DOI:10.7544/issn1000-1239.20201058
59(6):1213-1239,2022

连续学习研究进展

韩亚楠　刘建伟　罗雄麟
(中国石油大学(北京)信息科学与工程学院　北京　102249)
(857182813@qq.com)

Research Progress of Continual Learning

Han Yanan, Liu Jianwei, and Luo Xionglin
(*College of Information Science and Engineering, China University of Petroleum(Beijing), Beijing* 102249)

Abstract　In recent years, with the continuous development of information technology, all kinds of data have shown explosive growth. Traditional machine learning algorithms can only achieve better performance when the distribution of testing data and training data is similar. In other words, it is impossible to continuously and adaptively learn in dynamic environment. However, this ability that can learn adaptively in dynamic environment is very important for any intelligent systems. Deep neural networks have shown the best learning ability in many applications. However, when we apply these methods to incrementally update the model parameters, the model would face catastrophic interference or forgetting problems, which can cause the model to forget the old knowledge after learning a new task. The research of continual learning alleviates this problem. Continual learning is a process of simulating brain learning. It learns continual non-independent and identically distributed data streams in a certain order, and incrementally updates the model according to the results of task. The significance of continual learning is to efficiently transform and use the knowledge that has been learned to complete the learning of new tasks, and to greatly reduce the problems caused by forgetting. The study of continuous learning is of great significance for intelligent computing systems to adaptively learn changes in the environment. In view of the application value, theoretical significance and future development potential of continual learning, the article systematically reviews the research progress of continual learning. Firstly, this paper outlines the definition of continual learning. Three typical continual learning models are introduced, namely learning without forgetting, elastic weight consolidation and gradient episodic memory. Then, the key problems and solutions of continual learning are also introduced. After that, the three types of methods based on regularization, dynamic framework, memory replay and complementary learning systems have been introduced. At last, this paper points out potential challenges and future directions in the field of continual learning.

Key words　continual learning(CL); catastrophic forgetting; incremental learning; regularization; dynamic framework; memory replay

摘　要　近年来,随着信息技术的不断发展,各种数据呈现爆炸式的增长,传统的机器学习算法只有当测试数据与训练数据分布类似时,学习算法才能取得较好的性能,换句话说,它们不能在动态环境中连续

收稿日期:2020-12-25;**修回日期**:2021-07-05
基金项目:中国石油大学(北京)科研基金项目(2462020YXZZ023)
This work was supported by the Research Fund of China University of Petroleum(Beijing) (2462020YXZZ023).
通信作者:刘建伟(liujw@cup.edu.cn)

自适应地学习,然而,这种自适应学习的能力却是任何智能系统都具备的特性.深度神经网络在许多应用中显示出最好的学习能力,然而,使用该方法对数据进行增量更新学习时,会面临灾难性的干扰或遗忘问题,导致模型在学习新任务之后忘记如何解决旧任务.连续学习(continual learning, CL)的研究使这一问题得到缓解.连续学习是模拟大脑学习的过程,按照一定的顺序对连续非独立同分布的(independently and identically distributed, IID)流数据进行学习,进而根据任务的执行结果对模型进行增量式更新.连续学习的意义在于高效地转化和利用已经学过的知识来完成新任务的学习,并且能够极大程度地降低遗忘带来的问题.连续学习研究对智能计算系统自适应地适应环境改变具有重要的意义.基于此,系统综述了连续学习的研究进展,首先概述了连续学习的定义,介绍了无遗忘学习、弹性权重整合和梯度情景记忆3种典型的连续学习模型,并对连续学习存在的关键问题及解决方法进行了介绍,之后又对基于正则化、动态结构和记忆回放互补学习系统的3类连续学习模型进行了分类和阐述,并在最后指明了连续学习进一步研究中需要解决的问题以及未来可能的发展方向.

关键词 连续学习;灾难性遗忘;增量学习;正则化;动态结构;记忆回放

中图法分类号 TP391

近年来,随着机器学习(machine learning, ML)领域的快速发展,机器学习在自然图像分类、人脸识别等领域取得了一定的成果,深度学习的成功使机器学习的发展达到了另一个新的高度.然而,在现实世界中,机器学习系统总是会遇到连续任务学习问题,因此,如何对连续任务进行有效学习是当前研究的重点之一.现有的机器学习方法虽然可以在任务上取得较高的性能,但只有当测试数据与训练数据概率分布类似时,机器学习才能取得较好的性能.换句话说,目前的机器学习算法不能在动态环境中持续自适应地学习,因为在动态环境中,任务可能会发生显著变化,然而,这种自适应的学习能力却是任何智能系统都具有的能力,也是实现智能生物系统学习的重要标志.

目前,深度神经网络在许多应用中显示出非凡的预测和推理能力,然而,当通过基于梯度更新的方法对模型进行增量更新时,模型会出现灾难性的干扰或遗忘问题,这一问题将直接导致模型性能的迅速下降,即模型在学习新任务之后,由于参数更新对模型引起的干扰,将使得学习的模型忘记如何解决旧任务.人类和动物似乎学到了很多不同的知识,并且总是能不遗忘过去学到的知识,并将其应用在未来的学习任务中,受人和动物这种学习方式的启发,很自然地将这种想法运用到机器学习领域,即随着时间的推移,模型能够不断学习新知识,同时保留以前学到的知识,这种不断学习的能力被称为连续学习.连续学习最主要的目的是高效地转化和利用已经学过的知识来完成新任务的学习,并且能够极大程度地降低灾难性遗忘带来的问题.近年来,随着深度学习的不断发展,连续学习的研究已经受到极大的关注,因为连续学习主要有2点优势:

1) 不需要保存之前任务上学习过的训练数据,从而实现节约内存,同时解决了由于物理设备(例如机器内存)或学习策略(例如隐私保护)的限制,导致数据不能被长期存储这一问题.

2) 模型能够保存之前任务所学习的知识,并且能够极大程度地将之前任务学习到的知识运用到未来任务的学习中,提高学习效率.

1 连续学习概述

1.1 连续学习的形成与发展

在现实世界中,机器学习系统处于连续的信息流中,因此需要从不断改变的概率分布中学习和记住多个任务.随着时间的推移,不断学习新知识,同时保留以前学到知识,具备这种不断学习的能力称为连续学习或终身学习.因此,使智能学习系统具备连续学习的能力一直是人工智能系统面临的挑战[1-2].灾难性遗忘或灾难性干扰一直是连续学习所研究的重点,即当模型对新任务进行学习时会遗忘之前任务所学习的知识,这种现象通常会导致模型性能的突然下降,或者在最坏的情况下,导致新知识完全覆盖旧知识.因此,克服灾难性遗忘是人工智能系统迈向更加智能化的重要一步.

早期学者们曾尝试为系统增加一个存储模块来保存以前的数据,并定期对之前所学的知识与新样本的交叉数据进行回放来缓解灾难性遗忘这一问题[3],这类方法一直延续至今[4-5].然而,基于存储模

块连续学习方法的一个普遍缺点是它们需要显式存储旧任务信息,这将导致较大的工作内存需求,此外,在计算和存储资源固定的情况下,应设计专门的机制保护和巩固旧的知识不被新学习的知识所覆盖.在此基础上,Rusu 等人[6-7]尝试在新任务到来时,分配额外的资源来缓解灾难性遗忘.然而,这种方法随着任务数量的不断增加,神经网络架构将不断增加,进而直接降低模型的可伸缩性.由于连续学习场景中不能预先知道任务数量和样本大小,因此,在没有对输入训练样本的概率分布做出很强的假设情况下,预先定义足够的存储资源是不可避免的.在这种情况下,Richardson 等人[8]提出了针对连续学习模型避免灾难性遗忘的 3 个关键方面:1)为新知识分配额外的神经元;2)如果资源是固定的,则使用新旧知识的非重叠表示;3)把旧的知识叠加到新的知识上作为新的信息.在此基础上,受神经科学理论的启发,基于正则化策略、动态结构策略以及记忆策略等一系列连续学习的方法相继被提出.

1.2 连续学习的定义

目前,连续学习的研究仍然处于发展阶段,还没有明确一致的定义,本文对有监督连续学习给出定义.

在一个典型的学习环境中,我们的目标是使用一个具有独立同分布(independently and identically distributed, IID)并带有标签的训练数据集 $D=\{(x_t^i,y_t^i)\}_{t=1}^{T}$ 来完成对模型参数 $\boldsymbol{w}$ 的学习,进而实现对未来数据 (x^*,y^*) 的准确预测,即学习 $p(y^*\mid \boldsymbol{w},x^*)$.

而在连续学习环境中,数据集 D 中的数据所服从的概率分布不再是典型的 IID 概率分布,而是将它们分解成若干个不相交的子集,即 $D=\{(x_t^i,y_t^i)\}_{t=1}^{T}$,同时,假设这些集合是从 T 个不同的 IID 概率分布中抽样所得的,其中每个概率分布都代表一个任务.目前,大多情况下都假设,任务是明显不相交的,尽管在实际情况下通常不是这样.

假设有 T 个学习任务 $1,2,\cdots,T$,形成一个任务序列 $t\in\{1,2,\cdots,T\}$,其中每个任务 t 上的数据集为 $D_t=\{(x_t^1,y_t^1),(x_t^2,y_t^2),\cdots,(x_t^{N_t},y_t^{N_t})\}$,每个数据集包含的数据个数为 N_t.连续学习过程中,模型按照顺序从 $1\sim T$ 完成对每个任务的训练,且不能重新访问之前的任务.也就是说,假如学习者已经完成了 t 个学习任务 $1,2,\cdots,t$,当面对第 $t+1$ 的任务和它的数据 D_{t+1} 时,学习者能够不访问前 t 个任务的数据来完成对第 $t+1$ 个任务的学习,且不遗忘之前所学习的知识.连续学习的目的是能够学习一个单一的模型,它能够很好地预测来自任何任务的数据.连续学习的过程如图 1 所示:

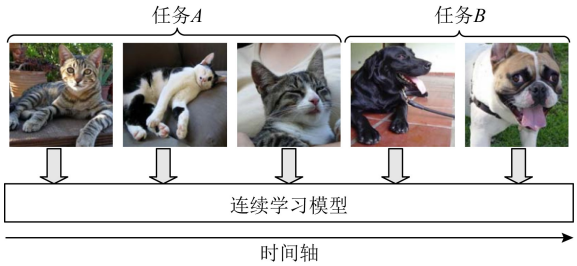

Fig. 1 Illustration of continual learning

图 1 连续学习示意图

如图 1 所示,在连续学习过程中,智能体逐个对每个连续的非独立均匀分布流数据示例进行学习,并且该智能体对每个示例只进行一次访问.这种学习方式与动物学习过程更为接近.如果我们忽略各个任务的先后次序问题,单独训练每个任务,这将导致灾难性遗忘,这也是连续学习一直以来所面临的最大问题.因此,连续学习的本质,是通过各种手段高效地转化和利用已经学过的知识来完成新任务的学习,并且能够极大程度地降低遗忘带来的问题.

1.3 连续学习场景

连续学习的问题是指模型能够连续学习一系列任务,其中,在训练期间,只有当前任务数据可用,并且假设任务间是有明显的分界[9].近年来,对这一问题,研究者们已展开积极的研究,提出了许多缓解连续学习过程中灾难性遗忘的方法.然而,由于各实验方案的不同,因此直接对各方法进行比较评估显然不可行.尤其是模型任务标识不可用等问题,这将直接影响模型实现的难易程度.因此,为了使评价更加标准化,并且也为了使实验结果比较更具意义,在此首先对连续学习过程中的 3 个学习场景进行简要概括[10],如表 1 所示:

Table 1 Three Continual Learning Scenarios

表 1 3 种连续学习场景

学习场景	测试要求
Task-IL	解决当前任务,提供 Task-ID
Domain-IL	解决当前任务,不提供 Task-ID
Class-IL	解决当前任务,推断 Task-ID

在第 1 个学习场景中,模型总是被告知需要执行哪些任务,这也是最简单的连续学习场景,将其称为任务增量学习(task-incremental learning, Task-IL).近年来,提出的大部分连续学习方法在此场景

都是适用的,且都具有较好的实验效果,例如正则化方法和动态结构方法等.

在第 2 个学习场景中,通常将其称之为域增量学习(domain-incremental learning, Domain-IL),任务标识不可用,模型只需要解决手头的任务,模型也不需要推断这是哪个任务.文献[11]的实验结果证明,基于情景记忆的方法在该场景下有较好的实验结果,例如 GER,DGR,RtF 等,然而基于正则化方法,例如 EWC,LwF,SI 等,模型学习的准确率相对较差.

在第 3 个学习场景中,模型必须能够解决到目前为止所看到的每个任务,并且还能够推断出它们所面临的任务,将此场景称为类增量学习(class-incremental learning, Class-IL),在该场景中包含一个很常见的实际问题,即增量地学习对象的新类.此场景是这 3 个场景中最为复杂的,也是最接近现实中的学习场景,近年来,针对此场景下的连续学习方法也相继提出.例如,通过存储之前任务数据的样本,缓解系统遗忘方法:文献[5]提出一种 iCarl 的连续学习方法,该方法通过在每个类中找出 m 个最具代表性的样本,那么其平均特征空间将最接近类的整个特征空间,最后的分类任务是通过最接近样本均值的分类器来完成的;文献[12]介绍了对遗忘和不妥协量化的度量方法,进而提出一种称为 RWalk 方法,完成类增量场景下的学习;文献[13]提出一种动态网络扩展机制,通过由所学习的二进制掩码动态确定网络所需增加的容量,以确保足够的模型容量来适应不断传入的任务.

1.4 连续学习相关领域研究

连续学习相关的领域研究主要包括多任务学习和迁移学习.

1) 多任务学习.多任务学习的目的是能够结合所有任务的共同知识,同时改进所有单个任务的学习性能,因此,多任务学习要求每个任务与其他任务共享模型参数,或每个任务有带约束的模型参数,别的任务能够给当前学习任务提供额外的训练数据,以此来作为其他任务的正则化形式.也就是说,多任务学习的良好效果依赖于单个函数的共享参数化以及对多个损失同时进行估计和求平均.当同时训练多个任务的共享层时,必须学习一个公共表示,从而有效地对每个任务进行交叉正则化,约束单个任务的模型.

对于神经网络而言,Caruana[14] 对多任务学习进行了详细的研究,指出网络的底层是共享的,而顶层是针对于特定任务的,多任务学习需要所有任务的数据,此外,多任务学习随着时间的推移,不会积累任何知识,也就是说没有持续学习的概念,这也是多任务学习的关键问题所在.

2) 迁移学习.迁移学习是使用源域来帮助另一个任务完成目标域学习的一种学习方式[15].它假设源域 S 中有大量的标记训练数据,而目标域 T 只有很少或没有标记的训练数据,但有大量未标记的数据.迁移学习可以利用被标记的数据来帮助完成目标域中的学习.然而迁移学习与连续学习,主要有 4 个不同:①迁移学习不是连续的,它仅仅是使用了源域来帮助完成目标域学习;②迁移学习并没有将过去所学的知识进行积累;③迁移学习是单向进行的,也就是说,迁移学习仅可使用源域来帮助完成目标域的学习,然而,连续学习是可以在任何方向上进行学习的;④迁移学习假设源域与目标域非常相似,且这种相似性是人为决定的,然而在连续学习中并没有做出这样一个很强的限制性假设.

2 连续学习的典型模型

2.1 无遗忘学习

Li 等人[16]在 2017 年提出了一种由卷积神经网络(convolutional neural network, CNN)组成的无遗忘学习(learning without forgetting, LwF)方法,该方法将知识蒸馏(knowledge distillation, KD)[17]与细调方法[18]相结合,其中,利用知识蒸馏策略来避免对之前知识的遗忘.

假设给定一个 CNN 神经网络,$\boldsymbol{\theta}_{\text{share}}$ 为网络的共享参数,$\boldsymbol{\theta}_{\text{old}}$ 是任务特定的参数.我们的目标是为一个新任务增加一个任务特定的参数 $\boldsymbol{\theta}_n$,并且只利用新的数据和标签(不使用已经存在任务的标签数据)对特定的任务参数 $\boldsymbol{\theta}_n$ 进行学习,使得它能够对新的任务和之前的任务都有好的预测效果.无遗忘学习方法的示意图如图 2 所示:

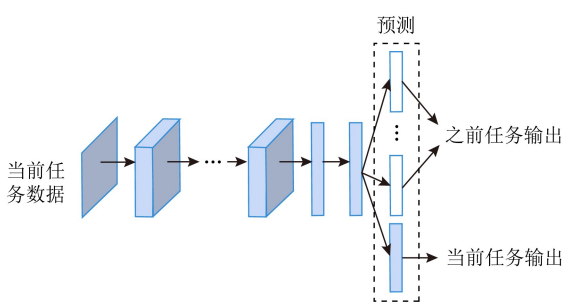

Fig. 2 Illustration of learning without forgetting

图 2 无遗忘学习方法示意图

训练阶段,为了使所有任务的损失最小化,使用随机梯度下降对正则化项进行训练.其中,预训练阶段,固定参数 $\boldsymbol{\theta}_{\text{share}}$ 和 $\boldsymbol{\theta}_{\text{old}}$,然后训练 $\boldsymbol{\theta}_n$ 直到模型收敛.然后在联合最优化步骤中,联合训练 $\boldsymbol{\theta}_{\text{share}}$,$\boldsymbol{\theta}_{\text{old}}$,$\boldsymbol{\theta}_n$ 直到模型收敛.假设给定一个新任务的训练数据 $(\boldsymbol{X}_n,\boldsymbol{Y}_n)$,对于新的输入数据,在旧任务分类器上的输出为 $\boldsymbol{Y}_{\text{old}}$,对于新的参数 $\boldsymbol{\theta}_n$ 进行随机初始化,则对参数 $\boldsymbol{\theta}^*_{\text{share}},\boldsymbol{\theta}^*_{\text{old}},\boldsymbol{\theta}^*_n$ 进行更新:

$$\boldsymbol{\theta}^*_{\text{share}},\boldsymbol{\theta}^*_{\text{old}},\boldsymbol{\theta}^*_n \leftarrow \underset{\hat{\boldsymbol{\theta}}_{\text{share}},\hat{\boldsymbol{\theta}}_{\text{old}},\hat{\boldsymbol{\theta}}_n}{\arg\min}\left(\lambda L_{\text{old}}(\boldsymbol{Y}_{\text{old}},\hat{\boldsymbol{Y}}_{\text{old}})+L_{\text{new}}(\boldsymbol{Y}_n,\hat{\boldsymbol{Y}}_n)+R(\hat{\boldsymbol{\theta}}_{\text{share}},\hat{\boldsymbol{\theta}}_{\text{old}},\hat{\boldsymbol{\theta}}_n)\right), \tag{1}$$

其中,$L_{\text{old}}(\boldsymbol{Y}_{\text{old}},\hat{\boldsymbol{Y}}_{\text{old}})$ 和 $L_{\text{new}}(\boldsymbol{Y}_n,\hat{\boldsymbol{Y}}_n)$ 是损失函数,分别使用参数 $\boldsymbol{\theta}_{\text{share}},\boldsymbol{\theta}_{\text{old}},\boldsymbol{\theta}_n$ 来最小化在新旧任务之间的预测值 $\hat{\boldsymbol{Y}}$ 和真值 $\boldsymbol{Y}$ 之间的差异;λ 是新旧任务之间的权衡因子;R 是模型的正则化项,用于防止过拟合.

2.2 弹性权重整合

Kirkpatrick 等人[19]在 2017 年提出了一种结合监督学习和强化学习方法,即弹性权重整合(elastic weight consolidation, EWC)方法.在提出的模型目标函数中,包括了对新旧任务之间模型参数的惩罚项,从而有效缓解对先前学习的知识中与当下任务相关知识遗忘.弹性权重整合示意图如图 3 所示:

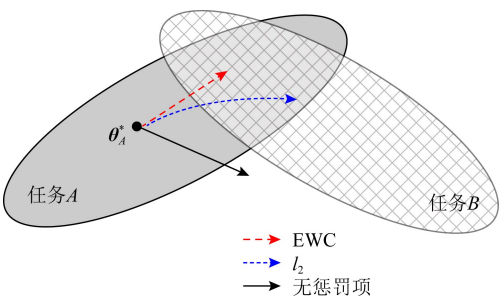

Fig. 3 Illustration of elastic weight consolidation

图 3 弹性权重整合示意图

图 3 为训练模型在模型参数空间中的轨迹展示,弹性权重整合可以保证在训练任务 B 的同时记住任务 A.在模型参数区域上任务 A(灰色背景)和任务 B(网格背景)的学习性能均表现良好.在学习第 1 个任务 A 后,学习的模型参数为 $\boldsymbol{\theta}^*_A$.此时,如果在任务 B 上独立地采用梯度学习步骤(密虚线箭头),可以最小化任务 B 上的损失,但是会破坏先前对任务 A 学到的东西.另外,如果用相同的权衡参数约束每个任务上的模型参数(实线箭头),所施加的权衡参数限制将会太过严重,这时虽然可以记住任务 A 上的模型,但是不能有效地学习任务 B 的模型.EWC 通过考虑权衡参数对任务 A 的重要性,把权衡参数的重要性与损失函数相关联,进而找到既能解决任务 B 又不会导致任务 A 重大损失的权衡参数(稀虚线箭头).

具体而言,通过一个模型参数 $\boldsymbol{\theta}$ 的后验概率分布 $p(\boldsymbol{\theta}|D)$ 对于任务训练数据集 D 的模型参数 $\boldsymbol{\theta}$ 进行建模.假设学习场景有 2 个独立的任务 $A(D_A)$ 和任务 $B(D_B)$,那么根据贝叶斯规则,模型参数 $\boldsymbol{\theta}$ 的后验概率的对数值表示为

$$\log p(\boldsymbol{\theta}|D)=\log p(D_B|\boldsymbol{\theta})+\log p(\boldsymbol{\theta}|D_A)-\log p(D_B), \tag{2}$$

其中,模型参数 $\boldsymbol{\theta}$ 的后验概率 $\log p(\boldsymbol{\theta}|D_A)$ 中包含了关于之前任务 A 上的所有信息,然而,由于估计该后验概率分布有困难,所以在 EWC 方法中将其近似为一个数学期望为 $\boldsymbol{\theta}^*_A$ 的高斯分布,由费雪尔信息矩阵(fisher information matrix, FIM)对角元素给出 $\boldsymbol{\theta}^*_A$ 的精度参数.因此,EWC 的损失函数为

$$L(\boldsymbol{\theta})=L_B(\boldsymbol{\theta})+\sum_i\frac{\lambda}{2}\boldsymbol{F}_i(\theta_i-\boldsymbol{\theta}^*_{A,i})^2, \tag{3}$$

其中,$L_B(\boldsymbol{\theta})$ 为任务 B 的损失函数,λ 表示新旧任务之间的相关性权衡参数,i 表示参数的下标索引,$\boldsymbol{F}$ 表示 FIM.因此,这种方法需要对学习任务的模型参数进行对角线加权,该加权值与 FIM 的对角线元素值成比例.

2.3 梯度情景记忆

Lopez-Paz 等人[20]在 2017 年提出梯度情景记忆模型(gradient episodic memory, GEM),该模型能够实现知识正向迁移到先前任务的功能,以及将先前任务学习的知识正向地迁移到当前任务上.GEM 模型最主要的特征是为每个任务 k 存储一个情景记忆模型 M_k 来避免灾难性遗忘,该模型不仅能够最小化当前任务 t 的损失,而且可以将任务 $k<t$ 情景记忆模型的损失 $l(f_{\boldsymbol{\theta}},M_k)$ 作为不等式约束来避免损失函数 $l(f_{\boldsymbol{\theta}},D_t)$ 的增加,但允许损失函数 $l(f_{\boldsymbol{\theta}},D_t)$ 的减少.对于任务 t,GEM 模型的目标函数为

$$\begin{aligned}&\min_{\boldsymbol{\theta}} l(f_{\boldsymbol{\theta}},D_t),\\&\text{s.t. } l(f_{\boldsymbol{\theta}},M_k),\forall k<t,\\&l(f_{\boldsymbol{\theta}},M_k)=\frac{1}{|M_k|}\sum_i l(f_{\boldsymbol{\theta}}(x_i,k),y_i),\end{aligned} \tag{4}$$

其中,$f^{t-1}_{\boldsymbol{\theta}}$ 表示直到任务 $t-1$ 时训练神经网络得到的模型预测值.为了及时对损失 $l(f_{\boldsymbol{\theta}},D_t)$ 是否增加进行检测,GEM 模型计算记忆任务损失的梯度向量 $\boldsymbol{g}_k$ 与当前任务损失的梯度向量 $\boldsymbol{g}_t$ 之间的夹角.

对于$\forall k<t$,$\boldsymbol{g}_k$中任意一个与当前任务$\boldsymbol{g}_t$的梯度向量之间的夹角大于90°时,也就是说,与记忆任务存在较大差异时,预测过程以l_2范数作为度量函数,将梯度$\boldsymbol{g}$投影到离梯度$\tilde{\boldsymbol{g}}$最近的点,从而使夹角保持在使损失$l(f_\theta, D_t)$增加的区间内.通过对其损失进行约束使其不增加,进而使得模型避免对之前任务的遗忘.

2.4 分析比较

LwF方法仅需要使用新任务的数据,对新任务进行优化,以提高新任务上模型预测的准确性,并保持神经网络对以前任务的预测性能.这种方法类似于联合训练方法,但是该学习方法不使用旧任务的数据和标签数据.实验表明,LwF方法可以极大地提高算法的分类性能以及计算效率,简化了学习过程,一旦学习了一个新的任务,训练过的数据将不需要再被保存或者回放.然而,这种方法的缺点是学习的性能高度依赖于任务的相关性,并且单个任务的训练时间随着学习任务的个数线性增加.虽然蒸馏方法为多任务学习提供了一个潜在的解决方案,但它需要为每个学习任务持久存储数据.另外需要注意,LwF方法不能被直接运用到强化学习场景中;EWC方法通过使用FIM对网络参数进行约束,降低模型对以前所学知识的遗忘程度,此外,该方法在训练过程中不增加任何计算负担,但这是以计算FIM为代价的,需存储FIM的值以及以前学习模型参数的副本;Lopez-Paz等人[20]的实验结果表明GEM模型,相较于LwF和EWC方法具有较好的实验效果,但是,该方法在训练时,由于对于每个任务都需要进行情景记忆,因此需要更多的内存空间,所需的内存是EWC用于保存过去信息大小的2倍,与其他方法相比内存开销较大,并且随着学习任务数量的增加,训练成本急剧增加,此外该方法也不能增量地对新的类别进行学习;同时提高性能也将加大计算负担.

3 连续学习的关键问题

3.1 灾难性遗忘

灾难性遗忘是连续学习面临的最大挑战.避免灾难性遗忘的问题,也就是说,在不断完成有序到达的新任务学习的同时,也能够在之前学习过的任务中表现得足够好.

Venkatesan等人[21]在2017年设计了一种结合生成式模型和知识蒸馏技术的全新采样策略,用其来产生来自过去学习任务概率分布上的"幻觉数据",使模型在不访问历史数据的前提下,缓解连续学习过程中的灾难性遗忘问题;文献[22]从序列贝叶斯学习规则出发,假定数据序列到达时,用前一个任务模型参数的后验概率分布作为新任务模型参数的先验概率分布,为缓解连续学习过程中的灾难性遗忘问题提供一种解决方案;文献[19]提出的正则化方法在模型参数更新时增加约束,以此在保持已有知识的前提下,实现对新任务的学习,来缓解灾难性遗忘等.

3.2 知识的正向迁移

连续学习过程中的知识正向迁移,即连续学习应该能够在学习新任务的同时,利用以前的任务中学习到的知识来帮助新任务的学习,从而提高学习的效率和质量.

文献[23]实验证明简单的细调可以实现知识的正向迁移;文献[24]提出保留训练好的模型基类信息编码,可将其知识迁移到模型要学习的新类中;文献[16]提出的LwF方法中,使用蒸馏损失来保存基类信息,进而使用保存的基类信息用于新数据的训练;文献[6]通过继承之前任务所学的知识,完成对新任务的学习;LGM模型是基于学生-教师的双重体系结构[25],教师的角色是保存过去的知识并帮助学生学习未来的知识,该模型通过优化一个增广的ELBO目标函数很好地帮助完成师生知识的正向迁移;文献[26]提出一种符号程序生成(symbolic program synthesis, SPS)的方法,来实现知识的正向迁移等.

3.3 知识的正向和反向迁移

知识在反向传播过程中的正向迁移,即如何利用当前任务所学到的知识来帮助之前任务的学习是连续学习模型研究的重点之一.

在连续学习场景中提出的LwF模型或者具有更为复杂正则化项的EWC模型,虽然可以在一定程度上缓解灾难性遗忘这一问题,然而却无法实现利用当前任务知识来帮助之前任务的学习.Li等人[27]在2019年提出一种连续结构学习框架,当网络进行结构搜索时,l层被选择"重用",即第l层能够学习到一个与先前的某个任务非常相似的表示,这要求l层的2个学习任务之间存在语义相关,因此,在第l层上使用正则化项对模型进行相应的约束来帮助之前任务的学习,该模型的提出为解决利用当前任务知识来帮助之前任务的学习提供了思路;Lopez-Paz等人[20]提出梯度情景记忆模型,实现

知识正向迁移到先前任务功能,进而提高模型对之前任务学习的学习能力.

3.4 可伸缩性能力

连续学习方法应该具有可伸缩性或扩展能力,也就是说,该方法既能完成小规模数据任务的训练,也能够可伸缩地实现大规模任务上的训练学习,同时需要能够保持足够的能力来应付不断增加的任务.

Schwarz 等人[28]在 2018 年提出一种进步和压缩框架(progress and compress framework, P&C)的连续学习模型,P&C 模型是由知识库(knowledge base)和活动列(active column)两部分组成,这个由快速学习和整合组成的循环结构,使模型不需要结构的增长,也不需要访问和存储以前的任务或数据,也不需要特定的任务参数来完成对新任务的学习,此外,由于 P&C 模型使用了 2 个固定大小的列,所以可以扩展到大规模任务上;文献[9]提出一种动态生成记忆模型(dynamic generative memory, DGM),在 DGM 模型中,利用一个生成对抗结构来替代之前模型的记忆模块,来缓解灾难性遗忘问题.其中,该模型中还结合一个动态网络扩展机制,以确保有足够的模型容量来适应不断传入的新任务;Yoon 等人[29]在 2018 年提出了一种新型的面向终身连续学习的深度网络结构,称为动态可扩展网络(dynamically expandable network, DEN),它可以在对一系列任务进行训练的同时动态地确定其网络容量,从而学习任务之间紧密重叠的知识共享结构,进而有效地对各任务间的共享和私有知识进行学习,不断学习新任务的同时有效地缓解灾难性遗忘.

4 连续学习方法研究进展

本节将具体介绍多个代表性的连续学习方法,本文将把目前的连续学习分为基于正则化方法、基于动态结构方法和基于情景记忆方法三大类,并阐明不同方法之间的关系,还比较了这些方法在减轻灾难性遗忘性能的差异性.图 4 是对近年来提出的一些流行的连续学习策略韦恩图总结.

连续学习中各个子类的分类图如图 5~7 所示.图中从模型引出到下一模型的箭头,代表了下一模型是在上一模型的基础上发展演变得来.

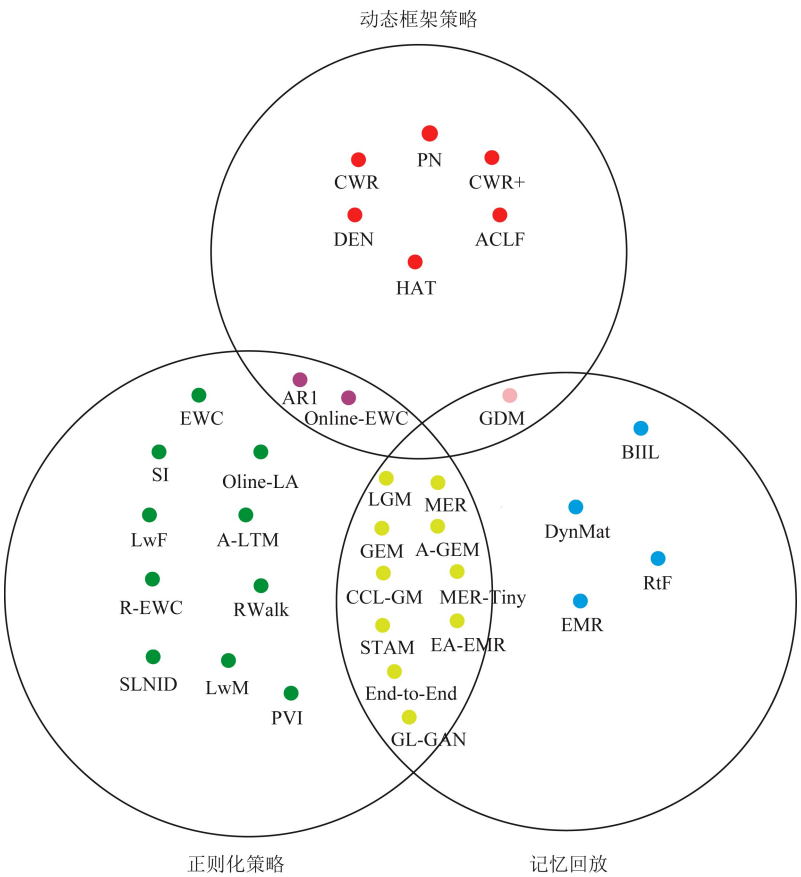

Fig. 4 Venn graph of the approaches for continual learning

图 4 连续学习方法 Venn 图

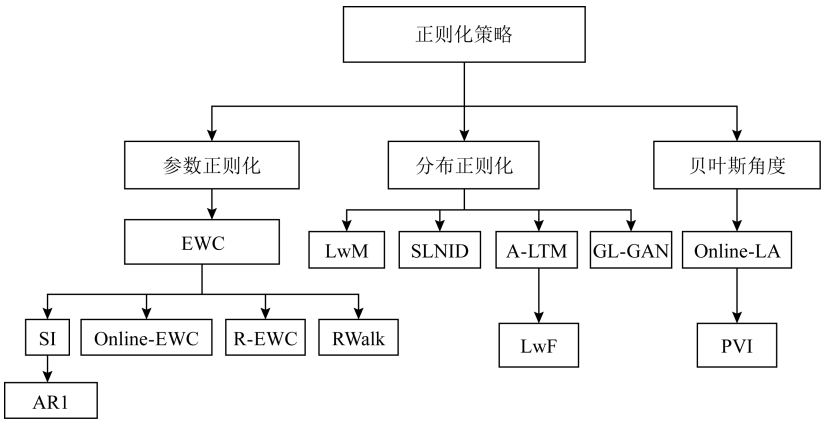

Fig. 5 Illustration of the classification for regularization model

图 5 正则化模型分类示意图

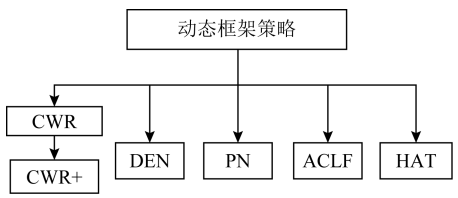

Fig. 6 Illustration of the classification for dynamic structural models

图 6 动态结构模型分类示意图

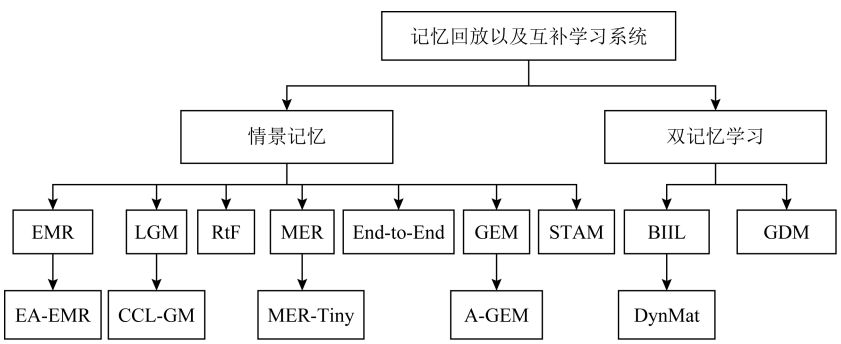

Fig. 7 Illustration of the classification for the memory replay and complementary learning methods

图 7 基于记忆回放以及互补学习方法分类示意图

4.1 正则化方法

在神经科学理论模型中,通过具有不同可塑性水平级联状态的触突刺激,来保护巩固学习的知识不被遗忘.受到这一机制的启发,从计算系统的角度来看,可以通过对模型施加正则化约束来避免遗忘.通过正则化方法在模型权重更新时加强约束,以此在保持已有知识的前提下实现对新任务的学习,来缓解灾难性遗忘这一问题.以下对近年来常见的基于正则化连续学习方法进行简要概括总结.

4.1.1 动态的长期记忆网络

在人工神经网络中,当不同的任务被依次学习时,连续学习会受到干扰和遗忘.Furlanello 等人[30]受到 McClelland 关于海马开创性理论[31]启发,在 2016 年提出一个新颖的基于知识蒸馏的主动长期记忆网络模型(active long term memory network, A-LTM),它是一种顺序多任务深度学习模型,能够在获取已知知识的同时,保持先前学习过的任务输入和行为输出之间的关联,也就是不遗忘之前所学习的知识.

A-LTM 模型主要由稳定的网络模块 N(neocortex)、灵活的网络模块 H(hippocampus)和双重机制 3 部分组成.其中,模块 N 用于保持对长期任务的记忆,当对新任务进行学习时,模块 H 的权重首先由模块 N 初始化,进而实现任务的学习,双重机

制则允许在不忽略新输入的情况下保持模块 N 的稳定性.

在模型训练发展阶段,首先对模块 N 进行训练,其中,模块 N 在一个受控环境下进行训练,也就是说,训练样例具有丰富的监督信息且服从一个稳定的概率分布.在进行训练时,利用该包含监督信息的训练样例训练网络模型,导致模型收敛.当学习任务发生改变时,模块 H 首先利用模块 N 的知识信息直接初始化,进而可以有效地利用之前任务的知识.通过动态地对梯度下降过程施加约束,实现在新旧任务间的权衡,进而快速地达到局部最优,也即是说,模块 H 具有快速适应新任务能力.

4.1.2 SI 模型

为缓解连续学习过程中 EWC 算法对 FIM 的计算实现较为复杂的问题,Zenke 等人[32] 在 2017 年提出了一种在线计算权重重要性的方法,即训练时根据各参数对损失贡献的大小来动态地改变参数的权重,如果参数 θ_i 对损失的贡献越大,则说明该参数越重要,该方法称为 SI(synaptic intellgence, SI)模型.权重的重要性计算为

$$F_k=\frac{\sum_k \Delta L_k}{T_k^2+\xi},\qquad \Delta L_k=\Delta\boldsymbol{\theta}_k\cdot\frac{\partial L}{\partial\boldsymbol{\theta}_k},\tag{5}$$

其中,$\Delta\boldsymbol{\theta}_k$ 是权重的更新量,$\frac{\partial L}{\partial\boldsymbol{\theta}_k}$是梯度,$\sum_k \Delta L_k$ 表示总的损失改变,T_k 是权重 $\boldsymbol{\theta}_k$ 总的运行轨迹,ξ 是一个小的常数,避免分母为 0.因此,由于计算 F_k 所需的全部数据在计算 SGD 期间是可用的,不需要额外的计算开销,有效地降低计算成本.

4.1.3 AR1 模型

Maltoni 等人[33] 提出一种结合结构(architectural)和正则化(regularization)策略的方法,简称为 AR1 模型,实现了单增量任务(single-incremental-task)场景的学习.AR1 模型是在 CWR(copy weights with re-init)模型[34] 的基础上,对权重 $\overline{\boldsymbol{\Theta}}$ 实行正则化约束,实现权重的跨批次调整.正则化方法倾向于逐步减少对每批权重更改的大小,其中大多数更改发生在顶层.在 AR1 中间层中,对权重的调整没有对遗忘产生负面的影响.AR1 方法在 CORe50 和 CIFAR-100 这 2 个数据集上的实验验证结果表明,AR1 可以被用来训练深度卷积模型,其遗忘率更低,性能也优于 LwF,EWC,SI 模型.

4.1.4 Online-EWC 模型

Schwarz 等人[28] 在 2018 年提出一种基于进步和压缩框架(progress and compress framework, P&C)的 Online-EWC 模型.该模型是一种结构可伸缩的连续学习方法,主要由知识库和活动列 2 部分组成,模型通过对这 2 部分进行交替优化,实现知识的正向迁移.这 2 部分可以被看作为网络层的列,在监督学习的情况下用于预测类的概率,在强化学习的情况下用于产生策略或奖励值(policies/values).图 8 表示将 P&C 框架应用于强化学习时知识库和活动列 2 个部分交替学习的过程.

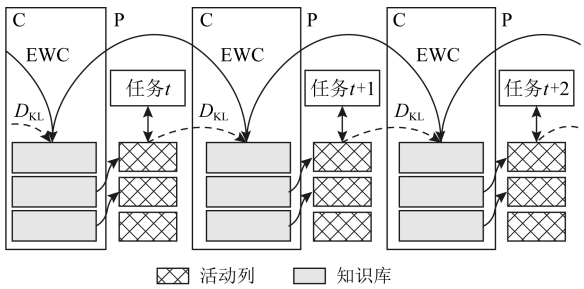

Fig. 8 Illustration of P&C

图 8 P&C 学习过程示意图

如图 8 所示,在对新任务进行学习时,也就是在“progress”阶段,首先固定知识库(灰色背景)模块,对活动列(网格背景)模块参数进行优化,其中在该优化过程中没有施加任何约束或者正则化项.值得注意的是,在该过程中可以通过一个面向知识库的简单分层适配器来实现对过去已学习到的知识(知识库)进行重用.

在“compress”阶段,模型需要进行知识蒸馏,也就是说,模型需要将新学习到的知识,正向地迁移到知识库中.该阶段的执行过程与经典的 EWC 相似,但是不同的是,该模型通过使用在线逼近算法来近似对角 FIM,将克服 EWC 随着任务个数的增加,计算量线性增加的问题.

4.1.5 R-EWC 模型

Liu 等人[35] 考虑到在 EWC 算法中,假设 FIM 是对角的,但是这一假设在实际中通常是不成立的,所以提出了一种基于旋转空间的 EWC 算法(EWC in the rotated space, R-EWC).该算法是通过对参数空间的因式旋转,即让$\frac{\partial \log p}{\partial\boldsymbol{\theta}}$与坐标轴对齐,如图 9 所示,对网络参数进行重参数化,以此实现对网络参数的 FIM 近似对角化.在旋转空间下的 EWC 算法能够更好地降低连续学习过程中的遗忘问题.

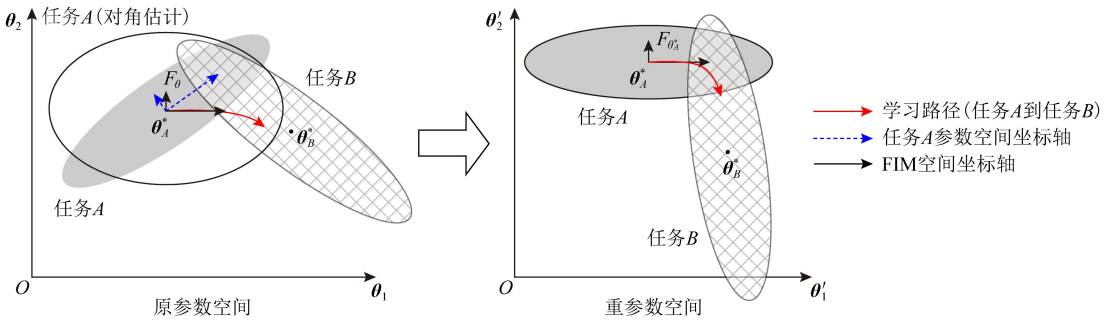

Fig. 9 Illustration of R-EWC

图 9 R-EWC 示意图

因此,R-EWC 算法主要是对参数空间 $\boldsymbol{\theta}$ 进行再参数化,具体而言,需要找到一个不改变网络前馈响应的重参数化空间,并能够更好地满足对角 FIM 的假设.在重新参数化以后,假设在新的参数空间上可以进行对角 FIM 估计,最后,在这个新的参数空间中,通过梯度下降对新任务进行最小化学习.

4.1.6 RWalk 模型

连续学习模型在增量学习过程中,除了面临遗忘问题之外,还容易遭受不妥协(intransigence)问题,即模型无法有效地对新任务学习的知识进行更新,Chaudhry 等人[12]对此问题进行权衡,提出 RWalk (Riemannian walk)模型.RWalk 模型主要有 3 个关键的组成部分:1)基于 KL-散度的条件似然正则化 $p_{\boldsymbol{\theta}}(y|x)$,这是经典 EWC 模型[19]的改进版本,也称其为 EWC++;2)基于 2 个概率分布的 KL 散度大小实现对参数的重要性打分;3)记忆模块,即从以前的任务中存储一些有代表性的样本策略.前 2 个组成部分缓解了模型灾难性遗忘的问题,而第 3 个部分对模型不妥协问题,即模型无法有效地对新任务学习的知识进行更新处理.

首先,关于当前任务学习的参数,要求新的条件似然函数应该与之前任务所学习的条件似然函数尽可能相近,即两者的 KL 散度尽可能小.为了实现该过程,在该模型中利用新旧任务分布的 KL 散度对新任务的条件似然分布 $p_{\boldsymbol{\theta}}(y|x)$引入正则化约束:

$$D_{\mathrm{KL}}(p_{\boldsymbol{\theta}} \parallel p_{\boldsymbol{\theta}+\Delta\boldsymbol{\theta}}) \approx \frac{1}{2}\sum_{i=1}^{P} F_{\theta_i}\Delta\theta_i^2. \tag{6}$$

因此,给定模型对第 $k-1$ 个任务学习后的参数,那么对第 k 个任务进行学习时的目标函数可以表示为

$$\arg\min_{\boldsymbol{\theta}} \widetilde{L}^k(\boldsymbol{\theta}) := L^k(\boldsymbol{\theta}) + \lambda D_{\mathrm{KL}}(p_{\boldsymbol{\theta}^{k-1}}(y|x) \parallel p_{\boldsymbol{\theta}}(y|x)), \tag{7}$$

其中,λ 为超参数,$\boldsymbol{\theta}^{k-1}$为对前 $k-1$ 个任务进行学习后的模型参数.至此这些思路与 EWC 思想是相同的,但是 RWalk 模型考虑到 EWC 需要离线计算 FIM,随着任务的逐渐增加,计算复杂度呈线性增加,因此,RWalk 模型提出一种动态更新策略.该方法只需要保存上一任务所计算的 FIM,减少内存消耗.当前任务的 FIM 为 $F_{\boldsymbol{\theta}}^m$,计算为

$$F_{\boldsymbol{\theta}}^m = \alpha F_{\boldsymbol{\theta}}^m + (1-\alpha)F_{\boldsymbol{\theta}}^{m-1}, \tag{8}$$

其中,m 表示训练迭代次数,$\alpha \in [0,1]$是一个超参数.

此外,一个值得注意的问题是,FIM 是通过对 L^k 求梯度,然后对 $\widetilde{L}^k$ 求局部最优计算得到的.这将导致当 $\widetilde{L}^k \approx L^k$ 时,FIM 几乎为 $\mathbf{0}$.因此,为了解决该问题,在该模型中使用一个正标量来对 FIM 的每个元素进行增强.

因此,该模型利用对权重的重要性评分来实现对 FIM 的增强.该评分可以被定义为参数空间损失函数的改变率到每步的条件似然分布的距离,具体而言,对于参数从 $\theta_i(m)\sim\theta_i(m+1)$的改变,把参数的重要性打分定义为损失的改变率对散度 $D_{\mathrm{KL}}(p_{\boldsymbol{\theta}(m)} \parallel p_{\boldsymbol{\theta}(m+1)})$的影响.直观而言,如果分布上一个小的改变可以对应于一个更优的损失改变,则说明该参数是更重要的.因此,该过程的权重重要性打分可以表述为

$$s_{m_1}^{m_2}(\theta_i) = \sum_{m=m_1}^{m_2} \frac{\Delta L_m^{m+\Delta m}(\theta_i)}{\frac{1}{2}F_{\theta_i}^m \Delta\theta_i\ (m)^2 + \varepsilon}, \tag{9}$$

其中,$\Delta\theta_i(m)=\theta_i(m+\Delta m)-\theta_i(m)$,$\varepsilon>0$.

最后,考虑到模型的测试通常是在目前所学习的整个任务上进行测试,而当下的模型仅是完成第 k 个任务的训练后的模型,因此为了进一步降低模型的困惑度,文献[12]的作者选择性地保存所有任务的部分代表性样本进行再训练.

RWalk 模型最终的损失函数为

$$\widetilde{L}^k(\boldsymbol{\theta})=L^k(\boldsymbol{\theta})+\lambda\sum_{i=1}^{p}(F_{\theta_i^{k-1}}+s_{m_0}^{m_{k-1}}(\theta_i))(\theta_i-\theta_i^{k-1})^2, \quad (10)$$

其中,$F_{\boldsymbol{\theta}}\in\mathbb{R}^{P\times P}$为参数$\boldsymbol{\theta}$的经验费雪矩阵,$s_{m_0}^{m_{k-1}}(\theta_i)$表示从第1个任务训练迭代$m_0$到最后的任务训练迭代$m_{k-1}$的分数积累.由于分数是随着时间累积的,正则化将变得越来越严格.为了缓解这种情况,并使任务能够进行连续学习,在每项任务训练完成后对分数进行平均:

$$s_{m_0}^{m_{k-1}}(\theta_i)=\frac{1}{2}(s_{m_0}^{m_{k-2}}(\theta_i)+s_{m_{k-2}}^{m_{k-1}}(\theta_i)), \quad (11)$$

其中,连续地平均不但使过去学习任务的影响力比最近学习的任务影响要小,而且同时保留$s_{m_0}^{m_{k-1}}(\theta_i)$和$F_{\theta_i^{k-1}}$这2项对任务的影响.

4.1.7 无记忆学习

Dhar等人[36]提出一种基于注意力机制映射的无记忆学习方法(learning without memorizing, LwM),该方法通过约束教师-学生模型之间的差异来帮助模型去增量地学习新的类别,此外,该模型对新类进行学习时不需要任何之前的信息.与之前研究方法不同的是,LwM模型考虑了教师-学生模型的梯度流信息,并利用梯度流信息生成注意力机制映射来有效地提高模型的分类准确性.在进行任务t的学习时,基于注意力机制的信息知识保存项L_{AD}可以有效防止学生模型与教师模型偏离太多.在学生模型进行学习时,为了有效利用教师模型中的"暗知识",施加蒸馏损失L_D惩罚项.LwM模型示意图如图10所示:

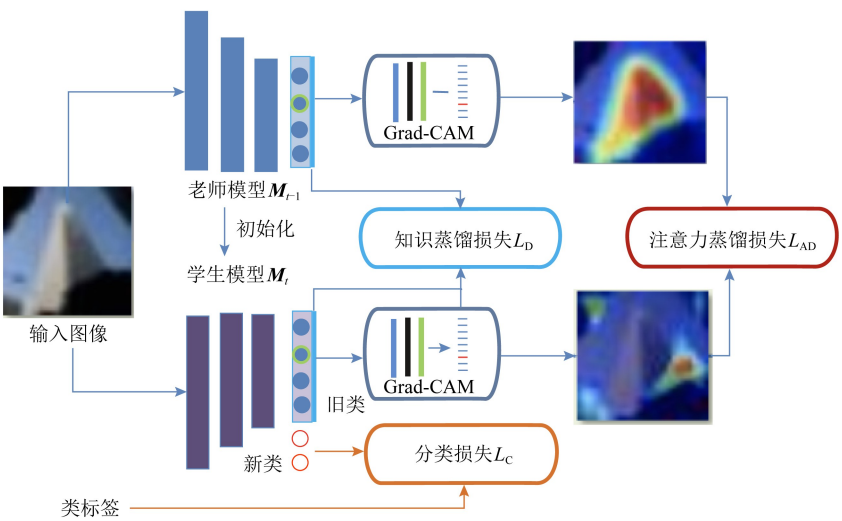

Fig. 10 Illustration of LwM

图10 LwM示意图

LwM模型的损失函数为

$$L_{LwM}=L_C+\beta L_D+\gamma L_{AD}, \quad (12)$$

其中,L_{AD}表示基于注意力机制映射的信息保存惩罚项,L_D表示蒸馏损失,L_C表示分类损失,β和γ分别表示L_D和L_{AD}的权重因子.

4.1.8 SLNID模型

Aljundi等人[37]研究了利用具有固定容量的网络进行序列学习的问题,在连续学习的背景下研究发现,相较于之前的网络参数层,在表示层施加稀疏性约束,将更有利于序列任务的学习.因此,受哺乳动物大脑侧抑制作用的启发,提出了一种新的基于正则化手段,即通过局部神经抑制和折扣的稀疏编码(sparse coding through local neural inhibition and discounting, SLNID),它通过抑制神经元来促进特征稀疏.施加该正则化的主要目的是对相同情况下的活跃神经元进行惩罚,进而产生一个更为稀疏和具有较低相关性的特征表示.同时考虑到,对于复杂任务的学习,一般在同一层需要多个活跃神经元来学习一个更强的特征表示,因此,只对局部的神经元进行惩罚.该模型通过局部神经抑制为未来的任务留出学习能力,进而有效地学习新任务,同时考虑到神经元的重要性来避免忘记以前的任务.

为了避免灾难性遗忘,基于重要性权重的方法,例如EWC或MAS方法,通过在网络中对每个参数$\boldsymbol{\theta}_k$引入重要权重$\boldsymbol{\Omega}_k$,虽然这些方法在如何估计重要参数上有所不同,但是在学习新任务T_n时,所有这些方法都使用l_2惩罚项对重要参数的变化进行惩罚,在局部神经抑制和折扣的稀疏编码中,通过增加

一个额外的正则项 R_{SSL},在每层 l 的激活中对隐特征表示施加稀疏性约束.其优化的目标函数为

$$T_n:\min_{\boldsymbol{\theta}}\frac{1}{M}\sum_{m=1}^{M}L(y_m,f(x_m,\boldsymbol{\theta}^n))+\lambda_{\boldsymbol{\Omega}}\sum_k\boldsymbol{\Omega}_k(\boldsymbol{\theta}_k^n-\boldsymbol{\theta}_k^{n-1})^2+\lambda_{\mathrm{SSL}}\sum_l R_{\mathrm{SSL}}(\boldsymbol{H}_l),\quad(13)$$

$$R_{\mathrm{SSL}}:=R_{\mathrm{SLNID}}(\boldsymbol{H}_l)=\frac{1}{M}\sum_{i,j}\mathrm{e}^{-(\alpha_i+\alpha_j)\mathrm{e}^{-\frac{(i-j)^2}{2\sigma^2}}}\sum_m\boldsymbol{h}_i^m\boldsymbol{h}_j^m,i\neq j.\quad(14)$$

其中,$\lambda_{\boldsymbol{\Omega}}$ 和 λ_{SSL}是权衡参数,当对第 1 个任务($n=1$)进行训练时,将向量 $\boldsymbol{\Omega}_k$ 初始化为零向量;对于隐含层 l,输入 $\boldsymbol{X}=(x^m)_{m=1}^M$ 的激活函数的输出为 $\boldsymbol{H}_l=(\boldsymbol{h}_i^m)_{i=1}^N$;$\sigma^2$ 是一个超参数,在此表示神经元相互影响的程度.

4.1.9 在线拉普拉斯近似

Ritter 等人[38]为了缓解灾难性遗忘,从贝叶斯理论的角度出发,提出一种 Kronecker 因子在线拉普拉斯近似(online Laplace approximation, Online-LA)方法.该方法是基于贝叶斯在线学习框架,在该框架中使用高斯函数递归逼近每个任务的后验函数,从而产生有关权重变化的二次惩罚项.拉普拉斯近似要求计算每个模式周围的海森矩阵,然而该种计算方式通常计算成本较高.因此,为了使该方法具有良好的伸缩性,引入块对角 Kronecker 因子逼近曲率,将该复杂的计算问题进行了转化.神经网络模型最大后验估计 MAP 形式为

$$\boldsymbol{\theta}^*=\arg\max_{\boldsymbol{\theta}}\log p(\boldsymbol{\theta}|D)=\arg\max_{\boldsymbol{\theta}}\log p(D|\boldsymbol{\theta})+\log p(\boldsymbol{\theta}),\quad(15)$$

其中,$p(D|\boldsymbol{\theta})$是数据的似然函数,$p(\boldsymbol{\theta})$代表先验信息.MAP 求解问题可以用损失函数加正则化项目标函数得到.例如,假设参数为零均值高斯先验的 MAP 问题,对应于交叉熵损失函数加模型参数 l_2 范数正则化项,使用标准的基于梯度的优化器可以很容易地找到该目标函数的局部最优形式.在某一模态附近,利用二阶泰勒展开式对后验函数进行局部逼近,得到以 MAP 参数为均值、负对数后验函数的 Hessian 为精度的正态分布,MacKay[39]在神经网络中使用拉普拉斯近似技术.因此,在 Online-LA 算法中,使用 2 个迭代步骤与贝叶斯在线学习类似,对于用高斯函数递归逼近每个任务的后验函数,进而可求得相应的均值和精度矩阵.

4.1.10 分离变分推理

变分推理(variational inference, VI)已成为许多现代概率模型拟合的常用方法,因此,Bui 等人[40]对此进行研究,提出一种分离变分推理算法(partitioned variational inference, PVI),文献[40]中的实验结果证明,该方法也可以很好地应用在连续学习的场景中,在该场景下新数据以非独立同分布的方式到达,任务可能随着时间发生变化,并且可能出现全新的任务.在这种情况下,PVI 框架既可以利用局部自由能不动点更新方法(local free-energy fix point update)来更新后验分布 $q(\boldsymbol{\theta})$,而且它也可以通过选择性重新访问旧数据来降低灾难性遗忘.其模型的更新步骤如图 11 所示:

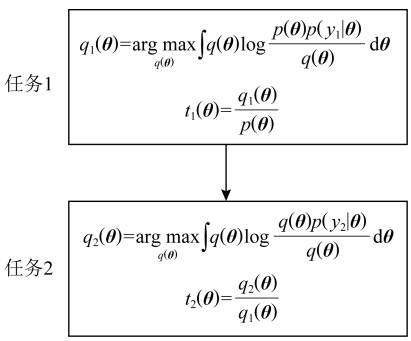

Fig. 11 PVI algorithm step

图 11 PVI 算法步骤

如图 11 所示,PVI 算法的过程主要分成 2 个步骤:1)利用局部(负)变分自由能不动点更新方法对近似后验函数进行更新;2)通过除法得到新的近似似然函数.PVI 算法在每个迭代步骤 i 中,选择一个近似的似然函数进行重新更新提炼;其中,似然近似值 $t_{b_i}^{(i-1)}(\boldsymbol{\theta})$是在上一次迭代中进行学习的,相应的数据组表示为 y_{b_i}.

4.1.11 分析比较

对于 A-LTM 模型在没有外部监督的情况下,通过知识蒸馏和回放机制,在接触了数百万个新例子之后,仍然能够保持之前对象的识别能力.然而,A-LTM 模型仅使用了一个小的数据集,例如,PASCAL,进行旧任务的训练,而使用较大数据集进行新任务的学习,例如 ImageNet,这将降低模型的准确性;SI 是在 EWC 基础上,进行在线计算权重重要性的方法,计算 F_k 所需的全部数据在 SGD 期间是可用,不需要额外的计算,有效地降低了计算成本;AR1 模型是对基于结构和正则化 2 种策略相结合,实验结果表明,将产生更低的遗忘;Online-EWC 模型是通过知识库和活动列 2 部分来完成对连续任务的学习,这个由进步学习和整合学习组成的循环结构,使得模型不需要框架的增长,也不需要访问和

存储以前的任务或数据,也不需要特定的任务参数来完成对新任务的学习,此外,由于 Online-EWC 模型使用了 2 个固定大小的列,所以可以扩展到大量任务.实验验证可得,该模型在最小化遗忘的同时实现知识的正向迁移,并且也可以直接应用到强化学习任务;R-EWC 通过对参数空间的因式旋转,更好地降低遗忘,然而,该方法为了实现对神经网络参数空间进行旋转,需要增加 2 个额外的卷积层,这将直接导致网络容量的增加;RWalk 相较于之前的基准模型具有更高的准确性,并且对于模型的遗忘和不妥协上有较好的权衡,此外,在训练过程中,RWalk 的空间复杂度是 $O(P)$,与任务的数量无关;LwM 模型对新类进行学习时不需要任何之前的信息,降低内存空间;Online-LA 从贝叶斯角度出发来降低遗忘,此外,模型也具有一定的伸缩性.

总之,正则化方法提供了一种在特定条件下减轻灾难性遗忘的方法.然而,该方法包含了保护巩固知识的额外损失项,这些损失项在资源有限的情况下,可能导致对旧新任务性能的权衡问题.

4.2 动态结构

基于动态结构的连续学习方法是通过动态地对网络结构进行调整以适应不断变化的环境,该训练方法可以选择性地训练网络,并在必要时扩展网络以适应新任务的学习.例如,使用更多的神经元或网络层进行再训练,从而有效提取新任务信息.以下为针对近年来常见动态结构的连续学习方法所进行的概括总结.

4.2.1 重新初始化复制权重

Lomonaco 等人[34]在 2017 年提出一种使用重新初始化复制权重(copy weights with re-init, CWR)的连续学习方法,该方法可以作为一种基准技术来实现对连续任务的识别.

为了不干扰对不同任务间权重的学习,CWR 方法为输出分类层设定了 2 组权重:$\boldsymbol{\theta}_{\mathrm{cw}}$是用于进行长期记忆的稳定权重,$\boldsymbol{\theta}_{\mathrm{tw}}$是对当前任务进行快速学习的临时权重.其中,$\boldsymbol{\theta}_{\mathrm{cw}}$在第 1 个任务进行训练前初始化为 $\mathbf{0}$;而 $\boldsymbol{\theta}_{\mathrm{tw}}$在每个任务训练前进行随机重新初始化,例如高斯分布抽样初始化.在多任务连续学习场景下,由于不同任务间存在一定差异,所以在每个任务训练结束时,$\boldsymbol{\theta}_{\mathrm{tw}}$中对应于当前任务的权重将会复制到 $\boldsymbol{\theta}_{\mathrm{cw}}$中.换句话说,$\boldsymbol{\theta}_{\mathrm{cw}}$可以被看作是一种进行长期记忆学习的机制,而 $\boldsymbol{\theta}_{\mathrm{tw}}$则是一种短期工作记忆机制,用来学习新任务知识而不遗忘之前所学习的任务知识.

此外,为了避免对神经网络较浅层连接边的权值矩阵和偏置向量改变过于频繁,在第 1 个任务训练完成之后,所有神经网络浅层级的权重将会被冻结.

4.2.2 CWR+方法

Maltoni 等人[33]2019 年在 CWR 方法的基础上进行改进,提出一种 CWR+的方法,该方法主要在 CWR 基础上引入了均值偏移(mean-shift)和零初始化(zero initialization)技术.均值偏移是对每批权重 $\boldsymbol{w}_i$ 进行自动补偿,即用在每个任务中学习到的权重减去在所有任务上的全局平均值实现归一化,这样将不再需要对网络权重进行重新归一化,实验发现,相较于其他形式的归一化,该方法可以取得较好的实验效果.此外,CWR+还引入了零初始化过程,即用 0 对权重进行初始化替代原来典型的高斯分布抽样初始化或 Xavier 初始化.实验结果证明,在连续学习的情况下,引入这些精细化的归一化和初始化方法,即使是像零初始化这样简单方法,也能在一定程度上提高实验效果.

4.2.3 渐进式网络

Rusu 等人[6]考虑通过分配具有固定容量的新子网络来防止对已学习知识的遗忘,这种通过分配具有固定容量的新子网来扩展模型的结构,称为渐进式网络方法(progressive networks, PN),该方法保留了一个预先训练的模型,也就是说,该模型为每个学习任务 t 都对应一个子模型.给定现有的 T 个任务时,当面对新的任务 $t+1$ 时,模型将直接创建一个新的神经网络并与学习的现有任务的模型进行横向连接.为避免模型灾难性的遗忘,当对新的任务 $t+1$ 的参数 $\boldsymbol{\theta}^{t+1}$ 进行学习时,将保持已经存在的任务 t 的参数 $\boldsymbol{\theta}^{t}$ 不变.

实验表明,在各种各样的强化学习任务上都取得了良好的效果,优于常见的基准方法.直观地说,这种方法可以防止灾难性的遗忘,但是会导致体系结构的复杂性随着学习任务的数量增加而线性增加.

4.2.4 动态扩展网络

Yoon 等人[29]在 2018 年提出了一种新的面向终身连续学习任务的深度网络模型,称为动态可扩展网络(dynamically expandable network, DEN),它可以在对一系列任务进行训练的同时动态地确定其网络容量,从而学习任务之间共享的压缩重叠知识.连续学习最主要的特征是,在对当前的任务 t 进行训练时,前 $t-1$ 个任务上所有的训练样例是不可用的,因此,在对任务 t 进行学习时,模型参数 $\boldsymbol{w}^{t}$ 的求解将转化为最优化问题:

$$\min_{\boldsymbol{w}^t} L(\boldsymbol{w}^t;\boldsymbol{w}^{t-1},D_t)+\lambda R(\boldsymbol{w}^t),t=1,2,\cdots,T, \quad (16)$$

其中,$D_t=\{x_i,y_i\}_{i=1}^{N_t}$表示训练集,$L$ 表示当前任务的损失函数,$\boldsymbol{w}^t$ 是任务 t 的参数,λ 是正则化参数,$R(\boldsymbol{w}^t)$是正则化项,例如 l_2 范数.对于目标网络,$\boldsymbol{w}^t=(\boldsymbol{w}_l)_{l=1}^{L}$是也称为权重矩阵.

对目标函数的求解过程,首先,DEN 模型通过选择性再训练,以在线的方式对训练样例进行高效训练;新的任务到达时,当已学的特征不能准确地表示新任务时,网络模型将进行动态扩展,换句话说,模型将引进额外的必要神经元来对新的任务特征进行表示.相较于之前的网络扩展模型,该模型能够动态地对网络容量进行扩展,进而使整个网络拥有恰当合适的神经元数量,完成对不同任务的学习.

4.2.5 面向任务的硬注意力机制

通常情况下,任务的定义或者任务描述对网络学习是至关重要的.如果对于 2 个任务训练数据是相同的,那么一个重要的不同就是任务的描述.例如,2 个同样都是猫和狗的训练数据集,第 1 个任务是区分猫和狗,第 2 个任务是区分毛的颜色.

因此,考虑学习使用任务鉴别器来对每个神经层进行约束,并且之后会利用这些所学知识去避免忘记过去的任务.进而,在不影响当前学习任务的前提下,保留先前的任务信息.基于此,Serra 等人[41]在 2019 年提出一种面向任务的硬注意力机制(hard attention to the task, HAT)模型,在该模型中,使用一个逐层注意力机制来设置当前的任务,如图 12 所示.所有的 $l=1,2,\cdots,L-1$ 网络层形式相同,都为一般的全连接结构,最后一层的处理与之不同.当给定 l 层的输出单元 $\boldsymbol{h}_l$,按顺序与 $\boldsymbol{a}_l^t$ 进行同或运算,即 $\boldsymbol{h}_l'=\boldsymbol{a}_l^t\odot\boldsymbol{h}_l$.该方法与普通的注意力机制的一个重要区别是代替原来形成概率分布的形式,$\boldsymbol{a}_l^t$ 是单层任务嵌入 $\boldsymbol{e}_l^t$ 的门控形式,形式为

$$\boldsymbol{a}_l^t=\sigma(s\boldsymbol{e}_l^t). \quad (17)$$

其中,$\sigma(x)\in[0,1]$是一个 Sigmoid 门函数,s 是正比例参数;除了最后一层,其他层 $l=1,2,\cdots,L-1$ 的学习过程是相同的;在第 l 层,$\boldsymbol{a}_l^t$ 是$\{0,1\}$的二进制编码形式[42],该层的学习过程相当于一个多头输出,以此来防止重要权重的更新,进而来防止灾难性遗忘[43].通过式(17)的门控机制可以形成二元注意机制掩码(binary attention masks)[44],以此可以选择性地激活或禁用每层神经元的输出.

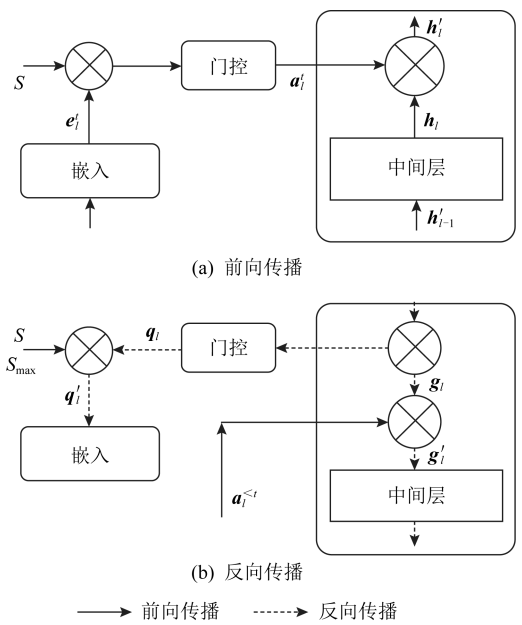

Fig. 12 Illustration of the forward-back propagation for HAT

图 12 HAT 模型前向-反向传播示意图

在网络训练过程中,模型根据之前所有任务学习的硬注意力机制来约束梯度更新.因此,为了获得直到当前任务的注意力机制向量,在学习了任务 t 并且获得 $\boldsymbol{a}_l^t$ 之后,递归地计算 $\boldsymbol{a}_l^{\leqslant t}=\max(\boldsymbol{a}_l^t,\boldsymbol{a}_l^{\leqslant t-1})$.为了对第 $t+1$ 任务训练,同时不遗忘之前的任务知识,在此使用当前层和之前层的累积注意力的最小值来对第 l 层参数进行更新:

$$\boldsymbol{g}'_{l,ij}=[1-\min(a_{l,i}^t,a_{l-1,j}^{\leqslant t})]\boldsymbol{g}_{l,ij}, \quad (18)$$

其中,下标 i,j 分别表示第 l 层的输入和第 $l-1$ 层的输出.通过式(18)创建的注意力机制模型,进而来避免对之前任务的重要参数的更新.这种方法在某种程度上与 PathNet 方法[45]类似,都是在不同层之间动态地创建路径或损毁路径达到不遗忘之前任务的知识,然而该方法的独特之处在于,HAT 不是基于模块而是基于单个神经元.因此并不需要事先分配一个模块大小或者为每个任务设置最大模块容量.

4.2.6 连续的结构学习框架模型

尽管在连续学习过程中,不同的任务具有一定的相关性,然而,对于所有任务共享一个网络结构,往往不是最优的.Li 等人[27]在 2019 年提出一个连续的学习框架模型(a continual learning framework, ACLF),该模型主要是由网络结构的优化和参数优化 2 部分组成,通过这 2 个部分能够显式地分离特定任务模型结构和模型参数的学习.该模型的损失函数为

$$L(\boldsymbol{\theta})=\frac{1}{T}\sum_{t=1}^{T}f(s(\boldsymbol{\theta});D_t)+\beta R^{\text{share}}(s)+\lambda R^{\text{split}}(\boldsymbol{\theta}), \quad (19)$$

其中,$s(\boldsymbol{\theta})$表示任务 t 的网络结构,式(19)中等号右边的第 1 项表示单个任务的损失;$\beta>0$ 和 $\lambda\geqslant 0$ 是正则化因子;R^{share} 和 R^{split} 分别表示任务共享网络结构参数的正则化项和特定分离模型参数的正则化项.在训练过程中,首先使用一个网络搜索框架为每个连续任务找到当前的最优结构,从而进行当前任务的学习,当模型的结构确定以后,使用基于梯度的方法完成对模型的参数学习.实验结果发现,相比于其他相同规模的网络框架模型,该模型将显著地降低灾难性遗忘问题,但是算法复杂性很高.

4.2.7 分析比较

生物学习机制既不需要存储流数据,也不需要以累积的方式学习知识,然而,生物却能有效地处理增量学习任务,其中不断学习和巩固新的知识,只有无用的知识被遗忘.CWR 方法的提出实现了对连续学习对象的识别,该方法作为一种基准方法为后续的研究开辟了道路.然而,CWR 和 CWR+方法的一个不足是:在每一个任务训练完后,为了避免对所学知识的遗忘,部分权重将被冻结,因此无法实现知识的反向传播,在一定程度上限制模型对新知识的学习能力;直观地说,渐进网络框架方法可以防止灾难性的遗忘,但是会导致体系结构的复杂性随着学习任务的数量增加而线性增加;DEN 通过显式地挖掘任务间的关联性,针对旧任务训练网络进行部分再训练,同时在需要时增加神经元个数以提高对新任务的解释能力,有效防止语义漂移;HAT 方法与 PathNet[45] 类似,当学习新任务时,通过动态地创建和删除跨层路径来保存新学的知识.然而,与 PathNet 不同,HAT 中的路径不是基于模块的,而是在单个神经元的,因此,不需要预先分配模块的大小,也不需要设置每个任务的最大神经元数量.当给定一个网络框架后,HAT 就可以学习并自动对单个神经元路径进行选择,进而影响单层的权重;为避免随着学习任务的数量增加模型结构线性增加问题,ACLF 方法使用一个网络搜索框架为每个连续任务找到当前的最优结构,从而进行当前任务的学习,当模型的结构确定以后,使用基于梯度的方法完成对模型的参数学习,在相同结构容量的情况下,模型将显著降低遗忘问题.然而,基于动态结构的方法,随着任务数量的不断增加,其模型结构也将不断变大,因此,无法应用到大规模数据,这也将是该模型应用于实际的重要限制.

4.3 记忆回放以及互补学习系统

在生物学上,互补学习系统(complementary learning systems, CLS)[46] 主要包括海马体和新皮质系统 2 部分,其中,海马体表现出短期的适应性,并允许快速学习新知识,而这些新知识又会随着时间的推移被放回到新皮质系统,以保持长期记忆.更具体地说,海马体学习过程的主要特点是能够进行快速学习,同时最小化知识间的干扰.相反,新大脑皮层的特点是学习速度慢,并建立了学习知识间的压缩重叠表示.因此,海马体和新皮质系统功能相互作用对于完成环境规律和情景记忆的学习至关重要.

如图 13 所示,CLS 包括用于快速学习情景信息的海马体和用于缓慢学习结构化知识的新皮质 2 部分,即海马体通常与近期记忆的即时回忆有关,例如短期记忆系统,新皮层通常与保存和回忆遥远的记忆有关,例如长期记忆.CLS 理论为记忆巩固和检索建模计算框架提供了重要的研究基础.

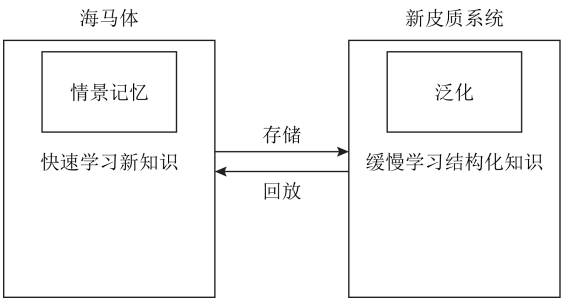

Fig. 13 CLS theory

图 13 CLS 理论

受该理论的启发,基于双记忆系统的神经网络模型的出现在一定程度上能够有效缓解连续学习过程中的遗忘问题,因此,受此生物学习系统的启发,基于情景记忆和生成模型等一系列连续学习模型相继提出,下文将对该类模型进行详细阐述.

4.3.1 BIIL 模型

Gepperth 等人[4] 受生物学习过程启发,在 2015 年提出了一种新的仿生增量学习框架模型(a bio-inspired incremental learning architecture, BIIL),当学习过程中数据具有非常高的维数(>1 000)时,仍然能有效地保持资源利用效率,同时在该模型中还增加一个短期记忆(STM)系统来提高模型性能,使其能够在连续任务学习的场景下,保持良好的分类准确性.具体而言,该模型研究了如何在不进行再训练的情况下将一个新的任务添加到一个经过训练的体系结构中,同时缓解众所周知的与此类场景相关的遗忘效应问题.该结构的核心是通过一种自组织的方法来对任务空间描述,进而在 2 维平面上近似估计该任务空间中的邻里关系.通过这种近似

方法,即使在非常高维的情况下,也允许通过有效的局部更新规则来进行增量学习.此外,增加的短期记忆系统还可以通过在特定的"睡眠"阶段对先前存储的样本进行回放来防止遗忘.该模型的结构图如图 14 所示:

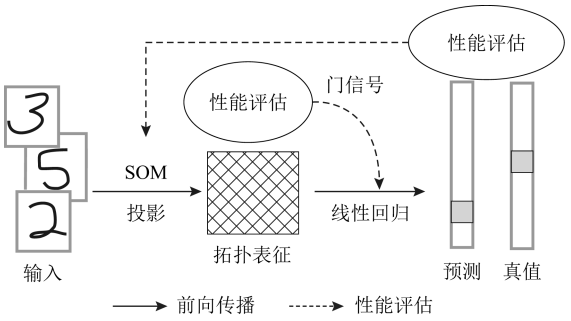

Fig. 14 Illustration of BIIL

图 14 BIIL 模型示意图

如图 14 所示,在该模型中使用了一个 3 层的神经网络结构完成对连续任务的学习.其中,使用改进的自组织映射(self-organizing map, SOM)算法来训练网络隐层的拓扑组织原型,通过线性回归完成从隐层到输出层的决策和学习;此外,该结构中引入调制机制来控制和限制隐层和输出层的学习.

4.3.2 增加的双记忆学习结构(GDM)

Parisi 等人[7]在 2018 年提出了一种适用于连续学习场景的双记忆自组织体系结构,将该方法称为增量的双记忆学习方法(growing dual-memory learning, GDM),该模型结构主要包括一个深度卷积特征提取模块和 2 个分层排列的递归自组织网络,模型原理示意图如图 15 所示:

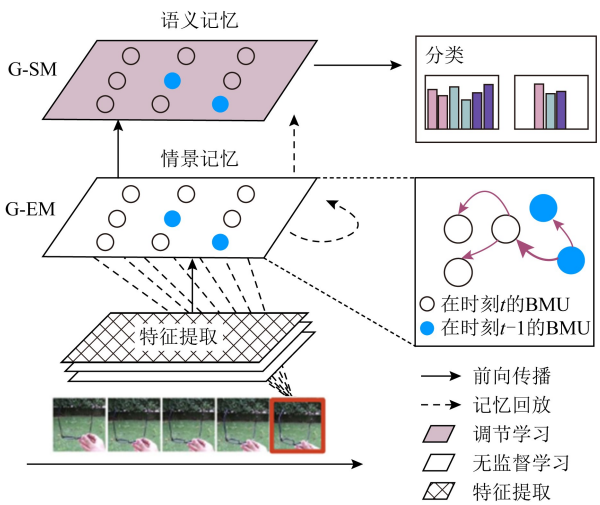

Fig. 15 Illustration of GDM

图 15 GDM 模型示意图

如图 15 所示,2 个递归网络是对 Gamma-GWR(Gamma grow-when-required)模型[47]基础上的扩展,该网络可以对模型按任务顺序输入动态地创建新的神经元和连接.不断增加的情景记忆(growing episodic memory, G-EM)以无监督的方式从任务中学习而来,其网络结构也将根据网络预测输入的能力来进行相应调节.相反,不断增加的语义记忆模块(growing semantic memory, G-SM)接收来自 G-EM 的神经激活信号,并使用与该任务相关的信号来调节神经元并进行神经元的更新,因此,该模型通过情景嵌入的方式形成一种更为压缩紧凑的知识统计表示.同时,情景记忆也将周期性地进行记忆回放,实现在没有外部输入情况下进行知识的巩固,防止对之前任务所学知识的遗忘.

4.3.3 LGM 模型

Ramapuram 等人[25]在 2017 年提出一种终生学习的生成模型(lifelong generative modeling, LGM),在该模型中通过一个学生-教师变分自编码器(student-teacher variational autoencoder, STVA)[48],不断地将新学习到的分布合并到所学的模型中,而不需要保留过去的数据或者过去的模型结构,实现模型对连续任务分布的学习.

同时,受贝叶斯更新规则的启发,在该模型中引入一种新的跨模型正则化(cross-model regularizer)方法,使得学生模型可以有效地利用教师模型的信息,此外,正则化器的使用还可以减少对分布序列学习过程中的灾难性遗忘或干扰.LGM 模型是一个基于学生-教师模型的双重体系结构.其中,教师的角色是保存以前所学知识的分布记忆,并将这些知识传递给学生;学生的角色是有效利用从老师那里获得的知识,进而有效地学习新输入数据的分布.因此,基于学生-教师模型的双重体系结构通过对教师模型和学生模型的联合优化训练,完成在学习新知识的同时不遗忘之前的知识.

4.3.4 CCL-GM 模型

Lavda 等人[49]在 2018 年提出一种基于生成模型的连续分类学习(continual classification learning using generative models, CCL-GM)方法,该方法是在 LGM 模型的基础上给目标函数增加额外的 KL-离差项,来保存之前所有任务的后验表示,以便加快模型的训练,加快来自关于教师模型中的隐表示和生成数据的负信息增益正则化项的收敛性.

4.3.5 平均梯度情景记忆

为了减轻经典 GEM 模型的计算负担,Chaudhry 等人[50]在 2018 年提出了平均梯度情景记忆模型

(averaged gradient episodic memory, A-GEM). GEM 模型的主要特征是确保在每个训练步骤中,每一个先前任务的损失不会增加,而在 A-GEM 模型中,为了降低计算复杂性,试图确保在每个训练步骤中,相对于先前任务的平均记忆损失不会增加,有效降低计算成本.在学习任务 t 时,A-GEM 的目标函数为

$$\min_{\boldsymbol{\theta}} l(f_{\boldsymbol{\theta}}, D_t)$$
$$\text{s.t. } l(f_{\boldsymbol{\theta}}, M) \leqslant l(f_{\boldsymbol{\theta}}^{t-1}, M), \text{where } M = \bigcup_{k<t} M_k. \tag{20}$$

式(20)优化问题可转化为

$$\min_{\tilde{\boldsymbol{g}}} \frac{1}{2} \| \boldsymbol{g} - \tilde{\boldsymbol{g}} \|_2^2,$$
$$\text{s.t. } \tilde{\boldsymbol{g}}^{\mathrm{T}} \boldsymbol{g}_{\mathrm{ref}} \geqslant \boldsymbol{0}, \tag{21}$$

其中,$\boldsymbol{g}_{\mathrm{ref}}$表示之前所有记忆任务参数的梯度,从情景记忆中随机抽取一批样本计算平均梯度.换句话说,A-GEM 用一个约束来替代 GEM 中的 $t-1$ 个约束,$\boldsymbol{g}_{\mathrm{ref}}$表示从情景记忆的随机子集计算出前一个任务梯度的平均值.因此,式(21)的约束优化问题可以更快地求解,更新规则为

$$\tilde{\boldsymbol{g}} = \boldsymbol{g} - \frac{\boldsymbol{g}^{\mathrm{T}} \boldsymbol{g}_{\mathrm{ref}}}{\boldsymbol{g}_{\mathrm{ref}}^{\mathrm{T}} \boldsymbol{g}_{\mathrm{ref}}} \boldsymbol{g}_{\mathrm{ref}}. \tag{22}$$

4.3.6 情景记忆回放

情景记忆回放(episodic memory replay, EMR)是在随机梯度下降算法的基础上进行改进的一种算法[51],它在对新任务进行训练时,从记忆中随机采样数据进行回放,进而将之前任务的知识保留在模型中.完成每个任务 k 训练之后,EMR 选择几个训练样本存储在记忆 M 中,即 $M \bigcap T_{\mathrm{train}}^{(k)}$.

为了解决伸缩性问题,EMR 随机回放记忆.具体而言,当对带有小批次 $D_{\mathrm{train}}^{(k)} \subset T_{\mathrm{train}}^{(k)}$ 的任务 k 进行训练时,EMR 从内存 M 中提取样本,来形成第 2 个小批次 $D_{\mathrm{replay}}^{(k)} \subset M$,然后对 $D_{\mathrm{train}}^{(k)}$ 和 $D_{\mathrm{replay}}^{(k)}$ 这 2 个小批次进行梯度更新.值的注意的是,EMR 可以对任何随机梯度优化算法进行模型优化,例如 SGD, AdaDelta, Adagrad 等.实验验证表明,这种记忆回放方式相较于之前的固定时间间隔记忆回放将显著降低遗忘问题.

针对固定时间间隔记忆回放问题,文献[51]的作者对 $D_{\mathrm{replay}}^{(k)}$ 采样方法提出了 2 种变种:1)基于任务级的抽样(task-level sampling)策略,该方法每次从前面的任务 j 中抽取样本,此时,$D_{\mathrm{replay}}^{(k)} \subset M \bigcap T_{\mathrm{train}}^{(k)}$;2)样本级(sample-level)采样策略,该方法对所有记忆进行采样,此时 $D_{\mathrm{replay}}^{(k)} \subset M$.这 2 种方法在任务样例抽样概率分布上存在差异.任务级方法在任务的样例概率分布上是服从均匀分布的,而样本级方法采样得到的是任务边缘分布,与 M 中训练数据的个数成正比.当任务满足平稳概率分布时,例如 MNIST 和 CIFAR 数据集,或者当不同任务在记忆中存储的数据满足平稳概率分布时,这 2 种方法是等效的.然而,由于样本级采样策略代码实现过程相对复杂,因此,EMR 算法中选用任务级采样策略.

4.3.7 嵌入对齐的 EMR

在嵌入对齐的情景记忆回放方法中(embedding alignment-episodic memory replay, EA-EMR),对于每一个任务 k,除了需要在记忆 M 中存储原来的训练样本$(x^{(k)}, y^{(k)})$之外,还需要存储它的嵌入表示信息.模型在对一个新的任务进行训练之后,模型参数将发生改变,因此,对于相同输入$(x^{(k)}, y^{(k)})$,嵌入表示包含的信息也将不同.直观地说,连续学习算法应该允许这样的参数变化,但要确保这些变化不会过多改变之前任务所学习的嵌入空间.

EA-EMR 算法的提出是为了防止在嵌入空间上发生的过大失真,EA-EMR 的想法为:如果在不同步骤中,嵌入空间并没有太大失真,那么应该存在一个足够简单的变换 a,例如线性变换,可以将新学习的嵌入空间变换为原始嵌入空间,而不会对之前任务存储的嵌入空间造成太大变化.因此,建议在原始嵌入的基础上增加一个变换 a,并自动学习基本模型 f 和嵌入空间的变换 a.具体而言,在第 k 个任务中,首先学习模型 $f^{(k-1)}$ 和变换 $a^{(k-1)}$,$f^{(k-1)}$ 和 $a^{(k-1)}$ 是由之前的 $k-1$ 个任务训练而来.进而,学习基本模型 f 和变换 a,以此来优化模型处理新任务和存储样例的性能,而不会对前面的嵌入空间造成太大的影响.在关系检测模型中加入嵌入对齐的方式如图 16 所示.

图 16 显示了如何在一个基本的关系检测模型上添加对齐模型的过程,在本例中为线性模型.其中,使用 2 个 BiLSTMs 模块[52]来对文本和关系进行编码,最后计算其嵌入对齐之间的余弦相似性进行打分.

最终,完成对模型的学习过程.通过最小化如式(23)所示的目标函数:

$$\min_{f(\cdot), a(\cdot)} \sum_{(x,y) \in D_{\mathrm{train}}^{(k)}} l(a(f(x)), y) + \sum_{(x,y) \in D_{\mathrm{replay}}^{(k)}} (l(a(f(x)), y) + \| a(f(x)) - a^{(k-1)} f^{(k-1)}(x) \|^2), \tag{23}$$

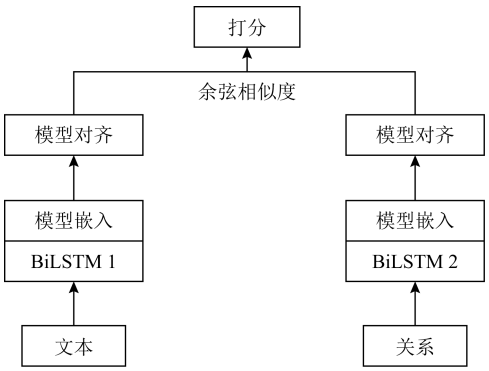

Fig. 16 Add the alignment model to the basic relationship detection model

图 16 基本关系检测模型上添加对齐模型

式(23)主要由 2 部分组成,前半部分是优化基本模型 f,在该步骤主要学习新任务,且不会对存储的样例造成性能下降.后半部分是优化变换 a,保持当前任务的嵌入空间,恢复之前存储的样本时的嵌入空间.

4.3.8 元经验回放

Riemer 等人[52]尝试通过梯度对齐来权衡连续问题中知识的迁移(transform)和干扰(interference)问题,因此提出一种元经验回放方法(meta-experience replay, MER).该方法与之前的连续学习方法最主要的一个不同是,在该模型中不仅考虑当前知识对之前知识的迁移,而且考虑到当前知识动态地前向迁移过程.该算法将经验回放与基于优化的元学习方法[53]相结合,使得该方法保持当前任务学习的参数对未来学习知识的干扰降到最小,而基于未来梯度的知识对当前任务知识的迁移更有可能发生,充分考虑了在连续任务学习场景中的迁移-干扰的平衡问题.

对于连续学习问题中迁移-干扰的平衡,即考虑在时间上的正向和逆向的权重共享和稳定性-可塑性平衡.在 MER 中,通过利用一个经验回放模块增强在线学习,实现了对到目前为止看到的所有样例的平稳分布的近似优化.同时,对于损失梯度计算困难的问题,使用元学习算法间接地将目标近似为一阶泰勒展开来解决这个问题.在线学习算法与元学习算法的结合,有效地实现知识的前向迁移.

4.3.9 小情景记忆回放

Chaudhry 等人[54]在 MER 模型的基础上进行研究,提出一种新的记忆回放方法,称其为小情景记忆回放(MER-Tiny),相较于之前的在特定时间进行记忆回放,联合训练当前任务中的样例和存储在记忆模块中的样例将获得更优的性能.此外,实验验证表明,对小情景记忆的重复学习并不会降低模型对过去任务的泛化能力.对于记忆内存的写入方法,实验验证发现,水库抽样(reservoir sampling)可以取得较优的效果,但是该方法往往需要较大的内存开销.然而,在内存非常小的情况下,牺牲随机性保证所有类平衡,即为每个任务存储特定个数的记忆样例.因此,新的小记忆回放方法可以实现对两者的权衡,提高模型性能.

与最简单的基准模型相比,MER-Tiny 模型主要有 2 个修改:1)它有一个小情景记忆,且每一步都会更新;2)通过将当前任务中的实际小批次记忆与从内存中随机抽取的小批次记忆叠加起来,以实现梯度下降的参数更新.实验结果表明:这 2 个简单的修改将使模型具有了更好的泛化性能,并在很大程度上降低了灾难性遗忘问题.

4.3.10 端到端增量学习

传统的神经网络体系结构要用到整个数据集,即之前类和新类的所有样本来更新模型,然而随着类的数量不断增加,该模型将无法连续学习.Castro 等人[55]在此研究基础上,提出一种增量的深度神经网络学习方法,称为端到端增量学习(end-to-end incremental learning),即只使用新任务数据和旧任务样例对应的小样本集来解决该问题.

端到端增量学习方法使用交叉熵和蒸馏损失来训练深度网络,使用蒸馏损失来保留从旧类中获得的知识,使用交叉熵作为损失函数来完成对新类的学习,由于该方法具有较好的通用性,所以网络的选取可以是基于任何为分类而设计的深层模型结构.增量式训练的整个框架是通过端到端的方式实现的,也就是,联合学习数据表示和分类器.其典型的带有分类层和分类损失的框架如图 17 所示.

在训练阶段,通过交叉熵蒸馏损失函数的对数计算梯度,更新网络的权值.交叉蒸馏损失函数的定义为

$$L(\boldsymbol{w})=-\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C}p_{ij}\log q_{ij}+\sum_{f=1}^{F}\left(-\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C}p_{ij}^{*}\log q_{ij}^{*}\right),\tag{24}$$

其中,式(24)的等号右侧第 1 项代表交叉熵损失函数,等号右侧第 2 项为在分类层 f 的蒸馏损失函数,F 表示总的分类层个数;q_{ij} 表示对样本 i 的分类层使用软最大函数得到的打分;p_{ij} 表示样本 i 的真值;N 和 C 分别表示样本个数和类数;p_{ij}^{*} 和 q_{ij}^{*} 分别是 p_i 和 q_i 的蒸馏修正形式[17].此外,该过程中还

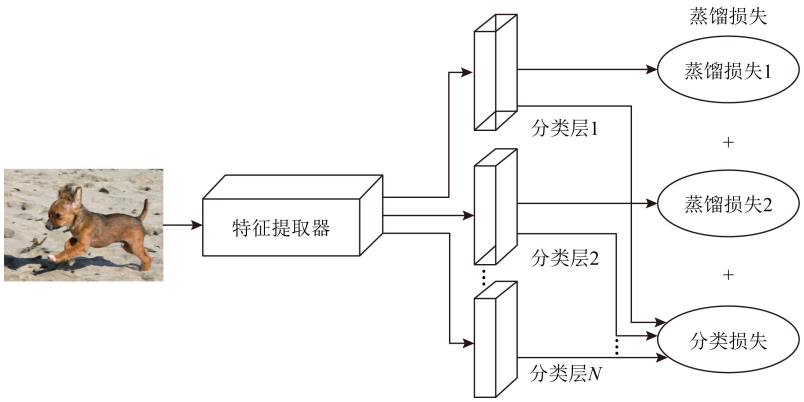

Fig. 17 An end-to-end learning framework with classification layers and classification losses

图 17 带有分类层和分类损失的端到端学习框架

使用了一个代表性记忆内存来存储和管理来自旧类的最具代表性的样本[56],进而帮助模型保留从旧类中获得的知识.

4.3.11 分析比较

受互补学习系统的启发,Gepperth 等人[4]在 2015 年提出 BIIL 模型,使用一个具有双记忆系统的网络结构来避免灾难性干扰,该框架虽然对灾难性干扰问题有所缓解,但是对系统如何实现快速学习的问题较少考虑;GDM 模型中,使用了 G-EM 和 G-SM 这 2 个模块,通过具有结构可塑性的自组织结构进行动态学习,对新输入的信息增加存储信息,缓解连续学习过程中的灾难性遗忘的问题;LGM 模型通过优化一个增广的 ELBO 目标函数很好地帮助完成师生学习知识的迁移,正则化的使用有效缓解了灾难性干扰这一问题.然而,该模型学习过程中的一个弊端是,该方法不能访问任何旧数据,并且必须将所有必要的信息提取到一个单独的最终模型中才能学习;CCL-GM 是在 LGM 基础上增加额外约束来保存之前所学的知识,降低遗忘可能性;A-GEM 模型相较于 GEM 模型,利用确保在每个训练步骤中,相对于先前任务的平均情景记忆损失不会增加,替代原来的每个单独的先前任务的损失不会增加,这使得 A-GEM 不仅内存效率高,此外,因为它不需要存储矩阵 $\boldsymbol{G}$,速度比 GEM 也快几个数量级;EMR 算法相较于 GEM 算法的一个明显不同是在进行无约束优化时,不需要进行梯度映射,例如求解 $\tilde{\boldsymbol{g}}$ 问题.EMR 算法的时间复杂度相当于每个前一个任务存储的样例数,而 GEM 方法需要计算存储在记忆中的所有数据的梯度,因此,随着任务数量的增加,该计算过程呈线性增长,GEM 将变得不再适用.此外,EMR 算法相较于 GEM 算法更为简单.文献[51]实验结果表明,通过显式地进行嵌入对齐,可以有效地通过减小先前嵌入空间的失真,提高模型的性能.MER 对于过去和将来的数据,能够在知识的迁移和干扰方面有较好的平衡,同时,MER 对未来的数据有较强的适应性.此外,MER 在各种监督下的连续学习和连续强化学习环境中显示出巨大的潜力;RtF 模型[11]通过在主模型上增加一个生成反馈连接,能够有效地加快计算效率,但是也会在一定程度上降低模型的性能.然而,该类模型需要存储之前观察到的例子,因此需要大量的记忆空间;文献[55]在 CIFAR-100 数据集的实验结果证明,end-to-end 方法由于依赖于一个次优的、独立学习的外部分类器,因此相较于一些其他最近的学习方法,例 iCarl,更具鲁棒性.

4.4 总 结

目前为缓解连续学习过程中的灾难性遗忘问题,主要集中在引入正则化策略、动态结构策略和基于情景记忆策略这 3 个方向进行研究.正则化方法在模型更新时,通过对权重进行约束,实现在保持已有知识的前提下,完成对新任务的学习,从而缓解灾难性遗忘这一问题,此外,这类方法通常不需要保存任何以前的数据,只需要对每个任务进行一次训练.然而,该类方法克服灾难性遗忘的能力是有限的,例如在类增量学习(class-incremental learning, Class-IL)场景下性能不佳,此外,随着任务数目的不断增加,对过去任务进行正则化处理,可能导致特征漂移.动态地改变模型结构以便在不干扰之前任务的学习知识的情况下学习新的任务,该类方法也可以成功地缓解灾难性遗忘这一问题,然而,该类方法不能从任务之间的正向迁移中获益,另外模型的大小随着观察到的任务数量的增加而急剧增长,这使得

它在实际问题中往往不可行.基于情景记忆的方法,通过保存一些以前任务的样例进行记忆回放来缓解对之前所学习知识的遗忘,该类方法在减轻灾难性遗忘方面显示出了巨大优势,然而,计算成本却随着先前任务的数量增加而快速增长,并且该方法需要保存之前样例,不利于数据安全保护.在基于情景记忆的方法中,为替代存储所学任务的样例数据,提出使用深层生成模型来记忆以前见过的数据分布,然而该类方法往往需要从头开始重新训练生成模型,训练效率低,此外,在每次生成以前任务的新的真实样本时,还极易造成"语义漂移",且随着时间推移,模型训练准确性逐渐下降.

5 实验数据集与评价准则

本节将对近年来连续学习实验分析过程中常用的实验数据集以及公认的评价准则进行详细介绍.

5.1 实验数据集介绍

表 2 和表 3 对连续学习过程中常用的分类数据集以及其主要特征进行总结.MNIST 数据集[57]是对 0～9 这 10 个数字进行手写样本的数据集,其中每个样本的输入是一个图像,标签是图像所代表的数字.为了在该数据集上进行连续学习问题的评估,提出 3 种用于连续学习场景下的 MNIST 数据集:1)排列的 MNIST 数据集[19],该数据集是参考某个固定的排列,通过重新排列像素来创建任务,即通过 K 个不同的排列来生成 K 个不同的任务;2)旋转的 MNIST 数据集,其中每个任务都是通过对数字旋转固定的角度创建的,即选择 K 个角度来创建 K 个任务;3)分离的 MNIST 数据集,将原始的 MNIST 数据集分成 5 个训练任务得到分离手写字体数据集.此外,其他常见的连续学习数据集包括:Fashion-MNIST 数据集由相同大小的灰度图像组成[58];Traffic Signs 数据集包含交通标志图像,其中使用来自 Udacity 自动驾驶汽车 github 存储库的数据集[59];Bulatov 等人[60]从公共可用字体中提取出的字形而创建的与 MNIST 类似的 Not MNIST 数据集;Netzer 等人[61]在谷歌街景图像中截取的房号创建了 SVHN 数据集;CIFAR10 数据集和 CIFAR100 数据集[62]是由 32×32 像素的彩色图像组成.

Table 2 Introduction of Distributions for Seven Classified Datasets

表 2 7 种分类数据集属性介绍

数据集	提出年份	类个数	训练集数	测试集数
MNIST[57]	1998	10	60 000	10 000
Fashion-MNIST[58]	2017	10	60 000	10 000
TRAFFICSIGNS[59]	2011	43	39 209	12 630
NotMNIST[60]	2011	10	16 853	1 873
SVHN[61]	2011	100	73 257	26 032
CIFAR10[62]	2009	10	50 000	10 000
CIFAR100[62]	2009	100	50 000	10 000

Table 3 Introduction of Distributions for Six Object Recognition Datasets

表 3 6 种对象识别数据集属性介绍

数据集	提出年份	类个数	识别对象数	总图像数
iCubWorld-T[63]	2016	20	10	约 200 000
CORe50[64]	2017	10	50	164 866
EMMI[65]	2017	12	30	约 2 300 000
iLab-20M[66]	2016	15	20～160	21 798 480
RGB-D[67]	2011	51	3～14	25 000
3D Object[68]	2007	8	10	约 7 000

iCubWorld 变换数据集(iCubWorld transformation, iCub-T)[63]和 CORe50 数据集是连续学习对象识别实验中最复杂,也是较为常用的 2 个数据集.这 2 个数据集是专为连续学习图像而设计,是从某个人作为移动对象的一系列帧中生成的一系列图像,例如,CORe50 数据集包括在不同的条件下同一对象的多个视图(不同的背景、对象的姿态和遮挡程度)的 10 个类别内的 50 个对象.数据集收集了 11 个具有不同背景和亮度的图像,其中,对于在每个场景下的每个对象使用 Kinect 2.0 传感器[64]录制一个 15 s 的视频(20 Hz).最终数据集是包含 164 866 张 128×128 RGB-D 的 11 个场景 50 个对象的图像.因此,这 2 个数据集是评估连续学习的理想数据集,因为当学习算法识别该对象时,该流数据不是 IID 形式,因此,很好地满足连续学习过程的要求.Wang 等人[65]提出一个以自我为中心、手工的以及多图像的数据集(egocentric, manual, multi-image, EMMI),EMMI 中的图像来自可穿戴式摄像机记录的常见家用物品和玩具被手动操作以进行结构化转换,如旋转和平移等,该数据集收集的目的是,视觉体验的外观相关和分布特性如何影响学习的结果等.表 3 对常见的 6 个用于对象识别的数据集的主要特征进行总结[66-68].

5.2 评价准则

连续学习算法可以从一系列连续的流数据中不断地学习,进而实现对模型增量式更新.对于连续学

习算法的性能可以从多方面进行评估,目前大多集中于模型学习知识的准确性和对之前所学知识的遗忘程度2方面[41].Lopez-Paz等人[20]认为连续学习问题常涉及知识的正向以及反向迁移能力,因此,需要对模型的知识迁移性能进行评估;Díaz-Rodríguez等人[69]考虑到连续学习算法往往还涉及模型框架的大小、内存记忆的占用以及计算效率等问题,因此提出一系列更为全面的评价指标,从多个方面对连续学习算法性能进行评估.以下从模型学习的准确性、知识的遗忘、反向迁移、正向迁移、模型规模度量和计算效率这6个方面对近年来模型的学习性能评估进行总结.

5.2.1 准确性(accuracy)

文献[54]提出一种称为平均准确性的估计方法 $a_{ij}\in[0,1]$ 表示模型在任务 i 上训练完之后,模型在任务 j 的测试集上的分类准确性能.对于任务 T 的平均准确性 f_T^{A} 定义为

$$f_T^{\mathrm{A}}=\frac{1}{T}\sum_{j=1}^{T}a_{T,j}. \tag{25}$$

Díaz-Rodríguez等人[69]给定训练-测试样本精度矩阵 $\boldsymbol{R}\in\mathbb{R}^{T\times T}$,其中包含每个条目 $R_{i,j}$ 通过观察任务 i 的最后一个样本得到的模型在任务 j 上的测试分类精度[20].模型的准确性是通过考虑矩阵 $\boldsymbol{R}$ 的对角元素,对实现对训练集 D_i 和测试集 D_j 的平均精度进行考虑.准确性 f^{A} 为

$$f^{\mathrm{A}}=\sum_{i>j}^{T}R_{i,j}\Big/\frac{T(T+1)}{2}. \tag{26}$$

文献[54]最初定义该准则是为了在最后一个任务结束时评估模型的性能而定义的,而在文献[69]中,该准确性准则应该考虑到模型在每一时间点(every timestep)的性能的准确性指标,这样能够更好地考虑连续学习模型的动态性能.

5.2.2 遗忘

Joan等人[41]引入遗忘率来获得对模型遗忘量的测量.首先,对任务进行权衡并统一随机化它们的顺序,在训练任务 t 之后计算所有的测试任务集 $\tau\leqslant t$ 的精度.

因此,对于分类问题,当模型已经被增量训练至任务 $k(j\leqslant k)$ 之后,定义对于第 j 个任务的遗忘模型的量化形式为

$$\widetilde{f}_j^k=\max_{l\in\{1,2,\cdots,k-1\}}a_{l,j}-a_{k,j},\ \forall j<k, \tag{27}$$

其中,我们感兴趣的是量化对先前任务的遗忘,所以对于任务 $j\leqslant k$ 定义了 $\widetilde{f}_j^k\in[-1,1]$.因此,通过对先前看到的任务数量进行归一化,将第 t 次任务的平均遗忘记为 $\widetilde{f}_t$,其中,较低的 $\widetilde{f}_t$ 意味着较少的遗忘,具体表示形式为

$$\widetilde{f}_t=\frac{1}{t-1}\sum_{j=1}^{t}a_{t,j}. \tag{28}$$

5.2.3 反向迁移

反向迁移能力(backward transfer, BWT)是衡量模型学习了一个新的任务后对先前任务的影响.当需要在多任务或流数据背景下进行学习时,往往就需要模型对其反向迁移性能的评估.模型对之前任务学习能力的提高和不降低的性能对连续学习是至关重要,因此,在其学习的整个过程中都应该被评估.f_{BWT} 定义在学习了 i 之后,在同一测试集的最后一个任务结束时,对任务 $j(j<i)$ 计算准确度.在此,与准确性的度量准则类似,将其扩展到对每个任务的后向迁移求平均值:

$$f_{\mathrm{BWT}}=\sum_{i=2}^{T}\sum_{j=1}^{i-1}(R_{i,j}-R_{j,j})\Big/\frac{T(T-1)}{2}. \tag{29}$$

因为 f_{BWT} 最初的取值规则是为后向迁移取正值,为灾难性遗忘取负值,因此,为了将 f_{BWT} 映射到区间[0,1],同时更好地区分这2个不同语义的概念.

5.2.4 正向迁移

知识正向迁移(forward transfer, FWT)是衡量学习任务对未来任务的影响.根据之前Lopez-Paz等人[20]对准确性的度量准则,Díaz-Rodríguez等人[69]进一步修改为训练-测试准确度量,其中 $R_{i,j}$ 的平均准确性高于准确率矩阵 $\boldsymbol{R}$ 的主对角线.因此定义 f_{FWT} 为

$$f_{\mathrm{FWT}}=\sum_{i<j}^{T}R_{i,j}\Big/\frac{T(T-1)}{2}. \tag{30}$$

5.2.5 模型规模度量

根据每个任务 i 的参数 $\boldsymbol{\theta}$ 的数量来量化每个模型 h_i 的存储器的大小,记为 $f_{\mathrm{Mem}}(\theta_i)$,相对于第1个任务内存大小 $f_{\mathrm{Mem}}(\theta_1)$,随着时间推移,模型对任务不断地学习,模型规模大小不应该增长过快.

因此,模型的规模(model size, MS)f_{MS} 定义为

$$f_{\mathrm{MS}}=\min\left(1,\frac{1}{T}\sum_{i=1}^{T}\frac{f_{\mathrm{Mem}}(\theta_1)}{f_{\mathrm{Mem}}(\theta_i)}\right). \tag{31}$$

5.2.6 计算效率

由于模型的计算效率(computational efficiency, CE)受训练集 D_t 的乘法和加法运算总数的限制,因此,文献[41]定义任务之间的平均计算效率 f_{CE} 为

$$f_{\mathrm{CE}}=\min\left(1,\frac{1}{T}\sum_{t=1}^{T}\frac{Ops\uparrow\downarrow(D_t)\times\varepsilon}{Ops(D_t)}\right), \tag{32}$$

其中,$Ops(D_t)$ 是指学习 D_t 所需要的操作数;

$Ops\uparrow\downarrow(D_t)$是指在 D_t 进行一次知识的正向和反向传播所需要的运算次数;ε 的默认值是大于 1,该因子的使用使得 f_{CE}的计算更有意义,例如,避免了趋近于 0 的情况.

6 连续学习的应用

作为机器学习领域中的一个极具潜力的研究方向,连续学习方法已经受到学者的极大青睐.随着人工智能及机器学习不断的发展,基于连续学习的方法已经获得了较多应用,例如图像分类、目标识别以及自然语言处理等.以下将对近年来连续学习在各领域的主要应用进行介绍.

6.1 图像分类

Li 等人[16]在 2017 年提出了一种由卷积神经网络组成的无遗忘学习方法,该方法将知识蒸馏与细调方法相结合,利用知识蒸馏的方法来加强与当前学习任务相关的已经学习过的知识,提高分类的准确性;Kim 等人[70]提出基于 DOS 的最大熵正则化增量学习模型(maximum entropy regularization and dropout sample for incremental learning, MEDIL),该模型通过最大熵正则化来减少对不确定迁移知识的优化,以及利用 DOS 来通过从新任务中选择性地删除样例减少对旧类的遗忘,以此减少记忆样例中类的不平衡,有效地完成连续学习过程中的图像分类;Smith 等人[71]在 2019 年提出一种新颖的自学习联想记忆框架(self-taught associative memory, STAM),有效解决在连续学习过程中的无监督学习分类问题;Aljundi 等人[37]提出一种基于稀疏编码的正则化方法,实现利用具有固定容量的网络进行有序学习问题,在 CIFAR100 和 MNIST 数据集上进行分类的结果表明,该模型能够有效地提高模型的分类能力;Rostami 等人[72]考虑到基于自编码器的生成模型能够很好地对输入样例进行编码,获得较好的隐特征表示,同时受并行分布式处理学习和互补学习系统理论的启发,提出一种新颖的计算模型,该模型能够将新学习的概念与之前模型学习的概念经过统一编码,进而形成一个统一的嵌入空间表示,实现了利用之前学习的概念知识来有效地帮助只有少量标签样例的新领域知识的学习,从而完成在连续学习背景下的样例分类.

6.2 目标识别

Siam 等人[73]提出一种新颖的教师-学生自适应框架,在无需人工标注的情况下,完成人机交互(human-computer interaction, HCI)背景下的视频目标对象分割(video object segmentation);Parisi 等人[7]提出了一种适用于终身学习场景的双记忆自组织体系结构,该模型结构主要包括一个深度卷积特征提取模块和 2 个分层排列的递归自组织网络,进而实现在终身学习场景下的视频序列中的目标对象的识别;Tessler 等人[74]提出一种新颖的分层深度强化学习网络(hierarchical deep reinforcement learning network, H-DRLN)框架,该模型在 Minecraft 游戏场景中,通过重用在之前任务中学习到的知识,进而完成对未来任务场景的目标对象学习,提高效率,同时,该模型的实验结果也展示了在不需要额外学习的情况下在相关 Minecraft 任务之间迁移知识的潜力;Michiel 等人[10]将当前的基于任务标识已知的序列学习方法推向了在线无任务标识的连续学习场景中,首先,假设有一个无限输入的数据流,其中该数据流中包含现实场景中常见的逐渐或者突然的变化.文献[10]中提出一种基于重要权重正则化的连续学习方法,与传统的任务标识已知场景中不同,在该场景中,该模型需要有效地检测何时、如何以及在哪些数据上执行重要性权重更新,进而有效地在无任务标识场景下进行在线连续学习.该文中在监督学习和自监督学习过程中都成功地验证了该方法的有效性.其中,具体而言,相较于基准学习方法,在电视剧人脸识别和机器人碰撞等具体应用中,该方法的稳定性和学习性能都有所提高.Tahir 等人[75]考虑到当下最先进的有关食物识别的深度学习模型不能实现数据的增量学习,经常在增量学习场景中出现灾难性遗忘问题.因此,提出一种新的自适应简化类增量核极值学习机方法(adaptive reduced class incremental kernel extreme learning machine, ARCIKELM),进而完成目标食物对象的识别,其中在多个标准的食物数据集的最终分类准确性证明了该模型可以有效地进行增量学习.

6.3 自然语言处理

d'Autume 等人[76]介绍了一种连续学习背景下的自然语言学习模型,该模型实现了对在线文本数据的有效学习.在文献[76]中介绍了一种基于稀疏经验回放的方法有效地防止灾难性遗忘,具体而言,对于每 10000 个新的样本随机均匀选择 100 个样本在固定的时间间隔进行稀疏经验回放,实验表明,该模型在文本分类和问答系统等自然语言领域可以实现较好的应用.Li 等人[77]考虑到现有的方法大多集中在对输入和输出大小固定的标签预测连续学习任务

上,因此,提出了一个新的连续学习场景,它处理自然语言学习中常见的序列到序列的学习任务.实验结果表明,该方法比现有方法有明显的改进,它能有效地促进知识正向迁移,防止灾难性遗忘.Kruszewski等人[78]提出一种基于多语言和多领域背景下的语言建模基准,该基准可以将任何明确的训练样例划分为不同的任务.与此同时,提出一种基于产品专家(product of experts, PoE)的多语言连续学习方法,Kruszewski等人的实验结果证明,在进行多语言连续学习时,该模型可以有效地缓解灾难性遗忘.Hu等人[79]对个性化在线语言学习问题(personalized online language learning, POLL)进行研究,涉及到适应个性化的语言模型以适应随着时间发展的用户群体.为了有效地对POLL问题进行研究,文献[79]的作者收集了大量的微博帖子作为训练数据集,进而对近年来流行的连续学习算法进行了严格评估,并在此基础上提出一种简单的连续梯度下降算法(continual gradient descent, ConGraD),实验结果表明,该算法在Firehose数据集和早期基准测试数据集的实验结果优于之前的连续学习方法.

7 未来的研究方向

作为机器学习领域中的一个新兴方向,连续学习近几年受到研究者们的极大关注,目前来看,连续学习在未来的研究中有10个潜在的方向:

1) 基于经验回放(experience replay)的模型相较于其他连续学习模型有较好的性能,然而,容量的饱和是该类模型中所面临的重要挑战,因此如何在保持原有知识的同时,不断提高模型的能力是未来重要的研究方向.

2) 对于任务不可知场景下的连续学习算法尚需进一步研究.目前,大多连续学习算法要求在任务边界(task boundaries)已知的场景中来进行训练和预测,即当需要学习一个新的任务时,模型需要被告知有新的学习任务,例如,改变损失函数中的参数等,以便系统能够采取某些行动.然而,在任务之间没有明显边界,即任务的转变是逐渐的或者连续的,这些模型将不再适用.然而,在实际应用中,往往需面对的是任务边界不可知场景学习问题.文献[9]从贝叶斯的角度提出一种贝叶斯梯度下降算法(Bayes gradient desent, BGD),对没有明确定义的任务边界的连续学习问题提供一种解决思路,然而,基于此场景的连续学习算法仍相对缺乏,尚需进一步研究.

3) 利用多模态信息.现有的连续学习方法通常使用来自单一模态(如图像或文本)的知识进行建模,然而,虽然当下训练集有一些当前模态的样例,但是,样例可能还存在另一个模态.因此,来自多模态的知识可以为连续学习提供较为丰富的样例信息,进而提高模型的建模能力.因此如何有效地利用这些多模态信息也是未来研究的重要方向.

4) 在未来可以对当下连续学习模型应用的灵活性进行进一步扩展研究,例如多感知领域的扩展.文献[80]可以从视听流中不断学习任务的特征,使得连续学习的方法向更加广泛的应用迈进一步.因此,可以通过将连续学习方法部署在具体的代理中,通过与环境的主动交互,在持续的时间内可以增量地获取和提取知识,以此来更好地完成对对象的识别等任务.

5) 数据集太小也是连续学习过程所面临的挑战之一.例如,目前存在的iCub-T和CORe50数据集,只包含几十个常见的家庭对象类,缺乏大规模和多样性数据集.因此,创建一个更大的和更多样化的数据集,即可以包括数百个或数千个类,也可以包括不同类型的识别,如人脸、场景以及活动等,对未来的研究工作是至关重要的.

6) 在实际分类问题中,数据的不平衡时常发生,易于导致数据的错误分类,因此如何从不平衡的数据集中进行正确分类,也是未来连续学习研究的一个重要方向.

7) 在线学习.当前的连续学习方法多集中于对每个单独的任务进行离线训练,然而,在实际应用中数据往往以数据流的形式存在[81].因此,如何对连续的数据流进行学习是未来的一个重要的研究方向.

8) 正向迁移.在连续学习方法中,正向迁移即知识的正向迁移能力,也就是对新任务进行学习时,如何有效地利用之前所学习的知识来有效地加快对当前任务的学习.近年来,元学习方法的出现,为进一步提高知识的正向迁移提供了前景.因此,如何有效地利用元学习技术来尽可能地加快对当前任务的学习是未来的一个重要的研究方向.

9) 权衡模型的稳定性与可塑性.模型的可塑性,即模型对学习新知识的能力.模型的稳定性,即模型对已经学习知识的保留能力.在连续学习过程中,如何有效地对模型的稳定性和可塑性进行权衡是一个值得研究的问题.

10) 应用领域扩展.大多实际应用场景都涉及连续学习的问题,计算机视觉中图像分类是连续学习

最常用的实验平台之一.连续学习最近在许多其他应用中也引起了广泛关注,如机器人技术、自然语言处理和视频信号处理.总之,连续学习还有很多值得探索的领域和应用.

8 总 结

连续学习是近年来机器学习领域的一个重要的研究方向.连续学习是模拟大脑学习的过程,按照一定的顺序对连续的非独立同分布的流数据进行增量学习.连续学习的意义在于高效地转化和利用已经学过的知识来完成新任务的学习,并且能够极大程度地降低遗忘带来的问题.本文系统地对近年来提出的连续学习方法进行综述,首先详细阐述了连续学习的定义、学习场景以及其相关领域,然后详细指出了各模型提出的原因以及具有的优缺点、常用的实验数据集、评价指标以及近年来的应用,最后对未来的研究方向及其巨大的应用潜力进行了细致说明.总之,随着对连续学习研究的不断深入,未来势必将发挥越来越重要的作用.

作者贡献声明:韩亚楠负责文献调研、内容设计、论文撰写和论文校对;刘建伟负责提出论文的整体研究和分析思路、指导写作、修改论文以及最终审核;罗雄麟参与论文校对.

参 考 文 献

[1] Hassabis D, Kumaran D, Summerfield C, et al. Neuroscience-inspired artificial intelligence [J]. Neuron Review, 2017, 95(2): 245-258

[2] French R M. Catastrophic forgetting in connectionist networks [J]. Trends in Cognitive Sciences, 1999, 3(4): 128-135

[3] Robins A V. Catastrophic forgetting, rehearsal and pseudo rehearsal [J]. Connection Science, 1995, 7(2): 123-146

[4] Gepperth A, Karaoguz C. A bio-inspired incremental learning architecture for applied perceptual problems [J]. Cognitive Computation, 2015, 8(5): 924-934

[5] Rebuffi S A, Kolesnikov A, Sperl G, et al. iCarl: Incremental classifier and representation learning [C] //Proc of the 2017 IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 2001-2010

[6] Rusu A A, Rabinowitz N C, Desjardins G. Progressive neural networks [J]. arXiv preprint, arXiv: 1606.04671, 2016

[7] Parisi G I, Tani J, Weber C, et al. Lifelong learning of spatiotemporal representations with dual-memory recurrent self-organization [J]. Frontiers in Neurorobotics, 2018, 12(1): 78-86

[8] Richardson F M, Thomas M S C, Critical periods and catastrophic interference effects in the development of self-organising feature maps [J]. Developmental Science, 2008, 11(3): 371-389

[9] Aljundi R, Kelchtermans K, Tuytelaars T. Task-free continual learning [C] //Proc of the IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2019: 11254-11263

[10] Van D, Tolias A S. Generative replay with feedback connections as a general strategy for continual learning [J]. arXiv preprint, arXiv:1809.10635, 2018

[11] Hsu Y C, Liu Y C, Ramasamy A, et al. Re-evaluating continual learning scenarios: A categorization and case for strong baselines [J]. arXiv preprint, arXiv: 1810.12488, 2018

[12] Chaudhry A, Dokania P K, Ajanthan T, et al. Riemannian walk for incremental learning: Understanding forgetting and intransigence [C] //Proc of the 15th European Conf on Computer Vision. Berlin: Springer, 2018: 556-572

[13] Ostapenko O, Puscas M, Klein T. Learning to remember: A synaptic plasticity driven framework for continual learning [C] //Proc of the 15th IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2019: 11321-11329

[14] Caruana R. Multitask learning [J]. Machine Learning, 1997, 28(1): 41-75

[15] Parisi G, Kemker R, Part J L, et al. Continual lifelong learning with neural networks: A review [J]. Neural Networks, 2019, 113(1): 54-71

[16] Li Zhizhong, Hoiem D. Learning without forgetting [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 40(12): 2935-2947

[17] Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network [J]. Computer Science, 2015, 14(7): 38-39

[18] Girshick R J, Donahue T, Darrell J. Rich feature hierarchies for accurate object detection and semantic segmentation [C] //Proc of the IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2014: 580-587

[19] Kirkpatrick J, Pascanu R, Rabinowitz N, et al. Overcoming catastrophic forgetting in neural networks [J]. Proceedings of the National Academy of Sciences, 2017, 114(13): 3521-3526

[20] Lopez-Paz D, Ranzato M. Gradient episodic memory for continual learning [J]. arXiv preprint, arXiv: 1706.08840, 2017

[21] Venkatesan R, Venkateswara H, Panchanathan S, et al. A strategy for an uncompromising incremental learner [J]. arXiv preprint, arXiv:1705.00744, 2017

[22] Nguyen C V, Li Yingzhen, Bui T D, et al. Variational continual learning [C/OL] //Proc of the 6th Int Conf on Learning Representations. Amsterdam: Elsevier, 2018 [2021-05-05]. https://arxiv.org/pdf/1710.10628.pdf

[23] Delange M, Aljundi R, Masana M, et al. A continual learning survey: Defying forgetting in classification tasks [J/OL]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021 [2020-12-11]. https://arxiv.org/pdf/1909.08383.pdf

[24] Wu Yue, Chen Yinpeng, Wang Lijuan, et al. Incremental classifier learning with generative adversarial networks [J]. arXiv preprint, arXiv:1802.00853, 2018

[25] Ramapuram J, Gregorova M, Kalousis A. Lifelong generative modeling [J]. arXiv preprint, arXiv:1705.09847, 2017

[26] Yosinski J, Clune J, Bengio Y, et al. How transferable are features in deep neural networks [C] //Proc of the 27th Advances in Neural Information Processing Systems. Cambridge, MA: MIT Press, 2014: 3320-3328

[27] Li Xilai, Zhou Yingbo, Wu Tianfu. Learn to grow: A continual structure learning framework for overcoming catastrophic forgetting [C] //Proc of the 36th Int Conf on Machine Learning. New York: ACM, 2019: 3925-3934

[28] Schwarz J, Czarnecki W, Luketina J, et al. Progress & compress: A scalable framework for continual learning [C] //Proc of the 35th Int Conf on Machine Learning. New York: ACM, 2018: 4528-4537

[29] Yoon J, Yang E, Lee J, et al. Lifelong learning with dynamically expandable networks [C/OL] //Proc of the 6th Int Conf on Learning Representations. Amsterdam: Elsevier, 2018 [2020-12-11]. https://arxiv.org/pdf/1708.01547.pdf

[30] Furlanello T, Zhao Jiaping, Saxe A M, et al. Active long term memory networks [J]. arXiv preprint, arXiv: 1606.02355, 2016

[31] 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): 419-425

[32] Zenke F, Poole B, Gangguli S. Continual learning through synaptic intelligence [C] //Proc of the 34th Int Conf on Machine Learning. New York: ACM, 2017: 3987-3995

[33] Maltoni D, Lomonaco V. Continuous learning in single-incremental-task scenarios [J]. Neural Networks, 2019, 116(1): 56-73

[34] Lomonaco V, Maltoni D. CORe50: A new dataset and benchmark for continuous object recognition [C/OL] //Proc of the 1st Annual Conf on Robot Learning. 2017: 17-26 [2020-12-11]. https://arxiv.org/pdf/1705.03550.pdf

[35] Liu Xialei, Masana M, Herranz L, et al. Rotate your networks: Better weight consolidation and less catastrophic forgetting [C] //Proc of the 24th Int Conf on Pattern Recognition. Piscataway, NJ: IEEE, 2018: 2262-2268

[36] Dhar P, Singh R V, Peng, K C, et al. Learning without memorizing [C] //Proc of IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2019: 5138-5146

[37] Aljundi R, Rohrbach M, Tuytelaars T. Selfless sequential learning [J]. arXiv preprint, arXiv:1806.05421, 2018

[38] Ritter H, Botev A, Barber D. Online structured Laplace approximations for overcoming catastrophic forgetting [C] //Proc of the 31st Advances in Neural Information Processing Systems. Cambridge, MA: MIT Press, 2018 [2020-12-11]. https://arxiv.org/pdf/1805.07810.pdf

[39] MacKay D J. Apractical Bayesian framework for backpropagation networks [J]. Neural Computation, 1992, 4(3): 448-472

[40] Bui T D, Nguyen C V, Swaroop S, et al. Partitioned variational inference: A unified framework encompassing federated and continual learning [J]. arXiv preprint, arXiv: 1811.11206, 2018

[41] Serra J, Suris D, Miron M, et al. Overcoming catastrophic forgetting with hard attention to the task [C] //Proc of the 36th Int Conf on Machine Learning. New York: ACM, 2019: 4555-4564

[42] Bakker B J, Heskes T M. Task clustering and gating for Bayesian multitask learning [J]. Journal of Machine Learning Research, 2003, 4(1): 83-99

[43] Ng H W, Winkler S. A data-driven approach to cleaning large face datasets [C] //Proc of the 14th IEEE Int Conf on Image Processing. Piscataway, NJ: IEEE, 2014: 343-347

[44] McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity [J]. The Bulletin of Mathematical Biophysics, 1943, 5(4): 115-133

[45] Fernando C, Banarse D, Blundell C, et al. PathNet: Evolution channels gradient descent in super neural networks [J]. arXiv preprint, arXiv:1701.08734, 2017

[46] Masana M, Liu Xialei, Twardowski B, et al. Class-incremental learning: Survey and performance evaluation [J]. arXiv preprint, arXiv:2010.15277, 2020

[47] Parisi G I, Tani J, Weber C, et al. Lifelong learning of humans actions with deep neural network self-organization [J]. Neural Networks, 2017, 96(1): 137-149

[48] Kingma D P, Welling M. Auto-encoding variational Bayes [C/OL] //Proc of the 2nd Int Conf on Learning Representations. Amsterdam: Elsevier, 2014 [2020-12-11]. https://arxiv.org/pdf/1312.6114.pdf

[49] Lavda F, Ramapuram J, Gregorova M, et al. Continual classification learning using generative models [J]. arXiv preprint, arXiv:1810.10612, 2018

[50] Chaudhry A, Ranzato M A, Rohrbach M, et al. Efficient lifelong learning with A-GEM [J]. arXiv preprint, arXiv: 1812.00420, 2018

[51] Wang Hong, Xiong Wenhan, Yu Mo. Sentence embedding alignment for lifelong relation extraction [C] //Proc of the 2019 Conf of the North American Chapter of the Association for Computational Linguistics. Piscataway, NJ: IEEE, 2019: 796-806

[52] Riemer M, Cases I, Ajemian R, et al. Learning to learn without forgetting by maximizing transfer and minimizing interference [C/OL] //Proc of the 7th Int Conf on Learning Representations. Amsterdam: Elsevier, 2019 [2020-12-11]. https://arxiv.org/pdf/1810.11910.pdf

[53] Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks [C] //Proc of the 34th Int Conf on Machine Learning. New York: ACM, 2017: 1126-1135

[54] Chaudhry A, Rohrbach M, Elhoseiny M, et al. Continual learning with tiny episodic memories [J]. arXiv preprint, arXiv:1902.10486. 2019

[55] Castro F M, Marín-Jiménez M J, Guil N, et al. End-to-end incremental learning [C] //Proc of the 15th European Conf on Computer Vision (ECCV). Berlin: Springer, 2018: 233-248

[56] Welling M. Herding dynamical weights to learn [C] //Proc of the 26th Annual Int Conf on Machine Learning. New York: ACM, 2009: 1121-1128

[57] 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

[58] Xiao Han, Rasul K, Vollgraf R. Fashion-MNIST: A novel image dataset for benchmarking machine learning algorithms [J]. arXiv preprint, arXiv: 1708.07747, 2017

[59] Stallkamp J, Schlipsing M, Salmen J, et al. The German traffic sign recognition benchmark: A multi-class classification competition [C] //Proc of the 31st Int Joint Conf on Neural Networks. Piscataway, NJ: IEEE, 2011: 1453-1460

[60] Bulatov Y. NotMNIST dataset [DB/OL]. 2011 [2020-12-11]. http://yaroslavvb.blogspot.it/2011/09/notmnist-dataset.html

[61] Netzer Y, Wang Tao, Coates A, et al. Reading digits in natural images with unsupervised feature learning [C/OL] // Proc of NIPS Workshop on Deep Learning and Unsupervised Feature Learning. Cambridge, MA: MIT Press, 2011 [2020-12-11]. http://ufldl. stanford. edu/housenumbers/nips2011 _ housenumbers.pdf

[62] Krizhevsky A, Hinton G. Learning multiple layers of features from tiny images [J]. Handbook of Systemic Autoimmune Diseases, 2009, 1(4): 1-10

[63] Pasquale G, Ciliberto C, Rosasco L. Object identification from few examples by improving the invariance of a deep convolutional neural network [C] //Proc of the 2016 Intelligent Robots and Systems. Piscataway, NJ: IEEE, 2016: 4904-4911

[64] Steward J, Lichti D, Chow J, et al. Performance assessment and calibration of the Kinect 2.0 Time-of-Flight range camera for use in motion capture applications [C/OL] //Proc of the 2015 FIG Working Week. 2015 [2020-12-11]. https://www.fig.net/fig2021/workshops.htm

[65] Wang Xiaohan, Eliott F M, Ainooson J, et al. An object is worth six thousand pictures: The egocentric, manual, multi-image (EMMI) dataset [C] //Proc of the 16th IEEE Int Conf on Computer Vision Workshop. Piscataway, NJ: IEEE, 2017: 2364-2372

[66] Borji A, Izadi S, Itti L. iLab-20M: A large-scale controlled object dataset to investigate deep learning [C] //Proc of the 2016 IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 2221-2230

[67] Lai K, Bo L, Ren Xiaofeng, et al. A large-scale hierarchical multi-view RGB-D object dataset [C] //Proc of the 2011 Robotics and Automation. Piscataway, NJ: IEEE, 2011: 1817-1824

[68] Savarese S, Li Feifei. 3D generic object categorization, localization and pose estimation [C/OL] //Proc of the 11th Int Conf on Computer Vision. Piscataway, NJ: IEEE, 2007 [2020-12-11]. http://www. vision. caltech. edu/wikis/EE148/images/3/37/Caltech_meeting_dec_07_2.pdf

[69] 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, arXiv: 1810.13166, 2018

[70] Kim D, Bae J, Jo Y, et al. Incremental learning with maximum entropy regularization: Rethinking forgetting and intransigence [J]. arXiv preprint, arXiv:1902.00829, 2019

[71] Smith J, Baer S, Kira Z, et al. Unsupervised continual learning and self-taught associative memory hierarchies [C/OL] //Proc of the 7th Int Conf on Learning Representations. Amsterdam: Elsevier, 2019 [2020-12-11]. https://openreview.net/pdf?id=SJxakiC4u4

[72] Rostami M, Kolouri S, Pilly P, et al. Generative continual concept learning [J]. Proceedings of the AAAI Conf on Artificial Intelligence, 2020, 34(4): 5545-5552

[73] Siam M, Jiang Chen, Lu Weikai, et al. Video object segmentation using teacher-student adaptation in a human robot interaction (HRI) setting [C] //Proc of the 2019 Int Conf on Robotics and Automation. Piscataway, NJ: IEEE, 2019: 50-56

[74] Tessler C, Givony S, Zahavy T, et al. A deep hierarchical approach to lifelong learning in minecraft [C] //Proc of the 31st AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2017: 1553-1561

[75] Tahir G A, Loo C K. An open-ended continual learning for food recognition using class incremental extreme learning machines [J]. IEEE Access, 2020, 8(1): 82328-82346

[76] d'Autume C, Ruder S, Kong Lingpeng. Episodic memory in lifelong language learning [C] //Proc of Advances in Neural Information Processing Systems. Cambridge, MA: MIT Press, 2019: 13122-13131

[77] Li Yuanpeng, Zhao Liang, Church K. Compositional language continual learning [C/OL] //Proc of the 8th Int Conf on Learning Representations. Amsterdam: Elsevier, 2020 [2020-12-11]. https://openreview.net/pdf?id=rklnDgHtDS

[78] Kruszewski G, Sorodoc I T, Mikolov T. Class-agnostic continual learning of alternating languages and domains [J]. arXiv preprint, arXiv:2004.03340, 2020

[79] Hu Hexiang, Sener O, Sha F, et al. Drinking from a firehose: Continual learning with web-scale natural language [J]. arXiv preprint, arXiv:2007.09335, 2020

[80] Su Lixin, Guo Jiafeng, Zhang Ruqing, et al. Continual domain adaptation for machine reading comprehension [C] // Proc of the 29th ACM Int Conf on Information and Knowledge Management. New York: ACM, 2020: 1395-1404

[81] Li Zhijie, Li Yuanxiang, Wang Feng, et al. Online learning algorithms for big data analytics: A survey [J]. Journal of Computer Research and Development, 2015, 52(8): 1707-1721 (in Chinese)
(李志杰, 李元香, 王峰, 等. 面向大数据分析的在线学习算法综述[J]. 计算机研究与发展, 2015, 52(8): 1707-1721)

Han Yanan, born in 1991. PhD. Her main research interest is machine learning.
韩亚楠,1991 年生.博士.主要研究方向为机器学习.

Liu Jianwei, born in 1966. PhD, associate professor, PhD supervisor. His main research interests include machine learning, pattern recognition, intelligent system analysis and prediction, control of complex system, and algorithm analysis and design.
刘建伟,1966 年生.博士,副教授,博士生导师.主要研究方向为机器学习、模式识别、智能系统分析和预测、复杂系统控制以及算法分析和设计.

Luo Xionglin, born in 1963. PhD, professor. His main research interests include intelligent control analysis and prediction, controlling of complicated nonlinear system.
罗雄麟,1963 年生.博士,教授.主要研究方向为智能控制分析和预测、复杂非线性系统控制.