工程科学学报, 第 45 卷, 第 7 期: 1225−1231, 2023 年 7 月
Chinese Journal of Engineering, Vol. 45, No. 7: 1225−1231, July 2023
https://doi.org/10.13374/j.issn2095-9389.2022.05.03.006; http://cje.ustb.edu.cn

基于参数惩罚和经验回放的材料吸声系数回归增量学习

王弘业[1)], 钱　权[1,2)], 武　星[1,2)]✉

1) 上海大学计算机工程与科学学院, 上海 200444　2) 之江实验室, 杭州 311100
✉通信作者, E-mail: xingwu@shu.edu.cn

摘　要　材料数据具有分批次、分阶段制备的特点, 并且不同批次数据的分布也不同, 而神经网络按批次学习材料数据时会存在平均准确率随批次下降的问题, 这为人工智能应用于材料领域带来极大的挑战. 为解决这个问题, 将增量学习应用于材料数据的学习上, 通过分析模型参数的变化, 建立了参数惩罚机制以限制模型在学习新数据时对新数据过拟合的现象; 通过增强样本空间多样性, 提出经验回放方法应用于增量学习, 将新数据与从缓存池中采样得到的旧数据进行联合训练. 进一步地, 将所提方法分别应用在材料吸声系数回归和图像分类任务上, 实验结果表明采用增量学习方法后, 平均准确率分别提升了 45.93% 和 2.62%, 平均遗忘率分别降低了 2.25% 和 7.54%. 除此之外, 还分析了参数惩罚和经验回放方法中具体参数对平均准确率的影响, 结果显示平均准确率随着回放比例的增大而增大, 随着惩罚系数的增大先增大后减小. 综上所述, 本文提出的方法能够跨模态、任务进行学习, 且参数设置灵活, 可以根据不同环境和任务进行变动, 为材料数据的增量学习提供了可行的方案.

关键词　材料数据; 神经网络; 增量学习; 参数惩罚; 经验回放
分类号　TG142.71

Incremental learning of material absorption coefficient regression based on parameter penalty and experience replay

WANG Hong-ye[1)], *QIAN Quan*[1,2)], *WU Xing*[1,2)]✉

1) School of Computer Engineering and Science, Shanghai University, Shanghai 200444, China
2) Zhijiang Laboratory, Hangzhou 311100, China
✉ Corresponding author, E-mail: xingwu@shu.edu.cn

ABSTRACT　Material data are prepared in batches and stages, and data distribution in different batches varies. However, the average accuracy of neural networks declines when learning material data by batch, resulting in great challenges to the application of artificial intelligence in the materials field. Therefore, an incremental learning framework based on parameter penalty and experience replay was applied to learn streaming data. The average accuracy decline is due to two reasons: sudden variations of model parameters and a quite homogeneous sample feature space. By analyzing the model parameter variation, a mechanism of parameter penalty was established to limit the phenomenon of model parameters fitting toward new data when the model learns new data. The penalty strength of the parameters can be dynamically adjusted according to the speed of parameter change. The faster the speed of parameter changes, the higher the penalty strength, and vice versa, the lower the penalty strength. To enhance sample diversity, experience replay methods were proposed, which train the new and old data obtained by sampling from the cache pool. At the end of each incremental task, the incremental data were sampled and used for the update of the cache pool. Specifically, random sampling was adopted for the joint

收稿日期: 2022−05−03
基金项目: 国家重点研发计划资助项目(2022YFB3707800); 云南省重大科技专项(202102AB080019-3, 202002AB080001-2); 之江实验室科研攻关项目(2021PE0AC02); 上海张江国家自主创新示范区专项发展资金重大项目(ZJ2021-ZD-006)

training, whereas reservoir sampling was used for the update of the cache pool. Further, the proposed methods (i.e., experience replay and parameter penalty) were applied to the material absorption coefficient regression and image classification tasks, respectively. The experimental results indicate that experience replay was more effective than parameter penalty, but the best results were obtained when both methods were used. Specifically, when both methods were used, the average accuracy of the benchmark increased by 45.93% and 2.62% and reduced the average forgetting rate by 86.60% and 67.20%, respectively. A comparison with existing methods reveals that our approach is more competitive. Additionally, the effects of specific parameters on the average accuracy were analyzed for both methods. The results indicate that the average accuracy increases with the proportion of experience replay and increases and then decreases when the penalty factor increases. In general, our approach is not limited by data modalities and learning tasks and can perform incremental learning on tabular or image data, regression, or classification tasks. Further, owing to the quite flexible parameter settings, it can be adapted to different environments and tasks.

KEY WORDS material data; neural network; incremental learning; parameter penalty; experience replay

吸声材料因其机理复杂, 尚未形成成熟的研究体系[1], 通过神经网络[2−3] 预测吸声材料的吸声系数是一项重要的研究内容. 神经网络的局限性在于必须依托完备的数据才能获得泛化能力较好的模型. 这一要求对材料数据来说是苛刻的, 因为材料数据的获取具有分批次、分阶段制备[4] 的特点, 且不同批次数据的分布也不同. 具体来说, 材料数据的获取需要大量实验的支持[5], 但设备和人力资源的短缺使得这些实验被分成多组, 而较长实验周期使得一次性获取全部的实验数据变得困难. 然而, 神经网络分批次地对材料数据进行训练会存在平均准确率随批次下降[6] 的问题, 即对之前批次的样本的预测准确率下降, 这是人工智能应用于材料领域的一个极大挑战. “分批次”与神经网络中常用的“小批量学习”不同. “分批次”指的是将全部样本分成多个批次, 每个批次的样本是不可互通的, 即学习完一个批次的样本后便将这个批次的样本丢弃或者保留部分; “小批量学习”指的是出于内存限制等考虑将全部样本分成多个小批量样本进行滚动学习. 增量学习[7−8] 能够使一个学习系统不断地学习分批次的数据流, 在吸收新批次的知识的同时保留之前批次学习到的知识. 结合材料数据分批次、分阶段制备这一特点, 增量学习非常适合材料科学方面的研究, 即能够在不断变化的分批次材料实验数据中学习到新的知识.

1 背景

增量学习应用在神经网络上主要有基于正则化[9−10]、样本回放[11−15] 和参数隔离[16−17] 的三种方法. 基于正则化的方法的主要思想是对模型参数施加约束来保护旧知识不被新知识覆盖. Li 和 Hoiem [9] 利用知识蒸馏来保存多头设置中过去任务的知识. 基于样本回放的方法会保留部分之前批次的样本, 主要考虑的是如何保留样本以及如何将这些样本与当前批次的样本进行联合训练. Rebuffi 等[12] 通过混合内存缓冲区中的所有样本和当前任务样本来进行表征学习, 同时采用二元交叉熵损失来处理类别不平衡问题. 基于参数隔离的方法将模型参数进行隔离, 模型的参数被划分为多个区域并用于不同批次样本的训练. Mallya 和 Lazebnik [16] 等采用硬隔离的方法, 通过迭代剪枝向单个网络添加多个任务来缓解平均准确率随批次下降的问题. 此外, 基于正则化和参数隔离的方法不需要存储之前批次的样本, 是一种较为理想的增量学习方法. 而基于样本回放的方法在设计方法时具有特殊性, 会根据样本的特点进行设计, 这类方法可能只适合特定的数据集.

增量学习方法在分类问题的应用十分广泛[18], 即不同批次的数据包含不同类别的样本, 比如第一个批次包含狗和猫, 而第二个批次包含兔子和羊. Mai 等[19] 将对比学习与基于样本回放的方法相结合, 并采用类别平均分类器(Nereast class mean classifier, NCM)分类器应用于增量学习. 具体地, 对比学习用于学习不同类别的内在区别, 而 NCM 分类器为每一类的样本构建了类中心并用于损失函数的计算. 然而, 增量学习方法在回归任务的研究很少, Mai 等提出的方法也仅限于分类任务. 在回归任务中, 如果样本特征空间分布不均衡, 也会存在模型平均准确率随批次下降的问题. 为此, 我们提出了一种基于参数惩罚和经验回放的材料吸声系数回归增量学习方法, 利用增量学习为神经网络按批次学习材料数据提供指导, 在减少计算量的同时缓解神经网络按批次学习时平均准确率下降的问题. 本文的贡献在于:

(1) 提出跨数据模态且任务类型无关的增量学

习方法，该方法不仅适用于表格或图像数据，且适用于回归或分类的机器学习任务.

(2) 针对新、旧样本设计的参数惩罚和经验回放方法，可有效解决神经网络按批次学习时平均准确率下降的问题.

(3) 增量学习方法中超参数的灵活设置策略，适用于各种存储空间的应用场景.

2 基于参数惩罚和经验回放的材料吸声系数回归增量学习

2.1 数据预处理

我们采用的多孔吸声材料数据如图 1 所示，多孔吸声材料由四个矩形块组成，每个矩形块的上面都有一个小孔. 该多孔吸声材料在不同的声音频率下会产生不同的吸声系数. 因此，我们以四个矩形块的高度、孔的深度、孔的半径和频率累计 13 个变量作为自变量，以材料吸声系数作为因变量进行回归任务学习. 在模型进行学习前，需要对数据进行归一化，归一化后的数据往往能够提升模型的收敛速度和模型精度，并且可以防止模型梯度爆炸等问题. 我们采用的归一化方法是最大最小值归一化，公式如式 (1) 所示：

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

其中，特征x的取值范围为$[x_{\min}, x_{\max}]$.

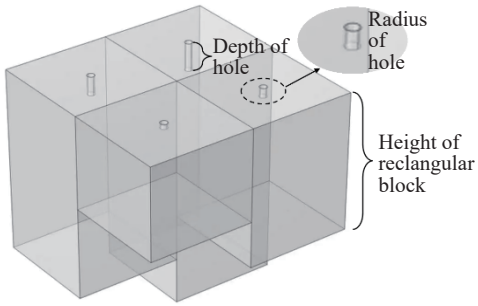

图 1 多孔吸声材料结构示意图

Fig.1 Structure of the sound-absorbing material

2.2 参数惩罚

模型平均准确率随批次下降的问题是因为模型会专注于学习当前批次数据的分布，具体表现为模型参数变化速度过快，使得模型对已经学习过的数据的准确率下降. 因此，我们建立参数惩罚机制（Parameters penalty, PP），通过限制模型参数的变化来缓解该问题. 我们认为每一个参数造成该问题的影响程度是不同的，在模型参数更新前后，若参数的变化较大，则认为该参数的影响程度较高，反之则较低. 由此得到限制模型参数变化的损失函数如式 (2) 所示：

$$\text{loss}_w = \lambda \sum_{i=1}^{N_w} p_i (w_i^{\text{new}} - w_i^{\text{old}})^2 \tag{2}$$

其中，w表示模型参数，λ表示惩罚损失系数，N_w表示参数的个数，w_i^{new}表示模型当前的参数值，w_i^{old}表示经过上一批次数据学习到的模型参数值，p_i表示第 i 个参数的惩罚系数，表示该参数对模型的重要程度，参数越大，则希望其在进行参数更新时的变化越小，即参数越不容易变化.

在公式 (2) 中，我们提到了惩罚系数，我们认为每个参数的惩罚系数是会不断变化的，在不同的模型学习阶段，参数对模型的重要程度发生变化，因此我们采用滑动平均的方式动态更新每个参数的惩罚系数. 除此之外，我们认为参数的重要性可以通过每个参数反向传播的梯度来表示，梯度越大，则表示参数的变化程度越明显，也表示该参数对模型的重要程度越大，由此得到惩罚系数的动态调整公式如式 (3) 所示：

$$p_i^t = \begin{cases} 0, & t = 0 \\ \alpha(\nabla p_i^t)^2 + (1-\alpha)p_i^t, & t > 0 \end{cases} \tag{3}$$

其中，参数 i 的惩罚系数初始化为 0，∇p_i^t表示参数 i 的惩罚系数在第 t 次反向传播时的梯度，α是平滑因子，文中设置为 0.9.

2.3 经验回放

除了通过参数惩罚机制减缓模型平均准确率随批次下降的问题外，我们还提出了基于经验回放（Experience replay, ER）的方法. 经验回放指在模型训练过程中从样本缓存池中挑选一些旧样本与当前训练样本进行联合训练，在训练结束时，在当前训练样本中选取一些样本用来更新样本缓存池. 经验回放本质上在于丰富样本空间的多样性，使模型在学习新样本的同时回顾旧样本中的知识. 增量学习的一大特点便是希望模型能够在有限的存储空间内实现对模型的持续性改进，因此基于经验回放的方法是属于增量学习范畴的. 基础的经验回放方法[20−21]是一个简单且有效的基准模型，样本缓存池中的样本是通过随机采样得到的，样本缓存池更新则是通过水库采样实现. 水库采样是一种典型的空间亚线性算法，能够在仅观察一次样本的条件下实现对全部数据的均匀采样. 我们提出的经验回放方法如算法 1 所示：

Algorithm 1:ER method

Input: Batch size b, learning rate α

Initialize: Memory $M \leftarrow \{\} * M$ Parameters θ;

for $t \in \{1,\ldots,T\}$ **do (1)**

for m epochs **do (2)**

for $B_n \sim D_t$ **do (3)**

$B_{\mathcal{M}} \leftarrow \text{MemoryRetrieval}(\mathcal{M})$

train θ by minimizing $\mathcal{L}_{\text{MSE}}(\theta, B_n \cup B_{\mathcal{M}})$

$\mathcal{M} \leftarrow \text{MemoryUpdate}(D_t, \mathcal{M})$

其中, T 表示批次的数量, m 表示样本训练轮次, $\mathcal{L}_{\text{MSE}}$表示均方误差损失函数. 为了使每个样本都得到充分的学习, m 的设置都大于 1, 因此缓存池更新的步骤应该放置在每个批次数据完成训练后, 即循环 (1) 下.

2.4 实现细节

我们提出的增量学习方法是基于开源的 python 机器学习库 PyTorch 实现的, 该开源学习库常用于人工智能领域的研究. 具体的实现细节, 我们分别从数据预处理和方法应用顺序展开.

数据预处理: 不同批次的数据分布是不同的, 具体表现在均值和方差上. 为了进一步降低在模型按批次学习时由于数据分布变化造成的影响, 我们将 2.1 节中的最大最小值归一化的数据范围预先固定, 不需要根据不同批次数据的分布变化而变化.

方法应用顺序: 从缓存池中采样得到的样本是属于旧数据范畴的, 可以帮助模型维护旧知识, 而参数惩罚机制是面向本批次的样本, 也就是增量样本, 因此从缓存池中采样得到的样本是不参与参数惩罚机制的. 在具体使用参数惩罚机制和经验回放方法时, 要对增量样本进行参数惩罚后再进行经验回放, 具体步骤如下:

1、对增量样本计算均方误差损失并生成梯度, 从而进行反向传播.

2、通过步骤 1 的反向传播, 模型的参数得到更新, 将更新后的参数和通过上一批次数据学习到的参数进行参数惩罚机制损失的计算, 生成梯度后进行反向传播. (参数惩罚)

3、在样本缓存池进行随机采样(经验回放), 对采样得到的旧样本计算均方误差损失, 生成梯度后进行反向传播.

4、当所有增量样本完成学习后, 通过水库采样对样本缓存池进行更新. (经验回放)

3 实验分析与讨论

为了说明我们提出的方法的有效性, 我们在 2.1 节中提及的多孔吸声材料数据集上开展了是否使用增量学习方法的对比实验. 除此之外, 为了说明方法的泛化性, 我们在开源的 CIFAR-10 图像数据集中也进行了对比实验. 从数据模态的角度来看, 多孔吸声材料数据集是结构化数据, 而 CIFAR-10 数据集则是图像数据. 从学习任务的角度来看, 前者是回归任务, 后者则是分类任务.

3.1 数据集

多孔吸声材料数据集: 多孔吸声材料数据集是在 7 次不同实验设置下生成的数据, 累计 157234 条. 其中第 1 批数据中的每个材料的四个矩形块的高度、孔的深度、孔的半径都是相同的, 不同材料间则不相同; 第 2 批数据中每个材料的四个孔的深度则不同. 从描述中就可以看出不同批次的数据分布是不同的.

CIFAR-10 图像数据集: CIFAR-10 图像数据集由 10 类(飞机、汽车、鸟、猫等)60000 张 32×32 的彩色图像组成, 每类有 6000 张图像. 为了让数据符合分批次、分阶段的特点, 我们将 CIFAR-10 数据集划分为 5 个批次. 每个批次的数据都包含 10 个类别的图像, 不同批次间图像的数量分布则不同, 比如第 1 个批次数据中飞机的数量占 50%, 而第 2 个批次数据中汽车的数量占 50%.

3.2 对比实验说明

对比实验分为不使用增量学习方法和使用增量学习方法两种, 使用增量学习方法可以细分为仅使用参数惩罚机制和仅使用经验回放方法和两者都用的三种情况. 在面对分批次、分阶段的数据时, 模型会按批次进行学习, 通过对当前批次数据的学习来更新之前批次学习到的模型参数, 不使用增量学习方法的具体步骤如下:

步骤 1: 若该批次属于第 1 个批次, 则模型进行随机初始化. 若不是, 则模型通过上一个批次数据学习到的最优模型参数进行初始化.

步骤 2: 通过任务对应的损失函数进行反向传播, 从而更新模型参数并保存最优模型参数.

使用增量学习方法时, 会在步骤 2 中加入参数惩罚机制和经验回放方法来进一步修正模型参数. 其中, 材料吸声系数回归任务以 4 层的多层感知机作为模型, 以均方误差作为损失函数; CIFAR-10 图像分类任务以残差网络 ResNet-18 作为模型, 以交叉熵作为损失函数. 多层感知机和残差网络都属于神经网络.

3.3 评价指标

除了模型在通过所有批次数据学习后的最终准确度之外, 评估一个模型的好坏, 还应该考虑模

型的遗忘程度和转移能力. 因此, 我们采用增量学习文献 [11,22−24] 中的四个标准指标来衡量模型的性能, 分别为模型的平均准确率、平均遗忘率、前向转移和后向转移能力. 其中, 平均遗忘率用来衡量模型对获得的知识的遗忘程度, 前向和后向转移用来评估模型知识转移的能力.

我们将$a_{i,j}$定义为模型在经过 1 到 i 个批次数据的学习后, 对第 j 个批次的保留测试集评估的准确率, 假设总共有 T 个批次的数据.

平均准确率可以被定义为公式 (4). 当$i = T$时, A_T代表所有批次的数据训练结束时的平均精度.

$$A_i = \frac{1}{i}\sum_{j=1}^{i} a_{i,j} \tag{4}$$

平均遗忘率可以被定义为公式 (5). $f_{i,j}$代表模型在第 i 个批次数据学习后模型对第 j 个批次数据的遗忘程度. 具体来说, $\max\limits_{l\in\{1,\cdots,k-1\}}\left(a_{l,j}\right)$表示在学习第 k 个批次数据之前模型在第 j 个批次数据上取得的最佳测试精度, $a_{k,j}$是学习第 k 个批次数据之后模型在第 j 个批次数据上的测试精度.

$$\begin{cases} F_i = \dfrac{1}{i-1}\displaystyle\sum_{j=1}^{i-1} f_{i,j} \\ f_{k,j} = \max\limits_{l\in\{1,\cdots,k-1\}}(a_{l,j}) - a_{k,j}, \forall j < k \end{cases} \tag{5}$$

前向转移(FWT)衡量学习一个新的批次数据后模型对未来数据表现的积极影响, 如公式 (6) 所示.

$$\mathrm{BWT} = \max\left(\frac{\displaystyle\sum_{i=2}^{T}\sum_{j=1}^{i-1}(a_{i,j} - a_{j,i})}{\frac{T(T-1)}{2}}, 0\right) \tag{6}$$

后向转移(BWT)衡量学习一个新的批次数据后模型对之前数据表现的积极影响, 如公式 (7) 所示.

$$\mathrm{FWT} = \frac{\displaystyle\sum_{i=1}^{j-1}\sum_{j=1}^{T} a_{i,j}}{\frac{T(T-1)}{2}} \tag{7}$$

3.4 实验结果与分析

我们分别对材料吸声系数回归任务和 CIFAR-10 图像分类任务开展了对比实验. 每个任务除了不使用增量学习方法(Base)外, 还包括使用参数惩罚机制(PP)、使用经验回放(ER)和两者都用(PPER). 此外, 我们与不遗忘学习[9](Learning without forgetting, LWF)和突触可塑算法[25](Memory aware synapses, MAS)做了对比. LWF 通过知识蒸馏来保留旧知识. MAS 与 PP 的方法类似, 不同的是 MAS 采用输出函数的敏感性来估算参数重要性, 并且它是一种无监督的方法.

多孔吸声系数回归任务的对比实验结果如图 2 所示, 准确率用决定系数表示. 从图中可以看出, 我们的方法在平均准确率、平均遗忘率、前向转移和后向转移的效果都优于 Base 方法, 其中 PPER 的方法最优, ER 其次, PP 再次. 具体表现为: 平均准确率、平均遗忘率和后向转移: PPER>ER>PP>Base; 前向转移: PP > ER > PPER > Base. 观察图 2(a) 发现 PP 的方法在初期时的平均准确率是高于 Base 的, 但是在学习第 5 个批次数据时, 准确率突然比 Base 小, 因此认为基于参数惩罚的方法更加适合于短期的按批次学习, 如果学习的任务过长, 该方法会丧失对初期学习到知识的保护. 此外, 我们还发现 PP、ER、PPER 这 3 组实验在前向转移和后向转移中的效果正好是相反的, 即前向转移好的后向转移差, 这是模型稳定性和扩展性相互制衡的体现, 当模型的稳定性较好时, 扩展性相对较差, 反之则扩展性越好. LWF 方法在该回归任务的表现不佳, 基本与 Base 平齐. MAS 方法在平均准确率的表现略差于 PP, 但在平均遗忘率和后向转移中略高于 PP, 在后向转移方面的能力是非常出色的. 但总体来说, MAS 还是逊色于我们所提出的方法.

CIFAR-10 图像分类任务的对比实验结果如表 1 所示, 我们统计了在该任务上进行的六组实验的评价指标的平均值. 从表中可以看出我们提出的方法依然有效, PPER 的方法在平均准确率和后向转移方面分别提高了 2.62% 和 2.82%, 平均遗忘率则降低了 7.54%.

3.5 灵敏度分析

为了更具体地分析参数惩罚和经验回放方法的有效性, 我们探讨了两种方法是如何影响按批次学习的平均准确率的. 图 3 表示了材料吸声系数回归任务在经验回放和参数惩罚方法下设置不同参数的平均准确率折线图. 从图 3(a)中可以看到随着经验回放方法的缓存池大小变大, 平均准确率增大. 从图 3(b)中可以看到随着惩罚参数方法中的惩罚项λ的变大, 平均准确率呈现先上升后下降的趋势.

4 结论

(1) 针对神经网络按批次训练材料数据时出现的平均准确率随批次下降的问题, 我们设计了一种基于参数惩罚和经验回放的增量学习方法, 可以缓解平均准确率的下降.

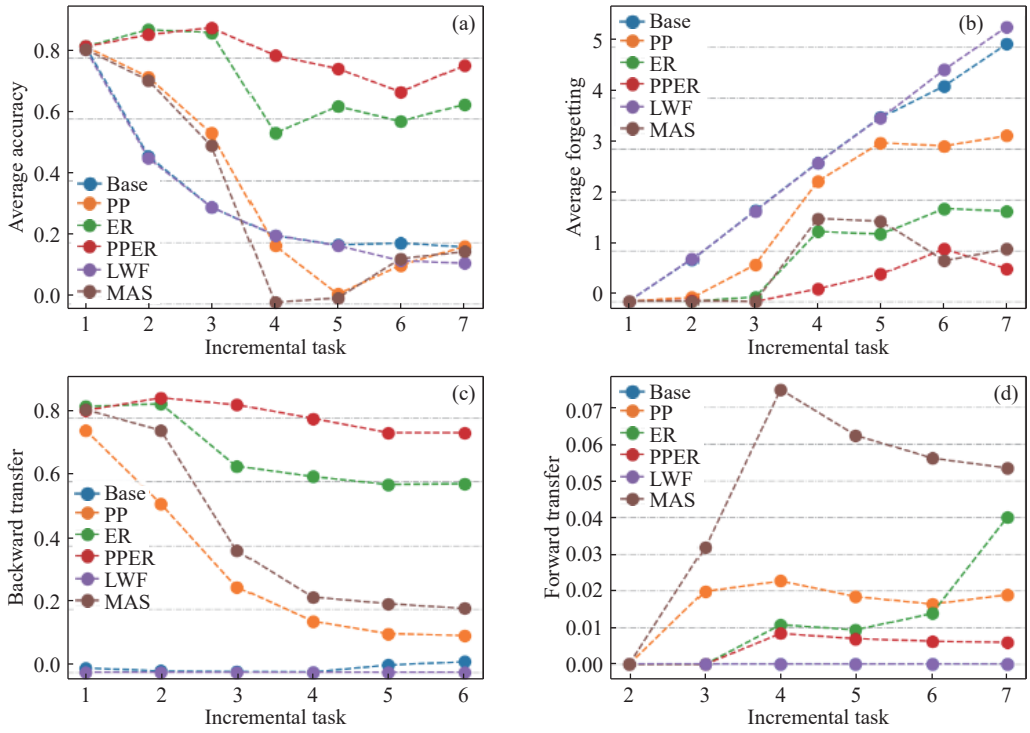

图 2　材料数据在不同方法下的按批次学习结果折线图. (a) 平均准确率; (b) 平均遗忘率; (c) 前向转移; (d) 后向转移

Fig.2　Line graph of incremental learning results for material data under different methods: (a) average accuracy; (b) average forgetting; (c) forward transfer; (d) backward transfer

表 1　CIFAR-10 上进行的四组实验的评价指标的平均值

Table 1　Mean values of the evaluation metrics for the four sets of experiments conducted on CIFAR-10

Method	Average accuracy	Average forgetting	Backward transfer	Forward transfer
Base	0.7278	11.2200	0.5579	0.4787
PP	0.7364	8.1500	0.5397	0.4630
ER	0.7392	8.0600	0.5619	0.4645
PPER	0.7540	3.6800	0.5861	0.4779
LWF	0.6378	4.4200	0.5297	0.4431
MAS	0.6397	24.9000	0.4566	0.4270

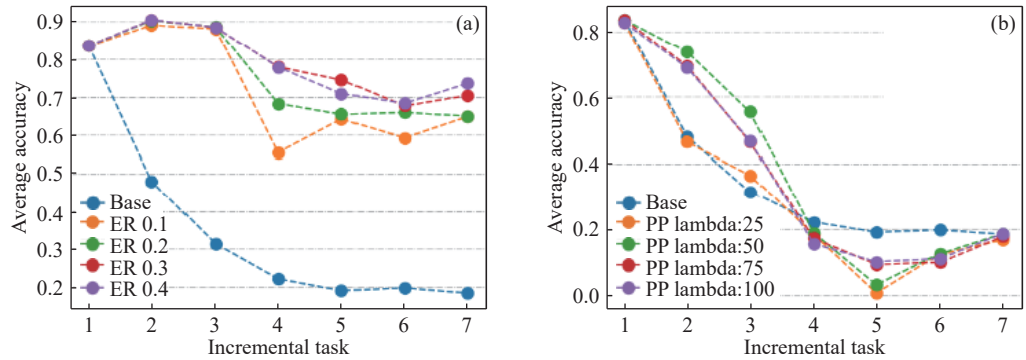

图 3　材料吸声系数回归任务在不同设置下的平均准确率折线图. (a) 材料吸声系数回归任务在经验回放方法下不同参数的平均准确率折线图; (b) 材料吸声系数回归任务在参数惩罚方法下不同参数的平均准确率折线图

Fig.3　Line graph of the average accuracy of incremental learning of material data for different settings: (a) line graph of the average accuracy of incremental learning of material data for different parameters under experiential replay; (b) line graph of the average accuracy of incremental learning of material data with different parameters under parameter penalty

(2) 针对不同模态或不同学习任务差异性较大导致方法不适用的问题，我们提出了一种通用方法可以不受数据模态、学习任务的限制进行按批次学习.

(3) 针对不同存储空间的应用场景和任务，我们提出的方法可以动态地调整参数大小以适应不同的场景.

参 考 文 献

[1] Liang L S, Guo W L, Ma H Y, et al. Research progress of sound absorption performance prediction and sound absorption model of porous sound-absorbing materials. *Mater Rep*, 2022(23): 1
(梁李斯, 郭文龙, 马洪月, 等. 多孔吸声材料吸声性能预测及吸声模型研究进展. 材料导报, 2022(23): 1)

[2] Ciaburro G, Iannace G, Ali M, et al. An artificial neural network approach to modelling absorbent asphalts acoustic properties. *J King Saud Univ Eng Sci*, 2021, 33(4): 213

[3] Iannace G, Ciaburro G, Trematerra A. Modelling sound absorption properties of broom fibers using artificial neural networks. *Appl Acous*, 2020, 163: 107239

[4] Zhai T T, Gao Y, Zhu J W. Survey of online learning algorithms for streaming data classification. *J Softw*, 2020, 31(4): 912
(翟婷婷, 高阳, 朱俊武. 面向流数据分类的在线学习综述. 软件学报, 2020, 31(4): 912)

[5] Dong J Y, Yang X Y. Integration and optimization of material data mining and machine learning tools. *Front Data & Comput*, 2020, 2(4): 105
(董家源, 杨小渝. 材料数据挖掘与机器学习工具的集成与优化. 数据与计算发展前沿, 2020, 2(4): 105)

[6] Kirkpatrick J, Pascanu R, Rabinowitz N, et al. Overcoming catastrophic forgetting in neural networks. *PNAS*, 2017, 114(13): 3521

[7] Mai Z D, Li R W, Jeong J, et al. Online continual learning in image classification: An empirical survey. *Neurocomputing*, 2022, 469: 28

[8] Parisi G I, Kemker R, Part J L, et al. Continual lifelong learning with neural networks: A review. *Neural Netw*, 2019, 113: 54

[9] Li Z Z, Hoiem D. Learning without forgetting. *IEEE Trans Pattern Anal Mach Intell*, 2018, 40(12): 2935

[10] Zenke F, Poole B, Ganguli S. Continual learning through synaptic intelligence // *Proceedings of the* 34*th International Conference on Machine Learning*. Sydney, 2017: 3987

[11] Chaudhry A, Dokania P K, Ajanthan T, et al. Riemannian Walk for Incremental Learning: Understanding Forgetting and Intransigence // *European Conference on Computer Vision*. Munich, 2018: 556

[12] Rebuffi S A, Kolesnikov A, Sperl G, et al. iCaRL: Incremental classifier and representation learning // *Conference on Computer Vision and Pattern Recognition*. Honolulu, 2017: 5533

[13] Aljundi R, Caccia L, Belilovsky E, et al. Online continual learning with maximally interfered retrieval // *Proceedings of the* 33*rd International Conference on Neural Information Processing Systems*. Vancouver, 2019: 11872

[14] Aljundi R, Lin M, Goujaud B, et al. Gradient based sample selection for online continual learning // *Proceedings of the* 33*rd International Conference on Neural Information Processing Systems*. Vancouver, 2019: 11817

[15] Prabhu A, Torr P H S, Dokania P K. GDumb: A simple approach that questions our progress in continual learning // *European Conference on Computer Vision*. Glasgow, 2020: 524

[16] Mallya A, Lazebnik S. PackNet: Adding multiple tasks to a single network by iterative pruning // 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, 2018: 7765

[17] Li X L, Zhou Y, Wu T, et al. Learn to grow: A continual structure learning framework for overcoming catastrophic forgetting // *International Conference on Machine Learning*. Long Beach, 2019: 3925

[18] Lange M D, Aljundi R, Masana M, et al. A continual learning survey: Defying forgetting in classification tasks. *IEEE Trans Pattern Anal Mach Intell*, 2022, 44(7): 3366

[19] Mai Z D, Li R W, Kim H, et al. Supervised contrastive replay: Revisiting the nearest class mean classifier in online class-incremental continual learning // *Conference on Computer Vision and Pattern Recognition*. Online, 2021: 1177

[20] Hayes T L, Cahill N D, Kanan C. Memory efficient experience replay for streaming learning // *International Conference on Robotics and Automation*. Montreal, 2019: 9769

[21] Liu Y Y, Su Y T, Liu A N, et al. Mnemonics training: Multi-class incremental learning without forgetting // 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, 2020: 12242

[22] Chaudhry A, Dokania P K, Ajanthan T, et al. Riemannian walk for incremental learning: Understanding forgetting and intransigence // *Proceedings of the European Conference on Computer Vision*. Munich, 2018: 556

[23] Lesort T, Lomonaco V, Stoian A, et al. Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges. *Inf Fusion*, 2020, 58: 52

[24] Lopez-Paz D, Ranzato M A. Gradient episodic memory for continual learning // *Proceedings of the* 31*st International Conference on Neural Information Processing Systems*. Long Beach, 2017: 6470

[25] Aljundi R, Babiloni F, Elhoseiny M, et al. Memory aware synapses: Learning what (not) to forget // *Proceedings of the European Conference on Computer Vision*. Munich, 2018: 144