

DOI:10.19651/j.cnki.emt.2518857

超材料领域的持续学习:动态数据生成与模型性能评估*

陈祺东¹ 徐国梁² 邓瑞翔³ 吴豪¹

(1. 无锡学院物联网工程学院 无锡 214105; 2. 南京信息工程大学计算机学院、网络空间安全学院 南京 210044; 3. 中国科学院上海硅酸盐研究所 上海 201899)

摘要: 数据集规模是影响深度学习模型性能的关键因素之一。由于深度学习模型性能高度依赖于数据集规模, 实现特定精度所需的数据量通常难以预估。该问题在超材料智能设计中同样存在, 成为制约建模精度和效率的重要因素。为此, 提出一种动态数据生成与模型性能评估框架, 以实现数据集规模与模型性能的动态监控。为提升模型动态评估效率并有效缓解灾难性遗忘现象, 设计了一种持续学习策略, 使模型在动态评估过程中仅需针对新数据进行学习, 同时保持对已有知识的记忆。实验结果表明, 基于该持续学习策略训练的模型预测平均准确率可达到 93.28%, 平均遗忘率为 3.68%, 充分验证了该模型在缓解灾难性遗忘问题方面的有效性。

关键词: 微波吸收超材料; 数据生成; 数据集规模; 灾难性遗忘

中图分类号: TP274+.2; TN955 **文献标识码:** A **国家标准学科分类代码:** 520.60

Continuous learning in metamaterials: Dynamic data generation and model performance evaluation

Chen Qidong¹ Xu Guoliang² Deng Ruixiang³ Wu Hao¹

(1. School of IoT Engineering, Wuxi University, Wuxi 214105, China; 2. School of Computer Science & School of Cyber-space Security, Nanjing University of Information Science & Technology, Nanjing 210044, China; 3. Shanghai Institute of Ceramics & Chinese Academy of Sciences, Shanghai 201899, China)

Abstract: The size of the dataset is one of the key factors affecting the performance of deep learning models. Since the performance of deep learning models is highly dependent on the size of the dataset, the amount of data required to achieve a specific accuracy is usually difficult to estimate. This problem also exists in the intelligent design of metamaterials, and has become an important factor restricting the accuracy and efficiency of modeling. To this end, a dynamic data generation and model performance evaluation framework is proposed to achieve dynamic monitoring of the size of the dataset and model performance. In order to improve the efficiency of dynamic evaluation of the model and effectively alleviate the catastrophic forgetting phenomenon, a continuous learning strategy is designed so that the model only needs to learn new data during the dynamic evaluation process while maintaining the memory of existing knowledge. Experimental results show that the average prediction accuracy of the model trained based on this continuous learning strategy can reach 93.28%, and the average forgetting rate is 3.68%, which fully verifies the effectiveness of the model in alleviating the problem of catastrophic forgetting.

Keywords: microwave absorbing metamaterials; data generation; dataset size; catastrophic forgetting

0 引言

微波吸收超材料 (microwave absorbing metamaterials, MAMs) 是一类人工材料, 通过合理设计其几何结构和空间

排布, 能够实现特定的电磁响应^[1]。这类材料在隐身、防护以及减少电磁干扰等方面发挥着关键作用, 并被广泛应用于军事领域^[2], 通信领域^[3] 和医学领域^[4-5] 等。传统的 MAMs 正向设计方法主要依赖于人工设计其几何结构, 并

收稿日期: 2025-05-19

* 基金项目: 无锡学院引进人才科研启动专项经费 (2024r004) 项目资助

通过高频结构模拟器(high-frequency structure simulator, HFSS)等仿真平台,采用有限元法或有限时域差分法等复杂电磁学计算方法对其吸波性能进行评估。通过不断尝试和仿真建模,迭代设计出一个吸波性能相对最优的设计方案。但多次仿真评估通常需要耗费大量时间资源,从而导致 MAMs 正向设计周期的增加。

随着深度学习技术的不断发展,人们开始尝试将深度学习模型应用于 MAMs 设计领域,通过深度学习模型构建 MAMs 吸波性能预测模型,以替代耗时的仿真过程,加速性能评估并提升设计效率。深度学习模型是由数据驱动的,其性能在很大程度上依赖于数据集的规模。一般来说,深度学习模型达到目标精度所需的数据集规模是先验未知的。当数据集规模不足时,模型可能欠拟合而无法达到目标性能;相反,当数据集规模超过模型需求阈值时,冗余数据可能会引入额外的仿真开销。这是由于在 MAMs 设计任务中,数据主要通过仿真生成,冗余数据意味着额外的计算资源消耗与生成成本。因此,探索一种有效的策略以确定最佳数据集规模成为关键问题。一种可行的方法是在数据集规模动态增长的过程中,动态评估模型性能,以明确模型达到目标性能所需的数据集规模。

在动态评估模型性能的过程中,每次使用全部数据集进行训练将导致高昂的时间和计算成本,而仅使用新获取的数据进行学习则可能引发“灾难性遗忘”现象^[6]。具体来说,由于输入参数的维度庞大,解空间呈指数级扩展,而解空间中不同区域的数据分布复杂多变,不同阶段生成的数据通常呈现非平稳分布的特征,仅基于新获取的数据进行训练的模型可能会对新数据产生过拟合,从而导致对旧数据的预测精度下降,出现“灾难性遗忘”现象。

持续学习在动态场景中已被广泛研究,是解决灾难性遗忘问题的关键技术路径之一。持续学习通过对新旧数据的平衡学习,避免模型在学习新知识时遗忘旧知识。其主要策略包括样本回放和正则化约束等,能够有效应对动态数据分布所带来的非平稳性挑战。基于回放的方法通过直接重用历史任务的数据,有助于更有效地缓解灾难性遗忘。然而,现有的回放策略难以有效捕捉样本在高维空间中的位置分布信息,所选择的样本在空间覆盖性和代表性方面存在不足。

针对以上问题,本文提出一种基于持续学习的动态数据生成与模型性能评估框架,主要贡献如下:

1) 提出一种集数据生成、仿真与模型性能评估于一体的动态数据生成与模型性能评估框架。该框架能够根据模型的性能反馈动态确定最佳的数据集规模,从而提高数据利用效率并降低不必要的仿真开销。

2) 提出一种持续学习策略用于解决模型性能动态评估过程中面临的灾难性遗忘问题,并在所构建的仿真生成数据集上开展实验,验证了所提出方法的有效性与适用性。

1 相关工作

1.1 深度学习技术在超材料智能设计领域中的应用

利用大量的标注数据训练后,深度学习模型能够学习到数据中复杂的特征关系,可以快速进行推理,在各个领域中具有广泛的应用场景。常见的深度学习模型包括反向传播神经网络、卷积神经网络和循环神经网络等。在 MAMs 智能设计领域中,深度学习模型能够通过分析材料的结构特性,在较低的计算成本下快速预测材料在不同频段下的吸波性能。Wang 等^[7]通过 MATLAB-CST 协同仿真收集了 20 000 个样本,这些样本被用于训练他们提出的基于卷积神经网络的代理模型。通过该代理模型,成功设计了多层超材料吸波器,并实现多层复合材料吸收性能的高效预测,避免了耗时的全波仿真。Zhang 等^[8]提出了基于深度卷积神经网络的 SAP-Net,用于预测元孔材料在特定频率下的吸声性能。训练数据由 50 000 个图像和有限元仿真获得的 6 个频率点吸声系数组成。该方法实现了快速准确的性能评估,为吸声材料的快速筛选与设计提供了新工具。Cao 等^[9]提出了一种基于 KAN 的数据驱动方法,通过随机生成包含 2 500 组多方向应力-应变数据的数据集对模型进行训练。该方法能够准确捕捉结构参数与 7 个方向力学响应之间的非线性映射关系,预测误差最低可达 0.81%。Zhang 等^[10]通过基于参数网格的随机抽样过程生成了包含 6 493 个样本的数据集,并将该数据集用于训练所提出的带有固定注意力机制的深度学习框架中,用于预测高自由度超材料的光学响应,从而有效解决了传统设计方法高计算量的问题。Chen 等^[11]通过固定参数范围及步长,生成满足结构要求的参数组合,仿真得到 10 934 组训练数据用于训练卷积神经网络,实现太赫兹超材料传感器的高效设计。岳朋等^[12]综合考虑样本质量和样本多样性,结合参数化曲线和膨胀函数生成 30 000 个弹性超材料样本用于训练生成对抗网络,实现弹性超材料带隙的逆向设计。

深度学习技术在 MAMs 结构设计中的应用大幅降低了传统试错法所需的时间和计算成本,但其具体的技术实现路径仍存在不足。当前研究普遍缺乏对数据集规模与模型性能之间耦合关系的探讨,针对不同类型的超材料吸收器,不同类别深度学习模型的性能与其所需的数据集规模之间尚未建立明确的关联机制。现有方法主要基于经验或约定设定数据集规模,缺乏系统化的优化策略。

1.2 持续学习

持续学习,也称为终身学习、增量学习,是一种机器学习方法,他允许模型使用新的数据训练而不遗忘从先前数据上学习的知识。与传统的批量学习方法不同,持续学习通过逐步引入新数据进行训练,而无需重新训练模型,从而有效应对数据量大且变化迅速的应用场景。此方法特别适用于数据流呈现时变性的动态环境,其中模型需在保证实时性和性能的前提下不断适应新信息。

典型的持续学习方法包括正则化方法和回放方法^[13],二者在解决灾难性遗忘问题中被广泛应用,具有较高的代表性。王家亮等^[14]考虑到钢铁行业数据体量大,轧制特征维度高的特点,基于正则化的方法提出了一种面向工况漂移的基于增量学习树模型的带钢精轧宽度预测方法,通过调整相关树叶节点权重,为当前面临的变工况环境造成预测不准的问题提供了一种解决思路。Dohare等^[15]为避免反向传播算法在持续学习环境中出现可塑性的丧失,提出通过随机重新初始化网络中使用较少的单元来保持模型的可变性和可塑性。在ImageNet、CIFAR-100以及强化学习任务上的实验证明,该方法能够在长序列的任务中保持较高的学习性能。同样的, Li等^[16]通过冻结对新数据不敏感的超参数而仅更新敏感的超参数,使得数字孪生预测模型在金属成型产品加工条件动态变化的任务中仍具有良好的适应性。Kleiman等^[17]提出了一种通过在训练期间将当前模型与旧任务的模型参数定期加权平均,有效缓解灾难性遗忘,适用于图像分类与跨领域语言任务。

受人类互补学习系统理论^[18]启发的基于回放学习的方法被认为是持续学习的有力工具^[19]。Lin等^[20]通过特征子空间学习和累积空间回放缓解灾难性遗忘,在图像分类任务中显著优于现有回放策略,展现了其有效性与优越性。Liu等^[21]为解决长尾分布下的持续学习问题,以图像分类为任务,通过不确定性引导的储层采样策略优先回放少数类数据,并结合原型约束和边界约束,缓解灾难性遗忘。Rahman等^[22]针对恶意软件分类中的持续学习问题,提出了一种基于回放方法的持续学习框架MADAR。该方法通过多样性感知的样本选择策略,结合代表性样本与离群样本进行回放,有效缓解灾难性遗忘,强调了在设计持续学习技术时理解领域特征的重要性。李虢等^[23]提出了一种基于置信度不确定性的回放样本选择算法,并结合数据增强以提升样本多样性。实验结果表明,该方法在图像分类任务中具有良好的有效性。

然而,在超材料智能设计的动态模型性能评估场景中,不同阶段生成的数据可能具有不同的分布特性,也存在与先前阶段分布相似的情况。在此类情形下,基于正则化的持续学习方法可能对模型施加过度限制,从而抑制其对新知识的适应能力。鉴于该场景下旧样本在整个训练过程中始终可用,本文采用样本回放方法,以构建适用于该类任务的持续学习方法。现有的回放策略多聚焦于分类任务,其回放样本选择通常依赖于分类不确定性或强调每个样本被选中的概率相等,该类策略难以直接迁移应用于超材料智能设计场景中。

2 基于持续学习的动态数据生成与模型性能评估

2.1 三层方形环状 MAMs

本文以基于电阻膜设计的一个三层方形环状 MAMs 为研究对象,其中每一层的顶部为电阻膜,中间为介质层,

三层结构的底部为金属底板,其单元结构呈现如图1所示。材料的结构参数包括3个方环的外尺寸 a_1, a_2, a_3 ,其取值范围为 $3\sim 15\text{ mm}$;方环环宽 w 的范围为 $0.2\sim 1.5\text{ mm}$;表面电阻 r_1, r_2, r_3 的取值范围为 $50\sim 1\ 000\ \Omega$;介质厚度 d_1, d_2, d_3 的范围为 $0.1\sim 8\text{ mm}$,排布周期恒定为 15 mm 。

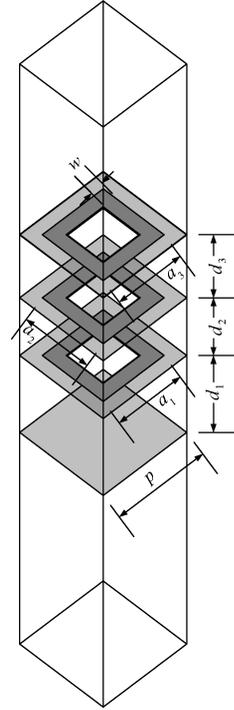


图1 吸波材料结构示意图

Fig.1 Schematic diagram of the structure of absorbing materials

2.2 动态数据生成与模型性能评估框架

不同吸波性能的超材料在结构参数空间中呈现差异化分布,特别是强吸波材料的结构参数具有显著的特异性和复杂性^[24],如图2所示。传统数据生成方法难以全面覆盖设计空间,单一的随机生成或拉丁超立方采样(latin hypercube sampling, LHS)等方法常导致对低吸波样本的过度采样,而对强吸波样本的捕获率较低。因此,应该结合多种生成策略以提高数据生成效率。此外,为更有效地生成强吸波样本,设计一种优化算法导向的数据生成方法。具体而言,使用当前模型作为评估模型,以评估种群中个体的吸波性能,并通过优化算法,如差分进化算法,搜索潜在的最优解,以获得强吸波样本。同时,在强吸波样本最优解的附近领域内执行数据生成,将优化算法获得的最优解与附近领域生成的结果合并形成新的数据生成集合。

在获取生成样本后,通过电磁仿真操作计算样本的反射率,并将其作为标签数据。创建HFSS接口对象,通过该对象加载三层方形环状MAMs模型,并根据结构参数赋值相应的变量参数。频率扫描的范围设置为 $2\sim 18\text{ GHz}$,步长设置为 1 GHz 。执行分析后,得到每个样本在频段

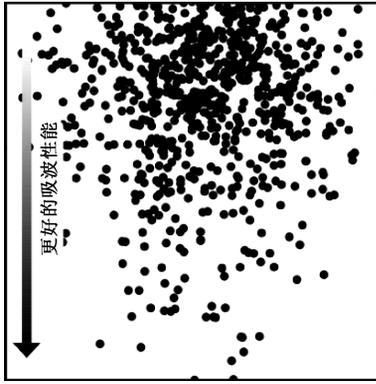


图 2 超材料结构参数空间

Fig. 2 Metamaterial structure parameter space

2~18 GHz 范围内每隔 1 GHz 等间隔采样点的 S 参数,用于表征其反射率。

在获取样本对应的反射率标签后,开展模型的训练与评估。考虑到三层方形环状 MAMs 结构参数存在不同量纲,这可能在模型训练过程中导致梯度消失或爆炸。因此,采用最大-最小归一化方法,将原始数据映射到统一尺度 [0,1] 区间,从而增强数据的可分性并优化模型训练效果。最大最小归一化的计算公式为:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

其中, x_{\max} 和 x_{\min} 分别代表参数的上边界和下边界。

将归一化处理后的样本结构参数作为特征向量,将电磁仿真获取的反射率作为标签向量,输入模型进行训练。当模型在预先划分的测试集上的损失值小于 1.5 时,表明其预测性能已满足需求,可终止数据收集流程;否则,需要重复执行数据生成、电磁仿真及模型训练与评估步骤。动态数据生成与模型评估整体流程如图 3 所示。

本文规定收集到的一系列 MAMs 仿真样本集 $D = \{D_0, D_1, \dots, D_i, \dots, D_T\}$, 其中 $D_i = \{x_1, x_2, \dots, x_j, \dots, x_n\}$, x_j 表示第 j 个 MAMs 样本, 包含 10 个特征, x_j 对应的标签 y_j 为一个 17 维向量, 表示 2~18 GHz 频段内 17 个采样点的反射率; n 表示样本集 D_i 中的样本数量。深度学习模型需要持续在样本集 D 上进行训练, 且在学习第 i ($i > 0$) 个样本集时, 需要确保模型在第 0 个到第 $i - 1$ 个样本集上的预测准确性不会显著下降。

2.3 一种持续学习策略

反向传播神经网络(back propagation neural network, BPNN)由输入层、输出层及隐藏层组成,能够学习并存储输入与输出之间的映射关系。由于其结构简单、参数可调且操作灵活,因此,本文选择 BPNN 作为用于数据集规模评估的深度学习模型,具体设计如下:输入层由 10 个结构参数组成,包含 6 层隐藏层(由实验经验得出),既能够确保模型的表达能力,又能有效降低由于参数过多而引发的过拟合风险。同时,在每一层后引入 ReLU 激活函数,以缓解

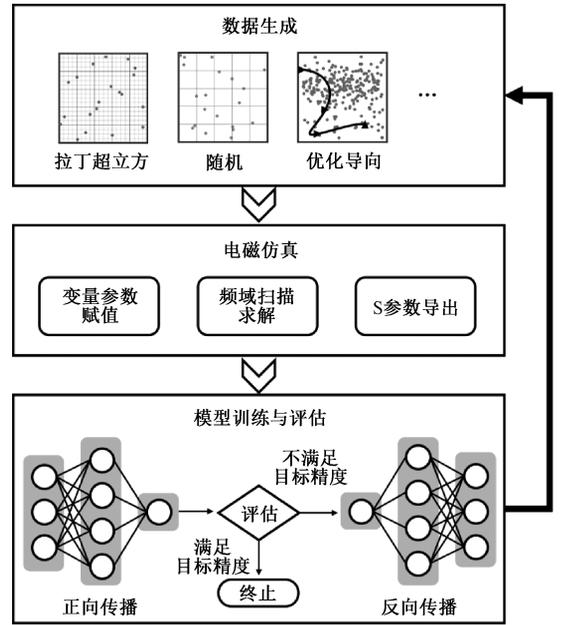


图 3 动态数据生成与模型性能评估框架

Fig. 3 Dynamic data generation and model performance evaluation framework

梯度消失问题;在第 3、第 4 和第 5 隐藏层后采用批归一化技术,以进一步提高网络的训练效率和稳定性。输出层包含 17 个参数,分别对应 2~18 GHz 上 17 个频段采样点的反射率。

为了降低模型性能动态评估过程中的时间和计算成本, BPNN 仅学习新获取的数据。同时,为了避免因对新数据过拟合导致“灾难性遗忘”现象,提出了一种持续学习策略。

当 BPNN 对第 T 组的数据 D_T 进行训练时,提出的持续学习策略可以分为 3 个训练阶段,即记忆选择阶段、学习巩固阶段和再选择学习巩固阶段。整体训练流程如图 4 所示。

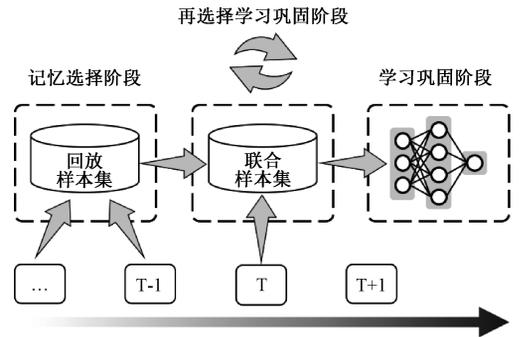


图 4 持续学习策略整体流程

Fig. 4 The overall process of continuous learning strategy

在记忆选择阶段,算法首先根据设定的回放样本数量 N 和衰减率 λ 计算在先前数据 $D_i, t \in \{1, \dots, T-1\}$ 上各自自选的回放样本数量 n_i 。通过这样的方式可以更好的

巩固最早阶段所学知识,最大程度的减少遗忘。然后进行回放样本的选择。LHS作为一种分层抽样策略被广泛应用于工程、计算机科学等众多领域,可以有效的减少计算密集型评估的数量^[25-26],确保样本在整个空间中的代表性。因此,使用LHS方法从各阶段的旧样本中均匀的选择回放样本,以确保样本的多样性和空间分布均匀性。

对于每个 D_i ,首先通过LHS生成 n_i 个候选样本 $L_i = \{L_i^0, \dots, L_i^j, \dots, L_i^{n_i-1}\}$,其中 $L_i^j, j \in \{0, \dots, n_i-1\}$ 表示 L_i 中第 j 个样本。LHS方法的核心思想是将每个特征维度均匀划分为 n_i 个区间,每个区间随机采样一个值。具体地,对于第 i 个特征:

$$x_{i,j} = x_{\min,i} + j \cdot \Delta x_i + u \cdot \Delta x_i, u \sim Uniform(0,1) \quad (2)$$

其中, $\Delta x_i = \frac{x_{\max,i} - x_{\min,i}}{n_i}$ 是每个区间的长度, u 是随机偏移量。

生成的候选样本集 L_i 提供了一种理论上的分布均匀性基础,但与实际样本分布 D_i 并不完全一致。为解决这一问题,引入KD树进行修正。具体而言,遍历 L_i ,判断每个样本 L_i^j 是否存在于 D_i 中。如果存在,则将其存入回放样本集合 R 中;如果不存在,则利用 D_i 构建KD树 $Tree_i$,通过最近邻搜索找到距离 L_i^j 最近的旧样本 $near_j$:

$$near_j = \operatorname{argmin}_{x \in D_i} \|L_i^j - x\|_2 \quad (3)$$

其中, $\|\cdot\|_2$ 表示欧几里得范数,用于度量候选样本和旧样本之间的距离。随后,将 $near_j$ 作为回放样本存入 R 中。

通过上述策略选择的回放样本集合 R 能够广泛覆盖高维参数空间,并在旧样本集合 D_i 的实际分布中保持均匀性,这既保留了实际物理场景下的代表性,又在整体上反映了样本的多样性和分布合理性。在材料设计中,样本分布通常反映了材料性能与结构参数之间的隐式关联,这种关联决定了不同参数区域可能对应不同的材料特性。上述策略确保了回放样本覆盖材料性能变化的关键区域,避免了样本集中在某些特定范围可能导致的物理信息丢失。

完成对第 $T-1$ 组数据的记忆选择后,进入学习巩固阶段。将 R 与第 T 组数据 D_T 合并,以联合训练模型的权重参数 θ ,最大限度的保证BPNN在学习当前知识的同时缓解对先前知识的遗忘。训练过程包括前向传播和反向传播两个阶段。在前向传播阶段,通过权重 θ 计算预测值 $O_{R \cup D_T} = f(X_{R \cup D_T})$,根据预测值 $O_{R \cup D_T}$ 和真实标签值 $y_{R \cup D_T}$ 计算均方误差(mean squared error, MSE)损失函数:

$$\text{loss}_{R \cup D_T} = \frac{1}{n} \sum_{i=1}^n (y_{R \cup D_T}^i - O_{R \cup D_T}^i)^2 \quad (4)$$

其中, n 表示 $R \cup D_T$ 中样本数量。

在反向传播阶段,首先计算损失对权重的梯度:

$$\frac{\partial \text{loss}_{R \cup D_T}}{\partial \theta} = \frac{\partial \text{loss}_{R \cup D_T}}{\partial O_{R \cup D_T}} \cdot \frac{\partial O_{R \cup D_T}}{\partial \theta} \quad (5)$$

然后,利用梯度下降法:

$$\theta' = \theta - \eta \frac{\partial \text{loss}_{R \cup D_T}}{\partial \theta} \quad (6)$$

更新各层的权重以最小化损失函数,其中,BPNN的参数由 θ 更新至 θ' , η 表示学习率。

再选择学习巩固阶段通过重复执行第1、2阶段 S 次,多次进行记忆选择、学习巩固,BPNN可以接触到更广泛的样本空间,从而强化对旧知识的记忆。

持续学习策略的伪代码如算法1所示。

算法1 持续学习策略

- 1: 输入:第 T 组数据 D_T
- 2: 输入:回放样本数量 N
- 3: 输入:再选择学习次数 S
- 4: 输入:衰减率 λ
- 5: 输出: θ
- 6: 由 N 和 λ 计算每组的回放样本数 n_1, \dots, n_{T-1}
- 7: for $i \leftarrow 0$ to S
- 8: 初始化 R 为空
- 9: for $t \leftarrow 1, \dots, T-1$ do
- 10: 由 D_t 和 n_t 生成LHS样本集 L_t
- 11: for $j \leftarrow 0$ to n_t-1 do
- 12: if L_t^j in D_t then
- 13: $R = R \cup L_t^j$
- 14: else
- 15: 使用 D_t 生成KD树 $Tree_t$
- 16: 在 $Tree_t$ 中找到 L_t^j 的最近邻点 $near_j$
- 17: $R = R \cup near_j$
- 18: end for
- 19: end for
- 20: $\theta' \leftarrow \text{MSE}(\theta, R \cup D_T)$
- 21: end for

3 实验

3.1 数据集

为验证持续学习策略的有效性,本研究开展了一系列实验。实验所使用的数据集由动态数据生成与模型性能评估框架构建,包括73 889组MAMs样本数据。数据由超材料设计领域中广泛使用的仿真操作生成。仿真操作能够准确模拟真实材料系统的物理特性和规律,从而确保了数据生成的合理性。此外,如图5所示,数据集中所有样本的特征在取值范围内呈均匀分布,充分体现了样本分布的全面性和多样性。

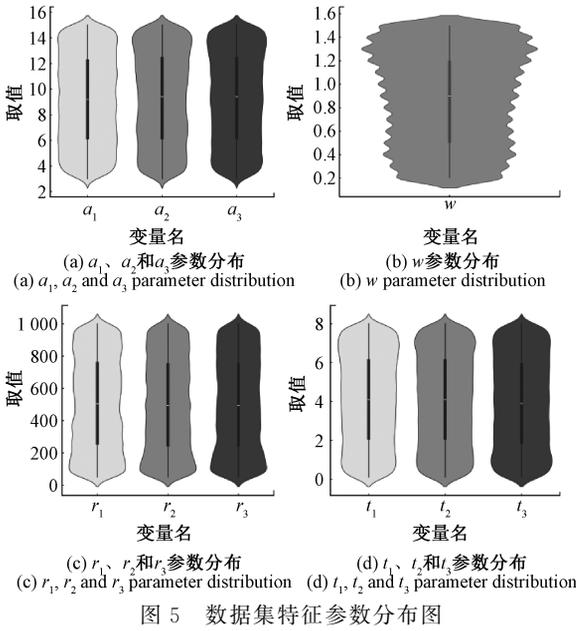


图 5 数据集特征参数分布图

将数据集静态划分为不同阶段,每个阶段内的样本具有相同的 w 取值,而不同阶段之间的 w 取值不同。从上述描述中可以看出不同阶段的数据分布是不同的。在每个阶段,数据按照 8 : 2 的比例划分为训练集和测试集。BPNN 在每个阶段的训练集上持续学习,而每阶段的测试集用于验证模型在该阶段上的性能。

3.2 评价指标

本文通过 5 个指标: MSE、平均准确率、平均遗忘率^[27]、前向迁移 (forward transfer, FWT) 和后向迁移 (backward transfer, BWT)^[13,28] 来评估方法的性能。值得注意的是,为适应回归任务场景,本文使用决定系数^[29] 替代 4 个指标中的准确率。判断性能时以决定系数接近 1 为目标,当决定系数小于 0 时,认为模型性能较差。为便于整体评价,将决定系数的最小值限制为 0。

通过计算模型在所有阶段测试集上的 MSE,以评估模型在整个学习后的整体性能。设 n 为所有测试集的样本总数, y_i 是实际值, \hat{y} 是模型的预测值,则 MSE 的数学表达为:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

平均准确率用于评估到目前为止模型的总体性能。设 $a_{k,j}$ 表示模型在经过第 1~ k 阶段的学习后,在第 j 阶段对应的测试集上的评估准确率,则平均准确率可以定义为:

$$A_k = \frac{1}{k} \sum_{j=1}^k a_{k,j} \quad (8)$$

平均遗忘率反映了模型在旧样本上的稳定性。设 $\max_{l \in \{1, \dots, k-1\}} a_{l,j}$ 表示模型学习到第 $k-1$ 个阶段时在第 j 阶段对应的测试集上的最佳准确率,那么最佳准确率和模型学习到第 k 个阶段时的准确率之差就是模型在经过 k 个阶

段的学习后对第 j 个阶段数据的遗忘程度,记为:

$$f_{k,j} = \max_{l \in \{1, \dots, k-1\}} a_{l,j} - a_{k,j} \quad (9)$$

则平均遗忘率为:

$$F_k = \frac{1}{k-1} \sum_{j=1}^{k-1} f_{k,j} \quad (10)$$

使用 FWT 评估模型在新样本上的学习能力,较大的 FWT 表明模型在新样本上具有较强的学习能力。使用第 j 个阶段的样本集训练一个随机初始化参考模型,其在第 j 个阶段对应的测试集上的准确率定义为 a_j^* ,则 FWT 定义为:

$$FWT_k = \frac{1}{k-1} \sum_{j=2}^k (a_{j-1,j} - a_j^*) \quad (11)$$

此外,通过 BWT 来评估新样本学习对模型在旧样本上表现的影响,当模型的 BWT 较大时,该模型被认为是可取的。BWT 被定义为:

$$BWT_k = \frac{1}{k-1} \sum_{j=1}^{k-1} a_{k,j} - a_{j,j} \quad (12)$$

3.3 实验设置

为验证模型性能,本文使用无持续学习策略的方法 (Base)、弹性权重巩固 (elastic weight consolidation, EWC)^[30]、不遗忘学习 (learning without forgetting, LwF)^[31]、记忆感知突触 (memory aware synapses, MAS)^[32]、基于经验回放的回归增量学习方法 (experience replay, ER) 和基于参数惩罚和经验回放的回归增量学习方法 (parameters penalty and experience replay, PPER)^[33] 以及全回放 (replay all memories method, RAMM) 方法进行对比。Base 方法作为基准方法,表示将上一个阶段学习得到的模型直接用于下一个阶段的学习而不采用任何持续学习方法;RAMM 方法对所有旧样本和新样本进行联合训练,代表持续学习的上限。

此外,为充分证明所提方法中各组件的有效性,本文进行了消融实验。在记忆选择阶段,“-LHS”表示以随机选择方式取代 LHS,进而评估 LHS 对结果的贡献;“-KD”表示当候选样本在旧样本中无法匹配时,用随机选择的方式来代替;“- n_i ”则表示直接从旧样本中选取 N 个回放样本,而不考虑每个阶段应选择的回放样本数,进而分析按阶段选择回放样本对结果的影响。在学习巩固阶段,研究新样本和回放样本联合训练与分别训练两种策略对结果产生的不同影响。在再选择学习巩固阶段,探讨在整个过程中仅进行一次回放样本选择对结果的影响。最后,本文通过实验分析,对最佳超参数的选择进行了深入分析。

为确保实验公平性,实验中模型的参数设置保持一致。具体包括:损失函数采用 MSE,优化器为 Adam 优化器,学习率设定为 0.000 1,训练批次大小设为 64。此外,为验证模型的泛化能力和鲁棒性,通过改变随机数种子,调整不同数据集的学习顺序,从而分析模型在不同数据集顺序下性

能的平均值和方差,研究模型的泛化性。对于超参数研究部分,本文仅在未设置随机数种子的条件下进行,以减少变量,使超参数调整的效果更加明显。

3.4 实验结果

1) 总体结果

在 MAMs 样本集上进行了与 7 种算法的对比实验,实验结果如表 1 所示。结果表明,在持续学习框架下,理想方法 RAMM 达到了理论性能的最优上限。然而,这种最佳性能是以高昂的计算成本为代价。随着数据集规模的线性增长,RAMM 的训练时间显著增加,限制了其在大规模数据处理中的应用。Base 方法的平均准确率仅为 0.4129 ± 0.0082 ,这表明不同阶段样本集的数据分布存在显著差异,进而导致模型在学习过程中出现灾难性遗忘现象。在基于正则化的方法中,包括 EWC、LwF 和 MAS,这三者的性能表现为 $MAS > EWC > LwF$;而在基于样本回放的方法中,包括 ER 和 PPER,这二者的性能表现为 $PPER > ER$ 。特别地,基于样本回放的方法整体性能显著优于基于正则化方法,表明样本回放策略在应对超材料领域的灾难性遗忘问题上具有更为直接且有效的作用。

本文提出的算法在所有测试集上的 MSE 为 0.2517 ± 0.0001 ,显著优于其他方法,且仅较 RAMM 高出 0.1229 ,表明该算法训练的模型在测试集上具有良好的拟合效果及

较强的泛化性能。同时,基于回放的方法,ER 和 PPER,包括本文提出的算法,在不同条件下 MSE 的方差都很小,三者的方差分别为 0.0045 、 0.0000 和 0.0001 ,这表明基于回放的方法在测试集上的表现较为稳定。进一步地,本文所提出的算法在平均准确率和平均遗忘率方面均优于其他对比算法,分别为 0.9328 ± 0.0001 和 0.0368 ± 0.0001 ,表明该算法在新知识学习和旧知识保持之间实现了良好的平衡,展现出较强的持续学习能力。在前向转移性能方面,本文算法的 FWT 达到了 -0.2768 ,表明该算法在学习新任务时能够较有效地利用先前任务中的知识。这主要归因于基于 LHS 的回放样本选择策略,该策略确保了回放样本在高维特征空间中的均匀分布,从而为模型提供了全面的先验知识支撑。然而,该算法的 FWT 方差为 0.0767 ,高于其他对比算法,反映出其在不同实验条件下的前向转移效果存在一定波动,这可能是由于本文算法训练的模型对任务间相关性差异较为敏感所致。尽管如此,综合其他性能指标,该算法仍展现出良好的综合能力,表明其在实际应用中具有较高的实用性与广阔的应用前景。此外,本文算法的 BWT 为 -0.0270 ± 0.0001 ,在所有对比算法中的表现最佳,表明其在学习新知识时对已学知识的干扰最小,进一步验证了其优异的持续学习能力。

表 1 不同对比算法的实验结果

Table 1 Experimental results of different comparison algorithms

算法	MSE	平均准确率	平均遗忘率	FWT	BWT
Base	4.8609 ± 8.0409	0.4129 ± 0.0082	0.5847 ± 0.0097	-0.6229 ± 0.0241	-0.5829 ± 0.0095
EWC	2.6603 ± 1.1016	0.4504 ± 0.0088	0.5468 ± 0.0101	-0.5724 ± 0.0006	-0.5452 ± 0.0100
LwF	4.2226 ± 3.0934	0.4179 ± 0.0089	0.5817 ± 0.0107	-0.6822 ± 0.0015	-0.5089 ± 0.0102
MAS	2.1061 ± 0.4087	0.4844 ± 0.0000	0.4390 ± 0.0001	-0.5434 ± 0.0055	-0.4383 ± 0.0001
ER	0.4626 ± 0.0045	0.7715 ± 0.0031	0.2022 ± 0.0037	-0.3583 ± 0.0016	-0.1951 ± 0.0037
PPER	0.3546 ± 0.0000	0.8198 ± 0.0003	0.1543 ± 0.0003	-0.3120 ± 0.0023	-0.1476 ± 0.0003
OUR	0.2517 ± 0.0001	0.9328 ± 0.0001	0.0368 ± 0.0001	-0.2768 ± 0.0767	-0.0270 ± 0.0001
RAMM	0.1288 ± 0.0000	0.9792 ± 0.0000	-0.0006 ± 0.0000	-0.1171 ± 0.0057	0.0191 ± 0.0000

2) 消融实验

消融实验结果如表 2 所示。结果表明,在记忆选择阶段,LHS 对模型性能提升最为显著。当使用随机选择回放样本时,模型的 MSE 为 0.3324 ,平均准确率仅为 0.8510 ± 0.0034 ;而基于 LHS 进行回放样本选择可以使 MSE 降低 0.0807 ,平均准确率提升 0.0818 。相比之下,KD 树和按阶段选择回放样本对模型性能的提升相对有限。尽管如此,通过 KD 树选择未匹配样本的最近邻样本,以及按阶段选择回放本来巩固先前知识,可以增加回放样本选择的合理性。

学习巩固阶段的联合训练对模型性能产生的影响最显著。由于新旧样本的分布可能存在不同,分离训练会导

致模型权重发生较大变化,各指标均受到显著负面影响。例如,MSE 由 0.2517 ± 0.0001 增至 0.5164 ± 0.0219 ,平均准确率从 0.9328 ± 0.0001 降至 0.7203 ± 0.0223 ,表明模型性能显著下降。此外,模型在不同条件下各指标的方差较大,例如 FWT 的方差为 0.0389 ,这表明分离训练削弱了模型的泛化能力,因此,联合训练对提升模型泛化性也有着积极影响。

再选择学习巩固阶段中的再选择同样对模型性能产生了显著影响。具体而言,MSE 增加了 0.2647 ,平均准确率由 0.9328 ± 0.0001 降至 0.8755 ± 0.0001 ,同时平均遗忘率增加了 0.0585 ,表明单次回放样本选择限制了模型的性能。此外,FWT 由 -0.2768 ± 0.0767 降至 $-0.4152 \pm$

0.115 2,这可能是由于在整个学习过程中持续使用相同的回放样本巩固记忆,导致对回放样本产生过拟合,进而影响了模型在新任务上的迁移能力。BWT 由 $-0.027 0 \pm$

$0.000 1$ 降至 $-0.087 3 \pm 0.000 1$,这可能是由于单次样本选择未能有效覆盖旧样本,导致模型难以有效保留旧任务的知识。

表 2 消融实验结果

Table 2 Ablation experiment results

阶段	组件	MSE	平均准确率	平均遗忘率	FWT	BWT
记忆选择阶段	-LHS	$0.332 4 \pm 0.003 4$	$0.851 0 \pm 0.003 4$	$0.121 0 \pm 0.003 9$	$-0.572 6 \pm 0.124 1$	$-0.112 9 \pm 0.003 9$
	-KD	$0.230 2 \pm 0.000 0$	$0.931 7 \pm 0.000 1$	$0.037 2 \pm 0.000 1$	$-0.273 8 \pm 0.034 9$	$-0.027 0 \pm 0.000 1$
	$-n_i$	$0.239 7 \pm 0.000 2$	$0.929 9 \pm 0.000 3$	$0.040 6 \pm 0.000 3$	$-0.191 3 \pm 0.023 8$	$-0.030 6 \pm 0.000 3$
学习巩固阶段	-联合训练	$0.516 4 \pm 0.021 9$	$0.720 3 \pm 0.022 3$	$0.263 5 \pm 0.026 1$	$-0.392 1 \pm 0.038 9$	$-0.253 3 \pm 0.026 1$
再选择学习巩固阶段	-再选择	$0.324 2 \pm 0.003 4$	$0.875 5 \pm 0.000 1$	$0.095 3 \pm 0.000 1$	$-0.415 2 \pm 0.115 2$	$-0.087 3 \pm 0.000 1$

3)超参数影响

再选择学习巩固次数 S 是影响持续学习效果的一个重要参数,因此,本文对不同的 S 值进行实验,实验结果如表 3 所示。

表 3 不同再选择学习巩固次数的实验结果

Table 3 Experimental results of different re-selection learning consolidation times

S	算法	平均准确率/%	平均遗忘率/%	FWT	BWT
50	EWC	27.87	71.55	-4.65	-71.55
	LwF	25.35	67.11	-12.68	-67.11
	MAS	31.55	67.58	-1.51	-67.58
	Base	33.81	65.12	-1.36	-65.12
	PPER	59.42	38.06	-0.66	-37.61
	Our	86.87	8.69	0.19	-8.42
	RAMM	97.57	0.30	1.35	1.71
100	EWC	22.37	78.28	-6.48	-78.28
	LwF	34.73	62.62	-7.73	-58.12
	MAS	20.90	79.45	-9.84	-79.45
	Base	21.76	78.97	-9.78	-78.97
	PPER	56.04	42.50	1.19	-42.05
	Our	90.57	5.39	-0.12	-4.93
	RAMM	97.94	0.26	2.75	1.41
150	EWC	31.21	68.99	-4.38	-68.99
	LwF	31.76	63.16	-8.01	-61.43
	MAS	29.66	70.72	-1.69	-70.72
	Base	38.42	61.71	-2.63	-61.7
	PPER	59.16	38.77	-0.65	-38.64
	Our	91.98	4.80	0.6	-4.01
	RAMM	98.17	0.12	2.64	1.67

从结果中可以明显看出,随着 S 的增加,大多数对比模型的各项指标逐渐趋向理想状态。RAMM 在不同设置

下的实验结果表明,S 的大小对 BPNN 的性能有显著影响。具体而言,较大的 S 值有助于提升模型的性能表现。此外,当 S 从 50 增至 150 时,算法的平均准确率提升了 5.11%,遗忘率降低了 3.89%。根据以上实验结果,最终设定 S 为 150。

此外,回放样本数 N 同样直接影响持续学习的效果。本文分别在 N 为 100、200、300 的条件下进行了实验,实验结果如表 4 所示。可以明显看出,随着 N 的增加,大多数指标逐渐趋向最优状态。实际上,较高的 N 能够为 BPNN 提供更多旧知识进行巩固。然而,随着 N 的增加,模型逐渐偏离持续学习的范畴,并且其时间和空间计算复杂度也呈线性增长。因此,综合考虑后,本文将 N 设定为 300。

表 4 不同回放样本数的实验结果

Table 4 Experimental results with different replay sample numbers

N	平均准确率/%	平均遗忘率/%	FWT	BWT
100	89.38	6.93	0.45	-6.48
200	90.18	6.29	0.19	-5.71
300	91.98	4.80	0.6	-4.01

衰减率 λ 控制了从最早阶段到最近阶段回放样本数量的衰减程度,从而影响对不同阶段记忆的巩固强度。当 $\lambda=0$ 时,每个阶段选择的回放样本数量相同;随着 λ 的增加,衰减程度逐渐加大。本文对 λ 的不同取值进行了实验研究,BPNN 在各阶段保留数据集上的平均准确率如图 6 所示。实验结果表明,当每个阶段选择的回放样本数量相同,即 $\lambda=0$ 时,由于对较早阶段记忆的巩固不足,模型在这些阶段的性能较差,平均准确率仅有 68.07%。随着 λ 的增加,更多的回放样本被分配到较早阶段以进一步巩固记忆,模型在这些阶段的性能也逐渐提高。然而,由于总回放样本数有限,当更多的样本被分配到较早阶段时,分配到最近阶段的样本数量将减少,从而导致模型在最近阶

段的性能下降。因此,选择合适的 λ 值至关重要。实验结果表明, $\lambda = 0.7$ 是较为合适的取值。

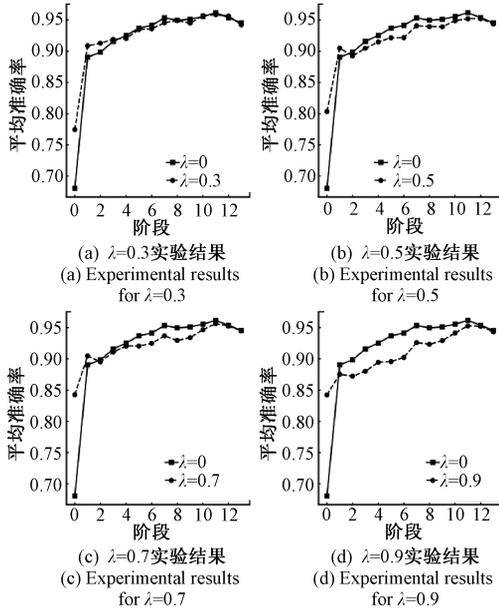


图6 不同 λ 的实验结果

Fig. 6 Experimental results with different λ

4 结 论

本文提出一种动态数据生成与模型性能评估框架,用与解耦数据集规模与模型性能之间的耦合关系,高效获取了目标精度所需的数据集规模。同时,设计了一种基于三阶段的持续学习策略用于解决模型性能动态评估过程中出现的灾难性遗忘问题。实验结果表明,本文提出的持续学习策略在实现较高预测平均准确率的同时,有效降低了平均遗忘率。未来研究将基于动态数据生成与模型性能评估框架生成的数据集,进一步探索深度学习模型的优化与设计。

参考文献

[1] QIAN CH, KAMINER I, CHEN H SH. A guidance to intelligent metamaterials and metamaterials intelligence[J]. Nature Communications, 2025, 16(1): 1154.

[2] SHIRKE N, GHASE V, JAMDAR V. Recent advances in stealth coating[J]. Polymer Bulletin, 2024, 81(11): 9389-9418.

[3] HOU S, WANG Y, GAO F, et al. A novel approach to electromagnetic wave absorbing material design: Utilizing nano-antenna arrays for efficient electromagnetic wave capture[J]. Chemical Engineering Journal, 2023, 471: 144779.

[4] BHARGAVA D, RATTANADECHO P, JIAMIROCH K. Microwave imaging for breast

cancer detection-a comprehensive review [J]. Engineered Science, 2024, 30: 1116.

- [5] JAIN P, CHHABRA H, CHAUHAN U, et al. Machine learning assisted hepta band THz metamaterial absorber for biomedical applications[J]. Scientific Reports, 2023, 13(1): 1792.
- [6] 石明宽,丁传仓,王锐,等.面向变工况下工业流数据故障诊断的持续迁移学习系统[J].仪器仪表学报,2024, 45(4):10-16.
- SHI M K, DING CH C, WANG R, et al. Continuous transfer learning system for fault diagnosis of industrial stream data under variable operating conditions [J]. Chinese Journal of Scientific Instrument, 2024, 45(4):10-16.
- [7] WANG J, FAN C, LIAO Y, et al. Pattern-to-absorption prediction for multilayered metamaterial absorber based on deep learning[J]. IEEE Microwave and Wireless Technology Letters, 2024, 34(5): 463-466.
- [8] ZHANG H, WANG Y, LU K, et al. SAP-Net: Deep learning to predict sound absorption performance of metaporous materials[J]. Materials & Design, 2021, 212: 110156.
- [9] CAO E, DONG Z, JIA B, et al. Inverse design of isotropic auxetic metamaterials via a data-driven strategy [J]. Materials Horizons, 2025, 12(13): 4884-4900.
- [10] ZHANG H, KANG L, CAMPBELL S D, et al. Fixed-attention mechanism for deep-learning-assisted design of high-degree-of-freedom 3D metamaterials [J]. Optics Express, 2025, 33(9): 18928-18937.
- [11] CHEN S, ZHAO C, WANG W, et al. Convolutional neural network-assisted design and validation of terahertz metamaterial sensor [J]. Materials & Design, 2025, 253: 13871.
- [12] 岳朋,张建刚,邱克鹏,等.基于深度学习弹性超材料带隙逆向设计研究[J].力学学报,2025,57(1):103-115.
- YUE P, ZHANG J G, QIU K P, et al. Research on reverse design of band gap for elastic metamaterials based on deep learning [J]. Chinese Journal of Theoretical and Applied Mechanics, 2025, 57(1): 103-115.
- [13] WANG L, ZHANG X, SU H, et al. A comprehensive survey of continual learning: Theory, method and application [J]. IEEE Transactions on pattern analysis and Machine Intelligence, 2024, 46(8): 5362-5383.
- [14] 王家亮,王景成,李继超.基于增量学习树模型的带

- 钢精轧宽度预测[J]. 锻压技术, 2024, 49(2): 152-160.
- WANG J L, WANG J CH, LI J CH. Prediction on strip finish rolling width based on incremental learning tree model[J]. Forging & Stamping Technology, 2024, 49(2): 152-160.
- [15] DOHARE S, HERNANDEZ GARCIA J F, LAN Q, et al. Loss of plasticity in deep continual learning[J]. Nature, 2024, 632(8026): 768-774.
- [16] LI J, WANG Z, ZHANG S, et al. Task incremental learning-driven digital-twin predictive modeling for customized metal forming product manufacturing process [J]. Robotics and Computer-Integrated Manufacturing, 2024, 85: 102647.
- [17] KLEIMAN A, DZIUGAITE G K, FRANKLE J, et al. Soup to go: Mitigating forgetting during continual learning with model averaging[J]. ArXiv preprint arXiv:2501.05559, 2025.
- [18] HAYES T L, KRISHNAN G P, BAZHENOV M, et al. Replay in deep learning: Current approaches and missing biological elements[J]. Neural computation, 2021, 33(11): 2908-2950.
- [19] ROLNICK D, AHUJA A, SCHWARZ J, et al. Experience replay for continual learning[C]. NeurIPS 2019: 33rd Conference on Neural Information Processing Systems. Vancouver: Curran Associates, 2019.
- [20] LIN H. ER-FSL: Experience replay with feature subspace learning for online continual learning[J]. ArXiv preprint arXiv:2407.12279, 2024.
- [21] LIU L, LIU L, CUI Y. Prior-free balanced replay: Uncertainty-guided reservoir sampling for long-tailed continual learning [C]. 32nd ACM International Conference on Multimedia, 2024: 2888-2897.
- [22] RAHMAN M S, COULL S, YU Q, et al. MADAR: Efficient continual learning for malware analysis with diversity-aware replay [J]. ArXiv preprint arXiv: 2502.05760, 2025.
- [23] 李斌, 郭辉. 基于多样性样本回放的类增量学习方法[J]. 计算机工程与设计, 2024, 45(5): 1391-1397.
- LI X, GUO H. Class-incremental learning method based on diversity sample replay [J]. Computer Engineering and Design, 2024, 45(5): 1391-1397.
- [24] FAN R, LI D, LIAO W, et al. A pattern design strategy for microwave-absorbing coding metamaterials with tortuosity and connectivity[J]. Composites Part B: Engineering, 2024, 287: 111880.
- [25] BOURCET J, KUBILAY A C C, DEROME D, et al. Representative meteorological data for long-term wind-driven rain obtained from latin hypercube sampling-application to impact analysis of climate change[J]. Building and Environment, 2023, 228: 109875.
- [26] VORECHOVSKY M. Hierarchical refinement of latin hypercube samples [J]. Computer-Aided Civil and Infrastructure Engineering, 2015, 30(5): 394-411.
- [27] CHAUDHRY A, DOKANIA P K, AJANTHAN T, et al. Riemannian walk for incremental learning: Understanding forgetting and intransigence [C]. ECCV 2018: European Conference on Computer Vision. Munich: Springer, 2018: 532-547.
- [28] LOPEZ-PAZ D, RANZATO M. Gradient episodic memory for continual learning [C]. NIPS 2017: Advances in Neural Information Processing Systems 30. Long Beach: Curran Associates, 2017.
- [29] CHICCO D, WARRENS M J, JURMAN G. The coefficient of determination r-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation [J]. PeerJ Computer Science, 2021, 7: e623.
- [30] KIRKPATRICK J, PASCANU R, RABINOWITZ N, et al. Overcoming catastrophic forgetting in neural networks[J]. National Academy of Sciences, 2017, 114(13): 3521-3526.
- [31] LI Z, HOIEM D. Learning without forgetting[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(12): 2935-2947.
- [32] ALJUNDI R, BABILONI F, ELHOSEINY M, et al. Memory aware synapses: learning what (not) to forget[C]. ECCV 2018: European Conference on Computer Vision. Munich: Springer, 2018: 139-154.
- [33] 王弘业, 钱权, 武星. 基于参数惩罚和经验回放的材料吸声系数回归增量学习[J]. 工程科学学报. 2023, 45(7): 1225-1231.
- WANG H Y, QIAN Q, WU X. Incremental learning of material absorption coefficient regression based on parameter penalty and experience replay[J]. Chinese Journal of Engineering, 2023, 45(7): 1225-1231.

作者简介

陈祺东, 博士, 讲师, 主要研究方向为自然语言处理、远程监督关系抽取、群体智能优化。

徐国梁, 硕士研究生, 主要研究方向为深度学习、关系抽取。

邓瑞翔, 博士, 副研究员, 主要研究方向为超材料、涂层材料。

吴豪(通信作者), 博士, 讲师, 主要研究方向为深度学习、计算机视觉。

E-mail: 860415@cwXu.edu.cn