

# 自编码器结合持续学习：现状、挑战与展望

吴美君<sup>1,2)</sup> 杨新<sup>1,3)</sup> 潘超凡<sup>1,3)</sup> 李天瑞<sup>4)</sup> 寇纲<sup>2)</sup>

<sup>1)</sup>(西南财经大学认知计算与群智协同创新实验室 成都 611130)

<sup>2)</sup>(西南财经大学工商管理学院 成都 611130)

<sup>3)</sup>(西南财经大学计算机与人工智能学院 成都 611130)

<sup>4)</sup>(西南交通大学计算机与人工智能学院 成都 610097)

**摘要** 近年来,许多研究利用自编码器进行增量式学习,以在面对新的数据分布、类别或任务时平衡模型的稳定性与可塑性。这些研究从多个角度推动了持续学习的发展。同时,持续学习的范式通过优化策略促进了自编码器架构的改进,实现了自编码器与持续学习之间的相互促进。目前,自编码器与持续学习的结合在多个领域都影响深远。本文对近五年来的相关研究进行了综述,概述了自编码器的类型与特点,持续学习的常见增量场景与主要挑战,并对二者在不同领域的应用情况进行了详细介绍。最后,本综述对当前研究的优点、局限性以及未来应用的前景进行了总结,旨在为推动持续学习与自编码器的结合与发展提供有价值的参考。

**关键词** 持续学习;自编码器;灾难性遗忘;知识传输;模型优化

中图法分类号 TP18 DOI号 10.11897/SP.J.1016.2025.00317

## Autoencoders Combined with Continual Learning: Development, Challenges, and Prospects

WU Mei-Jun<sup>1,2)</sup> YANG Xin<sup>1,3)</sup> PAN Chao-Fan<sup>1,3)</sup> LI Tian-Rui<sup>4)</sup> KOU Gang<sup>2)</sup>

<sup>1)</sup>(Innovation Laboratory of Cognitive Computing and Crowd Intelligence, Southwestern University of Finance and Economics, Chengdu 611130)

<sup>2)</sup>(Department of Business Administration, Southwestern University of Finance and Economics, Chengdu 611130)

<sup>3)</sup>(Department of Computing and Artificial Intelligence, Southwestern University of Finance and Economics, Chengdu 611130)

<sup>4)</sup>(Department of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 610097)

**Abstract** In recent years, many studies have employed autoencoders for continual learning, aiming to balance the stability and adaptability of models when faced with new data distributions, categories, or tasks. These studies have propelled the development of continual learning from various perspectives. At the same time, the paradigm of continual learning has facilitated improvements in autoencoder architectures through optimization strategies, thereby achieving mutual enhancement between autoencoders and continual learning. The integration of autoencoders and continual learning has shown a significant impact across various research. This

收稿日期:2024-04-07;在线发布日期:2024-10-15。本课题得到国家自然科学基金青年科学基金项目(62406259)、国家自然科学基金面上项目(62476228)、国家自然科学基金国际(地区)合作研究与交流项目(71910107002)、湖湘高层次人才聚集工程项目(2024RC4008)、四川省科技厅中央引导地方科技发展项目(2024ZYD0180)资助。吴美君,博士研究生,主要研究领域为终身持续学习、智能决策、金融风险预测。E-mail: wumeijun@smail.swufe.edu.cn。杨新(通信作者),博士,教授,博士生导师,中国计算机学会(CCF)杰出会员,主要研究领域为终身持续学习、可信联邦学习、多粒度认知计算。E-mail: yangxin@swufe.edu.cn。潘超凡,博士研究生,主要研究领域为终身持续学习、强化学习。李天瑞,博士,教授,博士生导师,中国计算机学会(CCF)杰出会员,主要研究领域为数据挖掘与知识发现、人工智能、粒度计算与粗糙集。寇纲(通信作者),博士,教授,博士生导师,长江学者以及国家杰出青年科学基金入选者,主要研究领域为大数据与决策、商务智能、信息系统、信用评级、应急管理。E-mail: kougang@swufe.edu.cn。

paper reviews the relevant research over the past five years, summarizing the types and characteristics of autoencoders, the incremental scenarios, and the main challenges of continual learning. Additionally, it offers a detailed overview of their applications in different industries. Finally, this review summarizes the advantages, limitations, and prospects, aiming to provide valuable insights for advancing research on developing continual learning and autoencoders.

**Keywords** continual learning; autoencoder; catastrophic forgetting; knowledge transfer; model optimization

## 1 引 言

传统深度学习方法通常在静态数据集上进行训练,并且假设训练集和测试集的数据分布相同。然而,动态环境中的现实数据通常是非独立同分布的,这使得探索深度学习模型的可持续性和长期稳定性成为极其重要的研究课题<sup>[1]</sup>。近年来,模拟人类终身认知学习方式而发展起来的持续学习范式,为攻克这一重要课题提供了新的解决方案,并因持续性累积知识和高适应性的特点引起了学术界的广泛关注。在众多基本模型架构中,自编码器网络不仅具备结构的简洁性以及稳定性等优点,还兼具特征提取与压缩、分布识别、样本生成等功能,在众多领域为持续学习范式的发展提供了重要支持。同时,受益于持续学习范式对模型泛化能力的关注,自编码器架构的性能得以逐步提升。当前,自编码器和持续学习在众多领域相互影响、相互促进,在动态大数据环境中,不断提升模型的类人认知稳定性和可塑性能力,系统性地归纳和总结它们之间的关系、研究现状以及未来发展趋势具有深远意义。

持续学习(Continual Learning)起源于终身学习(Lifelong Learning),由于二者研究重点的不同,持续学习这一术语逐渐独立起来。终身学习概念强调模型在整个生命周期中不断积累和利用知识的能力,类似于人类的认知过程。然而,随着研究的深入,持续学习这一术语被专门指代深度学习和机器学习中的这些技术和方法。持续学习不仅关注知识的积累,还特别强调在学习新任务时保持对旧任务的记忆,避免灾难性遗忘。这种设定使得持续学习能够更好地应对现实世界中数据不断变化和不可预见的环境,提供更具鲁棒性和灵活性的解决方案。相比于增量学习(Incremental Learning),持续学习在设定上更为严格,先前阶段的数据在后续阶段不可访问或仅部分可访问,这更符合实际应用场景的

需求,同时面临的挑战也更加严峻。

持续学习的研究目前面临两个主要挑战:(1)灾难性遗忘:在学习新任务时,模型可能需要调整网络参数甚至结构以适应新任务的特征分布,这可能会对先前学到的知识或技能产生严重干扰。(2)知识传输:持续学习要求模型能够从先前任务中获得通用知识,并传递给新任务以提升学习能力。自编码器及其变体,如可提取任务分布信息的变分自编码器、可结合约束信息生成伪样本的条件变分自编码器、可捕获时间依赖性的长短期自编码器等<sup>[2]</sup>,能够有效地执行分布识别、样本生成、数据去噪等任务,为持续学习的任务信息识别、克服灾难性遗忘和知识传输提供了重要支持。同时,自编码器借助持续学习范式,能够在数据分布动态变化的环境中逐渐积累知识,持续性地优化自身结构。

目前,自编码器与持续学习范式的结合已经在计算机视觉、生物医疗、语音文本、交通等领域取得了一系列优秀的理论研究成果和成功的行业应用。本综述对自编码器与持续学习的最新相关研究进行了整理,系统总结了自编码器在应对持续学习所面临的两大挑战方面所做的贡献,并探讨了优化自编码器模型以适应持续学习范式的具体策略。通过评述自编码器与持续学习之间的影响,本综述有助于研究者更全面地理解自编码器与持续学习的结合方式和应用场景,并为二者的未来发展提供新的启示和研究方向。

本综述的主要贡献在于综述了过去五年中自编码器与持续学习相结合的研究,明确阐述了它们的优势、不足以及应用前景。本综述与已有综述文献的差异在于,它系统性地探讨了自编码器与持续学习的结合,而目前尚无专门针对这一主题的综述文献。已有的相关综述主要聚焦于持续学习的总体方法和挑战,或是对单独的自编码器技术进行探讨,但缺乏对二者结合的深入分析和系统总结。本文的创新主要体现在以下几个方面:首次系统地将自编码

器与持续学习相结合进行综述,全面概括了自编码器在持续学习中发挥的作用;其次,对自编码器和持续学习的基本类型进行了概括,并探讨了自编码器的特点以及持续学习面临的挑战,为读者提供了全面而系统的视角;再次,详细分析了自编码器在克服持续学习面临的灾难性遗忘和知识传输挑战中的具体贡献,并探讨了自编码器如何基于持续学习范式不断优化模型性能;此外,根据具体的应用领域,评述了自编码器与持续学习结合的研究现状和前沿进展,为实际应用提供了参考;最后,指出了当前研究的不足,并提出了未来可能的研究方向,为学术界和工业界的进一步探索提供指导。

本综述的其余内容安排如下:第2节介绍自编码器及其变体的基本原理,从6个维度归纳自编码器的不同用途、实现方式和特点;随后探讨持续学习面临的主要挑战以及发展出的不同增量场景;第3节评述自编码器对持续学习克服灾难性遗忘与实现知识传输的影响;第4节总结自编码器如何在持续学习范式中随着任务推进而优化模型;第5节将根据应用领域的不同,评述自编码器与持续学习结合的现状;最后一节将对全文进行总结,并展望未来的研究方向。

## 2 背景与挑战

### 2.1 自编码器的类型与特点

自编码器是一种神经网络模型,结构如图1所示。自编码器的概念最早由Rumelhart等人<sup>[3]</sup>提出,其结构包含编码器(Encoder)和解码器(Decoder)两个部分。编码器将输入数据 $x$ 映射为

潜在空间中的特征 $z$ ,解码器则将其解码还原为 $\hat{x}$ ,并使其尽可能与原始输入近似。

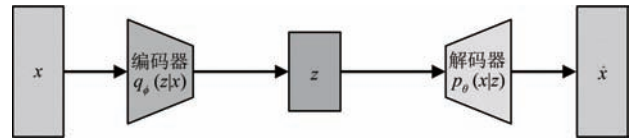


图1 自编码器

为满足不同领域和任务的需求,基础自编码器<sup>[4]</sup>已发展出多种类型,例如稀疏自编码器<sup>[5]</sup>、变分自编码器<sup>[6]</sup>及其变体<sup>[7-9]</sup>、对抗自编码器<sup>[10]</sup>、解纠缠自编码器<sup>[11]</sup>、降噪自编码器<sup>[12]</sup>、卷积自编码器<sup>[13]</sup>、堆叠自编码器<sup>[14]</sup>、序列自编码器<sup>[15]</sup>等。不同类型自编码器的关注重点、结构以及训练方式存在差异,为不同场景提供了多样化选择。

基础自编码器在自编码器体系中最简单<sup>[16]</sup>,由编码器映射函数 $f_\phi: x \rightarrow z$ 和解码器映射函数 $g_\theta: z \rightarrow x$ 组成,重构样本 $\hat{x}$ 则可由 $f_\phi(g_\theta(x))$ 表示,其中 $x$ 是输入样本, $z$ 是潜在表征。重构损失函数一般基于平方损失构建:

$$L_{AE} = \|x - \hat{x}\|^2 \quad (1)$$

或者基于交叉熵构建:

$$L_{AE} = -\sum x \log(\hat{x}) \quad (2)$$

为了提取满足特定需求的表征,在原有重构损失基础上引入正则化约束 $L_{Reg}$ 是常见办法,总损失函数可表示为

$$L = L_{AE} + L_{Reg} \quad (3)$$

稀疏自编码器在基础自编码器基础上,引入稀疏性约束作为正则化项,即通过 $L_{Reg} = L_{Sparse}$ 来限制被激活的隐藏层神经元个数,结构如图2所示。

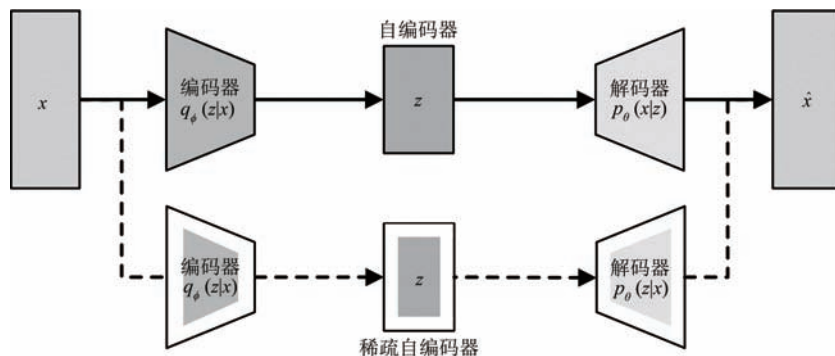


图2 稀疏自编码器

稀疏性约束可以采用L1正则化:

$$L_{Sparse} = \lambda \sum_j |h_j| \quad (4)$$

其中 $h_j$ 表示隐藏层第 $j$ 个神经元的输出, $\lambda$ 是正则化项的权重参数,控制了稀疏性的程度。

实现稀疏性约束也可以采用KL(Kullback-Leibler)散度:

$$L_{\text{Sparse}} = \beta \sum_j (\rho_0 \log \frac{\rho_0}{\hat{\rho}_j} + (1 - \rho_0) \log \frac{1 - \rho_0}{1 - \hat{\rho}_j}) \quad (5)$$

其中,  $\rho_0$  是稀疏性目标, 即所有隐藏单元的期望激活水平,  $\hat{\rho}_j$  是隐藏层神经元  $j$  的平均激活度。稀疏化激活值向量, 隐藏层的神经元在训练过程中的大部分时候处于非活跃状态, 特征的可解释性和泛化性将得到增强<sup>[5]</sup>。

变分自编码器引入了概率模型和隐变量来建模数据分布<sup>[17]</sup>, 结构如图3所示。

编码器产生一个均值向量  $\mu$  和一个方差向量  $\Sigma$ , 潜在表征的分布可表示为

$$q_\phi(z|x) = \mathcal{N}(z|\mu(x), \Sigma(x)) \quad (6)$$

解码器将潜在表征  $z$  映射回输入空间, 重构样本  $\hat{x}$  的分布可表示为

$$p_\theta(x|z) = \mathcal{N}(x|\hat{\mu}(z), \hat{\Sigma}(z)) \quad (7)$$

变分自编码器最大化对数边际似然的下界(即变分下界, ELBO), 其损失函数可表示为

$$L_{\text{VAE}}(\phi, \theta; x) = -\mathbb{E}_{z \sim q_\phi(z|x)} [\log p_\theta(x|z)] + D_{\text{KL}}(q_\phi(z|x) \| p(z)) \quad (8)$$

其中, 第一项表示从潜在空间采样得到的  $z$  经过解码器生成  $x$  的对数似然, 第二项是编码器生成的潜在表征的分布  $q_\phi(z|x)$  与先验分布  $p(z)$  之间的KL散度。通过最大化变分下界, 变分自编码器在保证生成样本质量的同时, 学习潜在表征的分布信息。为适应特定需求, 变分自编码器已发展出条件变分自编码器、韦伯变分自编码器等变体。

条件变分自编码器在变分自编码器基础上, 引入了与输入  $x$  相关的信息, 作为条件信息  $y$ , 使得所学分布在给定相关信息的情况下更具针对性。其从条件分布  $q_\phi(z|x, y)$  中采样得到  $z$ , 然后使用解码器基于分布  $p_\theta(x|z, y)$  生成与  $x$  接近的重构样本, 结构如图4所示。

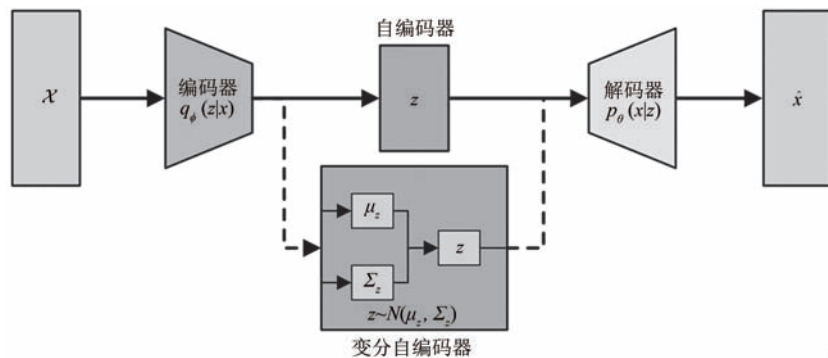


图3 变分自编码器

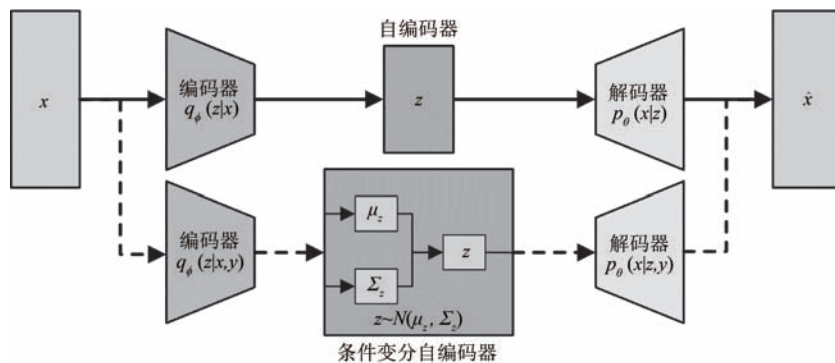


图4 条件变分自编码器

类似地, 条件变分自编码器的目标是最大化条件对数边际似然的下界(条件ELBO), 其损失函数可表示为

$$L_{\text{CVAE}}(\phi, \theta; x, y) = -\mathbb{E}_{z \sim q_\phi(z|x, y)} [\log p_\theta(x|z, y)] + D_{\text{KL}}(q_\phi(z|x, y) \| p(z|y)) \quad (9)$$

其中, 第一项表示重构误差, 即在给定  $z$  和  $y$  的情况下生成  $x$  的概率的对数。第二项是KL散度, 用于衡量近似后验分布  $q_\phi(z|x, y)$  与先验分布  $p(z|y)$  之间的差异。

韦伯变分自编码器, 是变分自编码器的另一种



变体,其通过学习数据的概率分布来重构样本。与变分自编码器通常假设潜在表征 $z$ 服从标准正态分布不同,韦伯变分自编码器将先验分布设定为韦伯分布,即在正实数范围内定义分布,该类分布的概率密度函数可表示为

$$p(z) = \frac{k}{\lambda} \left(\frac{z}{\lambda}\right)^{k-1} e^{-(z/\lambda)^k} \quad (10)$$

其中 $z$ 表示正实数范围内的潜在表征取值, $k$ 和 $\lambda$ 均为正数,是韦伯分布的形状参数和尺度参数,分别控制了分布的形状和尺度。韦伯变分自编码器将韦伯分布作为先验分布,在正实数范围内学习潜在表征的分布特性。与传统的基于正态分布先验的变分自编码器相比,韦伯变分自编码器更适用于建模正实数任务,例如提取文本数据特征<sup>[9]</sup>。

解纠缠自编码器由变分自编码器发展而来,旨在学习独立且可解释的特征表示<sup>[11]</sup>。相比于变分自编码器,解纠缠自编码器对潜在特征分布的控制更

为严格,倾向于遵循标准多变量正态分布的先验假设,整体的损失函数可表示为

$$L_{DAE}(\phi, \theta; x) = -\mathbb{E}_{z \sim q_\phi(z|x)}[\log p_\theta(x|z)] + \beta D_{KL}(q_\phi(z|x) \| p(z)) \quad (11)$$

其中,第一项表示解码潜在空间样本 $z$ 后生成 $x$ 的对数似然,第二项衡量近似后验分布 $q(z|x, y)$ 与先验分布 $p(z|y)$ 之间的差异,并通过参数 $\beta$ 控制解耦程度。通过增加 $\beta$ 值,潜在特征分布更加偏向标准的多变量正态分布,从而能降低特征之间的相关性。添加结构乘子 $\beta$ 后最大化变分下界,解纠缠自编码器能学习更独立、可解释性更强的特征。因此在需要深入理解数据特征或者实现高度可解释性的场景,解纠缠自编码器具有高度的应用价值。

对抗自编码器与变分体系的生成原理和训练方法并不相同,其不依赖变分推断<sup>[10]</sup>,而是通过训练自编码器和对抗网络来学习潜在表征,并将引入的对抗网络用于检测生成样本质量,其结构如图5所示。

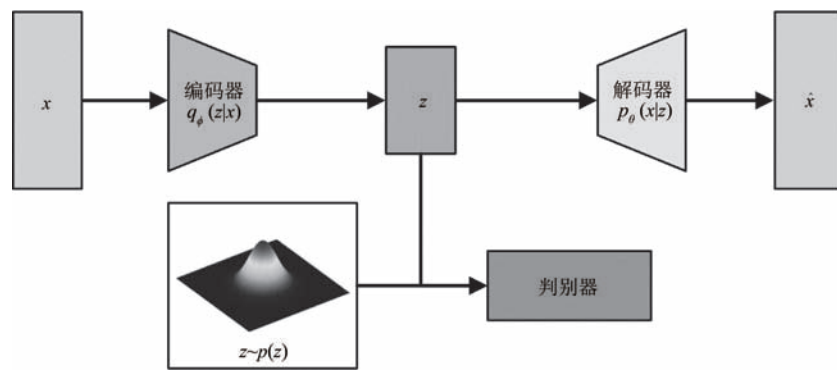


图5 对抗自编码器

生成器 $G$ 的目标是最小化生成样本与真实样本的分布之间的差异,通常使用负的判别器损失作为生成器的损失函数:

$$L_G = -\log(D(G(z))) \quad (12)$$

判别器 $D$ 的目标是正确地将生成样本与真实样本区分开,同时降低将真实样本错误分类为生成样本的可能性:

$$L_D = -\log(D(x)) - \log(1 - D(G(z))) \quad (13)$$

其中 $x$ 是输入数据, $z$ 是从潜在空间中采样的随机样本。对抗自编码器在自编码器重构损失的基础上,加入了生成器 $G$ 核判别器 $D$ 的对抗损失。整体的损失函数可表示为

$$L_{AAE} = L_{AE} + \lambda L_G + \gamma L_D \quad (14)$$

其中 $\lambda$ 和 $\gamma$ 是权衡生成器损失和判别器损失的超参数。对抗自编码器由于兼顾自编码器的重构目标和

生成对抗网络的变分推理,从而具备生成高质量样本的能力。

除了改变损失函数以满足特定需求外,调整特征提取方式以适应特定数据,拓展了自编码器的实际应用领域。降噪自编码器如图6所示,其在输入数据中引入噪声,然后通过解码器恢复为原始的干净数据,提升网络鲁棒性和泛化性能<sup>[18]</sup>。

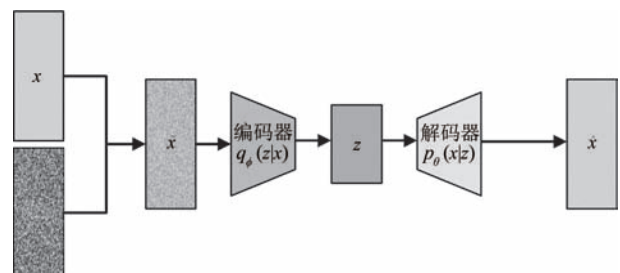


图6 降噪自编码器

在提取方式上,卷积自编码器通过卷积操作提取空间数据的局部特征,被广泛应用于图像特征提取任务<sup>[19]</sup>,其结构如图7所示。

时序自编码器基于循环神经网络<sup>[8]</sup>及长短期记忆网络<sup>[20]</sup>等结构,可从具有时间顺序特性的输入中提取信息,其基本结构如图8所示。网络状态沿时间步传递给下一时刻,因而时序自编码器在视频处理、语音识别以及动作识别等涉及时序关系的任务中被广泛采纳。

堆叠自编码器连接多个自编码器以提升特征处

理能力,其结构如图9所示,被应用于更为复杂的场景<sup>[14]</sup>。

依据自编码器的不同用途,自编码器的功能可归纳为特征提取与表征学习、数据重构与去噪、数据降维与压缩、生成模型与样本生成、分布识别、聚类、推荐系统、异常检测,以及分类9种。按照实现策略、目标以及是否需要进一步处理等因素,自编码器的用途及特征整理如表1所示。

不同类型的自编码器具备不同功能,选用的类型通常取决于具体的应用场景和数据特性。对于结

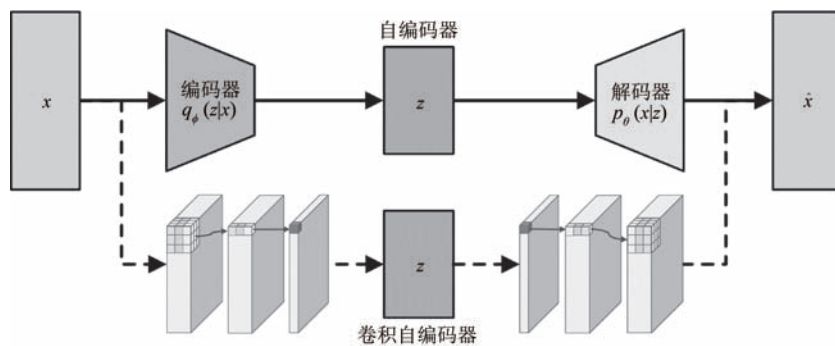


图7 卷积自编码器

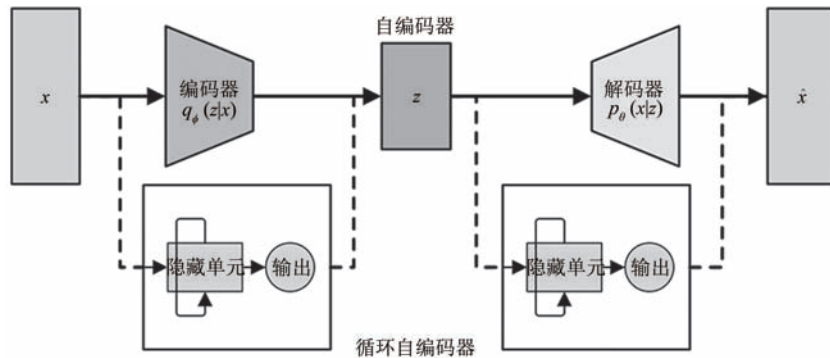


图8 循环自编码器

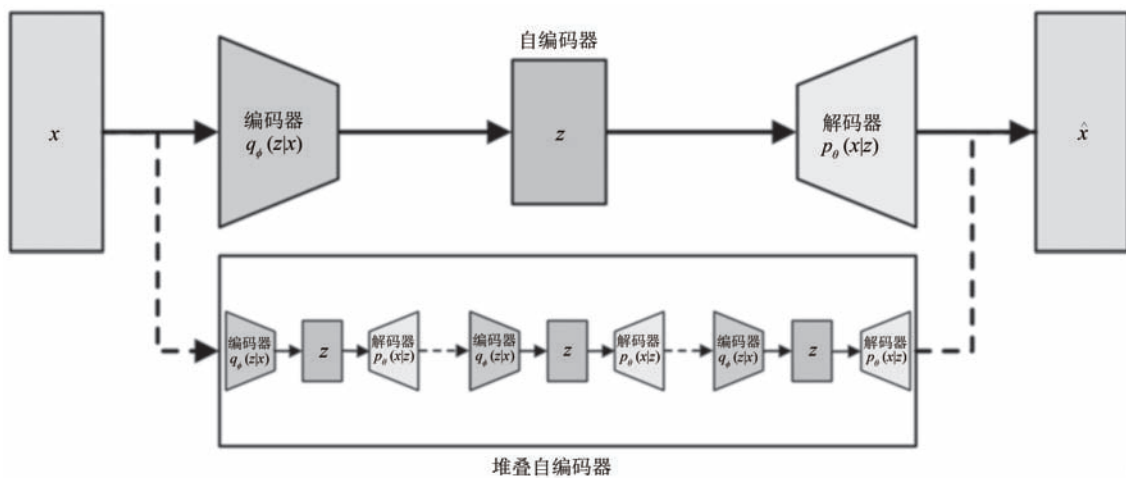


图9 堆叠自编码器

表1 不同类型自编码器六个维度的特点比较

目标	意义	策略	流程	维度	提取信息	需后续处理
特征提取与表征学习	提供紧凑、有意义的表示形式,促进特征工程和模型性能提升	最小化无监督的特征重构误差	$xz\hat{x}$	不限	$z$	×
数据重构与去噪	去除数据中的噪声或冗余信息,提高数据质量和模型鲁棒性	人为引入噪声,以最小化重构误差的方式学习如何还原带噪数据为干净数据	$xz\hat{x}$	不变	$\hat{x}$	×
数据降维与压缩	降低计算和存储成本,并提高后续任务的效率和模型的泛化能力	将高维输入数据映射到低维特征空间	$xz\hat{x}$	下降	$z$	×
分布识别	控制数据质量	最小化输入数据与重构输出之间的误差来学习数据的概率分布	$xz\hat{x}$	下降	$z$	×
生成模型与样本生成	作为样本生成工具,实现数据增强	基于学习到的概率分布,在特征空间中随机采样来生成具有相似特征和分布的合成样本	$xz\hat{x}$	不限	$\hat{x}$	×
聚类	服务于数据自动聚类任务,实现群体分析	基于隐藏层的特征或者学习到的特征分布,实现样本的聚类	$xz\hat{x}$	不限	$z$	√
推荐系统	提升推荐的个性化程度,提高推荐的准确性	基于自编码器学习到的隐藏特征之间的相似度,找到目标群体	$xz\hat{x}$	不限	比对 $x, \hat{x}$	√
异常检测	无监督场景中根据重构误差识别异常数据,为决策提供有力支持	通过学习正常特征或其分布,检测异常或离群点	$xz\hat{x}$	不限	比对 $x, \hat{x}$	×
分类	实现无监督特征提取与有标签分类的结合,充分利用数据资源	将编码器的隐藏特征连接到分类器模型,实现自动提取特征和分类预测	$xz\hat{x}, xzy$	不限	$y$	×

构化数据,自编码器主要用于数据的生成和重构任务<sup>[21]</sup>。而在图像处理中,自编码器的应用更为多样化,去噪自编码器消除图像噪声<sup>[18]</sup>,对抗自编码器检测图像任务<sup>[10]</sup>,变分自编码器生成图像<sup>[22]</sup>等。在处理序列数据时,自编码器被用于提取视频序列特征<sup>[23]</sup>、智能体任务推断<sup>[20]</sup>等。对于包含节点和边的图数据,自编码器可用于发现节点之间的重要联系和模式<sup>[24]</sup>以及生成伪样本<sup>[25]</sup>等。总体而言,自编码器在不同类型数据的任务中被多样化应用,目前已成为执行多样化任务的强有力工具。由于自编码器的广泛应用,学者们对自编码器在不同场景和数据类型下的性能进行了深入对比研究。关于不同类型自编码器在多种数据集上的对比实验,可参考综述<sup>[2,26-27]</sup>,这些文献中包含对不同类型自编码器在不同数据集上优势的详细分析。

自编码器作为传统的生成模型,尽管在生成能力上可能不如生成对抗网络和扩散模型,但其无需像生成对抗网络那样训练生成器和判别器,也无需进行扩散模型涉及的马尔可夫链反向过程,从而可以更好地避免训练过程中可能出现的不稳定和难以收敛的问题。随着自编码器体系的发展,在引入对抗网络后,基于自编码器的网络结构也能具备生成高质量样本的能力。此外,自编码器还具备特征提取与压缩表征、异常检测与分布识别等多项能力。这些特点使得自编码器在持续学习范式下具有独特

的优点和价值,特别是在解决灾难性遗忘和进行知识传输方面。通过利用自编码器的这些特点,模型能够在动态环境中压缩表征以减少存储需求、生成伪样本以回忆先前信息、识别任务分布以动态扩张网络分布、提取模型知识以指导后续任务等,从而构建出更为稳定和高效的持续学习系统。

## 2.2 持续学习增量场景与挑战

持续学习起源于终身学习<sup>[28]</sup>,是一种新兴的学习范式,以增量式更新模型的方式,在为新任务训练自适应模型的同时,注重保持模型在原有任务上的表现<sup>[29]</sup>,示意图如图10。

该范式在不断累积知识以提升模型有效性的同时,凭借增量建模方式节约计算资源和时间,为完成

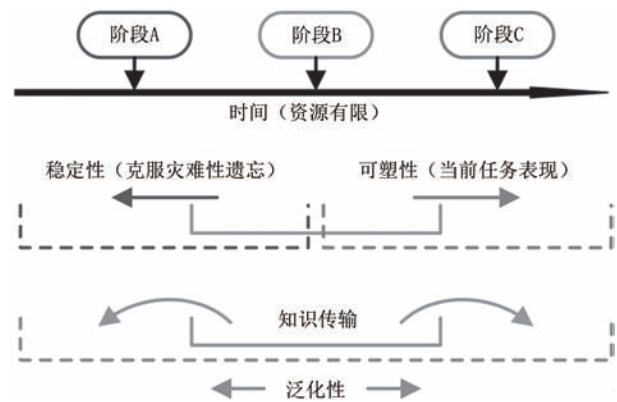


图10 持续学习示意图



资源有限背景下的一系列任务提供了切实可行的方案<sup>[30]</sup>。持续学习可数学化表示为<sup>[31]</sup>

$$\langle h_{k-1}, M_{k-1}, D_k \rangle \rightarrow \langle h_k, M_k \rangle \quad (15)$$

其中,  $h_{k-1}$  和  $h_k$  分别表示任务  $k-1$  和任务  $k$  的持续学习模型;  $M_{k-1}$  和  $M_k$  分别表示从任务  $k-1$  和任务  $k$  中提取的辅助信息;  $D_k$  表示任务  $k$  的增量数据。受到数据分布变化、任务类型多样以及模型复杂程度加深等原因影响, 持续学习范式面临多重挑战, 其中有两项引起广泛关注<sup>[32]</sup>, 一项是灾难性遗忘, 另一项是知识传输。

### 2.2.1 多种增量场景

随着研究问题不断深入, 持续学习已细化出多种增量场景<sup>[33]</sup>, 主要包括数据增量<sup>[8]</sup>、域增量<sup>[17]</sup>、类增量<sup>[34]</sup>、任务增量<sup>[35]</sup>、联邦增量<sup>[36]</sup>, 研究设定也从任务标签已知<sup>[37]</sup>, 探索至更有挑战性的任务标签未知场景<sup>[38]</sup>。基于任务分布的变化情况, 以及任务标签是否可得这两项指标, 数学化表示持续学习增量场景如表2所示。

数据增量是持续学习任务中最基本的场景, 新旧任务目的相同, 且数据分布一致<sup>[15]</sup>。域增量在领域或环境变化的情况下执行相同任务<sup>[39]</sup>。例如, 对于持续性区分猫与狗这一任务而言, 旧任务基于真实的猫与狗的图像训练模型, 而新任务的训练数据则与之前分布不同, 是卡通的猫与狗的图像。无论是数据增量还是域增量场景, 由于任务标签唯一, 无需为测试阶段指明所属任务。不同表数据增量分布相同的简单设定, 域增量场景下尽管任务目标相同, 模型还需适应分布变化<sup>[17]</sup>。

类增量更加复杂, 新的类别随着时间推移而出现。训练进程具备明确的任务标签, 新类别在指定阶段出现, 且数据分布改变仅由类别新增造成<sup>[40]</sup>。例如, 在持续性区分动物种类分类的进程中, 先前阶段基于牛、羊训练模型, 而新阶段则基于马、驴样本构建模型。尽管类别牛、羊对应先前阶段, 类别马、驴对应新阶段, 但测试进程并不指明样本所属阶段, 即反映样本所属阶段的任务标签无法获取, 这要求模型在能识别旧类的基础上, 还具备识别新类的能力<sup>[34]</sup>。为适应新类别出现, 模型需要动态调整与类别数量匹配的输出头数量, 甚至调整模型结构<sup>[41]</sup>。任务增量场景中, 新阶段的任务分布会发生改变<sup>[42]</sup>, 甚至出现新、旧任务目标不同的情况。例如, 旧任务是图像分类任务, 新任务则是图像识别, 为了适应新任务的输入输出特点, 模型结构可能需要较

大程度地调整<sup>[35]</sup>。尽管如此, 任务增量场景具备直接的任务指示性, 不仅训练进程明确任务内容, 即给定任务标签, 测试进程同样能在任务标签指引下, 直接选用对应的模块执行测试<sup>[4]</sup>。

当任务层面的标签不可知时, 任务增量的难度更大, 训练进程未知任务标签, 各阶段训练数据分布不同的成因可能仅仅是域改变, 也可能是增加新类别, 甚至是出现了新任务。因而往往需要模型具备识别任务信息的能力, 并以较高的灵活性学习新任务的特征和规律<sup>[43]</sup>。数据增量、域增量、类增量、任务增量, 以及任务未知的增量场景, 都基于集中式服务器进行。随着近年来保护数据隐私、分布式数据源、实时性要求及降低通信开销等现实需求的提出, 联邦学习借鉴持续学习的思路, 发展出了联邦增量学习<sup>[36]</sup>。不同于单一机构采用集中式方式开展不同类型的增量任务, 联邦增量学习的任务, 在数据不出本地的情况下, 由多个机构持续性共同执行<sup>[44]</sup>。客户端进行各自的任务, 在样本分布存在差异的情况下, 客户端在训练进程中进行多轮信息交互, 并基于返回的聚合信息更新本地模型。样本所在客户端被视为任务标签, 即明确测试进程任务标签, 客户端对本地样本进行预测。

### 2.2.2 克服灾难性遗忘

在持续学习任务中, 新任务适应性调整模型, 可能会导致模型在旧任务上性能下降, 甚至无法完成旧任务。由于学习新任务, 模型丧失先前任务的相关知识或能力, 这一现象被称为灾难性遗忘<sup>[33]</sup>。

在已知任务标签的增量场景中, 基于回放、正则和参数孤立是持续学习克服灾难性遗忘常见的方法<sup>[30]</sup>。基于回放的方法保存一部分历史数据<sup>[23]</sup>, 或者生成历史数据的伪样本<sup>[7]</sup>或特征<sup>[45]</sup>, 并将真实或者生成的这些信息添加到新阶段的训练进程中, 从而帮助模型回忆旧任务。回放方法尽管能在一定程度上保留旧任务的记忆, 但在高维数据和复杂增量场景中难以提高旧类的分类精度<sup>[46]</sup>。此外, 当新、旧任务的分布差异较大时, 回放历史样本会对新任务造成严重干扰<sup>[47]</sup>。设定恰当的样本挑选、伪样本生成机制是回放方法中关键却有挑战性的步骤<sup>[48]</sup>。基于正则的方法能在一定程度上避免上述问题, 其在模型训练过程中引入正则化项, 约束模型参数变动, 以防止旧任务信息的严重遗忘<sup>[49]</sup>。基于正则的方法, 不需要存储真实样本, 也不用训练生成器, 而是通过调节正则化项的权重来推动新、旧任务之间的平衡<sup>[24]</sup>。然而, 基于正则的方法可能会带来过度约



束的问题。由于参数灵活性受限,新任务的可塑性较低<sup>[50]</sup>。此外,正则化约束项随新任务出现而增加<sup>[51]</sup>,持续性保留旧任务知识的难度也越大<sup>[6]</sup>,因而在任务序列过长时难以得到推广。基于参数孤立的方法将不同任务的参数分隔开,避免任务之间的相互影响<sup>[52]</sup>。按照模型是否扩张这一指标,当前基于参数孤立的方法,可区分为静态孤立与动态孤立,前者有着固定的网络容量,后者逐渐扩张网络结构。尽管基于参数孤立的方法能避免任务之间的干扰<sup>[47]</sup>,却也阻碍任务传递共有信息。设立共享模块来提取任务之间的通用信息,已成为了解决这一问题的有效方法<sup>[35]</sup>,公共结构保留任务之间的共性,任务专属的子结构则维护任务特性<sup>[53]</sup>。

基于参数孤立方法为不同任务分配独立节点,实现任务隔离<sup>[11]</sup>,相比于克服灾难性遗忘的其他方法,有着独特的优势。与基于回放方法相比,该方法

无需存储历史样本,从而减少资源消耗<sup>[33]</sup>。而且,冻结旧任务参数,使得任务之间更加独立,避免了基于正则化方法中对新、旧任务的复杂权衡<sup>[28]</sup>。参数孤立尽管能保留旧任务知识,为后续任务提供回溯,然而,模型结构随任务增加而不断扩张,增加了突破资源承载力的可能性<sup>[54]</sup>。如何把控模型规模与扩张速度,是基于参数孤立实现持续学习需要解决的重要问题<sup>[48]</sup>。

然而,当任务标签未知时,持续学习面临着更大的挑战。模型需要在没有明确任务描述的情况下,快速适应并表现出良好性能<sup>[55]</sup>。基于回放<sup>[38]</sup>、基于正则化<sup>[56]</sup>、基于参数孤立<sup>[43]</sup>等方法同样是解决任务未知场景下灾难性遗忘问题的重要思路 and 典型策略。自编码器在任务未知场景下普遍被用作任务识别器,检测任务层面的信息<sup>[38]</sup>,或者作为样本生成器缓解灾难性遗忘<sup>[11]</sup>。

表2 持续学习常见增量场景<sup>[33]</sup>

场景	训练阶段	测试阶段
数据增量	$\{D_t, t\}_{t \in T}, \rho(X_i) = \rho(X_j) \& Y_i = Y_j \forall i \neq j,$ 任务唯一	$\{\rho(X_t)\}_{t \in T}$ , 不限定有、 无任务标签
域增量	$\{D_t, t\}_{t \in T}, \rho(X_i) \neq \rho(X_j) \& Y_i = Y_j \forall i \neq j,$ 不限定有、无任务标签	$\{\rho(X_t)\}_{t \in T}$ , 不限定有、无任务标签
任务增量	$\{D_t, t\}_{t \in T}, \rho(X_i) \neq \rho(X_j) \& Y_i \cap Y_j = \emptyset \forall i \neq j,$ 指定任务标签	$\{\rho(X_t)\}_{t \in T}$ , 指定任务标签
类增量	$\{D_t, t\}_{t \in T}, \rho(X_i) \neq \rho(X_j) \& Y_i \cap Y_j = \emptyset \forall i \neq j,$ 指定任务标签	$\{\rho(X_t)\}_{t \in T}$ , 任务标签无法获取
任务未知增量	$\{D_t, t\}_{t \in T}, \rho(X_i) \neq \rho(X_j) \& Y_i = Y_j \forall i \neq j,$ 任务标签无法获取	$\{\rho(X_t)\}_{t \in T}$ , 任务标签无法获取
联邦增量	$\{D_t, t\}_{t \in T}, \rho(X_i) \neq \rho(X_j) \& Y_i = Y_j \forall i \neq j,$ 多客户端联合训练	$\{\rho(X_t)\}_{t \in T}$ , 指定客户端

$T$ 代表总任务量,  $t, i$ 与 $j$ 表示任务序号。

$D_t$ 代表任务 $t$ 的训练数据集,  $X_t, Y_t$ 分别代表任务 $t$ 的数据集与标签集。

$\rho$ 代表数据分布。

### 2.2.3 实现知识传输

知识传输旨在从任务中提取一般性知识,并应用于新任务,起到加速新任务学习进程,或者提升新任务学习能力的效果<sup>[31]</sup>。如何定义知识并实现有效传输是知识传输的重点,已有方法包括层级特征提取<sup>[57]</sup>、元学习<sup>[36]</sup>、知识蒸馏<sup>[48]</sup>、原型<sup>[37]</sup>、知识库<sup>[58]</sup>等。

层级特征提取,将旧任务的部分架构纳入新任务的特征提取器,通过重用网络而直接复用先前模型知识<sup>[59]</sup>。通过共享底层网络,新任务从先前任务中继承得到底层特征的知识,基于增量数据仅学习任务特定的顶层特征<sup>[57]</sup>。这种方式能在新、旧任务

存在一定特征相似性时传递任务知识,并加速新任务学习进程。

元学习是一种学习如何学习的方法,其主要目的是提高模型快速适应新任务的能力<sup>[36]</sup>。与持续学习不同,其在训练阶段从这个任务分布中抽取多个分布的数据同时进行训练,在推理阶段则是接收一个新分布的任务,且不需要考虑在训练阶段包含的分布上的表现。元学习与持续学习的关联在于,元学习能够从先前分布中抽象出通用规律以指导新分布的任务,因此也常被用于实现持续学习中的知识迁移。相比于层级特征提取,元学习放松了对任务

相似性的要求,具备更大的灵活性和适用性,能较好地保证模型泛化能力<sup>[60]</sup>。知识蒸馏与之类似,提取抽象知识指导新任务,并没有直接复用模型参数。旧任务模型为教师模型,当前任务所构模型视为学生模型,受教师模型指导<sup>[40]</sup>。学生模型兼顾自身数据与教师模型软目标,从教师模型中学习更加丰富的知识表示和决策规则<sup>[48]</sup>。然而,知识蒸馏存在与层级特征提取相似的局限性,即要求新、旧任务具备一定的相似性。此外,受学生模型容量等因素制约,教师模型中的细微差异或决策边界等信息,会对学生模型产生较大影响。如何以较小损失将知识通过知识蒸馏传递给学生模型还需探讨<sup>[22]</sup>。

基于原型的方式注重提取样本中的高普适性信息,却依旧存在丢失局部信息的风险,有损任务的多样性<sup>[61]</sup>。知识库作为知识的显示存储工具,指导新任务建模,并不断更新。与其他知识传输方式不同,知识库为所学知识提供单独存储空间,避免知识遗忘和丢失<sup>[62]</sup>。新任务依据匹配关系,从知识库中提取相关知识作为指导<sup>[63]</sup>,并在任务完成后,依据最新信息更新知识库<sup>[64]</sup>。相较于层级特征提取、元学习和知识蒸馏,知识存储能提供多样化的知识资源,并具备完整的知识迁移和长期的知识保留等优势,能够有效提升模型的性能和适应性。

不同增量场景中的任务,克服灾难性遗忘的难易程度、传输知识的有效程度,都受任务之间相似度影响。新、旧任务之间的相似度越高,知识的重叠程度更高,因而任务之间的知识适用性逐渐加强。同时,相似度越高,学习新任务对旧任务的干扰越少,越不容易发生灾难性遗忘,因而依据任务之间的相似度制定合适方法,是克服灾难性遗忘和实现知识传输的关键<sup>[31]</sup>。任务之间的相似度较高时,回放样本与当前样本的分布更相近,并且正则化约束在任务之间产生冲突的概率更小,从而更能在新、旧任务之间取得平衡。然而,基于参数孤立的方式,则可能因为限制知识共享而无法提升任务表现。相反,任务之间的相似度较低时,方案带来的效果则不同。基于回放的方式,可能因引入分布有冲突的样本而降低当前任务的可塑性<sup>[65]</sup>,且也可能由于回放样本不具代表性或者数量不足而发生灾难性遗忘<sup>[11]</sup>。在任务之间差异程度较大时,基于正则的方式也并不适用,参数约束导致损失函数面临着难以收敛的风险<sup>[66]</sup>,对比起来,参数孤立反而具备更强的适用性<sup>[48]</sup>。在现有研究中,任务之间具备明确的关联程度时往往容易直接设计持续学习方案。例如,

CLELNet<sup>[35]</sup>在任务增量背景下,设定相邻任务具有相似之处,经由卷积自编码器提取任务之间的共享表征,并将其作为通用的图像基本信息供任务使用。GIM<sup>[39]</sup>则设定子任务之间存在分布偏移,每当出现下一个任务,直接以动态扩张模型结构的方式防止任务相互干扰。更难的是新、旧任务相似度尚未明确的情况,当前评估任务相似度的方法主要有基于概率距离<sup>[48]</sup>、基于统计偏差<sup>[67]</sup>与基于样本距离<sup>[53]</sup>三类。Ye等人<sup>[48]</sup>提出KIAM机制,通过比较任务样本与既有模型之间的概率距离,评估新、旧任务相似度,从而控制模型是否扩张;A-MJPF<sup>[67]</sup>中则采用 $3\sigma$ 原则判断新阶段样本的新颖性,即若样本分布信息与训练样本分布均值的偏差高于三倍标准差,则会将当前阶段的任务视为新任务,并触发新一轮模型训练进程;L-MVAE<sup>[53]</sup>通过知识的相近程度来定义任务的相似性。L-MVAE将子任务的生成样本作为对任务知识的反映,并通过计算不同子任务生成的伪样本与新阶段样本的L2距离判断任务之间的相似程度。当最低的相似程度低于既定阈值,新阶段的任务被视为新任务,然而人为设定阈值要求先验经验,自适应的阈值生成策略是未来的探索主题。

### 3 自编码器推动持续学习

自编码器在持续学习中既可以作为持续学习范式中数据预处理的工具,提升数据的表达能力,又可以通过样本回放、正则化、参数孤立等方式参与持续学习内部,为克服灾难性遗忘和实现知识传输提供支持,文献统计见表3。

#### 3.1 数据预处理

数据预处理可以提升数据质量、优化特征表示,从而提升模型性能、适应性和效率。持续学习范式中,自编码器在特征处理、降维压缩、去噪、数据增强等环节发挥重要作用。

##### 3.1.1 特征处理

自编码器基于编码器和解码器架构提取紧凑表征,有助于减轻持续学习范式的计算负担,并增强模型对新数据的适应性。

Jha等人<sup>[24]</sup>在数据增量任务中,以节点之间结构接近度构建重构损失,从而将生物医学二分网络压缩为低维表征。通过自编码器捕获压缩特征,该方法不仅具有高保真度,还能提升计算效率。Huang等人<sup>[89]</sup>先利用自编码器在有标签数据上训练特征提

取器,再基于无标签数据的对比损失调整模型,所提出的方法在类增量识别任务上取得不错效果。在任务增量场景下,Flashcards<sup>[74]</sup>借助自编码器提取任务层级知识,为通过回放任务信息克服灾难性遗忘提供了必要准备。类似的,LwF-KG<sup>[92]</sup>提取任务重要特征,为基于聚类算法评估图像质量提供高质量的任务信息。此外,基础自编码器发展出的变体,也有被用于特征处理的案例,例如,BGPO<sup>[20]</sup>借助变分自编码器从视频帧中提取具有高斯分布的低维抽象表征,mmDDLDA<sup>[9]</sup>通过韦伯变分自编码提取文本主题特征。

近年来自编码器已发展出多种类型架构,进一步丰富了持续学习提取和处理任务信息的方式。Guo等人<sup>[15]</sup>在自编码器基础上引入长短期记忆网络形成长短期自编码器架构,并将其用于提取时序数据特征,为持续评估设备性能提供了包含时序信息的特征提取方案。相比而言,引入卷积和池化架构以提取图像特征的卷积自编码器更为多见。在类增

量场景下,CALCS-CL<sup>[58]</sup>基于卷积自编码器提取空间局部特征以及特征之间的相关性,并将其映射进知识库,丰富了类增量场景下的图片分类方法。在任务增量场景中,CLELNet<sup>[35]</sup>为提取任务之间的相似信息,将卷积自编码器作为任务共享模块,而后为新任务添加专属任务特征,形成完整的特征提取器。通过合成任务共享、任务特定两个模块,CLELNet在包含共享知识的同时又包含任务的独特性,为提取任务增量的特征提供了新思路。此外,自编码器不仅能够在本服务器上提取任务特征,而且在聚合多个服务器信息的中心服务器上,即联邦增量学习场景下也有应用。例如,CISR<sup>[44]</sup>在自编码器架构上,依据参数的重要性限制参数变动幅度,提取医疗图像的形状和语义信息表征。类似地,MetaCL<sup>[36]</sup>依靠自编码器处理患者的时间序列生理信号,并通过一维的卷积层和一维的最大池化层将信号映射为潜在特征,进而开展联邦增量学习。

表3 自编码器推动持续学习在不同增量场景中克服挑战

文献	自编码器类型						持续学习增量类型						解决持续学习挑战	
	经典	变分	对抗	降噪	时空	其他	数据	域	类	任务	联邦	任务未知	灾难性遗忘	知识传输
Jha等 <sup>[24]</sup>	●	○	○	○	○	○	●	○	○	○	○	○	●	○
RAG <sup>[16]</sup>	●	○	○	○	○	○	○	●	○	○	○	○	●	○
GIM <sup>[39]</sup>	●	○	○	○	○	○	○	●	○	○	○	○	●	○
ECLA <sup>[68]</sup>	●	○	○	○	○	○	○	●	○	○	○	○	●	○
TsAE <sup>[40]</sup>	●	○	○	○	○	○	○	○	●	○	○	○	●	○
AQM <sup>[69]</sup>	●	○	○	○	○	○	○	○	●	○	○	○	●	○
BinPlay <sup>[70]</sup>	●	○	○	○	○	○	○	○	●	○	○	○	●	○
CoCo <sup>[71]</sup>	●	○	○	○	○	○	○	○	●	○	○	○	●	○
LogCL <sup>[72]</sup>	●	○	○	○	○	○	○	○	●	○	○	○	●	○
TAMiL <sup>[38]</sup>	●	○	○	○	○	○	○	○	●	●	○	○	●	○
RCR <sup>[73]</sup>	●	○	○	○	○	○	○	○	●	●	○	○	●	○
Khatib等 <sup>[4]</sup>	●	○	○	○	○	○	○	○	○	●	○	○	●	○
Generating PEs <sup>[21]</sup>	●	○	○	○	○	○	○	○	○	●	○	○	●	○
Flashcards <sup>[74]</sup>	●	○	○	○	○	○	○	○	○	●	○	○	●	○
FAE-based Task Detector <sup>[10]</sup>	●	●	●	○	○	○	●	○	○	○	○	○	●	○
NICAG <sup>[64]</sup>	○	●	○	○	○	○	●	○	○	○	○	○	●	●
GR-IG <sup>[8]</sup>	○	●	○	○	○	○	●	○	●	○	○	○	●	○
Generative Negative Replay <sup>[46]</sup>	○	●	○	○	○	○	●	○	●	○	○	○	○	○
A-MJPF <sup>[67]</sup>	○	●	○	○	○	○	○	●	○	○	○	○	●	●
ACS <sup>[17]</sup>	○	●	○	○	○	○	○	●	○	○	○	○	●	○
DEGM <sup>[47]</sup>	○	●	○	○	○	○	○	●	○	○	○	○	●	○
Sadhu等 <sup>[75]</sup>	○	●	○	○	○	○	○	●	○	○	○	○	●	○
Evolved Mixture Model <sup>[76]</sup>	○	●	○	○	○	○	○	●	○	○	○	●	●	○
KIAM <sup>[48]</sup>	○	●	○	○	○	○	○	●	○	○	○	●	●	●
Rahman等 <sup>[51]</sup>	○	●	○	○	○	○	○	●	●	●	○	○	●	○



续表

文献	自编码器类型						持续学习增量类型						解决持续学习挑战	
	经典	变分	对抗	降噪	时空	其他	数据	域	类	任务	联邦	任务未知	灾难性遗忘	知识传输
KRCL <sup>[7]</sup>	○	●	○	○	○	○	○	○	●	○	○	○	●	●
mmDDLDA <sup>[9]</sup>	○	●	○	○	○	○	○	○	●	○	○	○	●	○
OpenVAE <sup>[34]</sup>	○	●	○	○	○	○	○	○	●	○	○	○	●	○
UGCL <sup>[57]</sup>	○	●	○	○	○	○	○	○	●	○	○	○	●	○
Lesort 等 <sup>[77]</sup>	○	●	○	○	○	○	○	○	●	○	○	○	●	○
Wang 等 <sup>[78]</sup>	○	●	○	○	○	○	○	○	●	○	○	○	●	○
Smaller is Better <sup>[79]</sup>	○	●	○	○	○	○	○	○	●	○	○	○	●	○
Alqhtani 等 <sup>[80]</sup>	○	●	○	○	○	○	○	○	●	○	○	○	●	○
Tf-GCZSL <sup>[56]</sup>	○	●	○	○	○	○	○	○	●	○	○	●	●	○
Brain-like Replay <sup>[81]</sup>	○	●	○	○	○	○	○	○	●	●	○	○	●	○
Li 等 <sup>[18]</sup>	○	●	○	●	○	○	○	○	●	●	○	○	●	○
Robinson 等 <sup>[6]</sup>	○	●	○	○	○	○	○	○	○	●	○	○	●	○
Ma 等 <sup>[25]</sup>	○	●	○	○	○	○	○	○	○	●	○	○	●	○
LGR <sup>[45]</sup>	○	●	○	○	○	○	○	○	○	●	○	○	●	○
L-MVAE <sup>[53]</sup>	○	●	○	○	○	○	○	○	○	●	○	○	●	○
BooVAE <sup>[82]</sup>	○	●	○	○	○	○	○	○	○	●	○	○	●	○
S-TRIGGER <sup>[83]</sup>	○	●	○	○	○	○	○	○	○	●	○	○	●	○
SRM <sup>[84]</sup>	○	●	○	○	○	○	○	○	○	●	○	○	●	○
cVAE <sup>[85]</sup>	○	●	○	○	○	○	○	○	○	●	○	○	●	○
LGM <sup>[22]</sup>	○	●	○	○	○	○	○	○	○	●	○	○	●	●
CURL <sup>[11]</sup>	○	●	○	○	○	○	○	○	○	●	○	●	●	○
Bayesian Structural Adaptation <sup>[86]</sup>	○	●	○	○	○	○	○	○	○	●	○	●	●	○
ODDL <sup>[43]</sup>	○	●	○	○	○	○	○	○	○	○	○	●	●	○
MoVE Layers <sup>[55]</sup>	○	●	○	○	○	○	○	○	○	○	○	●	●	○
BGPO <sup>[20]</sup>	○	●	○	○	●	○	○	○	○	○	○	●	○	○
Clifer <sup>[37]</sup>	○	○	●	○	○	○	○	○	●	○	○	○	●	●
DEJA VU <sup>[12]</sup>	○	○	○	●	●	○	○	○	○	○	○	○	●	○
CLTP-MAN <sup>[23]</sup>	○	○	○	○	●	○	●	○	○	○	○	○	●	○
MetaCL <sup>[36]</sup>	○	○	○	○	●	○	●	●	○	●	●	○	●	○
ISYANA <sup>[49]</sup>	○	○	○	○	●	○	○	●	●	●	○	○	●	○
REMIND <sup>[87]</sup>	○	○	○	○	●	○	○	●	○	○	○	○	●	○
EPIE-Net <sup>[13]</sup>	○	○	○	○	●	○	○	○	●	○	○	○	●	○
ABC <sup>[52]</sup>	○	○	○	○	●	○	○	○	●	○	○	○	●	○
Progressive Clustering <sup>[88]</sup>	○	○	○	○	●	○	○	○	●	○	○	○	●	○
Huang 等 <sup>[89]</sup>	○	○	○	○	●	○	○	○	●	○	○	○	●	○
CALCS-CL <sup>[58]</sup>	○	○	○	○	●	○	○	○	●	○	○	○	○	○
CLIS <sup>[90]</sup>	○	○	○	○	●	○	○	○	○	●	○	○	●	●
CLELNet <sup>[35]</sup>	○	○	○	○	●	○	○	○	○	●	○	○	○	○
CISR <sup>[44]</sup>	○	○	○	○	●	○	○	○	○	○	●	○	●	○
NDL <sup>[19]</sup>	○	○	○	○	●	○	○	○	○	●	○	○	●	○
PGMA <sup>[59]</sup>	○	○	○	○	○	●	○	○	○	●	●	○	○	○
PRER <sup>[5]</sup>	○	○	○	○	○	●	○	○	○	○	○	○	●	○
Self-net <sup>[42]</sup>	○	○	○	○	○	●	○	○	○	○	○	○	●	○
ADCN <sup>[14]</sup>	○	○	○	○	○	●	○	○	○	○	○	○	●	○
KIERA <sup>[91]</sup>	○	○	○	○	○	○	○	○	○	○	○	○	○	○

注：●表示文献设定的研究类型，○表示文献不包含对应类型

### 3.1.2 降维压缩

持续学习范式处理的数据通常具备较大规模,其中可能包含大量高维特征。这些高维数据不仅会增加建模的计算复杂度,还可能引发维度灾难等问题。数据降维和压缩能够将高维数据映射到低维空间,提供更高效的数据表示,同时也有助于模型更好地捕捉数据的变化模式和分布。

Pelosin 等人<sup>[79]</sup>在有限的存储空间中,通过不同的降维方案分析了基于回放的持续学习框架下样本数量与质量的权衡。5个数据集上的实验结果表明,增加存储的样本数量能有效提升模型性能,特别是在低内存配置下,使用简单的图像压缩方案(例如将图像尺寸压缩至 $8 \times 8$ )实现了高达67%的性能提升。压缩样本来增加可存储的样本数量,从而提升模型性能的这一方式,为计算机视觉领域的图像分类任务提供了重要参考。SRM<sup>[84]</sup>不仅借助离散变分自编码器压缩图像表征,还为压缩后的表征开辟单独的内存空间来提升回放效率。在无监督场景下,Progressive Clustering<sup>[88]</sup>借助卷积自编码器将图像数据从高维空间映射到低维空间,并将其作为聚类任务的初始抽象表征。类似的,CLIS<sup>[90]</sup>将卷积架构引入到变分自编码器架构中,通过降维压缩图像的空间与分布信息来提升图像分割效果。

### 3.1.3 数据去噪

在持续学习中,数据的噪声和干扰会削弱模型提取知识的能力,导致性能下降,甚至对后续任务产生负面影响。降噪自编码器主动添加噪声,并将样本还原为原始输入,从而在一定程度上具备抗噪能力。在持续学习中,基于降噪自编码器提取的任务表征具有干净、有效等特点。

A-CZSL<sup>[93]</sup>在解码阶段加入噪声,训练鲁棒性更强的条件变分自编码器,从而提升自编码器学习干净表征的能力。类似的,Li 等人<sup>[18]</sup>在条件变分自编码器的解码阶段引入高斯噪声,增强生成模型的鲁棒性。随后,混合生成所得表征与新任务真实表征,并用于共同训练分类器。这一方式在医疗领域类增量和任务增量场景中,都实现了较高的准确率和较低的遗忘率。同样的,DEJA VU<sup>[12]</sup>以回放的方式克服灾难性遗忘,并在卷积自编码器架构中引入噪声来提升表征泛化性能。

### 3.1.4 数据增强

借助数据变换、扩充等技术,数据增强能提升样本多样性,有助于降低过拟合风险,提高模型泛化能力。自编码器作为一种强大的生成模型,已在多项

研究中被用作数据增强工具。

变分自编码器通过学习样本的概率分布,及其所在的潜在空间信息,能够生成多样性样本。LGM<sup>[22]</sup>采用变分自编码器提取任务分布信息,为教师-学生架构保留分布信息提供了必要准备。同时,为达到学生模型中新、旧任务样本量平衡,LGM将该变分自编码器作为生成模型回放样本。文章借助变分自编码技术,不仅促成了类别数量的平衡,还通过生成先前阶段的样本缓解了任务增量场景下的灾难性遗忘。cVAE<sup>[85]</sup>是任务增量场景下克服灾难性遗忘的另一种策略,其借助两套变分自编码器,将任务间和任务内的关系映射进共享的潜在空间中,并通过条件变分自编码器生成样本,在新任务中实现数据增强。卷积自编码器也是增强数据的常见架构,DEJA VU<sup>[12]</sup>从正态分布中随机初始化卷积自编码器参数,为域增量的图像分类任务提供历史样本信息。此外,DEJA VU还在训练中引入噪声,域变化,模型依旧能保持泛化能力。Clifer<sup>[37]</sup>则是受对抗网络启发而引入新的模块来增强样本。具体而言,在传统自编码器架构基础上,Clifer添加判别器网络来评估生成图像质量,保证持续学习任务能用高质量样本进行回放操作。

自编码器在当前作为数据预处理工具中已经展现出了显著的优势,未来可以通过优化结构和训练方法,进一步提高其精度和效率,为持续学习提供更加紧凑和有用的特征表示。此外,自编码器作为传统的生成模型,如何与先进的样本生成技术结合还需要进一步探索。尽管当前已有初步成果<sup>[37]</sup>,但未来研究可以进一步探索如何优化这些过程,以实现更高效的自动化数据增强。

## 3.2 有监督场景模型构建

当前,针对如图11所示的持续学习增量场景,自编码器及其各类变体凭借独特结构与功能,已从多种角度解决任务已知、未知场景中持续学习面临的难题。不同方法及其组合的模型如图12所示。

### 3.2.1 克服灾难性遗忘

任务标签已知时,模型无需推断任务层面的信息。目前,基于回放、基于正则、基于参数孤立及其组合的方法,是缓解灾难性遗忘的关键策略。

#### (1) 基于回放

基于回放的方法将旧任务信息添加至新任务的训练过程,使得模型在学习新任务的同时保持了对旧任务的记忆<sup>[23]</sup>。自编码器通过尽可能准确地重构输入样本,具备学习样本特征、生成伪样本的能

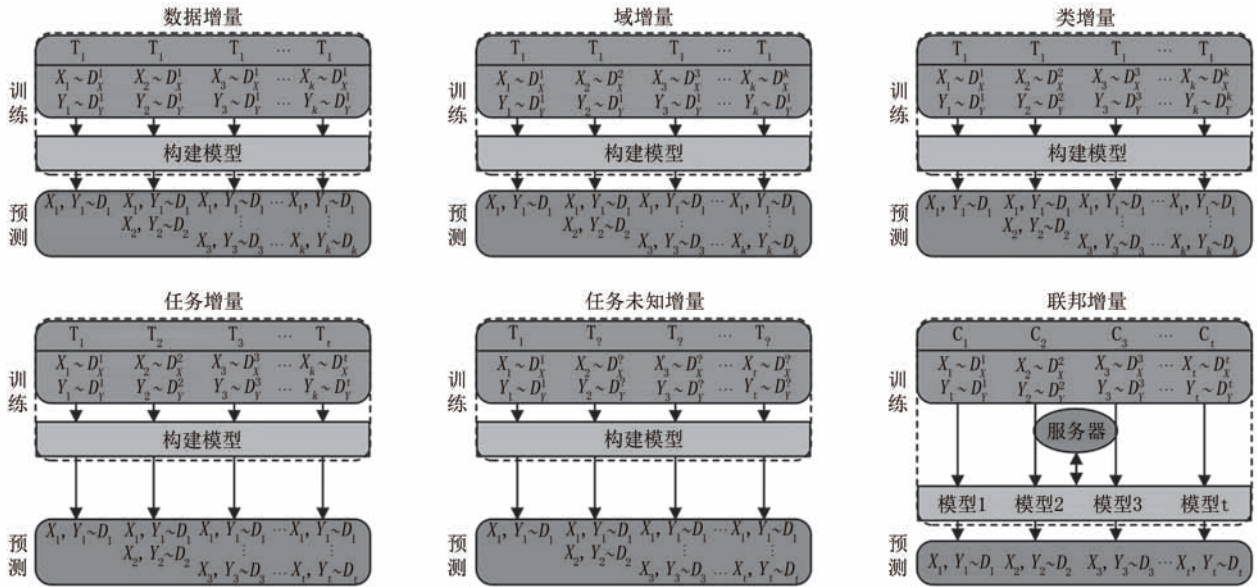


图 11 持续学习增量场景图(图中,  $T_k$ 代表第  $k$  项任务,  $C_k$ 代表第  $k$  个客户端执行的任务,  $X_k$ 与  $Y_k$ 分别代表不同阶段的样本与标签,对应的数据集分别为  $D_X^k$ 与  $D_Y^k$ ,共同构成  $D_k$ 。符号“ $\sim$ ”代表分布一致性)。

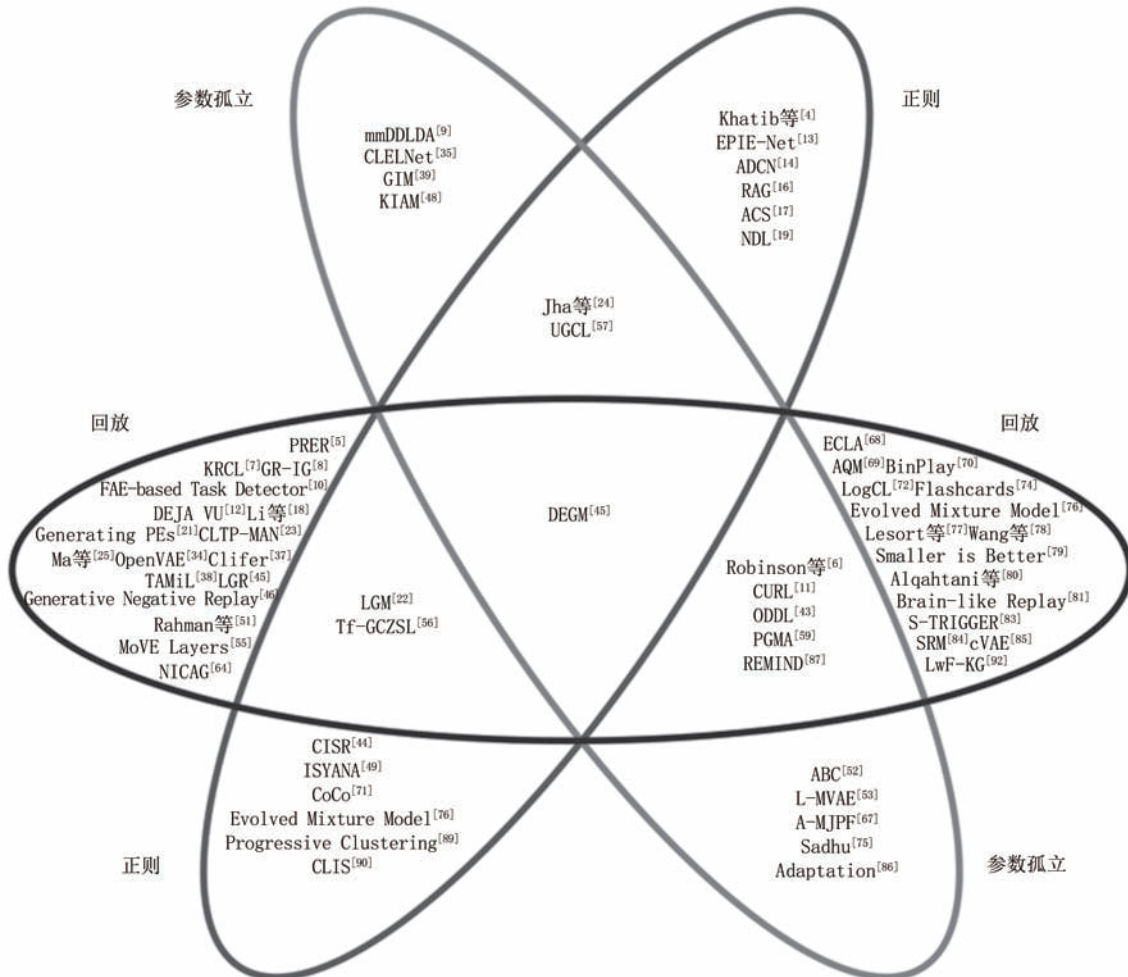


图 12 不同类型方法的组合应用



力。自编码器作为生成器,帮助持续学习回放先前信息,从而无需存储真实样本就能维持持续学习范式中的历史记忆<sup>[77]</sup>。

借助自编码器生成并回放的信息类型,除了伪样本外,伪特征也同样常见<sup>[45]</sup>。相比于保存并回放真实样本,生成回放策略尽管更加节约存储资源,但却降低了回放数据质量。为此,保存少量真实样本并将其与生成样本混合,从而实现存储开销与回放质量的平衡是一种可行方法<sup>[6]</sup>。

自编码器可生成多种类型的回放信息,其中伪原始数据最为常见。ECLA<sup>[68]</sup>将基础自编码器作为样本生成器,为域增量场景提供伪样本,缓解了图像分类任务中的灾难性遗忘。变分自编码器及其变体,具备提取任务分布的能力,并可依此生成高质量的伪数据,因此在持续学习中很受欢迎。S-TRIGGER<sup>[83]</sup>借助变分自编码器持续检测任务是否发生变化,并自动触发生成重放机制来防止灾难性遗忘。在强化学习的一系列任务中,S-TRIGGER执行快速且性能较好。KRCL<sup>[7]</sup>基于变分自编码器变体 Beta-VAE,提取任务之间的共享表征,并为新任务生成先前任务的伪本来缓解灾难性遗忘。为增强模型学习能力,KRCL引入了知识蒸馏模块,先前模型的知识被用于指导模型的后续更新。Clifer<sup>[37]</sup>在自编码器架构上借鉴生成对抗网络思想,加强对生成样本质量的把控。

同样是通过回放策略来克服灾难性遗忘, BinPlay<sup>[70]</sup>并没有将自编码器直接作为样本生成器,而是借助自编码器将原始数据映射为二进制编码,并采样这些二进制编码来确定伪回放样本。该方式无需在内存中保存样本,却能为持续学习提供原始数据信息。在此基础上,作者进一步探索如何在生成模型的复杂度与回放样本数量之间实现平衡,为此提出的 LogCL<sup>[72]</sup>能将回放样本数量控制在任务总样本量的对数范围内。

借助自编码器生成潜在特征以作为回放对象,是回放的另一种常见方案。PRER<sup>[5]</sup>在自编码器中加入稀疏化约束,学习更具意义和紧凑的表征,并将其作为回放对象。相比于回放原始类型的伪样本,生成模型的复杂度更低,且训练模型耗时更少。不足的是,回放伪特征可能会丢失更多信息。受限于数据集的复杂程度,如何在更复杂的数据集上增强 PRER 的适用性,还有待进一步研究。同样为减少内存占用、降低计算消耗而采用回放伪特征的还有 LGR<sup>[45]</sup>,其构建变分自编码器,生成低维表征并

进行回放。这一策略在 CK+、RAF-DB 以及 AffectNet FER 三个架构上都取得了较好的表现。类似的,Alqahtani 等人<sup>[80]</sup>和 Li 等人<sup>[18]</sup>采用回放策略来克服灾难性遗忘。借助变分自编码器生成回放所用的潜在表征,该方法的有效性分别在类增量和任务增量的分类任务得到验证。由于回放质量的高低直接影响任务表现,如何回放有代表性的特征引发了研究人员的关注。受大脑仅重放重要而非全部信息这一机制启发,Brain-like Replay<sup>[81]</sup>将网络自身信息作为代表性内容进行回放,指导模型在新任务中更新。该方式已在类增量场景下的自然图像分类任务上取得成效,为回放策略提供了新思路。

## (2) 基于正则

基于正则的方法通过增添正则化项,从约束参数的角度保持了模型的稳定性。集成了自编码器的正则化方法,引入了重构损失以持续性地限制参数变动,从而维护先前所学信息。在任务增量场景下,提取任务层信息十分关键。Khatib 等人<sup>[4]</sup>在传统判别模型基础上,将重构损失纳入总体损失,从而在一系列任务中提取到不只包含判别特性的高级特征,有效地缓解了任务增量场景下图像分类任务中的灾难性遗忘。

在一系列新、旧任务关联程度较高的图像分类任务中,神经元已学信息有助于模型更新。ISYANA<sup>[49]</sup>作为检测任务相关性的代表,在卷积自编码器基础上借助正则化约束调整学习率。在新任务相关性较低时,模型以降低节点学习率的方式缓解遗忘现象,新任务与旧任务的重合信息越少,学习率越低。此外,节点与任务之间的相关性,同样被用于调整学习率。节点与当前任务相关性越低,正则化约束将节点冻结的程度越大。模型在域、类和任务增量的场景下,既避免了新任务对已有知识的干扰,也防止了知识的负面传输。然而,持续性调整学习率这一方式可能发生过拟合。CLIS<sup>[90]</sup>在通过变分自编码器提取压缩后的图像表征后,使用贝叶斯非参建模提高模型的泛化能力,能在一定程度上降低模型过拟合的风险。CLIS 将先前任务所学知识整合为概率先验,并通过正则化约束参数更新。受益于先前知识,模型具备较好的适应性,在医疗图像分割任务上取得了较好表现。尽管如此,该方式仅针对单一站点开展持续分割任务有效,而无法在多个站点之间共享知识。

CISR<sup>[44]</sup>则提供了联邦持续的图像分割方案。该方案在数据隐私保护和存储受限的情景下,基于

多个站点来持续优化医学图像分割模型。依据参数对形状以及语义信息的敏感程度,CISR衡量模型的综合重要性。而后,CISR通过正则化约束参数的变动幅度,即对综合重要性更高的参数,予以较大的惩罚,而重要性较低的参数发生的波动,仅对应较低损失。模型在六个站点上实施联邦持续任务并取得较好效果,为保留先前站点的图形知识以及语义信息提供了可行策略。

### (3) 基于参数孤立

基于参数孤立的方法通过为每个任务分配物理或逻辑模块来调整网络结构,因此该方法也被称为基于结构的方法<sup>[52]</sup>。它旨在通过保持已有信息来尽可能维持先前任务的性能,同时通过调整节点信息或扩展新模块来提升新任务的表现<sup>[53]</sup>。

在固定网络容量条件下,Robinson等人<sup>[6]</sup>一方面借助变分自编码器生成伪样本防止灾难性遗忘,另一方面采用基于参数填充的方式不断扩充新信息来增加当前任务的可塑性。该方法评估节点的重要程度,并将重要程度较低节点的权重重置为零来存储新信息。与之类似,Jha等人<sup>[24]</sup>评估参数对损失函数的贡献来量化参数的重要性,并冻结较高重要性的参数以保存先前所学知识。遗憾的是,该框架仅适用于单增量任务,且网络容量的初始设定对任务表现有着较大影响。特征所包含的信息不仅受限于网络规模,随着任务增加,网络还难以实现稳定性与可塑性之间的平衡。

相较于固定网络容量下的通过参数填充以更新模型的方法,当网络容量可扩展时,在物理上隔离新、旧任务能够更好地保护旧任务信息,并防止旧任务对新任务的干扰。传统的方式将变分自编码器作为基本模型,通过动态扩张的方式持续性地为新任务训练子网络。每到来一个新任务,原有架构添加一个包含了条件自编码器的新模块,在保证先前信息不受影响的同时,适应性地学习新类别。尽管如此,由于不具备提取任务层面信息的能力,多数方法在固定网络容量的情况下仅能在单一任务场景中使用。L-MVAE<sup>[53]</sup>作为任务增量场景中的持续学习方案,借助变分自编码器识别任务的分布信息,具备判断任务层信息的能力。新任务构建专属变分自编码器,并以最大化对数似然性和整体变分下界的方式,联合训练整个模型。L-MVAE不仅隔离不同任务的知识以平衡稳定性与可塑性,还实现规模的自适应调整,另外在克服灾难性遗忘的同时,加速新任务的学习进程。

除了设计专属于子网络架构外,CLELNet<sup>[35]</sup>注重

任务间的信息共享,其在任务之间共用一部分网络结构作为通用知识。经由卷积自编码器提取任务之间的共享表征,模型将其作为任务之间通用的图像基本信息。而后,CLELNet为每个任务分配任务特定层,用于提取任务的专属特征。随任务增加,固有的任务共享层结构保持不变,而新任务对应的专属表征层随之增加。通过设定共享层与特定任务层,CLELNet既能保留任务的共性知识,防止灾难性遗忘,又能保证新任务学习专属知识,具备任务的特殊性。GIM<sup>[39]</sup>策略与之类似,基于循环神经网络来动态扩展网络,在新分布的任务到来时冻结模型参数,并在此基础上为新任务部署专属的循环神经网络。类似的还有文献[75],其采用扩展网络的方式持续开展任务。当分布发生变化时,该框架视当前任务为新任务,并为其分配单独的子网络结构。网络权重被视为任务知识,模型冻结先前的权重信息以防发生灾难性遗忘。尽管这一方式避免了任务的相互干扰,但由于需增加的数量随任务线性增加,任务量较多时模型扩张仍会带来较高的时空复杂度。此外,由于框架无法自动检测任务之间的相似程度,在更为一般的、任务未知的场景中,模型无法自主判断是否应该扩张。

任务未知时,模型保持先前任务的稳定性面临着更大的挑战。IRCL<sup>[94]</sup>为了缓解测试阶段任务信息未知时出现的灾难性遗忘问题,提出了一种基于回放的解决方案。由于未知任务标签,IRCL首先基于条件变分自编码器学习不同类别的通用表征,而后将其结合区分不同类别的判别表征来执行分类任务,解决了任务信息未知时克服灾难性遗忘这一难题。此外,为保留先前累积的知识,模型借助变分自编码器生成先前任务的伪样本,并以回放的方式将这些样本与当前任务样本共同训练。MoVe Layers<sup>[55]</sup>同样是通过变分自编码器生成回放样本的方式来解决灾难性遗忘。为应对任务标签未知的这种情况,该方法构建了多个子网络,并通过门控机制识别任务层信息。此外,其以贝叶斯前景理论为基础制定了分层信息论最优原则,保证了一系列任务的稳定性和可塑性。然而,受限于已有的子网络结构,该框架难以正确识别之前未出现过的类别。Mundt等人<sup>[34]</sup>提出的框架OpenVAE能较好地避免这一问题。OpenVAE运用极值理论,借助变分自编码器进行后验变分推断,通过检测已有分布外的样本,防止错误预测。此外,为了避免回放的样本信息量过小,该模型采用了一种回放样本机制,通过从聚合后验的高概率密度区域显式生成分布内样本,



从而以更高的回放质量克服灾难性遗忘。

探寻任务边界,是任务未知场景下克服灾难性遗忘的常见策略。Tf-GCZSL<sup>[56]</sup>基于两套变分自编码器提取图像语义与视觉特征,并将短期记忆和长期记忆相结合,明确了不同任务的边界。此外,Tf-GCZSL将提取的特征作为暗知识执行知识蒸馏,以正则化的方式保留先前任务的参数信息,并借助生成回放伪特征巩固任务的稳定性。然而,随着任务增加,基于正则化的方式可能造成对先前任务的过度保护,而损失当前任务的可塑性。为此,基于结构的方法被广泛采纳。类似于探寻任务边界,A-MJPF<sup>[67]</sup>利用变分自编码器统计分布信息,并通过聚类算法划分组别。每个组的质心、协方差和半径等信息被提取后,与提前设定的阈值进行比较,从而判断当前任务是否是新任务。若新任务出现,基于已有架构,模型增添专属子网络用于学习新任务特性。通过以聚类的方式提取组别信息,A-MJPF可应用于任务未知场景,受限于设置阈值伴随的较高主观性,模型难以在动态扩张的程度与表现之间取得平衡。LMC<sup>[41]</sup>同样通过阈值比较判断任务是否已知,对域、类增量场景下的持续学习任务开展研究。模型依据自编码器重构损失判断任务新、旧,并通过动态扩张网络来持续性存储新任务知识。当重构损失在阈值范围内,重构损失最小的分支预测样本类别,否则当前任务被视为新任务,新增专属子网络,进而完成持续性分类任务。然而,LMC缺乏对模型扩张规模的约束。随着任务序列的增加,模型会不断扩张,这在资源有限的背景下难以被广泛应用。

WEVAE<sup>[95]</sup>将模型剪枝考虑在内,为平衡模型表现与规模提供了解决办法。基于动态扩张策略,模型学习多样化信息,同时,衡量分支之间的相近程度,剪枝机制控制模型扩张速度,依次在模型规模与效果之间取得平衡。类似地,TAMiL<sup>[38]</sup>提出自适应的网络扩展方案,其将网络结构分为两个模块,一是任务之间的公共特征提取模块,被用于捕获任务的通用信息。二是用于捕获各任务具体信息的注意力模块,维护任务的可塑性。作为任务未知场景下的解决方案,模型根据变换系数选取与当前任务最相关的特征,并将其作为任务表征。此外,为防止灾难性遗忘,模型借助自编码器生成伪样本,通过回放策略保持先前任务的信息。结合动态扩张与回放,模型可塑性与稳定性得以巩固。采取类似策略的还有ODDL<sup>[43]</sup>。该方法通过衡量分布差异程度,构建未知任务信息时的处理框架,其包含样本回放与模型扩张两种策略。一方

面,模型把控回放质量,挑选离散程度较小的样本进行回放,另一方面,模型衡量缓冲区样本与任务分布之间的差异程度,控制模型扩张速度。回放真实样本的方式尽管能巩固先前记忆,存储真实样本带来的存储压力却会在长序列的持续学习任务中不断累加。EMM<sup>[76]</sup>在未知任务信息的场景下融合动态扩展、经验回放和正则技术三种策略,克服灾难性遗忘。模型设立任务专属的变分自编码器和分类器,并借助希尔伯特-施密特独立性准则评估已学分布信息与当前内存缓冲区分布信息之间的概率距离。依据分布差异程度,EMM判断当前任务是否为新任务,并为新任务划分以及训练专属子网络。当任务数量较多时,可控制扩展速度的经验性阈值,在很大程度上影响着模型表现。然而模型的阈值设置依赖于主观经验,这在一定程度上限制了EMM的推广。

### 3.2.2 实现知识传输

在任务之间共享通用知识可以提升新任务学习效率 and 性能。尽管知识传输对持续学习有着不可忽视的重要性,但由于任务特性和数据分布等方面的差异,实现知识传输面临诸多挑战。目前关于知识传输的研究还较少,集成了自编码器的知识传输,主要停留在知识蒸馏和知识库这两种方式上。

#### (1) 基于知识蒸馏

先前任务模型作为教师模型,将其输出以及在数据分布等方面的知识,作为附加的监督信号,指导新任务训练模型。在类增量图像分类任务中,TsAE<sup>[40]</sup>为了将先前任务的知识泛化到后续任务中,使用已有分类器作为教师模型,并通过蒸馏损失与分类损失共同约束新任务上的训练。该模型为每个任务训练一个自编码器分支,用于独立学习任务特有信息。重构损失最小的分支被选用于任务预测,预测类别判定为分类得分最高的类别。结构简单的TsAE框架,在基于知识蒸馏的类增量学习方法中,具备通用性与可移植性。KRCL<sup>[7]</sup>采取类似的策略,借助知识蒸馏技术传输任务知识,为每个任务都配备由条件变分自编码器和分类器构成的知识蒸馏模块。借助知识蒸馏添加监督信号,模型能够学习到先前图像分类任务的特征知识。在传输知识的同时,模型同样注重防止灾难性遗忘,KRCL借助变分自编码器生成伪样本,并以基于回放的方式避免灾难性遗忘。类似的,NICAG<sup>[64]</sup>在拟人化机器人手臂的一系列抓取任务中,借助知识蒸馏与样本回放保护和利用所学知识。内存缓冲区存储代表性样本,以基于回放的方式协助新模型更新。模型借助变分自编码器捕捉相邻任务的



KL 散度,并以知识蒸馏的方式将知识保留在参数空间中。通过累积和运用知识,NICAG在一系列拟人化机器手臂抓取任务中,以较低的平均损失和遗忘率取得较高的抓取成功率。

### (2) 基于知识库

建立并持续更新知识库,来为新任务提供直接指导,是实现知识传输的重要方法。基于知识库的方法,将先前任务的重要信息、模型经验等知识单独保存。新任务从知识库提取知识,以辅助模型构建,并依据提取的信息更新知识库。知识库的构建、应用以及更新,为知识保护和传输提供有力支持。CALCS-CL<sup>[58]</sup>借助卷积自编码器提取图像空间局部特征,并借助知识库指导模型增量更新。模型采用临时存储器保存编码后的特征,并从中提取泛化知识形成长期记忆存至知识库。通过表征压缩与转换,CALCS-CL实现了跨任务的知识传输,其在单增量任务与多增量任务上的有效性在3个基准图像数据集上得以验证。

存储在知识库中的知识,可以以多种形式存在。Clifer<sup>[37]</sup>将知识以原型形式存储,便于知识累积、共享与更新。此外,借助条件对抗自编码器生成伪样本,模型保留了先前任务的信息,有利于实现任务稳定性与可塑性之间的平衡。然而,随着新类别的出现,原型的数量将不断增加。在类别数多且任务量大的场景中,更新原型将会消耗大量计算资源,并且容易降低类间稀疏性。MetaCL<sup>[36]</sup>将不同类别之间的关系信息作为知识存储于知识库中。模型在知识库基础上,结合元学习增量提升表征学习能力的同时,实现知识有效更新。随时间推移,知识库累积的知识更全面,MetaCL在数据增量、域增量、任务增量与联邦增量时都取得突出表现。

为了进一步增强持续学习范式中的模型性能,在具备样本标签的情况下,借助标签信息提取和整合特征表示是实现这一目标的有效途径。cVAE<sup>[85]</sup>通过引入分类器模块,提取标签信息以完善整体架构,是对扩展自编码器结构的创新尝试。未来还需要探索其他的改造或扩展自编码器结构的方式,引入注意力模块以自动聚焦于与标签相关的重要特征或许值得尝试。

### 3.3 半监督、无监督等特殊场景模型构建

在有标签样本有限,或是样本不具标签情况下,模型需具备更强学习能力才能充分理解数据。在标签不足的情况下,持续学习借助自编码器开展的研究,还处于初步探索阶段,文献统计见表4。

#### 3.3.1 任务已知

当有标签样本的数量有限时,CoCo<sup>[71]</sup>制定了一种半监督学习方案。模型首先以监督学习方式,借助标签知识训练自编码器,而后通过大量的无标签图像来更新知识。通过对比学习的方式,CoCo为正、负样本对之间的相对信息构建了对比损失,以实现基于正则化约束的持续性表征学习。该半监督学习方案,不仅基于正则化约束克服了图像分类任务中潜在的灾难性遗忘问题,同时也为标签数量不足时持续性执行二分类任务提供了可行办法。然而,当前所提框架不适用于多分类任务,如何在半监督场景增强模型泛化能力还需进一步研究。

在无监督场景下,LGM<sup>[22]</sup>将变分自编码器作为任务基础模型,从未标记的数据中学习包含分布信息的表征。LGM将先前所学的分布的概率性知识存储在教师模型中,借助知识蒸馏向新任务传输所学分布信息,使得生成任务之间实现知识传递。学生模型将生成回放的伪样本与当前任务的样本混合,在适应教师模型分布的同时,提取高泛化性表征,从而在稳定性与可塑性之间达到平衡。

此外,一些研究不仅在无监督场景下可执行持续学习任务,还兼容于有监督学习。mmDDLDA<sup>[9]</sup>是一种类增量场景下的深层主题建模方案。为了保证文本特征的非负性,模型采用了变分自编码器的变体韦伯变分自编码器<sup>[96]</sup>,它将输入映射到符合韦伯分布的表征空间中,从而克服了高斯分布无法实现全部表征非负化的限制。同时,结合马尔可夫链式的蒙特卡罗与变分推断,mmDDLDA探索文本后验分布并生成后验样本,并通过调整梯度方向提升稀疏、非负和偏斜表征的有效性。模型首先在网络结构的每一层上引入共享主题,而后借助自编码器开展无监督学习。此外,为了适应有监督场景,模型以主题间最大间隔原则作为分类准则,指导分类任务。更有挑战性的是,cVAE<sup>[85]</sup>提供了一种任务增量场景下能执行有、无监督学习的可用方案。模型将任务间与任务内的信息映射到同一个表征空间,并训练条件变分自编码器作为生成器,生成并回放信息以克服灾难性遗忘。特殊的是,模型关注回放信息的质量,为了克服从复杂集合后验抽样的局限性,引入额外的近似阶段以生成高质量伪特征。cVAE<sup>2</sup>通过调整两阶段训练框架以适应新任务,既能处理无监督场景下的持续学习任务,也能通过判别损失处理类增量任务。Adaptation<sup>[86]</sup>同样为有、无监督场景制定不同策略,不仅能解决任务增量的

有监督图像分类任务,也能执行无监督场景下的图像生成任务。基于贝叶斯神经网络模型,模型依据判别损失学习样本标签,同时也基于变分自编码器的重构损失,执行无监督任务。在任务增量场景中,

模型以非参数贝叶斯和变分贝叶斯为建模原则,对神经网络的每个隐藏层建模。每个任务根据节点权重自主激活隐藏层中的节点数,从而动态调整网络连接以适应新任务。

表4 自编码器与持续学习在不同学习场景下的结合

场景	文献	克服遗忘							传输知识		
		回放样本	回放表征	回放伪样本	回放伪表征	正则	静态孤立	动态孤立	知识蒸馏	知识库	原型
有监督	CLTP-MAN <sup>[23]</sup>	●	○	○	○	○	○	○	○	○	○
	TAMiL <sup>[38]</sup>	●	○	○	○	○	○	○	○	○	○
	AQM <sup>[69]</sup>	●	○	○	○	○	○	○	○	○	○
	BinPlay <sup>[70]</sup>	●	○	○	○	○	○	○	○	○	○
	ODDL <sup>[43]</sup>	●	○	○	○	○	○	●	○	○	○
	NICAG <sup>[64]</sup>	●	○	○	○	○	○	○	●	●	○
	LogCL <sup>[72]</sup>	●	○	●	○	○	○	○	○	○	○
	Robinson等 <sup>[6]</sup>	●	○	●	○	○	○	○	●	○	○
	F AE-based Task Detector <sup>[10]</sup>	○	●	○	○	○	○	○	○	○	○
	Li等 <sup>[18]</sup>	○	●	○	○	○	○	○	○	○	○
	Smaller is Better <sup>[79]</sup>	○	●	○	○	○	○	○	○	○	○
	Alqahtani等 <sup>[80]</sup>	○	●	○	○	○	○	○	○	○	○
	Brain-like Replay <sup>[81]</sup>	○	●	○	○	○	○	○	○	○	○
	REMIND <sup>[87]</sup>	○	●	○	○	○	○	●	○	○	○
	GR-IG <sup>[8]</sup>	○	○	●	○	○	○	○	○	○	○
	DEJA VU <sup>[12]</sup>	○	○	●	○	○	○	○	○	○	○
	Generating PEs <sup>[21]</sup>	○	○	●	○	○	○	○	○	○	○
	Ma等 <sup>[25]</sup>	○	○	●	○	○	○	○	○	○	○
	OpenVAE <sup>[34]</sup>	○	○	●	○	○	○	○	○	○	○
	Generative Negative Replay <sup>[46]</sup>	○	○	●	○	○	○	○	○	○	○
	Rahman等 <sup>[51]</sup>	○	○	●	○	○	○	○	○	○	○
	MoVE Layers <sup>[55]</sup>	○	○	●	○	○	○	○	○	○	○
	ECLA <sup>[68]</sup>	○	○	●	○	○	○	○	○	○	○
	Flashcards <sup>[74]</sup>	○	○	●	○	○	○	○	○	○	○
	Lesort等 <sup>[77]</sup>	○	○	●	○	○	○	○	○	○	○
	Wang等 <sup>[78]</sup>	○	○	●	○	○	○	○	○	○	○
	S-TRIGGER <sup>[83]</sup>	○	○	●	○	○	○	○	○	○	○
	SRM <sup>[84]</sup>	○	○	●	○	○	○	○	○	○	○
	PGMA <sup>[59]</sup>	○	○	●	○	○	○	○	●	○	○
	KRCL <sup>[7]</sup>	○	○	●	○	○	○	○	○	●	○
	Clifer <sup>[37]</sup>	○	○	●	○	○	○	○	○	○	●
	PRER <sup>[5]</sup>	○	○	○	●	○	○	○	○	○	○
	LGR <sup>[45]</sup>	○	○	○	●	○	○	○	○	○	○
Tf-GCZSL <sup>[56]</sup>	○	○	○	●	●	○	○	○	○	○	
Khatib等 <sup>[4]</sup>	○	○	○	○	●	○	○	○	○	○	
EPIE-Net <sup>[13]</sup>	○	○	○	○	●	○	○	○	○	○	
RAG <sup>[16]</sup>	○	○	○	○	●	○	○	○	○	○	
ACS <sup>[17]</sup>	○	○	○	○	●	○	○	○	○	○	
NDL <sup>[19]</sup>	○	○	○	○	●	○	○	○	○	○	
CISR <sup>[44]</sup>	○	○	○	○	●	○	○	○	○	○	
ISYANA <sup>[49]</sup>	○	○	○	○	●	○	○	○	○	○	

续表

场景	文献	克服遗忘							传输知识		
		回放样本	回放表征	回放 伪样本	回放 伪表征	正则	静态孤立	动态孤立	知识蒸馏	知识库	原型
有监督	Evolved Mixture Model <sup>[76]</sup>	○	○	○	○	●	○	○	○	○	○
	Progressive Clustering <sup>[88]</sup>	○	○	○	○	●	○	○	○	○	○
	CLIS <sup>[90]</sup>	○	○	○	○	●	○	○	○	○	○
	Jha等 <sup>[24]</sup>	○	○	○	○	●	○	●	○	○	○
	CLELNet <sup>[35]</sup>	○	○	○	○	○	○	●	○	○	○
	GIM <sup>[39]</sup>	○	○	○	○	○	○	●	○	○	○
	ABC <sup>[52]</sup>	○	○	○	○	○	○	●	○	○	○
	L-MVAE <sup>[53]</sup>	○	○	○	○	○	○	●	○	○	○
	A-MJPF <sup>[67]</sup>	○	○	○	○	○	○	●	○	○	○
	Sadhu等 <sup>[75]</sup>	○	○	○	○	○	○	●	○	○	○
	TsAE <sup>[40]</sup>	○	○	○	○	○	○	○	●	○	○
	CALCS-CL <sup>[58]</sup>	○	○	○	○	○	○	○	○	●	○
MetaCL <sup>[36]</sup>	○	○	○	○	○	○	○	○	●	●	
有监督 无监督	cVAE <sup>[85]</sup>	○	○	●	○	○	○	○	○	○	○
	CURL <sup>[11]</sup>	○	○	●	○	○	○	●	○	○	○
	mmDDLDA <sup>[9]</sup>	○	○	○	○	○	○	●	○	○	○
	Bayesian Structural Adaptation <sup>[86]</sup>	○	○	○	○	○	○	●	○	○	○
无监督	KIERA <sup>[91]</sup>	○	○	●	○	○	○	○	○	○	○
	LGM <sup>[22]</sup>	○	○	●	○	●	○	○	●	○	○
	DEGM <sup>[47]</sup>	○	○	●	○	●	○	●	○	○	○
	ADCN <sup>[14]</sup>	○	○	○	○	●	○	○	○	○	○
	UGCL <sup>[57]</sup>	○	○	○	○	●	○	●	○	○	○
	KIAM <sup>[48]</sup>	○	○	○	○	○	○	●	●	○	○
半监督	CoCo <sup>[71]</sup>	○	○	○	○	●	○	○	○	○	○

注：●表示文献采用的方式，○表示文献未采用对应方式

### 3.3.2 任务未知

新任务缺少任务层面标签,即在任务未知场景中,模型基于半监督、无监督等特殊学习方式克服灾难性遗忘和实现知识传输面临更大的挑战。CURL<sup>[11]</sup>依据分布信息推断任务标签,且在扩张网络以存储新任务知识的同时,采用生成回放技术来保持对于先前任务的记忆。值得注意的是,CURL注重提取任务之间的通用表征以学习任务之间的共性知识,并通过扩张子网络这一方式来提取任务的专属信息,从而维持任务的可塑性。尽管如此,直接根据似然函数筛选样本构建新分支,难以适应多样化分布的样本,在任务中包含多种分布类型的样本时,模型面临失效风险。由短时和可扩展记忆系统组成的无监督持续学习框架KIAM<sup>[48]</sup>则更加关注任务之间的分布差异。模型借助变分自编码器提取未知任务的分布信息,并将分布之间的差异程度作为衡量新任务的指标。当分布的差异程度过大,当前

任务被视为新任务,且为此扩张网络结构。此外,模型包含两套记忆:一是将最新的训练样本作为短时记忆,二是将包含多任务信息的模型视为参数化的可扩展记忆。模型通过知识蒸馏获取先前任务知识,同时也从当前样本中提取近期信息。值得注意的是,KIAM具备子网络剪枝机制。模型依据分布信息评估分支之间的重叠程度,并在分支数量达到预设上限时剪去最大重叠程度的分支,这使得KIAM比无法约减网络规模的方法更加资源友好。

同样关注规模约减的还有ADCN<sup>[14]</sup>,其自适应调整网络结构完成域、类增量任务。模型借助堆叠自编码器,通过深度非线性变换提取复杂分布的任务特征,并在每个隐藏层上执行自聚类,动态增长以及修剪节点、隐藏层。此外,ADCN引入了正则化项,并且其约束强度会随类别数量的增加而增大,从而在提取当前任务重要表征时,动态地加强对以往知识的保护。类似的,UGCL<sup>[57]</sup>同样基于网络扩张



和正则约束构建无监督图像生成框架。不同的是,UGCL将变分自编码器作为基本模型,依据参数均值和标准差构建参数的信噪比,并在衡量节点重要性之后判断是否剪枝,从而避免参数冗余。此外,为累积知识,模型为新任务动态添加神经元,保证始终有足够的自由空间储存新任务信息。为保护先前所学的重要信息,UGCL评估参数的重要性,并限制关键参数的变动幅度,为任务未知场景下的无监督学习提供一种有效方法。不同的是,KIERA<sup>[91]</sup>借助聚类以避免发生灾难性遗忘。模型基于堆叠自编码器提取特征,随分布变化而适应性调整隐藏层的层数、节点数量、簇数。为避免灾难性遗忘,KIERA回放聚类后的质心,为回放代表性样本提供参考。

当有标签样本数量有限或样本无标签时,模型需要具备更强的学习能力才能充分理解和处理数据。通过整合多模态数据信息,模型可以借助来自不同模态的数据互补信息,提高学习效率和效果。当前借助自编码器构建的持续学习模型,主要还停留在单一数据类型上。然而,自编码器具有巨大的潜力,可以在多模态数据环境中发挥更大的作用。自编码器能够检测数据分布的变化,帮助模型判断任务层信息,并通过压缩和融合不同模态的数据特征,形成统一的特征表示。例如,自编码器可以处理图像、文本和音频等不同模态的数据,通过提取高维特征并融合这些特征,提高模型的学习能力和适应性。在无标签多模态数据中借助自编码器优势,这不仅能帮助模型更好地适应不断变化的任务,还有利于提升模型在复杂环境中的应用潜力。

## 4 持续学习优化自编码器

基于自编码器架构开展的工作,主要还停留在静态建模层面。然而在动态变化的现实环境中,仅依赖传统自编码器难以解决新分布、新环境下的问题。近年来,在持续学习范式的推动下,如何增量地优化自编码器以应对更加复杂的现实问题,已逐渐成为研究焦点。

AQM<sup>[69]</sup>以图像压缩和复原为任务,设计自适应量化模块确定图像的压缩程度,并以回放真实样本的方式,保持对先前类别图像的压缩与复原能力。在持续学习范式下,模型基于新增样本与回放样本增量式训练模型,该方式保证了自编码器在图像压缩和重构时具备较高的泛化性能。LDANE<sup>[97]</sup>通过动态扩展网络结构,在属性随时间变化的情况下,提

出了一种表征优化方案。LDANE借助自编码器提取网络中每个节点的低维特征表示,保证网络中邻近节点所学特征也接近,以此捕获网络的底层结构。此外,LDANE还注重捕获网络中的属性关系,即保证属性相似的节点对应的特征也相似。终身学习框架LDANE,随任务增加而自动扩展深度自编码器网络,并根据新任务信息动态优化模型结构和参数。然而,随着任务数量的增加,模型规模会不断扩张,在这种情况下捕获网络的底层结构与属性关系会导致模型的复杂度大幅上升。Guo等人<sup>[15]</sup>基于自编码器提取时序数据特征,在持续学习范式下,以较低计算开销和较高效率实时评估空调系统性能。由于输入是变长的时间序列,模型基于长短期自编码器将时间序列压缩至定长,实现特征对齐。此外,模型在持续优化的进程中引入正则化约束,避免自编码器模型的重要参数发生大幅度改变,从而保证出现先前任务的样本时,其能够快速精准识别。

为保证自编码器在学习一系列任务后,依旧能维持旧任务表现,动态扩展图方法DEGM<sup>[47]</sup>分析变分自编码器的遗忘行为,并以扩张网络结构的方式优化任务表现。模型包含任务间的共享网络与任务专属子网络两个部分,依据新任务与现有知识的相关性控制扩张进程。尽管如此,由于内存空间的限制,DEGM的扩张策略仍存在一定的局限性,如何更有效地平衡任务表现与模型参数量仍需深入研究。面临类似问题的还有ABC<sup>[52]</sup>。模型借助卷积自编码器,执行类增量场景下的生成任务。由于任务具备特殊性,模型将不同生成任务划分为独立分支,并在新任务到来时自动扩展模型。最终,模型选用重构损失最小的分支执行生成任务,否则视当前任务为新任务并为之训练新分支。然而,所提出的ABC框架仅适用于每次最多引入一个新类别的持续学习任务,对于更复杂的情况,如新任务中包含多个之前未出现的类别,仍需进一步探索解决方案。BooVAE<sup>[82]</sup>基于变分自编码器,使用一种带有熵正则化的贪婪增强方法来提取传输的后验知识,形成通用性更强的模型优化方案。为使得特征包含多样化知识,模型用尽可能小的网络结构处理当前任务,并将聚合后验作为先验知识。BooVAE通过最大化对数边际下界,减小潜在特征在近似与真实后验分布之间的差异,由于这一方式符合训练变分自编码器的常规进程,模型具备较高普适性。

更进一步地,为了能够更高效、更智能地应对复杂的现实问题,未来的研究还有必要从动态调整自

编码器结构着手,加强对自编码器规模大小的研究。当前基于参数孤立动态调整模型结构的方法,尽管能适应性动态扩张以及约减模型规模<sup>[95]</sup>,却缺少对自编码器模型扩张方式的深入探索。纵向扩张通过增加网络层数,使模型能够学习更加复杂的特征表示;横向扩张则通过增加每一层的神经元数量或增加平行分支,使模型能够并行处理更多信息。为了使模型灵活适应不同的任务需求,未来还需要基于任务复杂度和数据特征的变化制定适应性的扩张策略,从而提高学习效率并控制模型复杂度。

## 5 自编码器与持续学习结合的应用

基于现实场景的需要,自编码器与持续学习已在多个领域得到结合应用。本节旨在梳理这些应用场景,分析不同领域中的自编码器与持续学习的结合点及其应用效果,因此将按领域对相关研究进行归纳和总结,文献及其实验信息见表5。

### 5.1 计算机视觉

#### 5.1.1 图像

Lesort 等人<sup>[77]</sup>关注在持续学习场景下的图像生成任务,并将自编码器等多种生成模型作为伪样本生成工具,探索不同类型的持续学习技术在 MNIST、Fashion MNIST 和 CIFAR10 这三个常见数据集上的适用性。该模型利用自编码器生成并回放伪样本,相比直接回放真实样本或采用正则化等方法,显示出更低的遗忘率。然而,文章中选用的自编码器类型较为有限,针对不同特征的数据选择更适合的自编码器类型,可能更能凸显自编码器在样本生成方面的优势。受大脑注重回放重要信息来增强记忆这一机制启发,Brain-like Replay<sup>[81]</sup>注重选取有代表性的信息进行回放。该模型并没有直接回放原始数据,而是借助变分自编码器生成回放重要表征。该脑启发式生成回放策略,在不依赖于存储图像、特征等信息的情况下,实现了类增量分类任务的较好性能。类似地,AQM<sup>[69]</sup>将自编码器用作生成模型来产生先前任务的样本。混合这些代表性样本与新任务样本,即使任务分布变化,模型依旧具备先前记忆。图像数据集 CIFAR-100、样本量更大的图像数据集 Lidar 和 Atari,经由 AQM 处理后,图像的生成质量得以较好保证。cVAE<sup>[85]</sup>从变分推理的角度再次审视伪样本的生成过程,克服了从复杂的集合后验中抽样的局限性,在有、无监督场景下通用。模型包含两套变分自编码器,将任务间和任务内的关

系捕获在一个共享的潜在空间中,避免了推断任务层面的信息,既能处理无监督场景下的持续学习任务,也能处理类增量场景下的图像分类任务。尽管生成模型能够减少因保存真实样本而带来的内存消耗,但针对具有不同特性的数据集训练生成模型会导致计算量较大。LGR<sup>[45]</sup>为了降低训练生成模型带来的成本,借助变分自编码器训练低维潜在表征的生成器,以资源有效的方式为图像分类任务提供了新见解。Smaller is Better<sup>[79]</sup>为防止生成的低质量样本降低图像分类表现,仍以保存并回放真实样本的方式,开展持续学习任务。在存储资源有限的情况下,模型借助变分自编码器压缩图像,以保存更多回放实例。模型分析类增量场景下,样本压缩程度对图像分类表现产生的影响。在 ImageNet-100 等 5 个数据集上的测试结果显示,虽然极度压缩真实样本会降低样本质量,但样本总量的增加仍然有利于提高分类准确率,这为持续图像处理任务提供了有价值的参考。

ODDL<sup>[43]</sup>调整网络结构,无需保存样本以及训练生成模型,同样缓解了灾难性遗忘现象。模型借助变分自编码器提取分布信息,并以此评估检测缓冲区样本与已有知识之间的差异程度。即使在任务未知场景下,ODDL 仍能通过差异程度判断任务新、旧,并控制图像分类模型的扩张,以及确定执行分类任务的子网络。其不足之处在于,设定是否进行结构扩张的阈值过度依赖于主观判断,因此需要进一步探索更具适应性的阈值设定方法。与之类似,TsAE<sup>[40]</sup>基于自编码器为不同的任务训练专属子网络,通过重构损失来判断样本的任务信息并完成分类。此外,该研究将上一任务训练的分类器作为教师模型,用知识蒸馏损失与分类损失共同约束当前任务,平衡了模型的稳定性与可塑性。然而,重构特征并未考虑任务间的知识融合,网络结构随着任务的增加而持续性扩张。FAEs<sup>[10]</sup>同样基于自编码器实现任务未知场景下任务识别与模型扩张。不同的是,模型在增加子网络的同时适应性地调整已有子网络规模,使网络能够在任务之间有效地学习和转移知识。TAMiL<sup>[38]</sup>则通过共享不同任务间的特征空间来实现知识融合。为确定任务信息,模型引入了注意力机制,并基于变分自编码器选取与当前任务最相关的特征作为任务表征。此外,结合回放与正则化约束,模型加强保护类增量和任务增量场景中的先前信息,在 3 个数据集上的图像分类任务中,均维持较低的遗忘率与较高的准确率。相似



的, LGM<sup>[22]</sup>结合正则化约束与回放伪样本两项技术缓解灾难性遗忘, 并通过知识蒸馏技术在变分自编码器之间传递任务分布的概率性知识, 因而该方法在克服图像分类任务中灾难性遗忘的同时, 实现了知识传输。CALCS-CL<sup>[58]</sup>在类增量场景中通过知识库的方式传递类之间的知识。模型首先基于卷积自编码器提取图像特征, 而后以知识库的形式累积知识。模型首先提取分类器系统中基于规则的知识, 并对这些规则进行编码, 从而形成临时记忆。随后从这些临时记忆中, 模型提取知识作为长期记忆, 并将其存储在知识库中, 以指导多任务场景和单增量场景的图像分类。Clifer<sup>[37]</sup>在人脸识别任务中同样关注知识传输, 以原型的形式存储情绪记忆和语义记忆, 并将其作为知识来指导生成伪样本。为提升伪图像质量, 模型在条件自编码器的架构上引入判别网络, 捕获高质量特征来处理图像分类任务。

### 5.1.2 视频

持续学习研究范式下, 将自编码器技术应用于视频处理任务的研究还较少。A-MJPF<sup>[67]</sup>在一系列视频序列的检测与预测任务中, 借助变分自编码器提取视频帧的压缩表征, 以减少视频帧的原始尺寸。变分自编码器提取分布信息, 为马尔可夫跳跃粒子滤波器提供了必要的参数信息, 实现了资源高效的视频异常检测和持续预测。在任务未知情况下, 处理视频分类任务更具难度。FAEs<sup>[10]</sup>为此提出解决方案, 其在训练阶段为不同任务训练专属自编码器, 并选用重构损失最小的分支执行异常检测任务。为了在数据增量的一系列任务中保持先前任务的表现, 模型基于回放策略巩固已有信息。此外, 基本自编码器、变分自编码器, 以及对抗自编码器分别被作为任务的基本检测单元, 当数据有严重噪声时, 变分自编码器适应能力最强。

### 5.1.3 三维数据

ABC<sup>[52]</sup>、NICAG<sup>[64]</sup>与BGPO<sup>[20]</sup>借助自编码器处理三维立体空间中的持续学习任务。ABC<sup>[52]</sup>在类增量场景下执行点云生成任务。该模型使用卷积自编码器将输入数据映射到特征空间, 并根据重构损失判断是否为新任务, 然后通过动态扩展网络结构的方式为新任务构建子网络。通过扩张网络, 模型避免了新任务对旧任务的干扰, 也逐渐累积了多样性的生成能力。然而, 网络结构会随着新任务的增加而无限扩张, 这限制了该方法在资源受限环境中的应用。为此, NICAG<sup>[64]</sup>在不扩大网络结构的情况下, 通过适应性调整参数以提供解决方案。模型将

变分自编码器作为基本的抓取模型, 以回放真实样本的方式来巩固先前任务所学信息。为适应最新任务, 模型基于已有架构调整参数, 实现了模型稳定性与可塑性之间的平衡。此外, 为不断累积知识, 模型采用知识蒸馏技术, 将上一个抓取模型中的知识向新任务传递。该方式无需扩张模型, 却能适应性学习新任务。尽管如此, 在任务数量较多时, NICAG仍难以通过有限网络容纳参数信息, 依靠单独的内存缓冲区存放回放样本, 同样限制了该方案的推广。BGPO<sup>[20]</sup>不仅关注抓取表现, 还注重提升机器人的推动能力。模型包含两套自编码器系统, 即变分自编码器和长短期自编码器。前者用于提取视频帧的低维表征, 后者基于时序智能体数据构建循环神经网络, 将知识从以前学习的任务转移到新任务, 从而使机器人概括和优化任务性能。

### 5.2 生物医疗

在生物医疗领域, 自编码器与持续学习的结合, 常见于图像处理任务。CLIS<sup>[90]</sup>在医疗成像分割任务中, 借助卷积变分自编码器提取压缩表征, 并将先前任务的知识作为概率先验, 正则化约束参数更新。模型创新性地将注释纳入图像分割系统, 交互式地改善模型在未见过图像上的分割表现。Li等人<sup>[18]</sup>在训练医疗图像的分类模型时, 不仅关注如何保持资源有效, 还注重降低噪声带来的干扰。为此, 该研究采用持续学习范式对模型进行增量更新, 并利用条件变分自编码器生成伪特征, 通过回放的方式来克服灾难性遗忘。与存储并回放原始样本相比, 该方式能够节约大量存储资源, 相比于训练伪样本的生成器, 耗费更少计算量。此外, 为了提高生成器的鲁棒性, 模型在自编码器的训练阶段引入了噪声, 这种做法通过增强模型的鲁棒性来确保生成的特征具有高质量。不同的是, CLELNet<sup>[35]</sup>采取动态扩张网络结构的方式, 缓解内窥镜图像病变检测任务中的灾难性遗忘。模型借助卷积自编码器提取图像特征, 并以参数孤立的方式来构建检测框架。为学习任务之间的通用知识, 模型借助编码器构建任务之间的共享结构, 并在此基础上训练单独的任务子网络来提取任务专属信息。动态扩张任务子网络, 既防止新任务对之前任务的灾难性干扰, 又保证了新任务的可塑性。尽管如此, 文章指出, 仅从一家医院获取的样本量较小, 模型的有效性需要进一步考证, 如何通过多家医院和机构提取具有高泛化性的知识是一个值得探索的问题。联邦学习能够在保



护数据隐私的要求下允许多个参与方共同训练模型,因此成为了解决分布式数据合作和隐私安全问题的重要工具<sup>[98]</sup>。

在联邦场景下,CISR<sup>[34]</sup>基于持续学习与自编码器,为6家医疗机构构建联邦学习框架以持续性提升图像分割任务的表现。文章借助自编码器提取了医疗图像的形状、语义等信息,并定义了本地参数的重要性权重。为了防止在模型更新的一系列任务中发生灾难性遗忘,文章对重要性权重大的参数加以正则化约束,限制其发生变动的幅度。类似的,为了提升各个医疗机构在生理信号分类任务上的表现,MetaCL<sup>[36]</sup>在医疗机构之间共享建模知识。该研究基于卷积自编码器提取生理信号数据的特征,并以联邦学习的方式在2家医疗机构之间进行知识融合和共享。此外,文章通过元学习来指导模型的持续性更新,使其能够更好地完成不同医疗机构的生理信号分类任务。

此外,持续学习同样被应用于处理图像标签不足的分类任务。CoCo<sup>[71]</sup>将自编码器与持续学习范式结合,为解决生物学领域标签样本不足的分类问题提供了解决方案。模型基于深度自编码器构建任务网络,并在此基础上引入对比损失,从有标签数据学习表征。为减少对标记数据的需求,模型在增量学习阶段通过对比损失进行无监督训练。CoCo以半监督学习的方式充分利用样本信息,解决了有标签样本量不足时的建模困境,在肺癌、结肠癌、乳腺癌这三种图像的分类任务上体现出优越性。

Jha等人<sup>[24]</sup>则是基于图结构的生物医疗数据,开展持续性的表征学习任务。模型借助自编码器增量提取新任务信息,并以基于正则化的方式调整节点表征以反映最新知识,推动生物学双边网络的发展。自编码器被用于捕获节点邻近关系,从而保留网络的全局和局部结构特性。持续学习范式下,模型无需基于所有数据从头开始更新,而是以资源有效的方式,以高保真度实现了节点表征的更新。在更为现实的场景中,由于错误标签可能带来严重后果,医疗机构中存在着大量未标记的病例数据。在标签数据量有限这一背景下,如何训练有效分类模型,仍存在较大挑战。

### 5.3 语音与文本

Wang等人<sup>[78]</sup>研究类增量场景中的声音分类任务。为提取音频特征,时频频谱图首先经归一化处理,使得各类型音频值在区间 $[0,1]$ 范围内。规范化

后的时间序列音频数据,由一维卷积自编码器映射到潜在的表征空间中。在持续性音频分类任务中,变分自编码器作为生成器,回放先前任务的伪样本,以保证进行一系列任务后,模型依然保持对先前音频类型的处理能力。模型在包含10个类的音频数据集ESC-10上,仅用4%的原始数据训练生成器进行回放,其效果竟接近于存储20%原始数据可实现的效果。Sadhu等人<sup>[75]</sup>与之不同,以扩展模型结构的方式,增量地改进基于旧域训练的自动语音识别器。模型借助变分自编码器提取任务的分布信息,并据此评估语音特征属于已知类别的概率。对于新分布语音,模型以动态扩张的方式存储新信息。这一方式同样在持续学习范式下,帮助语音分类任务摆脱了传统从头开始训练的困境,有效避免了一系列任务的相互干扰。

类似地,mmDDLDA<sup>[9]</sup>以基于动态扩张的方式将持续学习范式应用到文本主题分类任务上,从而保证新、旧主题识别任务的良好表现。考虑到文本数据的非负、长尾等分布特性,模型并未选用处理音频帧常用的变分自编码器,而是通过韦伯变分自编码器进行变分推断来提取更符合文本特性的特征,从而提升了在文本处理任务中的表现。

### 5.4 交通

持续学习以增量更新模型的形式,不断更新模型知识,满足了交通领域对实时性和资源有效性的双重需求。Ma等人<sup>[25]</sup>将持续学习范式应用到多智能体轨迹预测任务中,并通过条件变分自编码器生成伪样本以减少任务中的灾难性遗忘。为了降低持续学习过程所需的内存要求,模型仅从当前任务中挑选有代表性的样本存至内存缓冲区,从而在保留重要信息与占用存储资源之间取得平衡。此外,模型基于变分自编码器生成回放样本,防止在一系列轨迹预测任务中发生灾难性遗忘。混合有代表性的样本子集与伪样本,不仅能防止遗忘过去任务所需知识,还能通过有效存储和检索,保证模型的响应速度。CLTP-MAN<sup>[23]</sup>更为关注交通轨迹数据自身的依赖关系,在环境分布发生改变的情况下持续预测行人轨迹。模型借助包含了门控机制的自编码器,捕获时序数据中的长期依赖关系,并基于预定义的样本采集间隔以及数量,以稀疏化回放的方式保持先前记忆,从而在学习新分布和保留重要知识之间实现平衡。

Guo等人<sup>[15]</sup>将持续学习应用于评估交通设备的运行情况。为了防止先前任务的性能下降,模型采

用了正则化约束来限制先前所学习参数的变动幅度。新颖的是,在任务的数据维度发生改变的一系列任务中,模型将长短期网络结构嵌入到自编码器中,在任务间对齐特征,因而针对变长的时间序列任务,模型仍具备处理能力。持续学习范式下,模型不仅为变长数据进行持续学习提供解决方案,同时也以较小的计算开销和较高的效率监测设备的运行,推动了交通领域“绿色低碳”技术的发展。

### 5.5 其他

采用持续学习范式并结合自编码器技术的研究方法,为各行业提供了创新性的实时数据处理和模型更新解决方案,展现出广泛的应用潜力。在强化学习任务中,S-TRIGGER<sup>[83]</sup>将持续学习范式应用于智能体识别周围环境的任务上,其借助变分自编码器压缩感知到的状态信息,并基于重构误差持续地检测环境是否发生变化。为了防止灾难性遗忘,该模型利用变分自编码器生成伪样本进行训练,为强化学习的发展提供了重要的参考。GR-IG<sup>[84]</sup>通过穿戴式传感器记录穿戴者的活动数据,实现对所执行任务的持续性识别。由于不同任务涉及的传感器类型和数量存在差异,因此不同任务的数据维度也不尽相同。为此,模型将传感器作为节点,构建完全连接的双向图神经网络。根据时间序列中相应的传感器数据调整参数,模型聚合相邻节点的信息后更新每个节点,为处理数据维度发生改变的增量式检测任务提供了可行方案。此外,模型借助变分自编码器生成用于回放的伪样本,在不存储数据的情况下保持对先前任务的记忆,较好地处理分类以及回归问题。

在地理信息科学领域,Alqahtani等人<sup>[80]</sup>借助持续学习范式进行遥感图像分类。模型将变分自编码器作为生成器回放之前任务的潜在表征,从而在防止遗忘过去任务信息的同时,减少了训练模型所需资源。EfficientNet作为分类网络架构,基于生成与真实表征训练模型,即使在类增量场景下,遥感图像分类模型也能以增量式更新的方式应对环境变化。类似的,Rahman等人<sup>[51]</sup>采用了增量式更新模型的方式来整合新信息并适应数据分布的变化。针对变化愈加复杂的恶意软件检测任务,为了能持续地抵御先前的恶意攻击,模型使用了变分自编码器生成先前分布的伪样本进行回放,该方法以资源高效的方式在域增量、类增量以及任务增量的场景下,实现了对新、旧恶意软件检测任务的平衡。

## 6 总结与展望

本综述归纳了近五年来结合持续学习与自编码器开展的研究。首先介绍了自编码器的类型与特点、持续学习面临的主要挑战以及增量场景。随后,本文探讨了两个主要方面:一是持续学习范式如何利用自编码器突破现有困境,二是自编码器如何借助持续学习范式开展自我优化。最后,本综述对二者结合的应用,按照领域以及特点进行归纳。

近年的研究表明,在资源有限的条件下,多种基于自编码器的方法已被成功应用于缓解灾难性遗忘。然而,受任务特性、数据分布和领域差异的影响,提取普适性知识并实现有效传输相对于克服灾难性遗忘而言更加复杂。尽管高效传输任务知识具有重大意义,但目前相关研究仍然较少。

在未来的研究中,自编码器与持续学习的结合有几个重要方向值得深入探索和研究。首先,尽管自编码器的生成能力相对较弱,但通过结合其他生成模型,如生成对抗网络和扩散模型,可以提高自编码器的生成能力。研究如何在保持自编码器简单稳定训练过程的同时,提升其生成质量,将是一个重要的方向。其次,自编码器在特征提取和压缩表征方面的优势为跨任务知识传输提供了基础。然而,如何在不同任务之间有效传输和共享知识仍是一个挑战。为了进一步推动持续学习的发展,未来的研究需要更深入地定义、有效地提取和高效地传输知识。更进一步地,自编码器在无监督学习中的优势明显,未来研究可以深入探索如何最大化利用自编码器以无监督或者半监督的方式进行持续学习。特别是,如何在缺乏样本标签的情况下,有效学习和传输知识,是一个值得深入研究的方向。

自编码器和持续学习的结合在多个领域已经展现出潜力,研究如何利用自编码器和持续学习范式解决实际问题,将具有广阔的前景。然而,当前大多数研究集中在单一模态的数据上,未来可以探索自编码器在多模态持续学习中的应用。例如,如何在图像、文本和语音等多种数据模态下,利用自编码器进行统一的特征提取和知识传输。在计算资源有限的应用环境中,自编码器的低计算资源需求和较短训练时间使其在持续学习范式下具有独特的优势。研究如何在资源受限的情况下,优化自编码器和持续学习的结合,将对实际应用具有重要意义。除此之外,未来还需要深入探索自编码器和持续学习结

合的理论基础。研究其数学模型、算法复杂度、收敛性和稳定性等,将为该领域的发展提供坚实的理论支持。通过进一步的研究和探索,自编码器与持

续学习的结合将不断推动人工智能的前沿发展,为构建更加智能和自主的学习系统提供新思路和方法。

表 5 不同应用领域的实验信息汇总

文献	应用领域	数据集	比较对象	评价指标	实验效果	编程语言	代码库
TsAE <sup>[40]</sup>	图像分类	CIFAR100、iLSVRC-small	Fine-tuning、LwF-MC、LwM	准确率	LwF-MC和LwM分别与TsAE结合后,能在iCIFAR100上分别提高6.3%和7%的Top-1准确率,并能在iLSVRC-small上分别提升6%和2%~4%的Top-5准确率。	/	/
TAMiL <sup>[38]</sup>	图像分类	Seq-CIFAR10、Seq-CIFAR100、Seq-TinyImageNet、Seq-core50	ER、FDR、DER++、Co2L、TARC、ER-ACE、CLS-ER1、DRI、PNN、CCLL、DEN、MNTDP、NISP、CLNP、PAE、Compacting、CPG、PackNet、LwF Sequential Fine Tuning、Coreset Sampling、Lower Bound、Continual VAE、AE+VAE、EWC、SI、LwF、AGEM、IMM、Generative Replay with VAE、ER、Regularization	准确率、遗忘率、预修误差	TAMiL在所有数据集上均优于现有的基于回放的基线方法,尤其在类增量场景中,TAMiL在Seq-TinyImageNet上的准确率比次优基线的准确率高出约10%。	PyTorch、Mammoth	<a href="https://github.com/NeurAI-Lab/TAMiL">https://github.com/NeurAI-Lab/TAMiL</a>
Flashcards <sup>[74]</sup>	图像分类	MNIST、Fashion MNIST、CIFAR10、SVHN、Omniglot	VAE、AE+VAE、EWC、SI、LwF、AGEM、IMM、Generative Replay with VAE、ER、Regularization	平均绝对误差、稳定性、可塑性、网络规模、存储容量	Flashcards在不增加额外内存和训练的情况下,取得与回放真实样本相当的性能,并且在某些情况下超过了生成式回放和正则化方法。	/	/
Generative Negative Replay <sup>[46]</sup>	图像分类	CORE50、ImageNet1000	No Replay、PR-OD、PR-GD、LwF、AR1	准确率	Generative Negative Replay在数据集CORE50和ImageNet-1000上的准确率尽管比回放真实样本的上界分别低2.39%与5.28%,却能在分类器、生成器和训练程序一致的情况下,比次优方法的分类准确率分别高出2.95%与0.83%。	/	/
Evolved Mixture Model <sup>[76]</sup>	图像分类	Split MNIST、Split CIFAR10、Split CIFAR100、Split MImageNet	Finetune、iCARL、CoPE、CURL、CNDPM、ER+GMED、ERa+GMED、MIR、MIR+GMED	准确率	Evolved Mixture Model在任务未知的场景中表现优异,特别是在Split MImageNet数据集上,取得的分类准确率超出了其他所有方法。	/	/



续表1

文献	应用领域	数据集	比较对象	评价指标	实验效果	编程语言	代码库
KRCL <sup>[7]</sup>	图像分类	Split MNIST、 Split FashionMNIST、 Split SVHN、 Split CIFAR10	Finetune、LWF、 IRCL、KRCL real	准确率、遗忘率	KRCL 在所有评估指标上的表现均优于 Finetune、LWF 和 IRCL 方法,尤其在缓解灾难性遗忘方面表现突出。此外,在使用真实样本参与训练时,KRCL 和 KRCL real 进一步提升了性能。	/	/
Smaller is Better <sup>[79]</sup>	图像分类	MNIST、CIFAR10、CIFAR100、ImageNet100、 Core50	AGEM、EWC、 ER、GEM、GSS、 ER-MIR、ASER、 ASER $\mu$ 、GDumb、 ELM、CutR、VAE	准确率	Smaller is Better 通过压缩图像来增加可存储样本量,在极端压缩图像尺寸的情况下,能取得比 Gdumb 的最终准确率高出 67% 的效果。Alqahtani 等人提出的模型在数据集 Merced 上获得了 94% 的最终准确率,这一结果比联合训练能达到的上界低 3%,但优于其他对比方法的效果。	Pytorch	<a href="https://github.com/francesco-p/smaller-is-better">https://github.com/francesco-p/smaller-is-better</a>
Alqahtani 等 <sup>[80]</sup>	图像分类	Merced	Joint Learning、 EWC、LwF	准确率	Robinson 等人通过模拟睡眠三种阶段对持续学习造成的影响,发现整合了全部阶段的模型在最大程度上提高准确率。	/	/
Robinson 等 <sup>[6]</sup>	图像分类	CIFAR100	Downscaling	准确率、KL 散度	cVAE <sup>[85]</sup> 通过生成重放策略缓解灾难性遗忘,并在允许存储真实样本的情况下进一步提高了性能,该方法在有、无监督的场景中都取得了有竞争力的结果。	/	/
cVAE <sup>[85]</sup>	图像分类	MNIST、FashionMNIST、 Flower、KM-NIST、 CIFAR10、CIFAR100	VAE、SI、EWC、 DGR、VCL、Open- VAE、CURL	准确率、负对数似然、训练时长	ODDL-S 在 Split MNIST、Split CIFAR10 和 Split CIFAR100 这三个数据集上,显示出优秀的持续学习能力和泛化性能,取得的准确率几乎比其他方法都好。	Pytorch	/
ODDL <sup>[48]</sup>	图像分类	Split MNIST、 Split CIFAR10、 Split CIFAR100、 Split MImageNet、 Permuted MNIST	finetune、GEM、 iCARL、reservoir、 MIR、GSS、CoPE、 ER+GMED、 ERa+GMED、 CURL、CNDPM、 Dynamic-OCM	准确率		/	<a href="https://dtuzi123.github.io/ODDL/">https://dtuzi123.github.io/ODDL/</a>

续表2

文献	应用领域	数据集	比较对象	评价指标	实验效果	编程语言	代码库
ADCN <sup>[14]</sup>	图像分类	MNIST、KM-NIST、CIFAR10、SEA、hyperplane、credit card default	DCN、AE+、KMEANS、SCARGC、STAM、LwF、SI	准确率、稳定性、可塑性	ADCN通过自适应网络结构和聚类机制,在处理数据流的非监督学习和非监督持续学习场景中,有效缓解了灾难性遗忘问题,并在多数情况下实现了更好的分类精度。	/	<a href="https://github.com/andriash001/AutonomousDCN.git">https://github.com/andriash001/AutonomousDCN.git</a>
KIERA <sup>[91]</sup>	图像分类	Permutted MNIST、Rotated MNIST、Split MNIST、Split CIFAR10	DCN、AE+、KMeans、STAM、LwF、SI	准确率、稳定性、可塑性	KIERA在数据集Rotated MNIST和Split CIFAR10上表现出色,取得了优于所有对比方法的准确率与稳定性。	/	<a href="https://researchdata.ntu.edu.sg/dataset.xhtml?persistentId=doi:10.21979/N9/P9DFJH">https://researchdata.ntu.edu.sg/dataset.xhtml?persistentId=doi:10.21979/N9/P9DFJH</a>
REMIND <sup>[87]</sup>	图像分类	ImageNet ILSVRC-2012、CORe50、TDIUC、CLEVR	Fine-Tuning、Exchange、SLDA、iCaRL、Unified、BiC、Offline	准确率、计算复杂度	REMIND在数据集ImageNet和CORE50上的分类准确率高于所有的比较方法,并在视觉问答任务中同样展现了优越的持续学习能力。	Pytorch	<a href="https://github.com/tyler-hayes/REMIND">https://github.com/tyler-hayes/REMIND</a>
ISYANA <sup>[49]</sup>	图像分类	splitMNIST、permutted MNIST、rotatedMNIST、omniglot	EWC、onlineEWC、SI、LWF、DGR、DGR+distill、A-GEM、XDG	准确率、稳定性、可塑性	ISYANA在permuttedMNIST上取得最好的准确率,并在前向迁移方面表现突出。然而,由于对类别依赖性的处理,ISYANA在splitMNIST上的表现不及A-GEM、EWC、o-EWC和SI。	/	<a href="https://github.com/ContinualAL/ISYANAKBS">https://github.com/ContinualAL/ISYANAKBS</a>
EPIE-Net <sup>[13]</sup>	图像分类	CIFAR10、CIFAR100、ImageNet-Sub	iCaRL、UCIR、BiC、WA、POD-Net、RPS-Net、DER、DyTox、RM、GDumb、EWC、Rwalk	准确率、遗忘率	EPIE-Net在CIFAR100的Blurry10-Online设置中,EPIE-Net的准确率达到60.31%,比次优结果高出8.94%。	/	/
CALCS-CL <sup>[58]</sup>	图像分类	CORe50、iCub-World28、STL-10	GDM with replay、Cumulative、LwF、EWC、Naive、SLDA、HPStream、CluStream、DenseNet、Xception、Inception、ResNet、MobileNet、AlexNet、VGG11、SqueezeNet	准确率、精度、召回率、F1	CALCS-CL在多任务和单增量任务场景下,相较于基线方法和其他最先进方法,在所有三个数据集上实现了更好的分类准确率、精度、召回率和F1。	TensorFlow	/

续表3

文献	应用领域	数据集	比较对象	评价指标	实验效果	编程语言	代码库
NDL <sup>[19]</sup>	图像分类	MNIST、CORe50、STEAD	EWC、LWF、Latent Replay	准确率、内存占用、训练时长	NDL与EWC和Latent Replay相比平均准确率略低,却在新类检测方面表现出色,并比LWF更具适应性优势。	/	/
Self-net <sup>[42]</sup>	图像分类	Permuted MNIST、Split MNIST、Split CIFAR10、Split CIFAR100、 Atari games、CORe50	EWC、Online EWC、Progressive Networks、Dynamically Expandable Networks、Context-Dependent Gating、Experience Replay、iCarl、Variational Continual Learning、Learning without Forgetting	准确率、余弦相似度、重构得分	Self-net在数据集Permuted MNIST、Split MNIST和CIFAR100上的性能与独立训练网络的性能相当,且在CIFAR100上取得的准确率超出了EWC等现有方法的表现20%以上。	/	<a href="https://github.com/jmandivarapu1/Self-Net-Lifelong-Learning-viaContinual-Self-Modeling">https://github.com/jmandivarapu1/Self-Net-Lifelong-Learning-viaContinual-Self-Modeling</a>
PRER <sup>[5]</sup>	图像分类	MNIST、KM-NIST、SVHN、CIFAR10	Naive、LWF、EWC、GEM、ER	准确率、遗忘率、内存占用、训练时长	PRER在四个数据集上都取得了缓解灾难性遗忘的最好表现,并且准确率几乎超过了所有的其他方法。此外,PRER在缩减内存占用上表现优异,尤其是在MNIST数据集上,PRER的内存占用不足LWF所需内存的1/10。	/	/
GIM <sup>[39]</sup>	图像分类、音频分类	MNIST、Devanagari、Audio-set	LSTM、LMN、EWC-LSTM、EWC-LMN、GIM-LSTM、GIM-LMN	准确率	GIM在大部分实验上都超越了LSTM、LMN及其相应的EWC版本,模型优势在数据集Devanagari上最为明显,LSTM、LMN在与GIM结合后,准确率分别领先次优方法2%与4%。	/	<a href="https://github.com/AndreaCossu/ContinualLearning-SequentialProcessing">https://github.com/AndreaCossu/ContinualLearning-SequentialProcessing</a>
Open-VAE <sup>[34]</sup>	图像分类、音频分类	MNIST、FashionMNIST、AudioMNIST、KMNIST、SVHN、CIFAR10、CIFAR100	EWC、DGR、VCL、VGR、Supervised VAE、PixelVAE、OpenPixelVAE、IntroVAE、OpenIntroVAE、Fine-tuning、MLP upper bound、WRN upper bound	准确率、负对数似然	OpenVAE在多个数据集上显示出更高的分类准确率和更低的负对数似然结果,其中,尤其是在MNIST上取得高于VGR等方法4.62%的分类准确率,并在负对数似然结果上取得有竞争力的结果。	Pytorch	/



续表4

文献	应用领域	数据集	比较对象	评价指标	实验效果	编程语言	代码库
PGMA <sup>[59]</sup>	图像分类、文本分析	MNIST、CIFAR10、DBPedia ontology、THUCNews	EWC、IMM、GR	准确率、训练时长、测试时长、内存占用	PGMA 在所有数据集上取得了最高的分类准确率,其中在数据集 CIFAR10 上的训练与测试时长之和不超过 GR 的 1.28 倍,却取得了高于次优方法 GR4% 的准确率。	/	<a href="https://github.com/morning-dews/PGMA">https://github.com/morning-dews/PGMA</a>
BooVAE <sup>[82]</sup>	图像分类、图像生成	MNIST、notMNIST、Fashion MNIST、CelebA	Standard VAE、EWC、VCL、MoG、Random Coreset、Multihead+EWC、Multihead+VCL	负对数似然、KL 散度、FID 距离	BooVAE 在生成样本的多样性和质量方面,在数据集 MNIST 和 CelebA 上的表现明显优于 EWC、VCL 等方法,取得了最好或者强有竞争力的效果。	/	<a href="https://github.com/AKuzina/BooVAE">https://github.com/AKuzina/BooVAE</a>
Bayesian Structural Adaptation <sup>[86]</sup>	图像分类、图像生成	SplitMNIST、Split notMNIST、Permuted MNIST、Split fashionMNIST、Split CIFAR100、MNIST、notMNIST	Pure Rehearsal、EwC、IMM、DEN、RCL、HAT、VCL、Naive、CURL	准确率、训练时长、对数似然	Bayesian Structural Adaptation 在监督学习的多重基准测试中,与对比方法相比展现出可比或更好的性能。在非监督的生成任务中,Bayesian Structural Adaptation 重用网络的子集并在每项任务中激活最小数量的节点,保证任务的高效执行。	/	<a href="https://github.com/npbcl/icml21">https://github.com/npbcl/icml21</a>
MoVE Layers <sup>[55]</sup>	图像分类、图像生成	MNIST、CIFAR10、CIFAR100、PyBullet	EWC、UCL、SAC、PPO	准确率、IS 分值	MoVE Layers 与离线模型相比还有一定差距,有监督场景下模型在 CIFAR100 上的准确率比离线模型的效果低 20% 左右,却能超过大部分对比方法。	/	<a href="https://github.com/hhnh/HVCL/">https://github.com/hhnh/HVCL/</a>
ECLA <sup>[68]</sup>	图像生成	Permuted-MNIST、USPS	FR	UMAP 可视化	ECLA 持续更新深度网络以提取抽象概念,比直接回放先前样本有更好的 UMAP 可视化效果。	/	/
BinPlay <sup>[70]</sup>	图像生成	MNIST、Fashion-MNIST、CIFAR10	GEM、iCARL、ER、ER-MIR、GEN-MIR、OCD-VAE、GR、GR+distill、RTF、AQM	准确率、内存占用	BinPlay 在所有三个数据集上的平均准确率均优于其他生成式重放方法,且在内存占用方面与对比方法持平甚至更小。	/	<a href="https://github.com/danielm1405/BinPlay">https://github.com/danielm1405/BinPlay</a>

续表5

文献	应用领域	数据集	比较对象	评价指标	实验效果	编程语言	代码库
LogCL <sup>[72]</sup>	图像生成	MNIST、Omniglot、CIFAR100	Generative Replay、Naive Generative Replay、Replay Buffer、BinPlay	准确率	LogCL 在数据集 MNIST 和 Omniglot 上的准确率高于所有生成回放式对比方法的结果，并在 CIFAR100 取得了高于次优方法 9.85% 的准确率。	/	<a href="https://github.com/e7mul/LogCL">https://github.com/e7mul/LogCL</a>
KIAM <sup>[48]</sup>	图像生成	MNIST、SVHN、Fashion、IFashion、RMNIST、CIFAR10、CelebA、3D-chair	Reservoir、Lifelong Teacher Student、Lifelong Generative Modelling、CN-DPM	FID 距离、IS 分值	KIAM 在不同的设定下均取得最优表现，在设定 MS-FIRC 下，模型在数据集 MNIST、Fashion、Ifashion、RMNIST 上的 FID 距离分别比次优结果减少 33.2、23.1、60.8、30.4。	/	<a href="https://github.com/dtuzi123/CGKD">https://github.com/dtuzi123/CGKD</a>
DEGM <sup>[47]</sup>	图像生成	Split MNIST、Split Fashion、COFMI、CelebA、CACD、3D-Chair、Omniglot、ImageNet、Car、Zap-Pos、CUB	ELBO-GR、IWELBO-GR-K'、LIMix、CN-DPM*、DEGM-2、BE、LGM、DEGM、CN-DPM	负对数似然、平方损失	DEGM 在数据集 Split MNIST、Split Fashion 和 COFMI 上的无监督持续学习表现超过了所有的对比方法。在 8 个数据集形成的序列任务中，DEGM 仅使用较少的参数，实现了为每个任务独立训练变分自编码器所能产生的效果。	/	<a href="https://github.com/dtuzi123/Expansion-Graph-Model">https://github.com/dtuzi123/Expansion-Graph-Model</a>
Brain-like Replay <sup>[81]</sup>	图像生成	CIFAR100	Joint、BI-R、LwF、EWC、SI、None、GR	准确率	Brain-like Replay 在类增量与任务增量场景中都优于现有方法，特别是在类增量场景中接近联合训练的上界性能。	/	<a href="https://github.com/GMvandeVen/brain-inspired-replay">https://github.com/GMvandeVen/brain-inspired-replay</a>
Lesort 等 <sup>[77]</sup>	图像生成	MNIST、Fashion MNIST、CIFAR10	VAEs、GANs、CVAEs、CGANs、WGANs、WGAN-GP、Finetuning、EWC、GR	准确率、FID 距离、类别兼容性	Lesort 等人探究哪种生成模型最适合图像生成任务，原始 GAN 结合生成重放策略能在数据集 MNIST 和 Fashion MNIST 上有效克服灾难性遗忘。然而，所有方法在 CIFAR10 上的表现并不理想，这表明生成模型在此类真实图像数据集上进行持续学习仍然是一个挑战。	/	<a href="https://github.com/TLESORT/Generative_Continual_Learning">https://github.com/TLESORT/Generative_Continual_Learning</a>

续表6

文献	应用领域	数据集	比较对象	评价指标	实验效果	编程语言	代码库
UGCL <sup>[57]</sup>	图像生成	MNIST、notMNIST、notMNIST+	EWC、LP、SI、VCL	负对数似然、分类器不确定性	UGVCL 在所有数据集上的整体性能上优于 VCL、EWC 和 SI, 特别是在数据集 notMNIST+MNIST 上的长任务序列上能保持稳定表现。	/	/
SRM <sup>[84]</sup>	图像生成	MNIST-Rotations、CIFAR10、CIFAR100、Omni-glot	GEM、Replay、EwC、LwF、iCaRL	准确率、存储空间	SRM 比其他先进方法如 GEM 和 Replay 更节约存储空间, 却能取得与回放真实样本相近的准确率。	/	/
LGM <sup>[22]</sup>	图像生成	MNIST、Fashion MNIST、Permuted MNIST、SVHN、Celeb-A	EWC、VCL、SI、Laplace propagation、full batch VAE、naive sequential VAE	负对数似然、负变分下界、Frechet 距离、图像重构	LGM 在数据集 FashionMNIST 和 MNIST 上取得有竞争力的负对数似然结果与 Freche 距离, 并能生成高质量样本。此外, LGM 在多样性方面的优势在 FashionMNIST 上得以验证, 任务所需的真实样本数量随时间推移能逐渐减少。	Pytorch	<a href="https://github.com/jramapuram/Lifelong-VAE_pytorch">https://github.com/jramapuram/Lifelong-VAE_pytorch</a>
DEJAVU <sup>[12]</sup>	图像生成	MNIST、SVHN、MNIST-M、SYN-D、USPS、PACS、DomainNet	CoTTA、AuCID、SHOT、GSFDA、BMD、Tent、T3A、L2D、PDEN、SNR、PCL、EFDN	TDG 分值、TDA 分值、遗忘率	DEJAVU 在不同数据集上取得优异表现, 所有情况下的泛化性相比次优方法的结果提升 3.1% 到 10.3%, 适应性与稳定性也在大多数情况下表现最佳, 或与最佳结果相近。	/	<a href="https://github.com/SonyAI/RaTP">https://github.com/SonyAI/RaTP</a>
ACS <sup>[17]</sup>	图像分割	2018 Medical Segmentation Decathlon challenge dataset、Scientific Data dataset、Alzheimer's Disease Neuroimaging Initiative dataset	U-Net block of ACS、Standard U-Net、OL-KD、BS-MAS	交并比、DSC 相似系数	ACS 在海马体分割任务上超越了所有的对比方法。	/	<a href="http://github.com/MECLabTUDA/ACS">http://github.com/MECLabTUDA/ACS</a>
CLIS <sup>[90]</sup>	图像分割	Cityscapes、MaStr1325、Melanoma	LwF、EWC、PackNet	交并比	CLIS 在大多数情况下优于对比的基线方法, 且能在确定边界方面表现更好。	/	<a href="https://github.com/ritmininglab/CLIS">https://github.com/ritmininglab/CLIS</a>



续表7

文献	应用领域	数据集	比较对象	评价指标	实验效果	编程语言	代码库
CISR <sup>[44]</sup>	图像分割	RUNMC、BMC、HCRUDB、UCL、BIDMC、HK	EWC、SI、MAS、FT	DSC相似系数、ASD距离	CISR在指标DSC相似系数和ASD距离上,超越了所有的对比方法。	/	<a href="https://github.com/jingy Zhang/CISR">https://github.com/jingy Zhang/CISR</a>
CURL <sup>[11]</sup>	图像聚类	MNIST、Omni-glot	VAE、DirVAE、SBVAE、VaDE、EWC、SI、MAS、LwF、GEM、DGR、iCARL	准确率、K近邻错误率	CURL在数据集MNIST和Omni-glot上取得无监督场景下的优越表现,在有监督的类增量设定下,仅被iCARL超越。	/	<a href="https://github.com/deepmind/deepmind-research/">https://github.com/deepmind/deepmind-research/</a>
Progressive Clustering <sup>[88]</sup>	图像聚类	MNIST、Fashion-MNIST	K-Means、SEC、SAE+k-means、CAE+k-means、DEC、IDEC、DEC-DA、DCEC	准确率、互信息、ARI指数	Progressive Clustering在数据集MNIST和Fashion-MNIST上以更少的重新聚类次数,取得了与当前最先进技术的持平结果。	/	/
F AE-based Task Detector <sup>[10]</sup>	图像修复、游戏模拟	MinAtar、CelebA、Atari	PNN、MoE、EWC、Task Replay、DQN	准确率、均方误差、F1	模型在MinAtar游戏模拟和图像修复任务中表现出色,取得了高于99%的准确率,在Atari游戏识别中的准确率高达99.96%。即使在引入噪声的情况下,模型准确率也能取得99.89%,相比次优方法提升6.7%。	PyTorch	<a href="https://github.com/arccosin/Task_Detector">https://github.com/arccosin/Task_Detector</a>
AQM <sup>[69]</sup>	图像压缩	CIFAR10、Imagenet、Kitti、Atari	iid online、iid offline、GEM、iCarl、fine-tuning、ER、ER-MIR、Gumbel AE、RS	准确率、遗忘率、均方根误差、F1	AQM在内存参数M=20、M=50时,在数据集CIFAR10上的准确率较ER能提升16.0%、13.9%,并在数据集Imagenet上的离线评估中取得最优表现,准确率较传统的回放方式RS高出18%。	/	/
mmDDL-DA <sup>[9]</sup>	文本分析	20News-groups、IMDB、RCV1、MNIST	LDA、DLDA、DocNADE、OR-Softmax、AVITM、FNN、SAVITM、MedLDA、wv-LSTM	准确率、主题一致性、测试时长	DDLDA及其变体在发现高度区分性的主题表示和实现更高分类准确率方面比现有的无监督和监督主题模型都更有效。	/	/

续表8

文献	应用领域	数据集	比较对象	评价指标	实验效果	编程语言	代码库
Wang 等 <sup>[78]</sup>	音频分类	ESC-10	Rehearsal, VAE, AE+GMM	准确率	Wang 等人仅使用4%的训练样本构建生成模型,该模型在数据集 ESC-10的准确率与添加20%先前样本进行回放的效果持平,并在整体上取得有竞争力的效果。	/	/
Sadhu 等 <sup>[75]</sup>	语音识别	WSJ,Reverb, Librispeech和 Chime4	Stand-alone ASR, Multi-condition ASR	错误率、稳定性、可塑性	Sadhu 等人提出的自动语音识别方法不再需要训练原始分类器的数据,却保证了模型在先前任务上的性能,并在域改变的情况下实现与离线训练方法持平的效果。	PyTorch, Kaldi	/
LGR <sup>[45]</sup>	表情识别	CK+,RAF-DB,AffectNet, MNIST,CIFAR10	Naive Rehearsal, Latent Replay, DGR,DGR with Distillation	准确率、内存占用、训练时长、CPU、GPU使用率	LGR以较低的内存和资源消耗在数据集 CK+、RAF-DB和 AffectNet上取得有竞争力的结果,并在数据集 MNIST和 CIFAR10上显示出了较好的性能和资源效率。	PyTorch	<a href="https://github.com/samuilstoychev/latent-generative-replay">https://github.com/samuilstoychev/latent-generative-replay</a>
Clifer <sup>[37]</sup>	表情识别	iCV-MEFED, RAVDESS, MMI,BAUM-1	MLP,GDM, GDM+Replay	准确率、F1	在加入Replay机制和CLIFER框架后,GDM借助情景记忆与语义记忆,在数据集RAVDESS上分别将F1分数提升至0.98与0.75。	/	/
Huang 等 <sup>[89]</sup>	身份识别	Market-1501, DukeMTMC-reID,CUHK03	No Forgetting, Fine-tuning, LwF, EWC, Encoder Lifelong, MAS, L2-SP, DELTA (without ATT), E-EWC	准确率、精度	EC和 Auto-Weighted EC 在新、旧任务之间取得平衡,在准确度和精度上表现优于其他方法。	Pytorch	/
RCR <sup>[73]</sup>	点云分类	ModelNet 10, ModelNet 40, ShapeNet, 以及 MNIST 的点云版本	Fine-tune, Incremental, LwF, EWC, iCaRL, I3 DOL, RCR	准确率	RCR在多个数据集上展现出与监督学习相似的分类准确率,有效缓解了灾难性遗忘问题,尤其在任务数量增多时,相比其他方法显示出更好的性能。	Pytorch	/

续表9

文献	应用领域	数据集	比较对象	评价指标	实验效果	编程语言	代码库
ABC <sup>[52]</sup>	点云生成	ShapeNet	EWC、PNN、ABC、PackNet、SupSup	准确率、Jensen-Shannon 散度、覆盖率、最大平均差异、最小匹配距离、Chamfer 距离、EMD 距离	经典的参数隔离技术如 PNN 在单增量任务场景中表现良好，但 ABC 在处理任务重叠和新旧概念混合时更为有效，能够生成更接近原始类别的点云对象。	/	/
MetaCL <sup>[36]</sup>	生理信号处理	MIT-BIH-AR、Wrist PPG During Exercise、INCART、Sleep-EDF Expanded	SVM、CNN、LSTM、Bi-LSTM	准确率、精度、召回率、F1、遗忘率	MetaCL 在多个数据集上展现了优越的分类性能，与其他方法相比遗忘率更低。	Pytorch	/
CoCo <sup>[71]</sup>	病变诊断	Breast Cancer Histology Dataset、Lung+Colon Cancer Histology Dataset	Contrastive Predictive Coding、Contrastive Multiview Coding、end-to-end、memory bank、MoCo	准确率	CoCo 在肺癌、乳腺癌、结肠癌三种组织学癌症数据集上均取得了最优表现，其准确率分别高于次优对比方法 3.7%、2.8% 和 1.5%。	/	/
CLELNet <sup>[35]</sup>	病变诊断	食管内镜图像数据集，收集自 2016 年至 2019 年间澳门镜湖医院的患者	AlexNet、VGG-16、ResNet-18、DenseNet-121、Inception-v3、RegNetX、FCN、UNet、SegNet、UNet++、HRNet、DDRNet	准确率、精度、灵敏度、特异度、负预测率、AUC、交并比、Dice 相似系数	CLELNet 在分类任务中表现优于对比方法，尤其在准确率、灵敏度、特异度和负预测率这四项指标上略微超过 Inception-v3，并取得能优于四位内镜医师平均水平的诊断性能。在分割任务中，CLELNet 在交并比和 Dice 相似度这两项指标上优于其他分割模型。	Pytorch	/
Jha 等 <sup>[24]</sup>	病理分析	BioGRID、PubTator、Disease-Symptom	DeepWalk、LINE、Node2Vec、SDNE、Metapath2Vec、DynGEM、BiNE	精度、F1、AUC	Jha 等人以在线方式更新自编码器捕获的邻近关系节点表示，在网络重建、链接预测和关联推荐这三项任务上均取得了整体最优的表现。	/	/
Li 等 <sup>[18]</sup>	医疗图像分类	Medmnist、Organmnist	DGR、EWC、Fine-tuning、Offline	准确率、遗忘率、IS 分值、FID 距离	Li 等人在任务增量和类增量的持续学习任务中表现突出，相较于对比方法能取得更高的平均准确率和更低的平均遗忘率。Li 等人提出的模型与 DGR 相比，IS 分值高出 5.7，FID 距离减少 1312.4。	/	/

续表10

文献	应用领域	数据集	比较对象	评价指标	实验效果	编程语言	代码库
Ma等 <sup>[25]</sup>	交通流预测	INTERACTION, InD, Round	LSTM, GNN, S-GAN, Trajec-tron++	平均位移误差、最终位移误差、平均误差、遗忘率	Ma等人提出的方法在所有场景中取得比基线方法更低的平均误差。 CLTP-MAN 在处理长期轨迹预测方面表现优异,通过稀疏经验回放策略在不同场景下具备泛化能力。 RAG 受益于域专属的神经单元,在 DrivingStereo 上实现了最佳性能,甚至在 DrivingStereo 和 Virtual KITTI 上超越了借助所有数据的 Joint Training 的效果。	/	/
CLTP-MAN <sup>[23]</sup>	交通流预测	ETH, UCY, SDD	SGAN, STGAT, MANTRA, Social-STGCNN, GTPPO, SGCN	平均位移误差、最终位移误差、平均误差、遗忘率	CLTP-MAN 在处理长期轨迹预测方面表现优异,通过稀疏经验回放策略在不同场景下具备泛化能力。 RAG 受益于域专属的神经单元,在 DrivingStereo 上实现了最佳性能,甚至在 DrivingStereo 和 Virtual KITTI 上超越了借助所有数据的 Joint Training 的效果。	Pytorch	/
RAG <sup>[16]</sup>	交通视差预测	DrivingStereo, KITTI raw, Virtual KITTI	Incremental Finetuning, EWC, iCaRL, Expert Gate, Learn to Grow, Joint Training	错误率、平均重用率	CLTP-MAN 在处理长期轨迹预测方面表现优异,通过稀疏经验回放策略在不同场景下具备泛化能力。 RAG 受益于域专属的神经单元,在 DrivingStereo 上实现了最佳性能,甚至在 DrivingStereo 和 Virtual KITTI 上超越了借助所有数据的 Joint Training 的效果。	PyTorch	<a href="https://github.com/chzhang18/RAG">https://github.com/chzhang18/RAG</a>
Generating PEs <sup>[21]</sup>	结构化样本分类	合成数据集、Hypothyroid	STL, MTL	准确率、散度评估、最大平均差异	Generating PEs 在维持先前任务知识方面的表现,能与依赖真实样本产生的表现相当,甚至在某些情况下比随机生成样本的效果更好。	/	/
L-MVAE <sup>[53]</sup>	图网络生成	MNIST, Fashion, SVHN, CIFAR10, CelebA, CACD, 3D-chairs, Omniglot	DGR, LGM, CURL, BatchEnsemble	准确率、负对数似然、IS 分值、均方误差、结构一致性相似因子、峰值信噪比	L-MVAE 在保持之前任务性能的同时,能快速学习新任务,与对比方法相比,具有较低的遗忘率和较好的新任务适应能力。	/	<a href="https://github.com/dtuzi123/Lifelong-MixtureVAEs">https://github.com/dtuzi123/Lifelong-MixtureVAEs</a>
NICAG <sup>[64]</sup>	智能体行为研究	YCB dataset, EGAD!	EWC, SI, ER, ER-RM, Fine-tune, IID-Offline	测试损失、行动成功率、遗忘率、穿透深度、穿透体积	NICAG 及其变体在抓取对象的成功率和穿透深度指标上优于对比方法,显示出其在非平稳数据流下的强大适应能力。此外, NICAG 缓解遗忘的表现,在缓冲区设定为 1K 时最为优异,相对于 EWC, SI, ER, ER-RM 能分别降低 90.1%、87.4%、90.8%、40.7% 的遗忘率。	Pytorch, Avallanche, MuJoCo	<a href="https://github.com/Wanyili/NICAG">https://github.com/Wanyili/NICAG</a>



续表 11

文献	应用领域	数据集	比较对象	评价指标	实验效果	编程语言	代码库
BGPO <sup>[20]</sup>	智能体行为研究	机器人操作任务的模拟环境 CopeliaSim 中生成的演示数据	MAML、Dist-NLL	总外部奖励	BGPO 相较于对比方法在测试任务中获得的奖励更高且更加稳定，最终实现超过 80% 的成功率。	TensorFlow	/
S-TRIGGER <sup>[83]</sup>	智能体环境识别	Flatland 环境中构建的 2D 实验环境，包括固定障碍和可食用物品的房间，以及随机生成的迷宫序列	Fine-tuning、Source Only、Upperbound	均方误差、标准化平均奖励	S-TRIGGER 在持续学习场景中能有效避免灾难性遗忘，尤其是在环境变化时，效果领先于 Fine-tuning 方法。	/	/
GR-IG <sup>[8]</sup>	传感器信号传输	DSADS、HAR、Turbofan Engine	Task-Specific、Fine-tune、Common Generator、Supermasks with Transfer、Upper Bound	错误率、均方根误差	GR-IG 在处理多变量时间序列任务时，无论是固定输入维度还是可变输入维度，均优于对比方法的效果。	/	/
Tf-GCZSL <sup>[56]</sup>	地理遥感	CUB、aPY、AWA1、AWA2、SUN	Offline、Seq-Tf-GCZSL、AGEM+ CZSL、EWC+ CZSL、MAS+ CZSL、GRCZSL、CZSL-CV+res	准确率、调和准确率	Tf-GCZSL 能接近离线联合训练这一上界，表现出较小的灾难性遗忘。	/	<a href="https://github.com/ChandanITI/Tf-GCZSL">https://github.com/ChandanITI/Tf-GCZSL</a>
Rahman 等 <sup>[51]</sup>	软件分类	Drebin、EMBER	EWC、EWC Online、SI、LwF、GR、GR with Distillation、RtF、BI-R、ER、A-GEM、iCaRL	准确率、AUC	Rahman 等人在域增量、类增量与任务增量的场景中对比了 11 种持续学习技术的表现，iCaRL 取得的准确率在整体上最为接近联合重放的效果。	PyTorch	<a href="https://github.com/msrocean/continual-learning-malware">https://github.com/msrocean/continual-learning-malware</a>

## 参 考 文 献

- [1] Hadsell R, Rao D, Rusu A A, et al. Embracing change: Continual learning in deep neural networks. *Trends in Cognitive Sciences*, 2020, 24(12): 1028-1040
- [2] Li P Z, Pei Y, Li J Q. A comprehensive survey on design and application of autoencoder in deep learning. *Applied Soft Computing*, 2023, 138: 110176
- [3] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors. *Nature*, 1986, 323(6088): 533-536
- [4] Khatib A E, Karray F. Toward durable representations for continual learning. *Advances in Computational Intelligence*, 2022, 2(1): 7
- [5] Scardapane S, Uncini A. Pseudo-Rehearsal for continual learning with normalizing flows//Proceedings of the 4th Lifelong Machine Learning Workshop. Vienna, Austria, 2020
- [6] Robinson B S, Lau C W, New A, Nichols S M, et al. Continual learning benefits from multiple sleep stages: NREM, REM, and Synaptic Downscaling//Proceedings of the International Joint Conference on Neural Networks. Padua, Italy, 2022: 1-9
- [7] Chen X Y, Liu J W, Li W T. Continual learning based on knowledge distillation and representation learning//Proceedings of the 31st International Conference on Artificial Neural Networks. Bristol, UK, 2022: 27-38
- [8] Gupta V, Narwariya J, Malhotra P, et al. Continual learning for multivariate time series tasks with variable input dimensions//Proceedings of the International Conference on Data Mining. Auckland, New Zealand, 2021: 161-170
- [9] Chen W C, Chen B, Liu Y Q, et al. Max-margin deep diverse latent Dirichlet allocation with continual learning. *IEEE*

- Transactions on Cybernetics, 2021, 52(7): 5639-5653
- [10] Jacobson M J, Wright C Q, Jiang N, et al. Task detection in continual learning via familiarity autoencoders//Proceedings of the International Conference on Systems. Prague, Czech Republic, 2022: 1-8
- [11] Rao D, Visin F, Rusu A, et al. Continual unsupervised representation learning//Proceedings of the Advances in Neural Information Processing Systems. Vancouver, Canada, 2019: 7647-7657
- [12] Liu C X, Wang L X, Lyu L J, et al. DEJA VU: Continual model generalization for unseen domains//Proceedings of the 11th International Conference on Learning Representations. Kigali, Rwanda, 2023
- [13] Du F, Yang Y, Zhao Z Y, et al. Efficient perturbation inference and expandable network for continual learning. Neural Networks, 2023, 159: 97-106
- [14] Ashfahani A, Pratama M. Unsupervised continual learning in streaming environments. IEEE Transactions on Neural Networks and Learning Systems, 2022, 34(12): 9992-10003
- [15] Guo F Z, Chen Z J, Xiao F, et al. Real-time energy performance benchmarking of electric vehicle air conditioning systems using adaptive neural network and Gaussian process regression. Applied Thermal Engineering, 2023, 222 (34) : 119931
- [16] Zhang C H, Tian K, Fan B, et al. Continual stereo matching of continuous driving scenes with growing architecture//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA, 2022: 18901-18910
- [17] Memmel M, Gonzalez C, Mukhopadhyay A. Adversarial continual learning for multi-domain hippocampal segmentation//Proceedings of the Domain Adaptation and Representation Transfer, and Affordable Healthcare and AI for Resource Diverse Global Health. Strasbourg, France, 2021: 35-45
- [18] Li X J, Li H F, Ma L. Continual learning of medical image classification based on feature replay//Proceedings of the International Conference on Signal Processing. Beijing, China, 2022: 426-430
- [19] Harris J, Kinkead S, Fox D, et al. Continual learning for pattern recognizers using neurogenesis deep learning. Albuquerque, USA: Sandia National Lab. Technical report: No. SAND2021-11382R, 2022
- [20] Hafez M B, Wermter S. Behavior self-organization supports task inference for continual robot learning//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Prague, Czech Republic, 2021: 6739-6746
- [21] Silver D L, Mahfuz S. Generating accurate pseudo examples for continual learning//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle, USA, 2020: 256-257
- [22] Ramapuram J, Gregorova M, Kalousis A. Lifelong generative modeling. Neurocomputing, 2020, 404: 381-400
- [23] Yang B, Fan F C, Ni R, et al. Continual learning-based trajectory prediction with memory augmented networks. Knowledge-Based Systems, 2022, 258: 110022
- [24] Jha K, Xun G X, Zhang A D. Continual representation learning for evolving biomedical bipartite networks. Bioinformatics, 2021, 37(15): 2190-2197
- [25] Ma H B, Sun Y F, Li J C, et al. Continual multi-agent interaction behavior prediction with conditional generative memory. IEEE Robotics and Automation Letters, 2021, 6(4): 8410-8417
- [26] Berahmand K, Daneshfar F, Salehi E S, et al. Autoencoders and their applications in machine learning: A survey. Artificial Intelligence Review, 2024, 57(2): 28
- [27] Janjua J, Patankar A, Shetty M, et al. Experimental comparison of autoencoder variants in content-based image retrieval//Proceedings of the 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). Delhi, India, 2023: 1-6
- [28] Parisi G I, Kemker R, Part J L, et al. Continual lifelong learning with neural networks: A review. Neural networks, 2019, 113: 54-71
- [29] Lomonaco V, Pellegrini L, Rodriguez P, et al. CVPR 2020 continual learning in computer vision competition: Approaches, results, current challenges and future directions. Artificial Intelligence, 2020, 303: 103635
- [30] Mundt M, Hong Y W, Pliushch I, et al. A wholistic view of continual learning with deep neural networks: Forgotten lessons and the bridge to active and open world learning. Neural Networks, 2023, 160: 306-336
- [31] Lesort T, Lomonaco V, Stoian A, et al. Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges. Information Fusion, 2020, 58: 52-68
- [32] Lange M D, Aljundi R, Masana M, et al. A continual learning survey: Defying forgetting in classification tasks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(7): 3366-3385
- [33] Wang L Y, Zhang X X, Su H, et al. A comprehensive survey of continual learning: Theory, method and application. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(8): 5362-5383
- [34] Mundt M, Pliushch I, Majumder S, et al. Unified probabilistic deep continual learning through generative replay and open set recognition. Journal of Imaging, 2022, 8(4): 93
- [35] Tang S G, Yu X Y, Cheang C F, et al. CLELNet: A continual learning network for esophageal lesion analysis on endoscopic images. Computer Methods and Programs in Biomedicine, 2023, 231: 107399
- [36] Sun L, Wu J, Xu Y, et al. A federated learning and blockchain framework for physiological signal classification based on continual learning. Information Sciences, 2023, 630: 586-598.
- [37] Churamani N, Gunes H. Clifer: Continual learning with imagination for facial expression recognition//Proceedings of the International Conference on Automatic Face and Gesture Recognition. Buenos Aires, Argentina, 2020: 322-328
- [38] Bhat P, Zonooz B, Arani E. Task-Aware information routing from common representation space in lifelong learning//

- Proceedings of the International Conference on Learning Representations. Kigali, Rwanda, 2022
- [39] Cossu A, Carta A, Bacciu D. Continual learning with gated incremental memories for sequential data processing// Proceedings of the International Joint Conference on Neural Networks. Glasgow, UK, 2020: 1-8
- [40] Kim E S, Kim J U, Lee S, et al. Class incremental learning with task-selection//Proceedings of the International Conference on Image Processing. Abu Dhabi, United Arab Emirates, 2020: 1846-1850
- [41] Ostapenko O, Rodriguez P, Caccia M, et al. Continual learning via local module composition//Proceedings of the Advances in Neural Information Processing Systems. New York, USA, 2021: 30298-30312
- [42] Mandivarapu J K, Camp B, Estrada R. Self-net: Lifelong learning via continual self-modeling. *Frontiers in Artificial Intelligence*, 2020, 3: 19
- [43] Ye F, Bors A G. Task-Free continual learning via online discrepancy distance learning//Proceedings of the Advances in Neural Information Processing Systems. New Orleans, USA, 2022: 23675-23688
- [44] Zhang J Y, Gu R, Wang G T, et al. Comprehensive importance-based selective regularization for continual segmentation across multiple sites//Proceedings of the 24th International Conference Medical Image Computing and Computer Assisted Intervention. Strasbourg, France, 2021: 389-399
- [45] Stoychev S, Churamani N, Gunes H. Latent generative replay for resource-efficient continual learning of facial expressions// Proceedings of the 17th International Conference on Automatic Face and Gesture Recognition. Waikoloa Beach, USA, 2023: 1-8
- [46] Graietti G, Maltoni D, Pellegrini L, et al. Generative negative replay for continual learning. *Neural Networks*, 2023, 162: 369-383
- [47] Ye F, Bors A G. Lifelong generative modelling using dynamic expansion graph model//Proceedings of the AAAI Conference on Artificial Intelligence. New Orleans, USA, 2022: 8857-8865
- [48] Ye F, Bors A G. Continual variational autoencoder via continual generative knowledge distillation//Proceedings of the AAAI Conference on Artificial Intelligence. Washington, USA, 2023: 10918-10926
- [49] Mao F B, Weng W W, Pratama M, et al. Continual learning via inter-task synaptic mapping. *Knowledge-Based Systems*, 2021, 222: 106947
- [50] Maschler B, Pham T T H, Weyrich M. Regularization-based continual learning for anomaly detection in discrete manufacturing. *Procedia CIRP*, 2021, 104: 452-457
- [51] Rahman M S, Coull S, Wright M. On the limitations of continual learning for malware classification//Proceedings of the Conference on Lifelong Learning Agents. Québec, Canada, 2022: 564-582
- [52] Sadowski M, Piczak K J, Spurek P, et al. Continual learning of 3D point cloud generators//Proceedings of the Conference on Neural Information Processing. Sanur, Indonesia, 2021: 330-341
- [53] Ye F, Bors A G. Lifelong mixture of variational autoencoders. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 34(1): 461-474
- [54] Ye F, Bors A G. Lifelong teacher-student network learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 44(10): 6280-6296
- [55] Hihn H, Braun D A. Hierarchically structured task-agnostic continual learning. *Machine Learning*, 2023, 112(2): 655 - 686
- [56] Gautam C, Parameswaran S, Mishra A, et al. Tf-gcysl: Task-free generalized continual zero-shot learning. *Neural Networks*, 2022, 155: 487-497
- [57] Liu G M, Guo Y, Yin C W S, et al. Unsupervised generative variational continual learning//Proceedings of the International Conference on Image Processing. Bordeaux, France, 2022: 4028-4032
- [58] Irfan M, Zheng J B, Iqbal M, et al. Knowledge extraction and retention-based continual learning by using convolutional autoencoder-based learning classifier system. *Information Sciences*, 2022, 591: 287-305
- [59] Hu W P, Zhou L, Liu B, et al. Overcoming catastrophic forgetting for continual learning via model adaptation// Proceedings of the International Conference on Learning Representations. New Orleans, USA, 2019
- [60] Banayeeanzade M, Mirzaiezhadeh R, Hasani H, et al. Generative vs. discriminative: Rethinking the meta-continual learning//Proceedings of the Advances in Neural Information Processing Systems. 2021: 21592-21604
- [61] Li Y J, Yang Y X, Gao Q, et al. Cross-regional fraud detection via continual learning (student abstract)//Proceedings of the AAAI Conference on Artificial Intelligence. Washington, USA, 2023: 16260-16261
- [62] Yang X, Li Y J, Meng D, et al. Three-way multi-granularity learning towards open topic classification. *Information Sciences*, 2022, 585: 41-57
- [63] Li M M, Zhu J Q, Yang X, et al. Cl-wstc: Continual learning for weakly supervised text classification on the internet// Proceedings of the ACM Web Conference. Austin, USA, 2023: 1489-1499
- [64] Li W Y, Wei W, Wang P. Neuro-inspired continual anthropomorphic grasping. *Iscience*, 2023, 26(6)
- [65] Ye F, Bors A G. Learning latent representations across multiple data domains using lifelong VAEGAN//Proceedings of the Computer Vision-ECCV 2020: 16th European Conference. Glasgow, UK, 2020: 777-795
- [66] Ye F, Bors A G. Dynamic scalable self-attention ensemble for task-free continual learning//Proceedings of the International Conference on Acoustics, Speech and Signal Processing. Rhodes Island, Greece, 2023: 1-5
- [67] Campo D, Slavic G, Baydoun M, et al. Continual learning of predictive models in video sequences via variational autoencoders//Proceedings of the International Conference on

- Image Processing. Abu Dhabi, United Arab Emirates, 2020: 753-757
- [68] Rostami M, Kolouri S, Pilly P, et al. Generative continual concept learning//Proceedings of the AAAI Conference on Artificial Intelligence. New York, USA, 2020: 5545-5552
- [69] Caccia L, Belilovsky E, Caccia M, et al. Online learned continual compression with adaptive quantization modules//Proceedings of the 37th International Conference on Machine Learning. Vienna, Austria, 2020: 1240-1250
- [70] Deja K, Wawrzyński P, Marczak D, et al. BinPlay: A binary latent autoencoder for generative replay continual learning//Proceedings of the International Joint Conference on Neural Networks. Shenzhen, China, 2021: 1-8
- [71] Chakraborti T, Gleeson F, Rittscher J. Contrastive representations for continual learning of fine-grained histology images//Proceedings of the Machine Learning in Medical Imaging: 12th International Workshop. Strasbourg, France, 2021: 1-9
- [72] Masarczyk W, Wawrzyński P, Marczak D, et al. Logarithmic continual learning. *IEEE Access*, 2022, 10: 117001-117010
- [73] Zamorski M, Stypułkowski M, Karanowski K, et al. Continual learning on 3D point clouds with random compressed rehearsal. *Computer Vision and Image Understanding*, 2023, 228: 103621
- [74] Gopalakrishnan S, Singh P R, Fayek H, et al. Knowledge capture and replay for continual learning//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. Waikoloa, USA, 2022: 10-18
- [75] Sadhu S, Hermansky H. Continual learning in automatic speech recognition//Proceedings of the Interspeech. Shanghai, China, 2020: 1246-1250
- [76] Ye F, Bors A G. Learning an evolved mixture model for task-free continual learning//Proceedings of the International Conference on Image Processing. Bordeaux, France, 2022: 1936-1940
- [77] Lesort T, Caselles-Dupré H, Garcia-Ortiz M, et al. Generative models from the perspective of continual learning//Proceedings of the International Joint Conference on Neural Networks. Budapest, Hungary, 2018: 1-8
- [78] Wang Z P, Subakan C, Tzinis E, et al. Continual learning of new sound classes using generative replay//Proceedings of the Workshop on Applications of Signal Processing to Audio and Acoustics. New Paltz, USA, 2019: 308-312
- [79] Pelosin F, Torsello A. Smaller is better: An analysis of instance quantity/quality trade-off in rehearsal-based continual learning//Proceedings of the International Joint Conference on Neural Networks. Padua, Italy, 2022: 1-8
- [80] Alqahtani A, Ammour N. Continual learning using efficientnet and data generation for remote sensing image classification//Proceedings of the 3rd International Conference on Distributed Sensing and Intelligent Systems. Sharjah, United Arab Emirates, 2022: 222-228
- [81] Ven G M V, Siegelmann H T, Tolia A S. Brain-like replay for continual learning with artificial neural networks//Proceedings of the International Conference on Learning Representations. Addis Ababa, Ethiopia, 2020: 47
- [82] Egorov E, Kuzina A, Burnaev E. BooVAE: Boosting approach for continual learning of VAE//Proceedings of the Advances in Neural Information Processing Systems. 2021: 17889-17901
- [83] Caselles-Dupre H, Garcia-Ortiz M O. S-TRIGGER: Continual state representation learning via self-triggered generative replay//Proceedings of the International Joint Conference on Neural Networks. Shenzhen, China, 2021: 1-7
- [84] Riemer M, Klinger T, Bouneffouf D, et al. Scalable recollections for continual lifelong learning//Proceedings of the AAAI Conference on Artificial Intelligence. Honolulu, USA, 2019: 1352-1359
- [85] Hong Y, Mundt M, Park S, et al. Return of the normal distribution: Flexible deep continual learning with variational auto-encoders. *Neural Networks*, 2022, 154: 397-412
- [86] Kumar A, Chatterjee S, Rai P. Bayesian structural adaptation for continual learning//Proceedings of the International Conference on Machine Learning. Vienna, Austria, 2021: 5850-5860
- [87] Hayes T L, Kafle K, Shrestha R, et al. Remind your neural network to prevent catastrophic forgetting//Proceedings of the European Conference on Computer Vision. Glasgow, UK, 2020: 466-483
- [88] Gunari A, Kudari S V, Tabib R A, et al. Progressive clustering: An unsupervised approach towards continual knowledge acquisition of incremental data//Proceedings of the International Conference Pattern Recognition and Artificial Intelligence. Paris, France, 2022: 355-367
- [89] Huang T J, Qu W W, Zhang J G. Continual representation learning via auto-weighted latent embeddings on person ReID//Proceedings of the 4th Chinese Conference Pattern Recognition and Computer Vision. Beijing, China, 2021: 593-605
- [90] Zheng E, Yu Q, Li R, et al. A continual learning framework for uncertainty-aware interactive image segmentation//Proceedings of the AAAI Conference on Artificial Intelligence. Vancouver, Canada, 2021: 6030-6038
- [91] Pratama M, Ashfahani A, Lughofer E. Unsupervised continual learning via self-adaptive deep clustering approach//Proceedings of the 1st International Continual Semi-Supervised Learning Workshop. 2022: 48-61
- [92] Zhang W X, Li D Q, Ma C, et al. Continual learning for blind image quality assessment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45(3): 2864-2878
- [93] Ghosh S. Adversarial training of variational auto-encoders for continual zero-shot learning (A-CZSL)//Proceedings of the International Joint Conference on Neural Networks. Shenzhen, China, 2021: 1-8
- [94] Sokar G, Mocanu D C, Pechenizkiy M. Learning invariant representation for continual learning//Proceedings of the AAAI-21 International Workshop on Meta-Learning for Computer Vision. Vancouver, Canada, 2021
- [95] Ye F, Bors A G. Wasserstein expansible variational autoencoder for discriminative and generative continual learning//Proceedings of the IEEE/CVF International Conference on Computer Vision. Paris, France, 2023: 18619-18629



- [96] Zhang H, Chen B, Guo D D, et al. WHAI: Weibull hybrid autoencoding inference for deep topic modeling//Proceedings of the International Conference on Learning Representations. Vancouver, Canada, 2018
- [97] Wei H, Hu G Y, Bai W, et al. Lifelong representation learning in dynamic attributed networks. *Neurocomputing*, 2019, 358: 1-9
- [98] Yang X, Yu H, Gao X, et al. Federated Continual Learning via Knowledge Fusion: A Survey. *IEEE Transactions on Knowledge and Data Engineering*, 2024, 36(8): 3832-3850



**WU Mei-Jun**, Ph. D. candidate. Her main research interests include continual learning, intelligent decision-making, and financial risk prediction.

**YANG Xin**, Ph. D., professor, Ph. D. supervisor. His main research interests include continual learning, trust-worthy federal learning, and multi-granular cognitive computation.

**PAN Chao-Fan**, Ph. D. candidate. His main research interests include continual learning and reinforcement learning.

**LI Tian-Rui**, Ph. D., professor, Ph. D. supervisor. His main research interests include data mining and knowledge discovery, artificial intelligence, granularity computing, and rough set.

**KOU Gang**, Ph. D., professor, Changjiang Distinguished Professor, Ph. D. supervisor. His main research interests include big data and decision-making, business intelligence, information systems, credit score, and emergency management.

## Background

This review refers to deep learning and continual learning in the field of artificial intelligence. Specifically, it focuses on the integration of autoencoders with continual learning, aiming to explore the applications, challenges, and prospects of this combination.

Currently, research on deep learning and continual learning has been gradually deepening internationally. Deep learning technologies have made significant strides in fields such as image recognition, speech recognition, and natural language processing, and are widely applied in practical scenarios. Continual learning, as a learning paradigm, has also attracted increasing attention in the field of deep learning. Although there has been considerable progress in both areas, research on the specific problem of integrating autoencoders with continual learning may still be in the exploratory and developmental stages. International research has proposed some methods and frameworks, but there are still challenges and unresolved issues in addressing practical problems and promoting technological applications. Therefore, there is considerable room for development and potential in research on this specific problem.

This paper reviews the relevant research over the past five years, summarizing the types and characteristics of

autoencoders, the main challenges of continual learning, and the incremental scenarios. Additionally, it offers a detailed overview of their applications in different industries. Finally, this review summarizes the advantages, limitations, and prospects, aiming to provide valuable insights for advancing research on developing continual learning and autoencoders.

Our group has been dedicated to research on continual learning, and we have published related studies in renowned conferences and journals such as AAAI (Cross-regional Fraud Detection via Continual Learning, Learning to Prompt Knowledge Transfer for Open-World Continual Learning), WWW (CL-WSTC: Continual Learning for Weakly Supervised Text Classification on the Internet), ACM (Debunking free fusion myth: Online multi-view anomaly detection with disentangled product-of-experts modeling) and *IEEE Transactions on Knowledge & Data Engineering* (Federated Continual Learning via Knowledge Fusion: A Survey).

The following funds supported this work: the National Natural Science Foundation of China (Nos. 62476228, 71910107002 and 62406259), the Science and Technology Innovation Program of Hunan Province (No. 2024RC4008), and the Sichuan Science and Technology Program (No. 2024ZYD0180). These funds provided important financial support and research resources.