

机器人持续学习进展与展望

赵超^{1,2}, 许杰^{1,2}, 陈星宇^{1,2}, 梅魁志^{1,2}, 兰旭光^{1,2}

(1. 西安交通大学人工智能与机器人研究所, 陕西 西安 710049;

2. 西安交通大学人工智能学院, 陕西 西安 710049)

摘要:当前机器人技术面临的较大限制之一是难以适应不断变化的任务,当机器人面对新环境或者学习新任务时,会不可避免地遗忘旧环境或旧任务的经验。为了总结机器人持续学习的研究和发展现状,首先介绍了持续学习的框架和评价基准,然后阐述了持续学习在机器人任务中的必要性和面临的挑战,并对持续学习的发展现状进行了回顾,最后展望了机器人持续学习的发展前景,提出了一些有价值的研究问题。

关键词:持续学习;终身学习;机器人

中图分类号:TP242.6

文献标志码:A

doi: 10.11959/j.issn.2096-6652.202235

A review of continual learning for robotics

ZHAO Chao^{1,2}, XU Jie^{1,2}, CHEN Xingyu^{1,2}, MEI Kuizhi^{1,2}, LAN Xuguang^{1,2}

1. Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China

2. College of Artificial Intelligence, Xi'an Jiaotong University, Xi'an 710049, China

Abstract: One of the limitations of robotics is that it is difficult for robots to adapt to fickle tasks. A robot will inevitably forget the knowledge from old environments or tasks when facing new environments or tasks. In order to summarize research in continual learning for robotics, firstly, the framework and evaluation protocols in continual learning were introduced. And then necessity and challenge of continual learning in the robotics were expounded. The research for continual learning was also summarized. Finally, the prospect of continual learning was predicted and some valuable research directions were put forward.

Key words: continual learning, lifelong learning, robotics

0 引言

机器人是一种融合了人工智能和自动控制领域知识的智能系统,能够代替人类在复杂的环境中进行作业。得益于机器人的快速发展和硬件设备的不断更迭,机器人已在众多领域获得不俗的表现。例如,仿造人类双手操作灵活性的机器人灵巧手、量产落地的自动驾驶汽车等。任何机器人都可以被

描述为从之前时刻的动作和感知信号到下一时刻的动作和感知信号的一组映射的实现^[1]。与人类类似,机器人需要在充满复杂物体和复杂社会环境(涉及人类、动物和其他机器人)的具有高度挑战性的环境中评估策略的优劣,并执行最优动作^[2]。

自主机器人技术最初的目标是建立可以在真实环境中无须人工干预即可完成特定任务的物理系统。尽管在过去10年中机器人技术取得了相当

收稿日期: 2022-06-01; 修回日期: 2022-06-30

通信作者: 梅魁志, meikuizhi@mail.xjtu.edu.cn

基金项目: 国家自然科学基金资助项目(No.62125305, No.62076193, No.62088102); 科技创新2030—“新一代人工智能”重大项目(No.2021ZD0112700); 中央高校基础研究基金资助项目(No.xtr072022001)

Foundation Items: The National Natural Science Foundation of China (No.62125305, No.62076193, No.62088102), The National Key Research and Development Program of China (No.2021ZD 0112700), The Fundamental Research Funds for the Central Universities (No.xtr072022001)

大的进展,但这一目标尚未实现^[3]。当前机器人技术较大的限制之一是缺乏对不断变化的环境和任务的适应性。虽然一些机器人已经成功地应用于动态和复杂感知的环境中,但它们的适应性和容错能力通常受到人类对机器人可能遇到的场景的预测的限制。由于机器人需要在真实和潜在未知的环境中工作,要预先编程机器人可能面临的所有环境是不现实的。应对这些挑战的唯一方法是让机器人拥有适应环境的能力,并发展恰当的策略,以适应可能不断变化的任务^[4]。此外,目前大多数机器学习方法是基于离线任务进行实验的,即可以同时获得所有任务的数据。在这样的设定下,研究者可以直接利用所有任务的数据设计出在当前所有任务上性能较佳的算法。然而实际场景中的很多情形不满足这种条件。首先,所有任务的数据可能不能同时获得,且在时间上呈现出序列化的特点。其次,由于隐私、学习效率和存储容量的限制,人们难以存储旧任务的数据。此外,由于任务属性的差异,不同任务的数据分布可能会发生变化,难以一次性学习到通用的表征。为了解决上述问题,研究者提出了持续学习(continual learning)的概念。

在机器人领域中,持续学习可以让机器人在学习序列任务的过程中,获得保留旧任务经验并加速新任务学习的能力。持续学习符合人类大脑学习事物的机制,例如婴儿从出生开始认知和学习这个世界,并不是一次可以获得所有信息的,而是逐渐地认知各种物体,学习各种技能。此外,人类大脑学习的一个重要特点是当学习新事物时不会遗忘过去所学的知识 and 技能,这些过去学习的知识和技能反而有助于新事物的学习。持续地学习新任务而不遗忘旧任务的经验,是持续学习区别于其他机器学习方法的特性,同时也是其他方法所不具备的能力。

机器人学习是一个非常适合持续学习方法的场景。许多机器人应用场景需要通过动态地探索来获取数据,因此场景中的任务和数据呈现序列化的特点,例如机器人寻路任务,如图1所示^[5],

机器人在新环境中探索时需要不遗忘旧环境的经验。由于机器人还需要和环境交互,机器人的行为可能会导致环境状态的变化,从而导致数据分布的偏移。此外,机器人的动作对环境产生的改变可能使得环境难以恢复到过去的状态,使得过去的数据不可得,同时隐私、学习效率和硬件限制也使得机器人设备难以存储过去所有任务的数据。因此,如何让机器人获得持续学习的能力已成为一个重要且有意义的研究方向。

虽然过去已有一些学者^[6-10]对持续学习的发展进行了总结,但是和本文有一些明显的区别。例如,Hadsell R等人^[6]和Parisi G I等人^[7]研究和总结了基于神经网络的持续学习方法,而没有关注机器人领域的持续学习;De Lange M等人^[8]则研究了如何在分类问题上解决遗忘问题,而没有考虑对任务类型具有普适性的持续学习;回放是一种将旧任务样本或表征选择性存储并加入新任务学习的持续学习方法,因其稳定的性能受到众多学者的研究,Verwimp E等人^[9]聚焦于基于回放的持续学习,总结了该类方法的发展,还详细地分析了回放方法的优点和缺陷,该综述虽然对回放方法研究得非常深入,但偏于一隅,不利于读者了解持续学习领域发展的全貌;Lesort T等人^[10]研究了机器人领域的持续学习,指出目前的研究对持续学习做了诸多限制(任务数量有限、任务标识可知和评价基准简单等),并对持续学习的发展给出了缓解限制因素的建议,但并没有提出持续学习在机器人领域未来的研究方向,此外近年来持续学习领域发展迅速,新方法体系层出不穷,该文献已不能全面概括机器人持续学习的发展现状。综上所述,笔者认为对机器人持续学习研究现状进行全面总结具有客观的必要性。因此,本文立足于机器人持续学习最新的研究现状,对该领域进行全面的总结与分析,并基于面临的挑战和启发性研究提出了一些有意义的研究方向。

为了更好地研究机器人持续学习,提供一个高

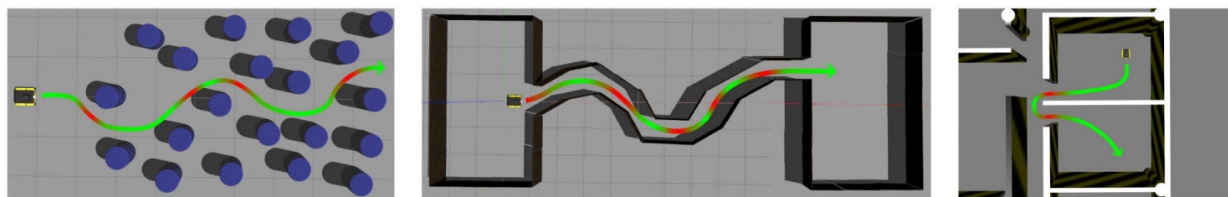


图1 机器人在不同环境下的寻路任务

效了解持续学习的途径,本文首先介绍了持续学习的背景以及面临的挑战,提出了持续学习的框架;然后总结了机器人持续学习方法;最后,对机器人持续学习未来的研究方向进行了展望,并给出了自己的见解以供读者参考。

1 持续学习

本节首先介绍了持续学习框架,然后阐明了持续学习与其他术语的区别,最后讲述了持续学习的评价基准与指标。本文使用的主要符号的含义见表1。

表1 本文使用的主要符号的含义

符号	含义
T_i	第 i 个任务
D_i	第 i 个任务的数据集
T	任务数量
$R_{i,j}$	模型学习完任务 T_i 后在任务 T_j 上的测试准确率
$\bar{b}_i$	随机初始化的任务 T_i 的测试准确率

1.1 持续学习框架

持续学习是指在无限的数据流中学习新知识的同时,不遗忘旧知识,从而让模型拥有持续获取知识的能力。这些旧知识通常还可以用来提高模型学习新知识的速度。在一些研究中,持续学习也被称为增量学习 (incremental learning)^[11]、终身学习 (lifelong learning)^[12]、永不终止学习 (never-ending learning)^[13]和序列学习 (sequential learning)^[14]。本文将这些同义的术语统一称为持续学习。

任务和数据流是持续学习中的两个比较重要的概念,不同学者对持续学习中任务和数据流的设定也不尽相同。本节对任务和数据流进行了详细阐述,以帮助理解和全面地对任务和数据流的设定进行分类。

1.1.1 任务

本文将某一时刻同时可以获得的数据和标签的集合称为任务。如图2所示,当获得当前任务 T_n 的数据集时,使用这些数据对模型进行训练,测试时需要在当前和先前所有任务 $T_1, \dots, T_n$ 的数据上对模型进行评估,并保证性能,即要求模型学习新任务时不遗忘旧任务的经验。

不同学者在研究持续学习时对任务的设定不尽相同,目前主要有3种类型:数据增量型、任务不可重复型和任务可重复型。数据增量型指持续学习中数据流对应的任务目标相同,但数据

的分布或者类型始终变化。例如,导航机器人在巡航时,场景的数据分布会改变。任务不可重复型指持续学习中先前出现过的任务不会在后续任务序列中重复出现。类别增量分类学习是典型的任务不可重复型的持续学习,其中每个任务学习一种或多种新的类别并进行分类。任务可重复型指同一种任务在未来的数据流中可能还会出现的持续学习情形。例如,机器人抓取物体可能会遇到不同类别的物体以及同类别不同形状的物体。此外,任务标识是用于区分不同任务的标签,通常由人工标注或者由任务的属性获得,然而一些学者^[15]认为已知任务的标识不符合持续学习的实际场景,在实际应用中任务的标识可能是不可知的,因此研究者还探索了任务标识不可知的持续学习。

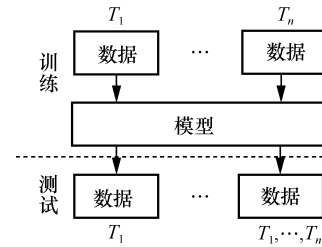


图2 持续学习

1.1.2 数据流

数据流 $S = \{D_i\}_{i=1}^{\infty}$ 是由序列任务的数据集组成的集合,根据任务的顺序依次获得。其中 D_i 是任务 T_i 的数据,在视觉任务中它可以是数据标签对的集合,在机器人任务中它可以是马尔可夫决策过程。持续学习中的数据流同样存在不同的设定。首先,数据流在设定上的主要分歧是旧任务的数据是否会在新任务中出现。例如在分割增量学习中,需要分割的图片会在多个任务中出现,但每个任务只分割图片中的某一类,其他类的区域被标注为背景;而在类别增量分类学习中,旧任务的数据通常不会在新任务中重复出现。此外,旧任务的数据是否可存储并用于之后的任务学习中也是一个存在分歧的地方。不可否认的是,存储每个任务中的代表性样本并用于后续任务的学习可以有效地兼顾模型在不同任务上的表现,但是这引入了隐私风险和额外的存储开销。

1.2 与其他术语的区别

1.2.1 多任务学习

多任务学习^[16]是指模型同时学习多个任务的

场景,如图3所示,所有任务 $T_1, \dots, T_n$ 的数据是一次性获得的,测试也是同时所有任务 $T_1, \dots, T_n$ 上进行的。模型可以兼顾多个任务进行学习,因此可以在多个任务上获得不错的表现。但多任务学习的可拓展性差,当需要学习新的任务时,难以兼顾在旧任务和新任务上的表现,虽然重新训练模型能解决这个问题,但这意味着极大的计算开销和时间成本。而持续学习无须从头开始训练模型,在新任务上训练的同时也能保证模型在旧任务上的性能。

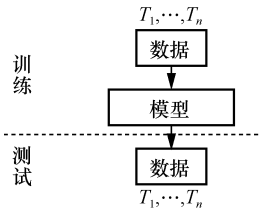


图3 多任务学习

1.2.2 迁移学习

迁移学习^[17]是指将模型在源任务上获得的经验迁移到目标任务上,帮助高效地学习目标任务的场景。如图4所示,源任务 T_1 和目标任务 T_2 通常具有相似性,迁移学习只关心模型在目标任务 T_2 上的表现,而不关心模型是否在源任务 T_1 上保持性能。而持续学习除了需要让模型在当前任务上表现较好,还需要保持在之前任务上的性能。

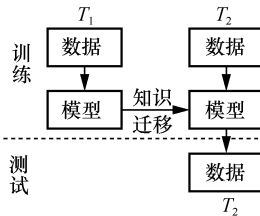


图4 迁移学习

1.2.3 在线学习

在线学习^[18]是指无须离线存储数据,让模型依次对获得的数据进行学习的场景,如图5所示。在线学习也是适用于序列数据流的学习方法,但是与持续学习的设定有3个明显的区别。首先,在线学习的数据都是独立同分布(independent identically distributed, IID)的,而持续学习中不同任务的数据可以来自不同的分布,数据分布之间存在偏移。其次,在线学习中所有的数据来自单个任务 T_1 ,每次学习一个样本,而持续学习通常需要学习不同的任务,每个任务有各自的数据集。最后,在线学习不存储先前的

样本,而一些持续学习方法设定可以存储旧任务的代表性样本,以防止遗忘旧任务的经验。

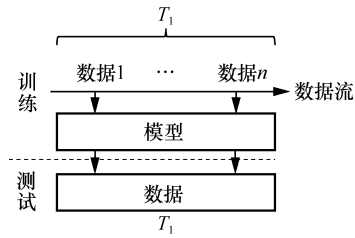


图5 在线学习

1.3 评价基准与指标

1.3.1 评价基准

持续学习是一种可以适用于各种机器学习任务的方法,目前研究者们主要在感知任务(图像分类等)和机器人操作任务上评价持续学习方法的优劣。

在感知任务中,常用的持续学习数据集有 Permuted MNIST^[19]、Rotated MNIST^[20]和 Incremental CIFAR^[21]等。如图6所示,这些数据集是常用的手写数字数据集经过像素打乱、旋转或者分割形成的具有多个分组的数据集,对应持续学习中的不同任务。一些持续域自适应方法将具有3个域31个类别的 Office 数据集^[22]作为基准,或者将多个手写数字数据集(MNIST^[23]、MNIST-M^[24]、USPS^[25]、SynNum^[24]和 SVHN^[26])结合以用于持续学习。



图6 手写数字数据集 (Rotated MNIST)

如图7所示,在机器人操作任务中,大多数方法在虚拟环境基准上进行持续学习方法的验证,比如雅达利游戏^[27],研究者^[28-30]将单个雅达利游戏作为一个任务,让智能体有序地学习多个雅达利游戏,并通过得分等指标评估方法的优劣。一些方法^[31-33]还在虚拟环境 MuJoCo^[34]和 Gym^[35]中的机器人场景下进行了持续学习方法的验证,通过修改环境的设定获得不同的任务。真实场景的基准主要以抓取^[36]和寻路^[15]等为主,例如将抓取不同物体作为持续学习的不同任务。

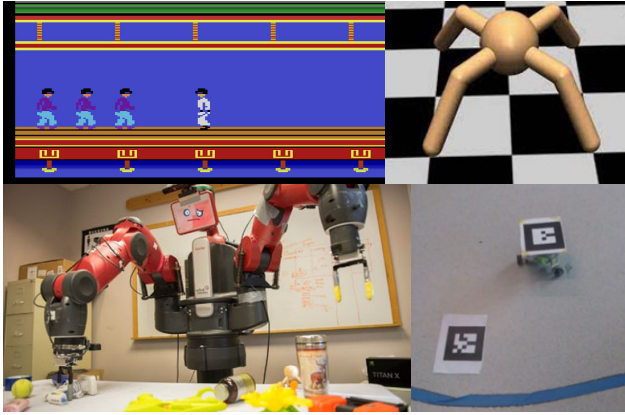


图7 机器人操作任务示例

1.3.2 评价指标

研究持续学习的目的是让智能体获得持续学习多任务的能力，同时不遗忘旧任务的经验。因此评价一个持续学习方法的优劣不能将模型学习完成后最终的性能作为唯一指标，还应衡量模型对旧任务的遗忘程度和遗忘速度，任务之间的知识转移效率，算法的稳定性以及时间和存储成本等因素。目前对持续学习方法的评价指标尚未达成共识，由于不同类型的持续学习方法高度异构，不同指标的评价结果也呈现出差异性。

为了评估持续学习算法的性能、前向迁移能力和后向迁移能力，Lopez-Paz D 等人^[20]分别提出了平均准确率（average accuracy, ACC）、后向迁移（backward transfer, BWT）和前向迁移（forward transfer, FWT）3种指标：

$$ACC = \frac{1}{T} \sum_{i=1}^T R_{T,i} \quad (1)$$

$$BWT = \frac{1}{T-1} \sum_{i=1}^{T-1} R_{T,i} - R_{i,i} \quad (2)$$

$$FWT = \frac{1}{T-1} \sum_{i=2}^T R_{i-1,i} - \bar{b}_i \quad (3)$$

平均准确率衡量了持续学习完所有任务后的模型在这些任务上的平均性能，能够直观地反映持续学习的效果；后向迁移衡量了持续学习方法是否有助于提升旧任务的性能，同时也衡量了旧任务的遗忘程度；前向迁移衡量了学习当前任务是否对未来任务有帮助。

Díaz-Rodríguez N 等人^[37]认为应该在持续学习算法的整个生命周期对其进行评估，而非仅仅在所有任务学习完成后进行评估。一个好的持续学习方法应该在持续学习每个任务时均保证好的性能以

及前向迁移和后向迁移的能力，因此3种新的评价指标被提出：准确率（accuracy, A）、修改后的后向迁移（mean backward transfer, mBWT）和修改后的前向迁移（mean forward transfer, mFWT）。

$$A = \frac{\sum_{i \geq j}^T R_{i,j}}{T(T+1)} \quad (4)$$

$$mBWT = \frac{\sum_{i=2}^T \sum_{j=1}^{i-1} (R_{i,j} - R_{j,j})}{\frac{T(T-1)}{2}} \quad (5)$$

$$mFWT = \frac{\sum_{i < j}^T R_{i,j}}{\frac{T(T-1)}{2}} \quad (6)$$

以上3种指标考虑了学习不同任务对其他任务的影响，评估贯穿了持续学习的整个过程，因此更加全面和客观地评估了持续学习方法的高效性和稳定性。此外，Díaz-Rodríguez N 等人^[37]还提出了模型大小（model size, MS）、样本储存效率（samples storage size, SSS）和计算效率（computational efficiency, CE）等指标用于评估持续学习方法的存储成本和计算成本。为了评估持续学习方法的抗遗忘能力，Chaudhry A 等人^[38]提出了遗忘度量（F）指标，其中第 k 个任务的遗忘度量定义为：

$$F_k = \frac{1}{k-1} \sum_{j=1}^{k-1} \max_{l \in \{1, \dots, k-1\}} R_{l,j} - R_{k,j}, \forall j < k \quad (7)$$

遗忘度量将特定任务的遗忘定义为过去模型在整个学习过程中获得的关于该任务的最大知识与在不同任务下拥有的该任务的知识之间的差异的平均值。由于该指标对遗忘程度的度量比较客观和全面，因此成为许多方法^[8,39]评估遗忘水平的第一选择。

除了上述常见的评价持续学习方法的指标，一些学者也提出了更多样化的指标。例如，Wolczyk M 等人^[40]提出了相对前向迁移（reference forward transfer, RT）来衡量持续学习方法前向迁移的能力；Jiang Y 等人^[41]使用单位时间平均回报来评价持续学习方法的优劣；Kemker R 等人^[42]使用知识保留度 Ω_{base} 、新任务准确率 Ω_{new} 和相对准确率 Ω_{all} 3个指标来评估模型在学习新知识的同时保留旧任务知识的能力；Riemer M 等人^[43]使用学习准确度（learning accuracy, LA）来评估模型在学习完每个

任务时的准确度均值; Henning C 等人^[44]使用得分和最终得分分别在每个任务学习完成后和所有任务学习完成后对持续学习方法进行评价。

2 机器人持续学习

机器人是持续学习的一个重要的实际应用场景, 而持续学习也可以帮助实现具备自主探索能力的机器人。本节首先强调了持续学习在机器人学习中的必要性, 然后分析了持续学习在机器人学习中面临的挑战, 最后分类介绍了不同的持续学习策略。

2.1 必要性

机器人的学习通常需要持续地与环境进行交互, 获取关于环境的数据, 这个过程的数据不是一次性获得的, 而且机器人的动作会影响并改变环境, 这种改变可能是不可逆的, 因此机器人很难回到先前的环境中进行重复学习。此外, 机器人的硬件条件限制使得其难以存储过去的完整数据。这些条件的约束使得机器人的学习场景正好与持续学习的设定相匹配, 这使得机器人学习成为持续学习最适合的应用场景。

机器人面临的任务是不断变化的, 任务会随着时间的推移而演变, 其原因要么是任务的环境发生了变化, 要么是出现了新的任务^[15]。而从头开始学习所有任务需要极大的计算开销和时间成本, 也受到环境不可恢复和存储容量的限制, 因此机器人学习不同任务需要依赖持续学习。例如, 抓取机器人需要学习抓取不同的物体, 为了在不遗忘抓取杯子的经验的前提下学会抓取盘子, 不得不引入持续学习方法。总之, 持续学习最适合的应用场景就是机器人学习, 而机器人学习不同的任务需要依赖持续学习。

2.2 机器人场景下的挑战

2.2.1 灾难性遗忘

持续学习面临的最大挑战是灾难性遗忘^[45]。灾难性遗忘是指模型在新任务上进行学习后, 过度拟合于当前任务的数据, 使得旧任务的经验产生偏移, 从而在旧任务上表现较差, 遗忘了旧任务的经验的现象。灾难性遗忘本质上是由数据在特征空间中的分布偏移引起的, 如图8所示, T_1 和 T_2 分别代表两个任务的特征分布, 每个任务的数据不是独立同分布的, 在当前任务上训练后的模型能较好地拟合对应数据在特征空间中的分布, 然而模型参数的更新使得模型对旧任务的数据的拟合程度下降, 从而产生了灾难性遗忘。

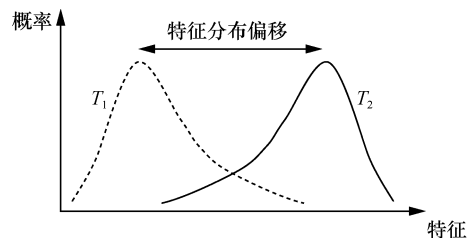


图8 特征分布偏移

2.2.2 硬件限制

机器人作为一种体积有限的嵌入式设备, 硬件条件通常受到诸多限制。首先, 机器人设备的内存资源十分有限, 难以存储大量的数据, 这使得机器人在序列化地学习多个任务时难以保存和利用旧任务的所有数据, 而且由于存储容量有限, 基于动态网络的持续学习方法也会随着任务数量增加而失效。其次, 机器人的计算性能也严重受到硬件设备的制约, 这使得机器人的学习需要更多的时间成本, 也使得从头学习所有任务变得难以实现, 因此更需要高效的持续学习方法。除了内存资源和计算能力的限制, 机器人还具有续航能力有限、实验成本高、需要人工辅助学习等问题。

2.2.3 虚拟到真实的迁移

尽管近年来取得了一些进展, 但机器人尚未得到预期的广泛应用。Grossmann A^[3]认为原因之一是缺乏足够的感知方法, 无法在真实环境中部署机器人。由于真实环境的复杂性和难以恢复的特性, 机器人在真实场景中学习序列任务十分困难, 因此目前大多数针对机器人应用的持续学习的研究主要集中在虚拟环境中(例如以虚拟场景为代表的强化学习平台 MuJoCo 和以虚拟游戏场景为代表的雅达利游戏), 通过让智能体持续地学习多个虚拟任务或者掌握不同的游戏来验证持续学习方法的优劣。这些方法虽然在虚拟环境中具有较好的表现, 但当迁移到真实场景时性能会严重下降。目前真实场景的持续学习方法也只是在一些简单的任务场景下进行验证, 例如三轮机器人二维寻路任务^[15]和简单的机器人抓取场景^[36]等。相比于虚拟场景, 真实场景的复杂性和干扰性极大, 这使得虚拟任务到真实任务的迁移成为机器人持续学习面临的一个仍未解决的问题。

2.3 持续学习策略

目前, 对持续学习的研究主要集中在感知领域(图像、深度、语言语音等)。对环境的感知是机器人在真实世界中运行必不可少的组成部分, 研究感

知领域的持续学习策略有助于帮助机器人更好地观测真实环境,从而更稳定地持续学习新任务和探索新环境,因此研究机器人的持续学习离不开讨论感知领域的持续学习策略。此外,基于与环境交互的持续学习策略也是机器人持续学习研究的重点课题,探索非稳态环境的能力是发展强智能体必需的要素。

2.3.1 基于正则化的持续学习

机器人通常是一种内存有限的嵌入式的移动设备,而持续学习需要不断地学习序列任务和数据,存储新任务的经验,能够高效地利用有限的存储空间的持续学习方法对于机器人学习来说具有重要的实际意义,因此基于正则化的持续学习方法受到广泛研究。常用于持续学习的正则化方法包括参数正则化、知识蒸馏和梯度正则化。

(1) 参数正则化

为了提高有限内存的利用效率,正则化方法对模型的更新施加约束,使其在参数空间中以合适的步长朝着更恰当的方向进行参数搜索。在持续学习中,正则化一般用来限制模型参数朝着不影响旧任务性能的方向进行更新,最小化参数更新对模型在旧任务上性能的影响^[46-50]。著名的弹性权重巩固(elastic weight consolidation, EWC)^[51]就是一种典型的参数正则化方法,依据参数相对于旧任务的重要性进行排序,对参数的变化施加惩罚以限制参数的更新量,重要性越大的参数惩罚越大。这种方法的基本假设是模型的每个参数对不同任务的重要性不同,通过限制重要参数的变化来对抗灾难性遗忘,同时积极调整不重要的参数以掌握新任务。这种思路成为后续许多持续学习工作的思想基础。提高不同任务的差异性以减少任务间的干扰也是一类常用的方法,例如通过正则项来拉大当前任务的数据在不同任务参数下的分布的差距^[52],或者使模型分类器输出概率的熵最大化^[53]。

(2) 知识蒸馏

知识蒸馏^[54]是一种特殊的正则化方法,其通过对网络的知识进行压缩,将旧任务的知识提取到另一个网络,以尽可能地保留前一个网络学到的经验,并利用过去的经验帮助学习新任务^[55-58]。如何使用知识蒸馏来最大化地保留旧任务的经验,同时保证知识正向迁移的稳定性是值得探索的课题^[59]。例如,Smith J等人^[60]建议通过改进交叉熵训练和加权重要特征,改善知识蒸馏的稳定性,更好地保留

旧任务的经验。然而上述基于知识蒸馏的持续学习方法在学习一个任务时需要至少同时使用两个网络,还需要额外的存储空间来存放专家网络的参数。为了减小计算成本和内存开销,一些方法仅使用单个网络进行知识蒸馏,例如提出一种利用共享特征提取网络,仅使用分类层的类别信息进行知识蒸馏^[61],或者将一个神经网络分割学习并重新连接以学习跨任务的知识^[62]。

(3) 梯度正则化

梯度方法是一类从梯度层面施加正则化的持续学习方法^[63-65]。梯度方法的目的是通过对梯度的更新加以约束,防止梯度更新对旧任务的经验产生干扰^[66-67],以此来对抗灾难性遗忘。例如正交梯度方法^[68]要求模型在旧任务的正交子空间中更新梯度以学习新任务,让梯度更新方向始终与旧任务梯度方向保持正交,防止因参数更新导致的灾难性遗忘。这类方法可以最大限度地解决遗忘问题,但忽视了持续学习的另一个目的是利用旧任务的经验改善和加速新任务的学习,以避免从头学习的低效性和不可实现性,而正交梯度方法对梯度正交性的约束无疑摒弃了任务间的正向迁移。为了解决这个问题,Lin S等人^[69]提出基于信任区域投影的梯度方法,在保证正交梯度不干扰旧任务的前提下,利用强相关任务的经验来改善新任务的学习。

总而言之,机器人设备通常因体积受限难以存储大量数据,因此基于正则化的持续学习探索通过对参数更新施加约束以高效地利用有限的内存,通过减少参数更新对旧任务的干扰以对抗灾难性遗忘。

2.3.2 基于回放持续学习

基于回放的持续学习方法是指在机器人设备的内存中单独开辟出少量的空间作为存储器来存储过去任务的部分样本或者表征,在学习新任务的同时回放存储器中的数据,以此来利用过去的经验减弱灾难性遗忘。基于回放的方法虽然需要少量的额外空间,但是带来的性能提升是十分显著的。常用于持续学习的回放方法包括样本回放、生成式回放、参数回放和原型(prototype)回放。

(1) 样本回放

最常见的基于回放的方法是选择性地存储旧任务的样本用于回放,默认情况下这些样本通过随机或者均匀采样获得,此外还可以是每个任务的代

表性样本或困难样本^[66]。这一类方法主要是研究如何动态添加和剔除存储器中旧任务的样本,并高效利用存储的样本。例如存储每个任务中最接近任务平均特征向量的样本^[21]或者存储困难样本,并在更新存储器时优先去掉最容易样本,而非最旧样本^[70]。除了简单地回放存储器数据用于扩充数据集,还可将其用于改进机器人的学习。Kim C D 等人^[71]和 Cha H 等人^[72]利用自监督回放存储器中的样本,能较好地避免噪声标注数据引起的错误训练信号问题,使得持续学习模型的性能更加稳定。

(2) 生成式回放

持续学习的一个理想假设是当机器人学习新任务时,旧任务的数据始终不可得,因为某些环境可能具有不可复现性,并且存储旧任务的样本也可能具有隐私泄露的风险,因此一些方法^[73-76]探索了基于生成式回放的方法。生成式回放是指使用生成模型生成的伪样本代替真实样本,只保留生成模型的参数或伪样本用于回放,这样可以避免上述问题。然而为每个任务保留一个单独的生成网络的存储开销极大,而可用于多个任务的生成网络也面临持续学习和灾难性遗忘的挑战,一种直接的解决思路是深度生成回放(deep generative replay, DGR)^[77],即混合旧任务的生成样本和新任务的真实样本对生成网络进行重复训练,但是这种方法的计算成本较大。为了解决这个问题,学习连接可塑性的动态生成存储(dynamic generative memory, DGM)^[78]利用参数层面的注意力机制,仅使用单个生成器就能够增量地学习新任务,而无须回放以前的知识。

(3) 参数回放

尽管基于生成式回放的方法没有隐私风险,但是受限于生成样本的质量,该方法仍面临性能不稳定的困扰,因此存储旧任务的表征或中间参数用于回放成为一种有效的替代方案^[79-80]。Singh P 等人^[81]使用空间和通道校准模块校准激活图,然后在存储器中存储任务适应性校准参数,这些参数包含特定任务的知识。此外,还可以通过存储一些中间层的激活值来代替存储样本,以减少回放所需的计算开销和存储成本^[82]。

(4) 原型回放

原型方法^[83-84]也是一类常见的基于回放的持续学习方法。原型是一个任务的数据在潜在分布空间的质心,更合理的原型能更好地表示当前任务,因此只需要存储原型即可较好地保留任务的

经验,原型方法是样本回放的一种高效的替代方案。例如,Stan S 等人^[85]学习了源任务域在特征空间中的原型分布,然后利用原型分布将目标域分布与源域分布进行对齐,可以保留源任务域的经验并将其迁移到目标域。然而仅使用静态的原型来保留旧任务的表征的适应性较差,难以在学习不同任务的过程中始终保持较好的判别力。对此,PASS 模型^[86]给出的解决方法是在存储旧任务原型的基础上,使用特征空间中的原型增强(protoAug)来保持旧任务的决策边界;De Lange M 等人^[87]则让原型在持续学习的过程中在潜在空间不断演化,以进行适应性调整,从而能够在任何时间点进行学习和预测。

总而言之,由于容易在模型更新后丢失过去任务的经验,基于回放的持续学习通过存储并回放过去任务的部分样本或表征来保留过去任务的经验。基于回放的持续学习得益于两个方面:首先,保留旧任务的部分样本或表征可以传递旧任务的经验,对抗灾难性遗忘问题;其次,存储器中的样本或表征能够起到平衡新旧任务重要性的作用,可以在一定程度上改善模型偏向于掌握新任务而忽视旧任务的偏差问题。

2.3.3 基于动态结构的持续学习

可拓展型机器人为基于动态结构的持续学习方法提供了可能,通过模块化的组件机器人让每个单元独立地调整其行为,组装系统能够在动态环境中学习并保持最佳表现^[88]。基于动态结构的持续学习是指机器人在持续学习序列任务的过程中,动态地调整模型结构以适应新任务,同时兼顾在旧任务上的表现的方法。常用的基于动态结构的持续学习方法包括显式动态结构和隐式动态结构。

(1) 显式动态结构

显式动态结构方法如图9所示,为了防止模型更新对旧任务的参数产生干扰,一种直接的思路就是参数独立,即每个任务有各自的神经元,每次学习新任务时,模型动态地横向扩展新的神经元及其连接,以容纳新任务的经验,避免灾难性遗忘的发生。例如,Rusu A A 等人^[29]提出渐进式神经网络,引入调节器在旧任务网络的基础上适应性地添加一系列神经元,并连接到旧任务的神经元以辅助新任务的学习。在学习当前任务时冻结先前的网络参数,仅学习新添加的参数。然而机器人的存储容量是有限的,即使大型设备也会随着任务数量的增加

而逐渐达到容量上限,因此对神经元数量的增加速度做出限制是有必要的,这也是目前显式动态结构方法的研究重点^[89-90]。例如, Xu J 等人^[91]通过对当前任务所需添加的神经元数量进行预测,根据任务需要紧凑地扩展网络。然而上述方法仅研究了扩展神经元以容纳新知识,没有考虑任务间的干扰性和旧任务的噪声, Wang L Y 等人^[30]受到突触扩展和收敛的启发,提出了主动遗忘与突触扩展收敛(active forgetting with synaptic expansion-convergence, AFEC)方法,该方法在新任务上扩展网络的同时还会选择性地遗忘与当前任务差异较大的任务经验。

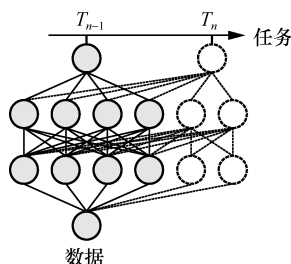


图9 显式动态结构方法

(2) 隐式动态结构

除了扩充网络的显式动态方法,还有一类常用的隐式动态结构,即预先设定好模型的最大容量,令每个任务的网络参数稀疏化,并在此基础上建立新的神经元连接,以学习新的任务。这类方法^[92-93]可以避免显式动态结构引入内存爆满问题,同时保留了动态结构的持续学习方法任务间互不干扰的优势。例如, Mazumder P 等人^[94]将网络参数分为对当前任务重要的部分和不重要的部分,仅冻结少数重要的参数,这有助于防止出现过拟合现象,而将其他参数用于后续任务的学习,最大限度地保证后续任务所需的参数量。

总而言之,固定结构的模型表示不同任务的能力是有上限的,因此基于动态结构的模型探索了通过适应性地扩展和收缩模型结构来容纳更多任务的经验,由于任务参数的独立性和冻结参数的设定,基于动态结构的方法不会产生遗忘问题。

2.3.4 基于元学习的持续学习

将元学习应用于持续学习的目的是对机器人进行适应性训练,让其获得记住旧任务和快速掌握新任务的能力^[95-96]。通过梯度下降来学习序列任务通常会导致较差的结果,这是因为旧任务的经验往

往由于干扰权重更新而被迅速遗忘^[97]。这种干扰可以通过元学习方法来减少,元学习方法使用当前和过去的任务优化基本学习算法,并将其保存在存储器中^[43,98],以此来帮助机器人快速地掌握新任务,同时保留了旧任务的经验^[99-100]。例如, Hurtado J 等人^[96]利用元学习来更新存储旧任务经验的知识基,将当前任务的数据和类别拆分成子任务,使用每个子任务和对应的分类层训练知识基。基于元学习更新的模型相比其他方法对不同任务更加具有鲁棒性,不容易过拟合于新任务的数据,从而减轻了灾难性遗忘。

总而言之,学习序列任务容易因权重干扰导致遗忘,因此基于元学习的持续学习利用先验知识来减弱遗忘,使得模型获得对不同任务鲁棒的初始化参数,同时可以加速新任务的学习。

2.3.5 基于模型压缩的持续学习

移动机器人通常被要求具有较高的灵活性,这使得机器人设备的体积受到限制,并进一步限制了机器人的计算性能和存储空间。而持续地学习新任务或探索新环境会使得模型体积不可避免地增加,带来了更多的计算开销和存储成本,这对于要求机器人进行实时推理的场景来说是不可接受的。因此如何解决持续学习导致的计算开销和存储成本增加的问题是一个值得研究的课题。对于上述问题,基于模型压缩的持续学习给出了一些可行的解决方案,这类方法的研究内容是通过选择、编码和压缩知识实现高效存储,通常用到的模型压缩方法有知识蒸馏^[72,101]、剪枝^[6,102]和量化^[103-104]等。需要说明的是,这里的知识蒸馏和第2.3.1节提到的知识蒸馏的目的不同。作为正则化手段的知识蒸馏通过将旧任务知识蒸馏到新任务的模型中来防止灾难性遗忘;而本节提到的知识蒸馏是为了减少持续学习模型的计算开销和存储成本。例如, Cha H 等人^[72]利用实例关系的自蒸馏来高效地存储和对比学习表征。而剪枝根据一定规则删减模型中不重要的神经元或连接来实现模型压缩。例如, Ebrahimi S 等人^[102]利用不确定性对权重进行剪枝,并存储任务的二值掩码,从而保留持续学习模型的性能。量化可以减少存储权重所需要的比特数,使得持续学习模型占用的存储空间更小。例如, Ravaglia L 等人^[103]对模型冻结阶段和潜在回放的权重进行量化,使得持续学习模型的存储成本大幅度下降。

总而言之,持续学习的设定使得机器人更容易

遇到计算瓶颈和存储上限问题, 因此模型压缩是机器人持续学习必须考虑的问题。提出适合于持续学习模型的压缩方法可以在不严重损害模型性能的前提下缩减计算开销和存储成本, 具有重要的实际意义。

2.3.6 基于强化学习的持续学习

强化学习是机器人系统最常用的学习方法, 机器人需要在与环境的不断交互中探索并学习策略以采取恰当的行动, 从而学会做出决策。强化学习的最新进展已显示出巨大的成功, 在各种具有挑战性的棋盘游戏和雅达利游戏中表现较好, 并可以通过反复试验获得复杂的机器人操作技能, 而无需事先了解任务。然而一般的强化学习方法在学会解决给定任务的策略后, 一旦需要探索新的环境或任务, 将不得不从头开始培训新策略^[105], 因此提出适应持续学习场景的强化学习方法意义重大。

将强化学习方法拓展到持续学习场景是一项复杂的任务。一个完整的强化学习算法通常包含对状态到动作的映射建模的策略函数、状态集合、动作集合、评估状态动作对的值函数和奖励值, 依据方法类型可能还会有经验回放和对环境建模的状态转移函数。持续的强化学习意味着需要强化学习算法的每一个要素都能适应不断变化的环境或任务, 复杂性远高于单一感知域下的持续学习。目前的持续强化学习方法通常聚焦于强化学习中的某一部分 (例如策略函数), 并使其适应于持续学习场景, 而假设其他部分可以随着该部分的持续学习自适应地调整以获得持续学习的能力。

改进奖励函数或值函数来适应持续学习场景是一种常见的解决思路^[106-107]。例如, 探索值函数的上界可以让值函数以贪婪的方式初始化, 帮助机器人更快速地掌握新任务, 同时更好地实现新旧任务间的值传递, 从而实现值函数的持续学习。Brafman R I 等人^[108]提出一种名为 R-MAX 的值函数的上界, 但是这个上界是非常松弛的, 不能高效地实现值传递, 因此后续的学者不断改进, 分别提出了更紧致の上界 MAXQINIT^[109]和 LRMax^[107]。

策略函数是强化学习中机器人进行决策的核心, 更优的策略函数能够使得智能体选择出更适合当前状态的动作, 因此改进策略函数^[15,32,110]能够直接帮助机器人适应持续学习场景, 例如利用旧任务积累的知识来辅助当前任务的策略学习的持续策略梯度方法^[31]。

虽然不断变化的策略和值函数等问题在有关持续学习的文献中得到了关注, 但对于动作集大小变化的情景, 仍然没有好的解决方法。为此, Chandak Y 等人^[111]提出一种可以适应规模不断变化的动作集的算法, 可以推断动作空间中潜在的结构。此外, 经验回放^[112]作为一种在传统强化学习任务中被提出的方法, 也被改进并应用于机器人的持续学习中^[113]。

在真实环境中研究机器人持续学习通常存在诸多限制。一个典型的问题是机器人需要与环境进行交互来探索环境, 而在真实环境运行机器人通常非常耗时, 这使得研究人员难以获得大量的训练数据来改进模型。例如, Julian R 等人^[36]仅使用离线数据集对真实环境的持续学习抓取机器人进行微调, 就需要尝试抓取超过 15 000 次, 花费近一个月的时间。此外, 真实环境的机器人面临硬件磨损、电池容量有限和性能故障等问题, 尤其在需要探索不断变化的环境来收集大量轨迹的情况下。因此, 目前基于强化学习的持续学习的研究主要集中在虚拟游戏环境^[29]或仿真场景^[33]。为了将基于强化学习的持续学习应用于真实环境的任务, 一种妥协的解决思路是让机器人在虚拟环境中进行探索和学习, 这样可以克服真实环境下的限制, 然后在真实环境的任务上验证算法的有效性。例如, Traoré R 等人^[15]通过在仿真环境中改变场景设置来研究机器人寻路任务中的持续学习问题, 并将模型迁移到真实环境的寻路机器人上进行验证。

总而言之, 机器人离不开与环境的动态交互, 让机器人在探索环境中获得持续适应不同环境或任务的能力, 这使得研究基于强化学习的持续学习方法具有必要性。

2.3.7 通用框架

持续学习是一个更贴合机器人应用场景的研究方向, 然而目前大多数对持续学习方法的研究主要集中于神经网络和深度学习领域。虽然这些方法在理论场景下取得了较好的结果, 但由于真实场景的复杂性, 将其迁移到实际的机器人应用场景可能面临性能严重下降甚至完全失效的问题。Yang F 等人^[33]总结了各类持续学习方法在机器人场景中应用的优缺点, 并提出了一个框架来修改监督学习算法, 使之与强化学习兼容, 为应用于机器人场景的基于神经网络的持续学习提供了一条通用的途径。机器人学习的优势是可以自

主地探索环境,并从不断变化的环境中学习经验,但目前大多数机器人系统产生的静态策略不会在部署时进一步调整,这使得这些方法难以应用于持续学习场景。Julian R 等人^[36]对持续学习中机器人端到端的任务适应性问题进行了研究,基于微调方法提出了一个可以将机器人静态策略轻松地转换为适用于持续学习的策略的框架。此外,持续学习需要让机器人持续适应不断变化的环境,但目前的强化学习方法在环境非稳态和互动非偶发性的情况下面临失效的困境。为了解决这个问题,Lu K 等人^[114]提出一个非偶发性持续学习算法框架,可以从离线数据中获得经验,从而减少机器人对与环境互动的依赖。

总而言之,研究如何拓展基于神经网络的持续学习和传统强化学习的通用框架可以将现有的只能在理论环境中生效的方法成功应用到机器人真实场景的持续学习中,充分利用已有方法的宝贵经验,推动机器人持续学习的进一步发展。

3 前景与展望

持续学习是机器人实际应用中一项有趣但极具挑战性的研究课题,由于真实世界的的数据通常是非平稳的且呈现序列性的特点,因此需要持续学习方法来帮助机器人适应不断变化的环境。现有的机器人持续学习方法分别从不同角度对持续学习进行了研究。基于正则化的方法可以有效地约束和引导模型参数朝着不遗忘旧任务经验的方向进行更新。基于回放的方法通过重用旧任务的数据或表征来辅助学习,可以有效地对抗遗忘。模块化可拓展的动态体系结构为解决干扰和灾难性遗忘提供了实用的方案,兼顾了系统的稳定性和可塑性。基于元学习的方法可以为持续学习模型学习先验知识,超越手工设计的算法和体系结构^[6]。基于模型压缩的方法能更好地平衡准确性和模型体积,可以容纳更多任务的知识。基于强化学习的方法探寻了序列任务环境下的强化学习策略。通用框架则为机器人持续学习提供了一系列高效的范式。

虽然目前持续学习方法取得了一些结果,在解决灾难性遗忘方面获得了较大进展,但大多数结果施加了许多限制^[8]。基于正则化的方法在任务数量较少的持续学习问题上表现出令人满意的性能,但是当任务数量增加时(诸如权值的累积漂移和它们之间的干扰等),这种方法难以扩展。基于回放的方法

需要额外的存储空间,此外由于存储器容量是固定的,随着任务数量增加,单个任务用于回放的样本或表征数量会减少,模型越来越容易过度拟合于存储样本,从而损害在旧任务上的泛化能力^[9]。基于动态结构的方法受限于设备存储容量,随着任务数量增加容易达到容量上限,此外还缺乏利用旧任务中关键信息的高效方法。基于元学习的方法通常需要更多的任务数量以获得好的元知识,当任务数量较少时较难拟合出合理的任务先验分布。基于模型压缩的方法聚焦于压缩模型体积,因此不得不稍微损害模型的性能。基于强化学习的方法目前尚未探索出一套高效的适用于真实持续学习场景的结构体系。

在此,本文通过总结和分析现有模型存在的优点和缺陷,提出了一些有挑战性和有价值的未来研究方向。

现有的机器人持续学习方法仅在少量任务上进行实验,这与人们希望的在大量任务甚至无限任务下持续学习的要求相去甚远。由于机器人(尤其是体积较小的移动机器人)设备容量受限,无论是动态模型方法还是其他持续学习方法都会在处理一定数量的任务后达到机器人承载能力的上限,模型容量趋于饱和,性能难以进一步提升,甚至面临在新任务上的性能下降的问题。因此如何将现有的持续学习方法推广到更多任务的学习,或者提出无限任务下的持续学习方法是一个很有价值的研究方向^[115-116]。

如何将实验环境下的持续学习方法应用到实际场景的机器人任务一直是一个棘手的问题。现有的持续学习研究仍集中在实验环境的神经网络领域,这些方法大多也是针对神经网络和诸如图像分类这样的离线(静态)数据集设计的,难以迁移到需要与真实世界的环境进行实时交互的机器人上。此外,目前大多数专用于机器人任务的持续学习方法仅仅是在虚拟环境下进行验证的,这些任务场景的设定通常比较简单,没有过多的干扰因素,当迁移到具有各种复杂因素的真实环境中时可能会失效。因此,设计出能够应用在与真实环境交互的机器人上的持续学习方法具有重要的现实意义^[15]。

Bing L^[117]提出了在职学习的概念,即模型在部署到机器人后,需要在应用过程中进行持续学习。目前持续学习的设定虽然是逐步学习一连串的任务,但在每个任务时仍然遵循传统的两阶段独

立训练和测试范式,同时这些任务是人为给定的。而在职业学习的目的是利用学到的知识来发现新的任务,与人类和环境互动并进行推断,在应用过程中以自监督和互动的方式逐步学习新的任务。在职业学习研究的是如何通过智能体自身不断发现新任务,同时在应用过程中收集新的训练数据,并逐步学习新任务,而不中断应用,因此这是一个具有实际意义的研究方向。

人类大脑能够在整个生命周期内不断获取的信息中持续地学习新的知识和技能,而这种能力理论上是没有上限的,换言之,人类大脑天然具有持续学习无限任务的能力,并且新任务的学习不会以牺牲旧的记忆为代价。因此生物的大脑工作模式仍然是持续学习的最佳模式,这使得受生物启发的持续学习成为一类有潜力的方法,对于实现强智能体具有重要的启示意义^[7]。例如, Flesch T 等人^[118]研究了人类持续学习的认知机制,并将发现用于模型设计。如图 10 所示, Barron H C^[119]观察到神经抑制可以调节神经网络的稳定性,控制大脑皮层可塑性和记忆恢复,以此保护记忆免受干扰,从而允许可持续地学习不同的任务。然而目前受生物启发的持续学习研究非常少,并且处于初级阶段,因此受生物启发的持续学习将会是未来实现强智能的通用机器人的有力途径。

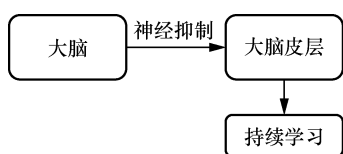


图 10 神经抑制作用于大脑皮层,避免记忆被干扰,从而允许持续学习

4 结束语

本文总结了机器人的持续学习的研究现状,首先介绍了持续学习的框架,阐述了持续学习的重要性,进一步分析和总结了机器人持续学习面临的挑战及研究方法。最后,通过对现有的持续学习策略存在的问题进行分析,提出了机器人持续学习未来有价值的研究方向。相信持续学习将来可以从实验走向真实,并成为一项研究机器人学习的有力工具。

参考文献:

[1] RING M B. Continual learning in reinforcement environments[M].

Austin: University of Texas, 1994.

- [2] TJOMSLAND J, KALKAN S, GUNES H. Mind your manners! A dataset and a continual learning approach for assessing social appropriateness of robot actions[J]. *Frontiers in Robotics and AI*, 2022, 9: 669420.
- [3] GROSSMANN A. Continual learning for mobile robots[D]. Birmingham: University of Birmingham, 2001.
- [4] NEHMZOW U, RIANO L. A proposal for continual learning in robotics[C]//*Proceedings of Towards Autonomous Robotic Systems*. [S.l.:s.n.], 2009.
- [5] LIU B, XIAO X S, STONE P. A lifelong learning approach to mobile robot navigation[J]. *IEEE Robotics and Automation Letters*, 2021, 6(2): 1090-1096.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] VERWIMP E, DE LANGE M, TUYTELAARS T. Rehearsal revealed: the limits and merits of revisiting samples in continual learning[C]//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision*. Piscataway: IEEE Press, 2021: 9365-9374.
- [10] 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.
- [11] LIU Y Y, SCHIELE B, SUN Q R. Adaptive aggregation networks for class-incremental learning[C]//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE Press, 2021: 2544-2553.
- [12] ZHAI M Y, CHEN L, HE J W, et al. Piggyback GAN: efficient lifelong learning for image conditioned generation[M]//*Computer vision – ECCV 2020*. Cham: Springer International Publishing, 2020: 397-413.
- [13] MITCHELL T, COHEN W, HRUSCHKA E, et al. Never-ending learning[J]. *Communications of the ACM*, 2018, 61(5): 103-115.
- [14] MCCLOSKEY M, COHEN N J. Catastrophic interference in connectionist networks: the sequential learning problem[J]. *Psychology of Learning and Motivation*, 1989, 24: 109-165.
- [15] TRAORÉ R, CASELLES-DUPRÉ H, LESORT T, et al. Continual reinforcement learning deployed in real-life using policy distillation and Sim2Real transfer[C]//*Proceedings of ICML Workshop on Multi-Task and Lifelong Reinforcement Learning*. [S.l.:s.n.], 2019.
- [16] ZHANG Y, YANG Q. A survey on multi-task learning[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2021.
- [17] PAN S J, YANG Q. A survey on transfer learning[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2010, 22(10): 1345-1359.
- [18] HOI S C H, SAHOO D, LU J, et al. Online learning: a comprehensive survey[J]. *Neurocomputing*, 2021, 459: 249-289.
- [19] GOODFELLOW I J, MIRZA M, XIAO D, et al. An empirical investigation of catastrophic forgetting in gradient-based neural networks[C]//*Proceedings of the 2nd International Conference on Learning Representations*. [S.l.:s.n.], 2014.

- [20] LOPEZ-PAZ D, RANZATO M. Gradient episodic memory for continual learning[J]. arXiv preprint, 2017, arXiv:1706.08840.
- [21] REBUFFI S A, KOLESNIKOV A, SPERL G, et al. iCaRL: incremental classifier and representation learning[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 5533-5542.
- [22] SAENKO K, KULIS B, FRITZ M, et al. Adapting visual category models to new domains[J]. Lecture Notes in Computer Science, 2010: 213-226.
- [23] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [24] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation[C]//Proceedings of the 32nd International Conference on Machine Learning. [S.l.:s.n.], 2015: 1180-1189.
- [25] HULL J J. A database for handwritten text recognition research[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(5): 550-554.
- [26] NETZER Y, WANG T. Reading digits in natural images with unsupervised feature learning[J]. Advances in Neural Information Processing Systems, 2011.
- [27] BELLEMARE M G, NADDAF Y, VENESS J, et al. The arcade learning environment: an evaluation platform for general agents[J]. Journal of Artificial Intelligence Research, 2013, 47: 253-279.
- [28] JUNG S, AHN H, CHA S, et al. Continual learning with node-importance based adaptive group sparse regularization[J]. Advances in Neural Information Processing Systems, 2020, 33: 3647-3658.
- [29] RUSU A A, RABINOWITZ N C, DESJARDINS G, et al. Progressive neural networks[J]. arXiv preprint, 2016, arXiv:1606.04671.
- [30] WANG L Y, ZHANG M T, JIA Z F, et al. AFEC: active forgetting of negative transfer in continual learning[J]. arXiv preprint, 2021, arXiv:2110.12187.
- [31] MENDEZ J A, WANG B Y, EATON E. Lifelong policy gradient learning of factored policies for faster training without forgetting[J]. Advances in Neural Information Processing Systems, 2020.
- [32] XU K, VERMA S, FINN C, et al. Continual learning of control primitives: skill discovery via reset-games[J]. Advances in Neural Information Processing Systems, 2020.
- [33] YANG F, YANG C, LIU H, et al. Evaluations of the gap between supervised and reinforcement lifelong learning on robotic manipulation tasks[C]//Proceedings of the Conference on Robot Learning. [S.l.:s.n.], 2022: 547-556.
- [34] TODOROV E, EREZ T, TASSA Y. MuJoCo: a physics engine for model-based control[C]//Proceedings of 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2012: 5026-5033.
- [35] BROCKMAN G, CHEUNG V, PETERSSON L, et al. OpenAI Gym[J]. arXiv preprint, 2016, arXiv:1606.01540.
- [36] JULIAN R, SWANSON B, SUKHATME G S, et al. Never stop learning: the effectiveness of fine-tuning in robotic reinforcement learning[J]. arXiv preprint, 2020, arXiv:2004.10190.
- [37] DÍAZ-RODRÍGUEZ N, LOMONACO V, FILLIAT D, et al. Don't forget, there is more than forgetting: new metrics for continual learning[J]. arXiv preprint, 2018, arXiv:1810.13166.
- [38] CHAUDHRY A, DOKANIA P K, AJANTHAN T, et al. Riemannian walk for incremental learning: understanding forgetting and intransigence[C]//Proceedings of the European Conference on Computer Vision. [S.l.:s.n.], 2018: 532-547.
- [39] YIN H, YANG P, LI P. Mitigating forgetting in online continual learning with neuron calibration[J]. Advances in Neural Information Processing Systems, 2021, 34.
- [40] WOŁCZYK M, ZAJĄC M, PASCANU R, et al. Continual world: a robotic benchmark for continual reinforcement learning[J]. Advances in Neural Information Processing Systems, 2021.
- [41] JIANG Y, BHARADWAJ S, WU B, et al. Temporal-logic-based reward shaping for continuing learning tasks[J]. arXiv preprint, 2020, arXiv:2007.01498.
- [42] 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.
- [43] RIEMER M, CASES I, AJEMIAN R, et al. Learning to learn without forgetting by maximizing transfer and minimizing interference[C]//Proceedings of the 7th International Conference on Learning Representations. [S.l.:s.n.], 2019.
- [44] HENNING C, CERVERA M, D'ANGELO F, et al. Posterior meta-replay for continual learning[J]. Advances in Neural Information Processing Systems, 2021, 34.
- [45] DONG J, CONG Y, SUN G, et al. I3DOL: Incremental 3D object learning without catastrophic forgetting[C]//Proceedings of the AAAI Conference on Artificial Intelligence. [S.l.:s.n.], 2021: 6066-6074.
- [46] ZHOU M, XIAO J, CHANG Y F, et al. Image de-raining via continual learning[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 4905-4914.
- [47] CHERAGHIAN A, RAHMAN S, RAMASINGHE S, et al. Synthesized feature based few-shot class-incremental learning on a mixture of subspaces[C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 8641-8650.
- [48] KUKLEVA A, KUEHNE H, SCHIELE B. Generalized and incremental few-shot learning by explicit learning and calibration without forgetting[C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 9000-9009.
- [49] LOO N, SWAROOP S, TURNER R E. generalized variational continual learning[C]//Proceedings of the International Conference on Learning Representations. [S.l.:s.n.], 2020.
- [50] GORUR D. Linear mode connectivity in multitask and continual learning[J]. arXiv preprint, 2020, arXiv:2010.04495.
- [51] 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.
- [52] VERMA V K, LIANG K J, MEHTA N, et al. Efficient feature transformations for discriminative and generative continual learning[J]. arXiv preprint, 2021, arXiv:2103.13558.
- [53] CHA S, HSU H, HWANG T, et al. CPR: classifier-projection regularization for continual learning[C]//Proceedings of the International Conference on Learning Representations. [S.l.:s.n.], 2020.

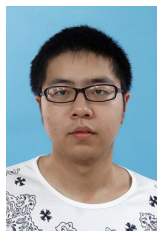
- [54] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network[J]. arXiv preprint, 2015, arXiv:1503.02531.
- [55] ZHAI M Y, CHEN L, TUNG F, et al. Lifelong GAN: continual learning for conditional image generation[C]//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 2759-2768.
- [56] TAO X Y, HONG X P, CHANG X Y, et al. Few-shot class-incremental learning[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 12180-12189.
- [57] MENG Q E, SHIN'ICHI S. ADINet: attribute driven incremental network for retinal image classification[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 4032-4041.
- [58] FINI E, LATHUILIÈRE S, SANGINETO E, et al. Online continual learning under extreme memory constraints[J]. Lecture Notes in Computer Science, 2020: 720-735.
- [59] SHMELKOV K, SCHMID C, ALAHARI K. Incremental learning of object detectors without catastrophic forgetting[C]//Proceedings of 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE Press. [S.l.:s.n.], 2017: 3420-3429.
- [60] SMITH J, HSU Y C, BALLOCH J, et al. Always be dreaming: a new approach for data-free class-incremental learning[C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 9354-9364.
- [61] CASTRO F M, MARÍN-JIMÉNEZ M J, GUIL N, et al. End-to-end incremental learning[J]. Lecture Notes in Computer Science, 2018: 241-257.
- [62] KIM J Y, CHOI D W. Split-and-bridge: adaptable class incremental learning within a single neural network[J]. arXiv preprint, 2021, arXiv:2107.01349.
- [63] DENG D R, CHEN G Y, HAO J Y, et al. Flattening sharpness for dynamic gradient projection memory benefits continual learning[J]. arXiv preprint, 2021, arXiv:2110.04593.
- [64] TANG S X, CHEN D P, ZHU J G, et al. Layerwise optimization by gradient decomposition for continual learning[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 9629-9638.
- [65] SAHA G, GARG I, ROY K. Gradient projection memory for continual learning[J]. arXiv preprint, 2021, arXiv:2103.09762.
- [66] JIN X S, DU J Y, REN X. Gradient based memory editing for task-free continual learning[J]. arXiv preprint, 2020, arXiv:2006.15294.
- [67] CAI Z P, SENER O, KOLTUN V. Online continual learning with natural distribution shifts: an empirical study with visual data[C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 8261-8270.
- [68] CHAUDHRY A, KHAN N, DOKANIA P K, et al. Continual learning in low-rank orthogonal subspaces[J]. Advances in Neural Information Processing Systems, 2020: 1-12.
- [69] LIN S, YANG L, FAN D L, et al. TRGP: trust region gradient projection for continual learning[J]. arXiv preprint, 2022, arXiv:2202.02931.
- [70] ALJUNDI R, KELCHTERMANS K, TUYTELAARS T. Task-free continual learning[C]//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 11246-11255.
- [71] KIM C D, JEONG J, MOON S, et al. Continual learning on noisy data streams via self-purified replay[C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 517-527.
- [72] CHA H, LEE J, SHIN J. Co2L: contrastive continual learning[C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 9496-9505.
- [73] KUZINA A, EGOROV E, BURNAEV E. BooVAE: boosting approach for continual learning of VAE[J]. Advances in Neural Information Processing Systems, 2021.
- [74] MARACANI A, MICHIELI U, TOLDO M, et al. RECALL: replay-based continual learning in semantic segmentation[C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 7006-7015.
- [75] AYUB A, WAGNER A R. EEC: learning to encode and regenerate images for continual learning[J]. arXiv preprint, 2021, arXiv:2101.04904.
- [76] ROSTAMI M, KOLOURI S, PILLY P, et al. Generative continual concept learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4): 5545-5552.
- [77] SHIN H, LEE J K, KIM J, et al. Continual learning with deep generative replay[J]. arXiv preprint, 2017, arXiv:1705.08690.
- [78] OSTAPENKO O, PUSCAS M, KLEIN T, et al. Learning to remember: a synaptic plasticity driven framework for continual learning[C]//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 11313-11321.
- [79] HU X T, TANG K H, MIAO C Y, et al. Distilling causal effect of data in class-incremental learning[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 3956-3965.
- [80] LI T J, KE Q H, RAHMANI H, et al. Else-Net: elastic semantic network for continual action recognition from skeleton data[C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 13414-13423.
- [81] SINGH P, VERMA V K, MAZUMDER P, et al. Calibrating CNNs for lifelong learning[J]. Advances in Neural Information Processing Systems, 2020.
- [82] PELLEGRINI L, GRAFFIETI G, LOMONACO V, et al. Latent replay for real-time continual learning[C]//Proceedings of 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2020: 10203-10209.
- [83] KUNDU J N, VENKATESH R M, VENKAT N, et al. Class-incremental domain adaptation[M]//Computer vision – ECCV 2020. Cham: Springer International Publishing, 2020: 53-69.
- [84] WEI K, DENG C, YANG X, et al. Incremental embedding learning via zero-shot translation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. [S.l.:s.n.], 2021: 10254-10262.
- [85] STAN S, ROSTAMI M. Unsupervised model adaptation for continual semantic segmentation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. [S.l.:s.n.], 2021: 2593-2601.
- [86] ZHU F, ZHANG X Y, WANG C, et al. Prototype augmentation and self-supervision for incremental learning[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 5867-5876.

- [87] DE LANGE M, TUYTELAARS T. Continual prototype evolution: learning online from non-stationary data streams[C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 8230-8239.
- [88] OLIVERI G, VAN LAAKE L C, CARISSIMO C, et al. Continuous learning of emergent behavior in robotic matter[J]. Proceedings of the National Academy of Sciences of the United States of America, 2021, 118(21): e2017015118.
- [89] YAN S P, XIE J W, HE X M. DER: dynamically expandable representation for class incremental learning[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 3013-3022.
- [90] VENIAT T, DENOYER L, RANZATO M. Efficient continual learning with modular networks and task-driven priors[C]//Proceedings of the 9th International Conference on Learning Representations. [S.l.:s.n.], 2021.
- [91] XU J, MA J, GAO X, et al. Adaptive progressive continual learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.
- [92] CHEN T, ZHENYU Z, SIJIA L, et al. Long live the lottery: the existence of winning tickets in lifelong learning[J]. International Conference on Learning Representations, 2021, 51(4): 273-287.
- [93] SHI Y J, YUAN L, CHEN Y P, et al. Continual learning via bit-level information preserving[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 16669-16678.
- [94] MAZUMDER P, SINGH P, RAI P. Few-shot lifelong learning[J]. arXiv preprint, 2021, arXiv:2103.00991.
- [95] CACCIA M, RODRIGUEZ P, OSTAPENKO O, et al. Online fast adaptation and knowledge accumulation: a new approach to continual learning[J]. Advances in Neural Information Processing Systems, 2020.
- [96] HURTADO J, RAYMOND-SAEZ A, SOTO A. Optimizing reusable knowledge for continual learning via metalearning[J]. Advances in Neural Information Processing Systems, 2021.
- [97] VON OSWALD J, ZHAO D, KOBAYASHI S, et al. Learning where to learn: gradient sparsity in meta and continual learning[J]. arXiv preprint, 2021, arXiv:2110.14402.
- [98] GUPTA G, YADAV K, PAULL L. La-MAML: look-ahead meta learning for continual learning[C]//Proceedings of the Advances in Neural Information Processing Systems. [S.l.:s.n.], 2020.
- [99] VOLPI R, LARLUS D, ROGEZ G. Continual adaptation of visual representations via domain randomization and meta-learning[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 4441-4451.
- [100] LUO Y D, HUANG Z, ZHANG Z, et al. Learning from the past: continual meta-learning with Bayesian graph neural networks[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4): 5021-5028.
- [101] ZHANG J, ZHANG J T, GHOSH S, et al. Regularize, expand and compress: nonexpansive continual learning[C]//Proceedings of 2020 IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE Press, 2020: 843-851.
- [102] EBRAHIMI S, ELHOSEINY M, DARRELL T, et al. Uncertainty-guided continual learning with Bayesian neural networks[J]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2019: 75-78.
- [103] RAVAGLIA L, RUSCI M, NADALINI D, et al. A TinyML platform for on-device continual learning with quantized latent replays[J]. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2021, 11(4): 789-802.
- [104] PUTRA R V W, SHAFIQUE M. lpSpikeCon: enabling low-precision spiking neural network processing for efficient unsupervised continual learning on autonomous agents[J]. arXiv preprint, 2022, arXiv:2205.12295.
- [105] BURHAN HAFEZ M, WERMTER S. Behavior self-organization supports task inference for continual robot learning[C]//Proceedings of 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2021: 6739-6746.
- [106] XU J, ZHU Z. Reinforced continual learning[J]. Advances in Neural Information Processing Systems, 2018: 899-908.
- [107] LECARPENTIER E, ABEL D, ASADI K, et al. Lipschitz lifelong reinforcement learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. [S.l.:s.n.], 2021: 8270-8278.
- [108] BRAFMAN R I, TENNENHOLTZ M. R-MAX - a general polynomial time algorithm for near-optimal reinforcement learning[J]. IJCAI International Joint Conference on Artificial Intelligence, 2001, 3: 953-958.
- [109] ABEL D, JINNAI Y, YUE G, et al. Policy and value transfer in lifelong reinforcement learning[C]//Proceedings of the 35th International Conference on Machine Learning. [S.l.:s.n.], 2018: 30-41.
- [110] LOMONACO V, DESAI K R, CULURCIELLO E, et al. Continual reinforcement learning in 3D non-stationary environments[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE Press, 2020: 999-1008.
- [111] CHANDAK Y, THEOCHAROUS G, NOTA C, et al. Lifelong learning with a changing action set[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4): 3373-3380.
- [112] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529-533.
- [113] ROLNICK D, AHUJA A, SCHWARZ J, et al. Experience replay for continual learning[J]. arXiv preprint, 2019, arXiv:1811.11682.
- [114] LU K, GROVER A, ABBEEL P, et al. Reset-free lifelong Learning with skill-space planning[C]//Proceedings of the International Conference on Learning Representations. [S.l.:s.n.], 2020.
- [115] ALJUNDI R. Continual learning in neural networks[J]. arXiv preprint, 2019, arXiv:1910.02718.
- [116] MORENO-MUÑOZ P, RAMÍREZ D, ARTÉS-RODRÍGUEZ A. Continual learning for infinite hierarchical change-point detection[C]//Proceedings of 2020 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2020: 3582-3586.
- [117] BING L. Learning on the job: online lifelong and continual learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(9): 13544-13549.

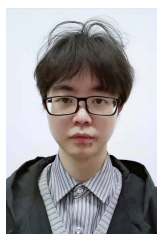
[118]FLESCH T, BALAGUER J, DEKKER R, et al. Comparing continual task learning in minds and machines[J]. Proceedings of the National Academy of Sciences of the United States of America, 2018, 115(44): 10313-10322.

[119]BARRON H C. Neural inhibition for continual learning and memory[J]. Current Opinion in Neurobiology, 2021, 67: 85-94.

[作者简介]



赵超（1996—），男，西安交通大学人工智能与机器人研究所博士生，主要研究方向为机器人持续学习。



许杰（1996—），男，西安交通大学人工智能与机器人研究所博士生，主要研究方向为人机交互、人体运动预测。



陈星宇（1990—），男，博士，西安交通大学人工智能与机器人研究所助理教授，主要研究方向为计算机视觉、机器人环境感知、人-机器人共融协作。



梅魁志（1978—），男，博士，西安交通大学人工智能与机器人研究所教授、博士生导师，主要研究方向为机器人场景感知与定位、深度学习/智能系统的软硬件协同设计。



兰旭光（1976—），男，博士，西安交通大学人工智能与机器人研究所教授、博士生导师，主要研究方向计算机视觉、统计学习与模式识别、深度强化学习与机器人学习、人-机器人共融协作。