网络的每一层输出都作为新网络相同层的额外输入. 尽管这样的结构在理论上会有着不错的训练效果, 但其缺点也是十分明显的. 随着任务数量的增加, PNN 中的参数量会随之呈现二次增长. 这种增长往往是非常不利的,

因为增量学习中的任务量通常较大, 并且理论上会无限增加, 这会导致参数量的爆炸式增长. 因此 PNN 能够完成的任务数量是有上限的, 这取决于物理条件的限制.

与 PNN 的设计思想类似, Aljundi 等<sup>[75]</sup>提出使用自动编码器设计多个独立的专家网络, 以适应不同任务. 通过自动编码器, 可以对不同的任务数据特征进行识别, 从而将其分隔开, 形成不同的“门”. 当有新数据输入时, 模型就可以判断出最适合处理该数据的专家网络, 从而提高模型的性能. 其他具有代表性的显式动态结构方法还有神经发生深度学习 (neurogenesis deep learning, NDL)<sup>[76]</sup>、动态可拓展网络 (dynamically expandable network, DEN)<sup>[77]</sup>等.

显式动态结构方法的设计思想较为简单, 理论上可以达到增量学习所设想的理想效果, 即系统能够不断学习新的技能, 并且还能保留原来的技能. 然而在实际训练过程中, 由于计算内存的限制, 要训练理论上无限制的动态拓展模型在现阶段仍然无法实现.

### 2.2.2 隐式动态结构方法

与显式动态结构方法不同, 隐式动态结构方法通常不会改变模型的架构. 在每个任务训练结束时, 隐式动态结构方法通常会选择冻结模型中的一部分参数权重, 使其无法再用于后向传播, 但仍然可以用于前向传播. 这样的机制能有效缓解灾难性遗忘的问题, 同时还可以预留足够多的参数用于学习新任务. 但其难点在于冻结权重的数量, 如果过少则不会起到明显的效果, 如果过多则会影响到新任务的学习. 隐式动态结构方法常用于固定模型最大容量不变的情况下, 使得模型可以自适应地进行增量式学习. 因此, 隐式动态结构方法也被称为网络自适应方法.

Mallya 等<sup>[78]</sup>提出了 PackNet 方法. PackNet 方法采用特殊掩码机制, 每个任务都有一个掩码来判别该任务需要使用的参数权重, PackNet 方法会冻结涉及到的权重, 以保护旧知识不被遗忘. 在新任务训练完成后, PackNet 方法会对该任务的子网络进行剪枝, 选出与该任务相关性较高的权重. PackNet 方法的主要缺点是可用参数的数量有限, 随着后续任务的添加, 模型的性能会逐渐下降.

为了解决 PackNet 的上述问题, Mallya 等<sup>[79]</sup>提出了 Piggyback 方法, 使用单个深度神经网络来适应多个任务的学习过程. 在 Piggyback 方法中, 同样也使用了掩码机制. 不同于 PackNet 方法, Piggyback 直接冻结了骨干网络的权重, 转而使用端到端的二进制

掩码作为骨干网络参数权重的附加信息,这样就避免了对原权重的直接修改.部分隐式动态结构方法在网络中引入了注意力机制. Serra等<sup>[80]</sup>提出了一种面向任务的硬注意力机制模型(hard attention to the task, HAT),通过设计逐层的硬注意力机制,在学习任务时HAT方法会限制相关权重的梯度更新,从而保护重要权重,避免灾难性遗忘的问题.

除了通过冻结部分权重,还可以为不同任务设计特定的传播路径来实现对新任务的学习. PathNet<sup>[81]</sup>是基于该思想设计的一种经典方法. PathNet模型中有多层深度神经网络,在训练不同任务时,通过遗传算法来确定不同任务数据在网络中传播的特定路径.实验表明,PathNet具有良好的防止遗忘的效果,并且模型的性能较好. PathNet的不足之处在于,其要求每个任务对应的输出层之间相互独立,这使得PathNet在类增量学习场景中的应用较为困难.

基于动态结构方法的思想, Zhang等<sup>[82]</sup>提出了一种自适应增量学习方法,用于Nadine社交机器人中的图像语义理解.该方法利用卷积神经网络对输入图像进行特征提取和表示,并使用增量学习算法引入了自适应控制机制,逐渐更新CNN模型的参数,能够根据当前任务和学习情况对学习速度和模型结构进行调整和优化,较好地提高了学习效率和准确度.

### 2.3 混合方法

混合方法是从参数和模型层面综合考虑的方法,事实上很多方法并不单纯地属于变模型和变参数方法,因为它们总是有一个缓慢的学习和一个快速的学习机制来增量式学习.因此,大多数持续学习方法并不依赖于单一策略来解决灾难性遗忘.正如上文所述,每种方法都有优缺点,但大多数情况下,混合方法可以找到一个效果最好的解决策略,用于平衡模型的稳定性与可塑性.

#### 1) 正则化与回放混合方法.

在混合方法中,结合正则化策略和回放策略是最为常见的一类方法. Rebuffi等<sup>[83]</sup>提出了基于增量分类器和表示学习的方法(incremental classifier and representation learning, iCaRL). iCaRL方法引入了一种新的训练模式——类增量学习的方式,只需要存储旧数据的少量类训练数据和少量新类的数据就可以进行学习. iCaRL的主要思路是首先构造增量分类器,通过样本的最近均值规则<sup>[84]</sup>进行分类;然后进行优先样本选择,构建一个具有代表性的样本集;最后基于知识蒸馏和原型排练的表征学习<sup>[85-86]</sup>进行特征提取,实现良好的分类效果.但是iCaRL仍未解决所

有类增量学习中的分类问题,尤其是当所有训练样本同时可用时, iCaRL方法的性能仍然有待提升.

与iCaRL同类型的方法还有Ramapuram等<sup>[87]</sup>提出的终身生成式建模(lifelong generative modeling, LGM), Castro等<sup>[88]</sup>提出的端到端增量学习(end-to-end incremental learning, EEIL), Douillard等<sup>[89]</sup>提出的PODNet等方法.

#### 2) 动态结构与正则化混合方法.

Maltoni等<sup>[90]</sup>在2019年提出了结合动态结构方法和正则化方法的AR1方法,适用于单增量任务场景.实验结果表明,AR1在CORE50和CIFAR100数据集上的表现优于LwF、EWC和SI方法. AR1方法的局限性在于其应用场景过于单一,不适用于现实环境中的复杂任务学习.

与AR1同类型的方法还有Jung等<sup>[91]</sup>提出的基于深度神经网络的少量遗忘学习(less-forgetting learning in deep neural networks, LLDNN), He等<sup>[92]</sup>提出的概念辅助反向传播(conceptor-aided backpropagation, CAB)等方法.

#### 3) 其他混合方法.

其他经典的混合方法还有Nguyen等<sup>[93]</sup>提出的变分持续学习方法(variational continual learning, VCL), Sprechmann等<sup>[94]</sup>提出的基于记忆的参数自适应方法(memory-based parameter adaptation, MPA)等.

混合方法的优势在于其良好的任务性能,但是,同样也存在模型结构复杂、计算成本高、内存占用大等不足.在增量学习的实际应用场景中,往往融合了多种算法,需要根据实际计算内存、实时性要求、精度要求等条件选择合适的混合方法.

## 2.4 小结

增量学习方法的种类较多,各有优劣势,各种增量学习方法的对比如表1所示.上述各类增量方法由于核心思想存在一定差异性,适用的场景也有所不同<sup>[36,40,62,74]</sup>,下面列举一些较为常见的可以应用增量学习方法的场景.

在变模型方法中,基于正则化的方法原理较简单,但其精度也存在一定不足.一个经典的应用场景是推荐系统,例如在广告推荐算法中,正则化方法可以用来更新推荐模型,并确保模型在适应新数据的同时不会忘记已有的用户偏好.基于回放的方法通过保存和重演历史数据来进行增量学习,通常具有较高的精度,适用于较为依赖历史数据的场景,例如语音识别系统、自动驾驶系统的训练.变模型结构方法允许模型的结构在学习过程中进行显式或隐式的动态

表1 增量学习方法优劣势对比

| 分类    | 子类      | 优点                | 缺点                            |
|-------|---------|-------------------|-------------------------------|
| 变参数方法 | 正则化     | 占用内存空间少,实现较简单     | 学习的任务数少,模型记忆存在退化,训练成本较高,精度较低  |
|       | 排练式回放   | 较好地保持模型记忆,学习效果较好  | 对数据量要求较高,数据隐私性差               |
|       | 生成式回放   | 可以生成数据,只需少量示教数据   | 依赖生成网络,不适用于数据特征复杂情形           |
|       | 基于元学习回放 | 对新任务学习能力强,所需数据量少  | 训练成本高,难收敛                     |
| 变模型方法 | 显式动态结构  | 原理简单,学习效果较好,可拓展性强 | 对计算内存要求高,随着任务和数据的增加,参数可能呈二次增长 |
|       | 隐式动态结构  | 模型容量小,学习速度快,容易收敛  | 可伸缩性差,在多类增量学习任务中表现差           |
| 混合方法  |         | 任务性能良好,鲁棒性强       | 模型结构复杂,计算成本高,内存占用大            |

调整,以适应新的数据和任务要求,一个适用的场景是自然语言处理,变模型方法可以在实时的文本数据流中不断调整模型结构,进行知识的积累.混合方法将多种增量学习技术结合在一起,以充分利用它们的优点和应对不同的学习场景,因此,混合方法适用于较为复杂的应用场景,例如机器人系统、金融风控系统等.

### 3 数据集与评价指标

本节主要介绍增量学习中常用的数据集与评价指标.

#### 3.1 数据集

目前,在分类任务中,增量学习领域的数据集大多来源于其他领域.对MNIST<sup>[95]</sup>数据集进行分割、人工修改(例如图像旋转),产生了Split MNIST<sup>[45,96]</sup>、Rotation MNIST<sup>[52]</sup>、Fashion MNIST<sup>[97]</sup>、Permutation MNIST<sup>[98]</sup>等数据集,可应用于不同的增量学习场景.一些比较著名的分类数据集也被应用于增量学习中,例如SVHN<sup>[99]</sup>、CIFAR10/100<sup>[100]</sup>、USPS<sup>[101]</sup>、CUB200<sup>[102]</sup>、LSUN<sup>[103]</sup>、ImageNet<sup>[104]</sup>等.

以上数据集都只适合于虚拟场景或理论场景中的增量学习任务,实际机器人场景下的数据集目前仍然较少.CORE50<sup>[105]</sup>是一个连续目标识别数据集,提供了多类别、多实例的数据,如图2所示.但其缺点在于数据规模太小,无法用于更加复杂任务的训练.MEx是一个人体运动数据集<sup>[106]</sup>,可用于机器人行为分析的训练,其数据包含多个模态,很适合于机器人增量学习任务.Push Dataset是一个机器人推动任务数据集<sup>[107]</sup>,在多个维度对实际机器人推挤动作的数据进行采集,对机器人多种物体推动技能的学习过程有较高的适用性.ACRONYM是一个机器人物品抓取数据集<sup>[108]</sup>,适用于机器人多类型物体抓取任务的训练.Pasquale等<sup>[109]</sup>和Azagra等<sup>[110]</sup>通过采集机器人操作时的实际图像制作数据集,用于分类和检测任务的训练.



图2 CORE50数据集示意图<sup>[105]</sup>

此外,还有一些在多种机器人应用场景中具有代表性的数据集.例如:适用于机器人问答与对话任务训练的SQuAD数据集<sup>[111]</sup>和Movie-Dialogs Corpus数据集<sup>[112]</sup>,适用于机器人视觉感知与自动驾驶的KITTI数据集<sup>[113]</sup>,适用于机器人动作识别的PAMAP2数据集<sup>[114]</sup>等.以上数据集在机器人增量学习的训练中较为常见,适用程度较高.

#### 3.2 评价指标

机器人增量学习的最终目的是让机器人学习序列化的任务,在掌握新技能的基础上不遗忘旧技能.对增量学习方法的评价指标可以从准确率、模型规模、计算效率几个方面考虑.

任务的准确率应当首先考虑.Lopez-Paz等<sup>[52]</sup>提出了平均准确率(average accuracy)、正向迁移能力(forward transfer)以及反向迁移能力(backward transfer),即

$$A_{c_k} = \frac{1}{k} \sum_{i=1}^k R_{k,i}, \quad (9)$$

$$F_{WT_k} = \frac{1}{k-1} \sum_{i=2}^k (R_{i-1,i} - \bar{b}_i), \quad (10)$$

$$B_{WT_k} = \frac{1}{k-1} \sum_{i=1}^{k-1} (R_{k,i} - R_{i,i}). \quad (11)$$

其中: $R_{i,j}$ 表示在任务 $t_i$ 的所有数据都输入模型之后任务 $t_j$ 的测试精度, $b_i$ 表示任务 $t_i$ 在经过随机初始化之后的测试精度, $\bar{b}_i$ 表示 $b_i$ 的平均值, $k$ 表示当前学习

的任务编号。

$A_{c_k}$  为平均准确率,代表模型在当前  $k$  个任务上的平均性能;  $F_{WT_k}$  为正向迁移能力,代表模型学习前  $k-1$  项旧任务之后对第  $k$  个任务的性能影响;  $B_{WT_k}$  为反向迁移能力,代表模型学习第  $k$  项任务之后对前  $k-1$  项旧任务的性能影响。在两个方法的平均准确率相近的情况下,应当考虑正向和反向迁移能力以衡量方法的优劣。

其次,模型的规模也值得重视,特别是对于嵌入式系统和机器人而言,某种方法占用的计算和内存资源大小是很重要的。Díaz-Rodríguez 等<sup>[115]</sup> 提出衡量模型容量的指标主要为模型大小,即

$$M_S = \min \left( 1, \frac{\sum_{i=1}^N \text{Mem}(\theta_i)}{N} \right). \quad (12)$$

其中:  $\text{Mem}(\theta_i)$  表示模型在任务  $i$  中所需的容量大小,  $\theta_i$  为任务  $i$  的内部参数,  $N$  为任务数量。

$M_S$  为模型大小,衡量了一个模型的总体容量,总体容量越大,说明对计算内存和硬件要求越高。在实际应用场景中,应根据实际情况选择合适的增量学习算法。

最后,计算效率也是不可忽视的性能指标。在现实场景下,往往需要机器人实时作业,因此,对机器人的反应速度有一定要求。具有较高计算效率的算法才能满足这一需求。Lesort 等<sup>[8]</sup> 认为,计算效率可以表示为

$$C_E = \min \left( 1, \frac{\sum_{i=1}^N \frac{O_{ps} \uparrow \downarrow (T_{r_i}) \cdot \varepsilon}{1 + O_{ps}(T_{r_i})}}{N} \right). \quad (13)$$

其中:  $T_{r_i}$  表示任务  $i$  的训练集;  $O_{ps}(T_{r_i})$  表示模型在  $T_{r_i}$  上训练的总操作数(相加操作);  $O_{ps} \uparrow \downarrow (T_{r_i})$  表示模型在  $T_{r_i}$  上进行一次前向和反向传播的操作数;  $\varepsilon$  为尺度因子,根据实际训练的回合数进行调整;  $N$  为任务数量。

$C_E$  为计算效率,其值越高,代表模型的实时性能越好,在复杂环境中的表现也会越好。

## 4 未来研究趋势

本节对机器人增量学习技术的未来研究趋势进行展望。

1) 知识的正确性判别对增量学习模型而言至关重要<sup>[4,116]</sup>。一个合格的增量学习模型应能判断出知识的正确性。因为增量学习需要利用过去的知识来帮助新知识的学习,而错误的旧知识显然会影响后续的学习过程。所以必须在很大程度上解决或缓解

这一问题,才能保证训练得到的增量学习模型是有效的。对知识的正确性进行判别这一研究方向显然是充满挑战性的。

2) 如何对知识进行表示和推理是一个值得重点关注的问题<sup>[4,117]</sup>。知识表示能力代表模型可以对知识进行抽象处理,并从中提取出一般规律,现在亟待解决的问题包括什么样的知识是重要的,如何表示它们,以及什么样的推理能力对增量学习模型有用。到目前为止,很少有研究在增量学习的背景下解决这些问题。现有增量学习模型中的知识主要基于特定学习算法或应用的直接需求来表示,不具备推理能力,这有待于进一步研究。

3) 增量学习模型应具有可伸缩性或扩展能力,即模型不仅可以在不同规模的数据集之间保持较好的性能,而且能够对不断增加的任务保持较好的鲁棒性。尽管如此,当前的主要方法都只能在一定范围内进行伸缩,而无法达到增量学习理想状态下的无限拓展能力<sup>[118-119]</sup>,这将是一个具有重大研究价值的方向。

4) 在增量学习过程中,模型的稳定性与可塑性往往不能兼得。稳定性差,会使得模型容易遗忘旧知识;可塑性差,会使得模型不太容易学习新的知识。近年来,部分研究工作<sup>[120-122]</sup>有了一定的进展,但远没有彻底解决该问题。因此,如何在模型的稳定性与可塑性之间进行权衡是一个比较重要的研究方向。

5) 近年来增量学习尽管取得了一些进展,但针对实际机器人工作场景应用的研究工作仍然较少。目前,大多数针对机器人应用的增量学习的研究主要集中在仿真环境中,真实场景的增量学习方法验证也只是在一些简单的任务场景下进行实验<sup>[123-125]</sup>。相比于虚拟场景,现实场景的复杂性和干扰性极大,因此,如何将增量学习应用于实际的机器人工作场景是一个颇具价值的研究方向。

6) 在特定的增量学习方法中,也有如下一些具有挑战性的问题:

① 正则化方法在增量学习中的应用较为广泛<sup>[126-128]</sup>,其普遍存在的问题是模型的精度较低,对于多任务工作场景的适应性较差<sup>[39-40]</sup>。因此,对于正则化方法的未来研究方向主要集中在如何提高模型精度以及自适应能力上。

② 回放方法通过存储过去的经验数据来实现更好的模型性能,但是,大量的经验需要更大的容量来存储,而容量的饱和和限制了这一类方法的进一步发展<sup>[129-130]</sup>,尤其是在实际应用于机器人时容量的限制是必须要考虑的现实因素。因此在容量有限的前提

下,尽可能提高模型的存储效率是一个亟待解决的问题。

③ 隐式与显式动态结构方法的兼容性问题一直存在<sup>[79-80]</sup>。二者在原理上的设计思路是完全不一致的。显式方法的局限性主要在于计算内存的限制,而隐式方法的局限性在于模型的表达能力有限。如果能将二者结合使用,则动态结构方法的模型性能会有较大的提升。

④ 混合方法在模型的训练效率上一直存在着较为明显的劣势<sup>[83,90]</sup>,此外,由于混合方法组合了多种算法,其模型也会更为复杂,导致实时性较差。因此,未来对于混合方法的研究应更加注重于如何提高模型的训练效率以及端到端的部署。

## 5 结 论

增量学习力图解决当前机器人在动态变化的环境中适应能力较差的问题,模拟人类的学习过程,使得机器人能够学习序列化且具有非独立同分布训练数据的多个任务,最终能够掌握新技能,并且不会遗忘旧技能。本文首先对增量学习进行了简要说明;其次,将当前的机器人增量学习主流方法从宏观上分为变参数方法、变模型方法、混合方法3类,并对每一类进行了详细论述;然后介绍了增量学习中常用的数据集和评价指标;最后整理了增量学习的关键问题,提出了一些有价值的未来研究方向。相信增量学习将会拥有越来越广阔的应用前景,在未来的研究和实际生产中必将发挥重要作用。

## 参考文献(References)

- [1] Abdel Kareem Jaradat M, Al-Rousan M, Quadan L. Reinforcement based mobile robot navigation in dynamic environment[J]. *Robotics and Computer-Integrated Manufacturing*, 2011, 27(1): 135-149.
- [2] Jiang Y Q, Walker N, Hart J, et al. Open-world reasoning for service robots[J]. *Proceedings of the International Conference on Automated Planning and Scheduling*, 2021, 29: 725-733.
- [3] Sun Y Y, Fox D. NEOL: Toward never-ending object learning for robots[C]. 2016 IEEE International Conference on Robotics and Automation. Stockholm, 2016: 1621-1627.
- [4] Chen Z, Liu B. Lifelong machine learning[J]. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 2018, 12(3): 1-207.
- [5] de Lange M, Aljundi R, Masana M, et al. A continual learning survey: Defying forgetting in classification tasks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(7): 3366-3385.
- [6] Parisi G I, Kemker R, Part J L, et al. Continual lifelong learning with neural networks: A review[J]. *Neural Networks*, 2019, 113: 54-71.
- [7] Qu H, Rahmani H, Xu L, et al. Recent advances of continual learning in computer vision: An overview[J/OL]. 2021, arXiv: 2109.11369.
- [8] Lesort T, Lomonaco V, Stoian A, et al. Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges[J]. *Information Fusion*, 2020, 58: 52-68.
- [9] van de Ven G M, Tolia A S. Three scenarios for continual learning[J/OL]. 2019, arXiv: 1904.07734.
- [10] Clegg B A, DiGirolamo G J, Keele S W. Sequence learning[J]. *Trends in Cognitive Sciences*, 1998, 2(8): 275-281.
- [11] Aljundi R, Rohrbach M, Tuytelaars T. Selfless sequential learning[J/OL]. 2018, arXiv: 1806.05421.
- [12] Thrun S, Mitchell T M. Lifelong robot learning[J]. *Robotics and Autonomous Systems*, 1995, 15(1/2): 25-46.
- [13] Mermillod M, Bugaiska A, Bonin P. The stability-plasticity dilemma: Investigating the continuum from catastrophic forgetting to age-limited learning effects[J]. *Frontiers in Psychology*, 2013, 4: 504.
- [14] Abraham W C, Robins A. Memory retention—The synaptic stability versus plasticity dilemma[J]. *Trends in Neurosciences*, 2005, 28(2): 73-78.
- [15] McClelland J L, McNaughton B L, O'Reilly R C. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory[J]. *Psychological Review*, 1995, 102(3): 419-457.
- [16] Kumaran D, Hassabis D, McClelland J L. What learning systems do intelligent agents need? Complementary learning systems theory updated[J]. *Trends in Cognitive Sciences*, 2016, 20(7): 512-534.
- [17] Hinton G E, Plaut D C. Using fast weights to deblur old memories[C]. *Proceedings of the 9th Annual Conference of the Cognitive Science Society*. Hillsdale, 1987: 177-186.
- [18] Pan S J, Yang Q. A survey on transfer learning[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2010, 22(10): 1345-1359.
- [19] Torrey L, Shavlik J. Transfer learning[C]. *Handbook of Research on Machine Learning Applications and Trends*. Hershey, 2010: 242-264.
- [20] Taylor M E, Stone P. Transfer learning for reinforcement learning domains: A survey[J]. *Journal of Machine Learning Research*, 2009, 10: 1633-1685.

- [21] Zhang Y, Yang Q. An overview of multi-task learning[J]. *National Science Review*, 2018, 5(1): 30-43.
- [22] Kivinen J, Smola A J, Williamson R C. Online learning with kernels[J]. *IEEE Transactions on Signal Processing*, 2004, 52(8): 2165-2176.
- [23] Herbster M, Pontil M, Wainer L. Online learning over graphs[C]. *Proceedings of the 22nd International Conference on Machine Learning*. Bonn, 2005: 305-312.
- [24] Bengio Y, Louradour J, Collobert R, et al. Curriculum learning[C]. *Proceedings of the 26th Annual International Conference on Machine Learning*. Montreal, 2009: 41-48.
- [25] Wang X, Chen Y D, Zhu W W. A survey on curriculum learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(9): 4555-4576.
- [26] Thrun S, Pratt L. *Learning to learn*[M]. Berlin: Springer Science & Business Media, 2012: 1-212.
- [27] Vilalta R, Drissi Y. A perspective view and survey of meta-learning[J]. *Artificial Intelligence Review*, 2002, 18(2): 77-95.
- [28] Vanschoren J. Meta-learning: A survey[J/OL]. 2018, arXiv: 1810.03548.
- [29] Mundt M, Hong Y, Pliushch I, et al. A wholistic view of continual learning with deep neural networks: Forgotten lessons and the bridge to active and open world learning[J/OL]. 2020, arXiv: 2009.01797.
- [30] Hadsell R, Rao D, Rusu A A, et al. Embracing change: Continual learning in deep neural networks[J]. *Trends in Cognitive Sciences*, 2020, 24(12): 1028-1040.
- [31] Wang L, Zhang X, Su H, et al. A comprehensive survey of continual learning: Theory, method and application[J/OL]. 2023, arXiv: 2302.00487.
- [32] 朱飞, 张煦尧, 刘成林. 类别增量学习研究进展和性能评价[J]. *自动化学报*, 2023, 49(3): 635-660.  
(Zhu F, Zhang X Y, Liu C L. Class incremental learning: A review and performance evaluation[J]. *Acta Automatica Sinica*, 2023, 49(3): 635-660.)
- [33] 韩亚楠, 刘建伟, 罗雄麟. 连续学习研究进展[J]. *计算机研究与发展*, 2022, 59(6): 1213-1239.  
(Han Y N, Liu J W, Luo X L. Research progress of continual learning[J]. *Journal of Computer Research and Development*, 2022, 59(6): 1213-1239.)
- [34] 宋锐, 李凤鸣, 权威, 等. 多约束条件下机器人柔性装配技能自学习[J]. *控制与决策*, 2022, 37(5): 1329-1337.  
(Song R, Li F M, Quan W, et al. Flexible assembly skill self-learning of robot under multiple constraints[J]. *Control and Decision*, 2022, 37(5): 1329-1337.)
- [35] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. New York: ACM, 2014: 580-587.
- [36] Donahue J, Jia Y Q, Vinyals O, et al. DeCAF: A deep convolutional activation feature for generic visual recognition[J/OL]. 2013, arXiv: 1310.1531.
- [37] Caruana R. Multitask learning[J]. *Machine Learning*, 1997, 28(1): 41-75.
- [38] Furlanello T, Zhao J, Saxe A M, et al. Active long term memory networks[J/OL]. 2016, arXiv: 1606.02355.
- [39] Li Z Z, Hoiem D. Learning without forgetting[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(12): 2935-2947.
- [40] Rannen A, Aljundi R, Blaschko M B, et al. Encoder based lifelong learning[C]. *2017 IEEE International Conference on Computer Vision*. Venice, 2017: 1329-1337.
- [41] Ayub A, Wagner A R. Tell me what this is: Few-shot incremental object learning by a robot[C]. *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. New York: ACM, 2020: 8344-8350.
- [42] Isele D, Luna J M, Eaton E, et al. Lifelong learning for disturbance rejection on mobile robots[C]. *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Daejeon, 2016: 3993-3998.
- [43] Kirkpatrick J, Pascanu R, Rabinowitz N, et al. Overcoming catastrophic forgetting in neural networks[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2017, 114(13): 3521-3526.
- [44] Kemker R, McClure M, Abitino A, et al. Measuring catastrophic forgetting in neural networks[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018, 32(1): 3390-3398.
- [45] Zenke F, Poole B, Ganguli S. Continual learning through synaptic intelligence[C]. *Proceedings of the 34th International Conference on Machine Learning (ICML)*. Sydney, 2017: 3987-3995.
- [46] Liu X L, Masana M, Herranz L, et al. Rotate your networks: Better weight consolidation and less catastrophic forgetting[J/OL]. 2018, arXiv: 1802.02950.
- [47] Schwarz J, Czarnecki W, Luketina J, et al. Progress & compress: A scalable framework for continual learning[C]. *Proceedings of the 35th International Conference on Machine Learning (ICML)*. Stockholm, 2018: 4528-4537.
- [48] Chaudhry A, Dokania P K, Ajanthan T, et al. Riemannian walk for incremental learning: Understanding forgetting and intransigence[J/OL]. 2018, arXiv: 1801.10112.
- [49] Valipour S, Perez C, Jagersand M. Incremental learning for robot perception through HRI[C]. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Vancouver, 2017: 2772-2777.

- [50] Calinon S, Billard A. Incremental learning of gestures by imitation in a humanoid robot[C]. Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction. Arlington, 2007: 255-262.
- [51] 王磊, 苏中, 乔俊飞, 等. 基于增量式学习的正则化回声状态网络[J]. 控制与决策, 2022, 37(3): 661-668.  
(Wang L, Su Z, Qiao J F, et al. Design of incremental regularized echo state network[J]. Control and Decision, 2022, 37(3): 661-668.)
- [52] Lopez-Paz D, Ranzato M. Gradient episodic memory for continual learning[C]. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, 2017: 6470-6479.
- [53] Chaudhry A, Ranzato M A, Rohrbach M, et al. Efficient lifelong learning with a-gem[J/OL]. 2018, arXiv: 1812.00420.
- [54] Masana M, Liu X L, Twardowski B, et al. Class-incremental learning: Survey and performance evaluation on image classification[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(5): 5513-5533.
- [55] Gepperth A, Karaoguz C. A bio-inspired incremental learning architecture for applied perceptual problems[J]. Cognitive Computation, 2016, 8(5): 924-934.
- [56] Liu B, Xiao X, Stone P. A lifelong learning approach to mobile robot navigation[J]. IEEE Robotics and Automation Letters, 2021, 6(2): 1090-1096.
- [57] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [58] Shin H, Lee J K, Kim J, et al. Continual learning with deep generative replay[C]. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, 2017: 2994-3003.
- [59] Wu C, Herranz L, Liu X, et al. Memory replay gans: Learning to generate new categories without forgetting[C]. Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS). Montreal, 2018: 5966-5976.
- [60] Ostapenko O, Puscas M, Klein T, et al. Learning to remember: A synaptic plasticity driven framework for continual learning[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, 2020: 11313-11321.
- [61] Ayub A, Wagner A R. Learning to encode and regenerate images for continual learning[J/OL]. 2021, arXiv: 2101.04904.
- [62] Kemker R, Kanan C. Fearnnet: Brain-inspired model for incremental learning[J/OL]. 2017, arXiv: 1711.10563.
- [63] Gopalakrishnan S, Singh P R, Fayek H, et al. Knowledge capture and replay for continual learning[C]. 2022 IEEE/CVF Winter Conference on Applications of Computer Vision. Waikoloa, 2022: 337-345.
- [64] Riemer M, Klinger T, Bouneffouf D, et al. Scalable recollections for continual lifelong learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 1352-1359.
- [65] Pfülb B, Gepperth A, Bagus B. Continual learning with fully probabilistic models[J/OL]. 2021, arXiv: 2104.09240.
- [66] Ye F, Bors A G. Learning latent representations across multiple data domains using lifelong VAEGAN[C]. European Conference on Computer Vision. Cham: Springer, 2020: 777-795.
- [67] Soh H, Demiris Y. Incrementally learning objects by touch: Online discriminative and generative models for tactile-based recognition[J]. IEEE Transactions on Haptics, 2014, 7(4): 512-525.
- [68] Liu X L, Wu C S, Menta M, et al. Generative feature replay for class-incremental learning[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle, 2020: 915-924.
- [69] van de Ven G M, Siegelmann H T, Tolia A S. Brain-inspired replay for continual learning with artificial neural networks[J]. Nature Communications, 2020, 11(1): 1-14.
- [70] Riemer M, Cases I, Ajemian R, et al. Learning to learn without forgetting by maximizing transfer and minimizing interference[J/OL]. 2018, arXiv: 1810.11910.
- [71] Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks[C]. Proceedings of the 34th International Conference on Machine Learning. Sydney, 2017: 1126-1135.
- [72] Henning C, Cervera M, D'Angelo F, et al. Posterior meta-replay for continual learning[C]. Proceedings of the 35th International Conference on Neural Information Processing Systems (NIPS). New York, 2021: 14135-14149.
- [73] Wang K, Liu X L, Bagdanov A, et al. Incremental meta-learning via episodic replay distillation for few-shot image recognition[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. New Orleans, 2022: 3728-3738.
- [74] Rusu A A, Rabinowitz N C, Desjardins G, et al. Progressive neural networks[J/OL]. 2016, arXiv: 1606.04671.
- [75] Aljundi R, Chakravarty P, Tuytelaars T. Expert gate: Lifelong learning with a network of experts[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 7120-7129.
- [76] Draelos T J, Miner N E, Lamb C C, et al. Neurogenesis deep learning: Extending deep networks to accommodate

- new classes[C]. 2017 International Joint Conference on Neural Networks. Anchorage, 2017: 526-533.
- [77] Yoon J, Yang E, Lee J, et al. Lifelong learning with dynamically expandable networks[J/OL]. 2017, arXiv: 1708.01547.
- [78] Mallya A, Lazebnik S. PackNet: Adding multiple tasks to a single network by iterative pruning[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 7765-7773.
- [79] Mallya A, Davis D, Lazebnik S. Piggyback: Adapting a single network to multiple tasks by learning to mask weights[J/OL]. 2018, arXiv: 1801.06519.
- [80] Serra J, Suris D, Miron M, et al. Overcoming catastrophic forgetting with hard attention to the task[C]. Proceedings of the 35th International Conference on Machine Learning (ICML). Stockholm, 2018: 4548-4557.
- [81] Fernando C, Banarse D, Blundell C, et al. Pathnet: Evolution channels gradient descent in super neural networks[J/OL]. 2017, arXiv: 1701.08734.
- [82] Zhang H, Wu P, Beck A, et al. Adaptive incremental learning of image semantics with application to social robot[J]. Neurocomputing, 2016, 173: 93-101.
- [83] Rebuffi S A, Kolesnikov A, Sperl G, et al. iCaRL: Incremental classifier and representation learning[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 5533-5542.
- [84] Mensink T, Verbeek J, Perronnin F, et al. Distance-based image classification: Generalizing to new classes at near-zero cost[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(11): 2624-2637.
- [85] Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(8): 1798-1828.
- [86] Saxena S, Verbeek J. Convolutional neural fabrics[C]. Proceedings of the 30th International Conference on Neural Information Processing Systems (NIPS). Barcelona, 2016: 4060-4068.
- [87] Ramapuram J, Gregorova M, Kalousis A. Lifelong generative modeling[J]. Neurocomputing, 2020, 404: 381-400.
- [88] Castro F M, Marín-Jiménez M J, Guil N, et al. End-to-end incremental learning[C]. Proceedings of the European Conference on Computer Vision (ECCV). Munich, 2018: 233-248.
- [89] Douillard A, Cord M, Ollion C, et al. PODNet: Pooled outputs distillation for small-tasks incremental learning[C]. European Conference on Computer Vision. Cham: Springer, 2020: 86-102.
- [90] Maltoni D, Lomonaco V. Continuous learning in single-incremental-task scenarios[J]. Neural Networks, 2019, 116: 56-73.
- [91] Jung H, Ju J, Jung M, et al. Less-forgetting learning in deep neural networks[J/OL]. 2016, arXiv: 1607.00122.
- [92] He X, Jaeger H. Overcoming catastrophic interference using conceptor-aided backpropagation[J/OL]. 2017, arXiv: 1707.04853.
- [93] Nguyen C V, Li Y, Bui T D, et al. Variational continual learning[J/OL]. 2017, arXiv: 1710.10628.
- [94] Sprechmann P, Jayakumar S M, Rae J W, et al. Memory-based parameter adaptation[J/OL]. 2018, arXiv: 1802.10542.
- [95] LeCun Y, Cortes C. MNIST handwritten digit database[EB/OL]. (1998-04-18)[2023-03-05]. <http://yann.lecun.com/exdb/mnist>.
- [96] Lesort T, Gepperth A, Stoian A, et al. Marginal replay vs conditional replay for continual learning[C]. International Conference on Artificial Neural Networks. Cham: Springer, 2019: 466-480.
- [97] Xiao H, Rasul K, Vollgraf R. Fashion-mnist: A novel image dataset for benchmarking machine learning algorithms[J/OL]. 2017, arXiv: 1708.07747.
- [98] Goodfellow I J, Mirza M, Xiao D, et al. An empirical investigation of catastrophic forgetting in gradient-based neural networks[J/OL]. 2013, arXiv: 1312.6211.
- [99] Netzer Y, Wang T, Coates A, et al. Reading digits in natural images with unsupervised feature learning[EB/OL]. (2011-01-01)[2023-03-11]. [http://ufdl.stanford.edu/housenumbers/nips2011\\_housenumbers.pdf](http://ufdl.stanford.edu/housenumbers/nips2011_housenumbers.pdf).
- [100] Krizhevsky A, Hinton G. Learning multiple layers of features from tiny images[R]. Toronto: University of Toronto, 2009: 3-47.
- [101] Hull J J. A database for handwritten text recognition research[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(5): 550-554.
- [102] Welinder P, Branson S, Mita T, et al. Caltech-UCSD birds 200[R]. California: California Institute of Technology, 2010: 1-15.
- [103] Yu F, Seff A, Zhang Y, et al. Lsun: Construction of a large-scale image dataset using deep learning with humans in the loop[J/OL]. 2015, arXiv: 1506.03365.
- [104] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [105] Lomonaco V, Maltoni D. CORe50: A new dataset and benchmark for continuous object recognition[J/OL]. 2017, arXiv: 1705.03550.
- [106] Wijekoon A, Wiratunga N, Cooper K. Mex: Multi-modal exercises dataset for human activity recognition[J/OL]. 2019, arXiv: 1908.08992.

- [107] Yu P K T, Bauza M, Fazeli N, et al. More than a million ways to be pushed: A high-fidelity experimental dataset of planar pushing[J/OL]. 2016, arXiv: 1604.04038.
- [108] Eppner C, Mousavian A, Fox D. ACRONYM: A large-scale grasp dataset based on simulation[C]. 2021 IEEE International Conference on Robotics and Automation. Xi'an, 2021: 6222-6227.
- [109] Pasquale G, Ciliberto C, Odone F, et al. Teaching iCub to recognize objects using deep Convolutional Neural Networks[J]. Machine Learning for Interactive Systems, 2015, 12: 21-25.
- [110] Azagra P, Golemo F, Mollard Y, et al. A multimodal dataset for object model learning from natural human-robot interaction[C]. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. Vancouver, 2017: 6134-6141.
- [111] Rajpurkar P, Jia R, Liang P. Know what you don't know: Unanswerable questions for SQuAD[C]. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne, 2018: 784-789.
- [112] Danescu-Niculescu-Mizil C, Lee L. Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs[J/OL]. 2011, arXiv: 1106.3077.
- [113] Geiger A, Lenz P, Stiller C, et al. Vision meets robotics: The KITTI dataset[J]. International Journal of Robotics Research, 2013, 32(11): 1231-1237.
- [114] Reiss A, Stricker D. Introducing a new benchmarked dataset for activity monitoring[C]. The 16th International Symposium on Wearable Computers. Newcastle, 2012: 108-109.
- [115] Díaz-Rodríguez N, Lomonaco V, Filliat D, et al. Don't forget, there is more than forgetting: New metrics for continual learning[J/OL]. 2018, arXiv: 1810.13166.
- [116] Chen Z, Liu B. Topic modeling using topics from many domains, lifelong learning and big data[C]. Proceedings of the 31st International Conference on Machine Learning (ICML). Beijing, 2014: 703-711.
- [117] Mitchell T, Cohen W, Hruschka E, et al. Never-ending learning[J]. Communications of the ACM, 2018, 61(5): 103-115.
- [118] Moreno-Muñoz P, Ramírez D, Artés-Rodríguez A. Continual learning for infinite hierarchical change-point detection[C]. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Barcelona, 2020: 3582-3586.
- [119] Aljundi R. Continual learning in neural networks[J/OL]. 2019, arXiv: 1910.02718.
- [120] Bailey C H, Kandel E R, Harris K M. Structural components of synaptic plasticity and memory consolidation[J]. Cold Spring Harbor Perspectives in Biology, 2015, 7(7): a021758.
- [121] Verbeke P, Verguts T. Learning to synchronize: How biological agents can couple neural task modules for dealing with the stability-plasticity dilemma[J]. PLoS Computational Biology, 2019, 15(8): e1006604.
- [122] Lee A, Gomes H M, Zhang Y Q. Balancing the stability-plasticity dilemma with online stability tuning for continual learning[C]. 2022 International Joint Conference on Neural Networks. Padua, 2022: 1-8.
- [123] Lebichot B, Marco Paldino G, Bontempi G, et al. Incremental learning strategies for credit cards fraud detection: Extended abstract[C]. 2020 IEEE 7th International Conference on Data Science and Advanced Analytics. Sydney, 2020: 785-786.
- [124] Traoré R, Caselles-Dupré H, Lesort T, et al. Continual reinforcement learning deployed in real-life using policy distillation and sim2real transfer[J/OL]. 2019, arXiv: 1906.04452.
- [125] Sadhu S, Hermansky H. Continual learning in automatic speech recognition[J/OL]. 2021, arXiv: 2112.09427.
- [126] Liu Y, Parisot S, Slabaugh G, et al. More classifiers, less forgetting: A generic multi-classifier paradigm for incremental learning[C]. European Conference on Computer Vision. Cham: Springer, 2020: 699-716.
- [127] Lee C C, Lin Y C, Lin H T, et al. SERIL: Noise adaptive speech enhancement using regularization-based incremental learning[J/OL]. 2020, arXiv: 2005.11760.
- [128] Gok E C, Yildirim M O, Kilickaya M, et al. Adaptive regularization for class-incremental learning[J/OL]. 2023, arXiv: 2303.13113.
- [129] Rolnick D, Ahuja A, Schwarz J, et al. Experience replay for continual learning[J/OL]. 2018, arXiv: 1811.11682.
- [130] Isele D, Cosgun A. Selective experience replay for lifelong learning[C]. Proceedings of the AAAI Conference on Artificial Intelligence. New Orleans, 2018: 3302-3309.

## 作者简介

马旭淼(1998—), 男, 博士生, 从事机器人技能学习、增量学习、强化学习、自动装配等研究, E-mail: maxumiao2021@ia.ac.cn;

徐德(1965—), 男, 研究员, 博士生导师, 从事机器人视觉测量、视觉控制、智能控制、视觉定位、显微视觉、微装配、技能学习等研究, E-mail: de.xu@ia.ac.cn.