

DOI:10.20176/j.cnki.nxdz.20251208

融合持续学习策略的近红外光谱煤炭挥发分检测方法

武治峰^{1,2}, 陈海林¹, 叶金燕³, 赵金秋¹, 邹亮^{2*}

(1. 中国检验认证集团河北有限公司, 河北 石家庄 050071; 2. 中国矿业大学 信息与控制工程学院, 江苏 徐州 221116;
3. 湛江海关技术中心, 广州 湛江 524022)

摘要:煤炭挥发分是评价煤质与燃烧特性的重要指标,其含量直接影响燃烧效率、反应活性及工艺适应性。近红外光谱作为一种快速、无损的检测手段,已被广泛用于煤炭挥发分的定量预测。然而,在跨批次应用中,光谱分布常发生系统性漂移,若直接以新数据重新建模,模型将丧失对旧样品的预测能力,进而加剧模型更新过程中的“灾难性遗忘”问题。针对这一挑战,提出一种面向煤炭挥发分预测的持续学习建模方法:特征提取阶段,构建融合密集连接与自注意力机制的网络结构,以充分捕获复杂光谱的局部细节与全局依赖;模型更新阶段,引入结合统计特征回放与知识蒸馏的持续学习策略,在无需访问原始历史样本的前提下,实现旧知识保持与新任务适应。基于3个批次煤样的多阶段实验表明,所提方法可有效缓解“灾难性遗忘”,在分布漂移条件下保持稳定预测性能,其整体精度接近单任务独立建模的最优水平,为复杂工业环境下煤炭挥发分的长期稳定检测提供了可行技术途径。

关键词:煤质;近红外光谱;持续学习;特征回放;知识蒸馏

分类号:(中图)O657.33;TQ533

文献标志码:A

煤炭作为我国一次能源的主体,长期在能源消费结构中占据50%以上的比重,在保障国家能源安全和经济社会发展中发挥着不可替代的作用^[1]。国家“十四五”规划和“双碳”战略均明确提出,要在确保能源安全的前提下,推动煤炭清洁高效利用,实现能源体系的绿色低碳转型^[2-3]。煤质分析是煤炭清洁利用的重要基础,其中挥发分含量不仅决定煤炭的燃烧特性和热值利用效率,还直接影响其在储运安全、气化、焦化等工艺环节中的适应性^[4]。因此,建立快速、准确的挥发分检测方法,对支撑煤炭合理分级及绿色高效利用具有重要意义。

传统的挥发分测定方法主要依据国家标准(如《煤的工业分析方法》(GB/T 212—2008)^[5])执行,即通过对煤样进行加热处理并计算其失重量来获取检测结果。该方法虽具备较高的测定精度,但检测周期长且会对样品造成不可逆的严重破坏,难以满足现场快速检测的实际需求。近年来,激光诱导

击穿光谱、X射线荧光光谱、微波等新兴检测技术被逐步应用于煤质分析^[6-10],这些技术虽在提升自动化程度和检测速度方面取得了一定进展,却普遍存在设备造价昂贵、对样品状态敏感以及现场适应性弱等短板。相比之下,近红外光谱技术兼具检测速度快、制样要求低和无损测量的突出优势,已在农业^[11-12]、生物医药^[13-14]、石油化工^[15-16]等多个领域被广泛应用,为实现煤中挥发分的高效检测提供了全新途径^[17]。然而,由于煤样颜色深、吸光性强且矿物组分复杂,其采集到的光谱信号常伴随谱峰混叠与非线性干扰问题,进而增加了光谱特征提取与建模的难度。传统机器学习方法在表征光谱数据的非线性关系和处理高维复杂光谱特征方面存在明显局限;深度学习方法虽展现出较强的自动特征提取能力和非线性拟合能力^[18-19],但在跨批次建模中仍难以应对数据分布差异引发的模型预测性能退化问题。

收稿日期:2025-09-21

基金项目:国家自然科学基金资助项目(62473368,62373360);中国检验认证集团河北有限公司(2025ZJHBYF004-1);海关总署科研项目(2023HK113)

作者简介:武治峰(1982—),男,工程师,硕士研究生,主要从事煤炭、矿石、金属等大宗商品的检验检测以及智慧检测、数字化转型等研究,(电子信箱)wuzhifeng1982@126.com。

*通信联系人:邹亮(1987—),男,副教授,博士,主要从事计算机视觉、智能语音及无损检测研究,(电子信箱)liangzou@cumt.edu.cn。

引用格式:武治峰,陈海林,叶金燕,等.融合持续学习策略的近红外光谱煤炭挥发分检测方法[J].宁夏大学学报(自然科学版中英文),2025,46(4):418-427.

在实际生产环境中,由于煤样来源多样且采集条件存在差异,其光谱分布往往会随着时间推移及应用场景变化而产生系统性漂移;同时,煤质数据多具有商业敏感性,历史数据难以共享或长期保存,导致模型在迭代更新时无法直接调用旧数据。若仅凭新数据顺序训练,模型易发生“灾难性遗忘”,即在旧任务上的知识与预测能力显著退化,从而削弱模型的长期可用性与稳定性。持续学习(又称增量学习、终身学习)为缓解该问题提供了有效途径:一类方法通过参数重要性约束抑制遗忘(如弹性权重巩固(elastic weight consolidation, EWC)与记忆感知突触(memory aware synapses, MAS));一类方法通过经验回放与蒸馏保持历史知识(如暗经验回放(dark experience replay, DER));另有研究尝试在无历史样本的条件下,以特征整合实现知识巩固(如弹性特征整合(elastic feature consolidation, EFC))^[20-23]。然而,如何将上述持续学习思想与煤炭近红外光谱建模有效结合,并在历史数据不可访问的约束下,同时兼顾旧知识保持与新任务适应,目前仍缺乏系统性、工程可行的解决方案。

针对上述问题,本文提出一种面向煤炭挥发分预测的持续学习建模框架。在特征提取环节,设计融合密集连接与自注意力机制的网络结构(dense-transformer network, DT-Net),可联合捕获光谱的局部细节与全局依赖,有效缓解谱峰混叠与特征信号微弱问题;在模型更新环节,提出结合统计特征回放与知识蒸馏的持续学习策略(statistical replay with distillation, SRD),在不依赖完整历史数据的前提下,实现旧知识留存与新分布数据的适配。基于多批次的煤炭近红外光谱数据的实验结果表明,该方法能在分布漂移条件下有效缓解遗忘,并在新任务上保持稳定预测性能,其整体表现接近单任务独立训练的最优基准。本研究不仅验证了该方法在煤炭挥发分预测中的有效性,更为复杂工业场景下的光谱持续建模与智能检测提供了理论与实践参考。

1 数据与方法

1.1 样本采集

为保证实验结果的可重复性与方法验证的有效性,本研究在制样与光谱采集环节严格控制环境和操作条件,使分布差异主要来源于煤样产地和批次,而非外部噪声因素。煤样采集与制备严格遵循《煤样的制备方法》(GB 474—2008)^[24]。具体过程包括破碎、干燥、研磨与筛分,最终使样品粒度小于0.2 mm,以确保样品均匀性与代表性。样品制备完

成后,依据《煤的工业分析方法》(GB/T 212—2008)对煤样进行工业分析^[5],测定其挥发分含量。

在工业分析的同时,利用近红外光谱仪同步采集煤样的光谱数据。测试条件为室温、环境湿度不大于40%,并避免日光直射。光谱采集波段范围为900~1700 nm,探头与样品表面保持5 mm固定距离,以减少散射效应和光程波动。为降低样品表面不均匀性带来的误差,每个煤样在5个不同位置独立扫描,并对测量结果取算术平均,作为该样品的最终光谱数据。

为模拟持续学习场景下数据来源差异引发的分布漂移,本研究将样本划分为3个批次,分别记作D0(284个样本)、D1(247个样本)和D2(232个样本)。图1给出了3个批次煤样的近红外光谱对比结果,可以看出光谱整体形态较为一致,主要吸收峰位置保持稳定,但在吸收强度和局部细节上存在一定差异,且不同批次在部分区间表现出重叠现象,说明样本分布既非完全一致,也非完全分离。

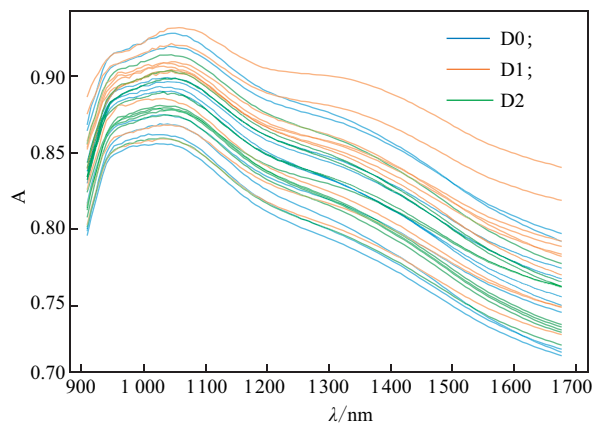


图1 3个批次样本的光谱对比图

进一步地,图2展示了3个批次煤样的挥发分真值分布情况,结果表明D0与D1的分布区间相对接近,而D2整体分布明显右移并更加分散,反映出跨批次条件下的参数差异。

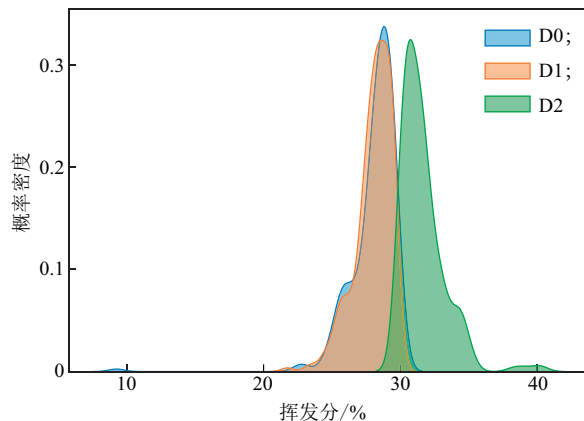


图2 3个批次样本的挥发分真值分布图

预测,实现从光谱数据到目标指标的端到端建模。

1.2.2 SRD 为应对煤近红外光谱建模中数据分布漂移与历史数据不可访问而导致的“灾难性遗忘”与模型性能下降问题,本文在DT-Net的基础上设计了一种持续学习策略SRD。该策略通过存储并重构历史样本的分布统计量,近似还原旧任务的

特征空间;同时在新任务训练过程中引入教师-学生式蒸馏约束,以维持模型在预测结果与中间表征上的一致性,从而在学习新任务的同时保持对旧任务的泛化能力。SRD的整体框架如图4所示,主要由特征提取器、统计特征生成回放模块、知识蒸馏模块以及回归优化模块组成。

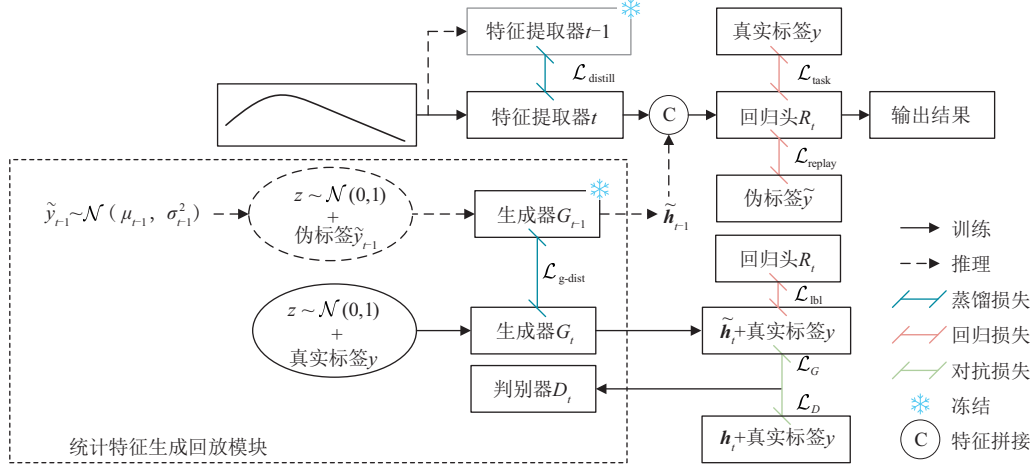


图4 统计特征回放与知识蒸馏协同的持续学习流程图

在持续学习场景下,煤炭光谱数据多受商业保密约束,历史样本难以直接获取,导致模型在更新过程中易遗忘先前分布。为此,SRD引入统计特征生成回放模块。该模块的核心思想是存储并重构历史数据的分布统计量,近似模拟旧任务的特征空间,从而实现知识回放。设当前为任务 t 阶段, $t-1$ 阶段对应标签为 $\{y_{t-1,j}\}_{j=1}^m$,记录标签分布的均值 μ_{t-1} 与标准差 σ_{t-1} , m 为 $t-1$ 阶段的样本数量。在新阶段训练时,从该统计量采样伪标签 $\tilde{y}_{t-1} \sim \mathcal{N}(\mu_{t-1}, \sigma_{t-1}^2)$,再与随机噪声 $z \sim \mathcal{N}(0, 1)$ 联合输入冻结的生成器 G_{t-1} ,以合成旧任务风格的深层特征 $\tilde{h}_{t-1} = G_{t-1}(z, \tilde{y}_{t-1})$ 。这一机制能够在无需访问原始样本的情况下,逼近旧任务的特征空间分布,为后续模型训练提供符合旧任务数据特征的数据支持。

为使合成特征在可分性和标签一致性上逼近历史分布,SRD在回放模块中引入投影条件判别器 D_t ,并结合WGAN-GP(Wasserstein GAN with gradient penalty)进行优化。WGAN-GP在Wasserstein距离的基础上,通过在判别器 D_t 中引入梯度惩罚函数,有效缓解了生成对抗训练的不稳定性,并提升了特征分布的逼近质量;生成器的优化目标则为最大化判别器对伪样本的评分。生成器 G_t 的WGAN-GP损失(\mathcal{L}_G)计算式为

$$\mathcal{L}_G = -\frac{1}{n} \sum_{i=1}^n D(\tilde{h}_{t,i}, y_i) \quad (1)$$

为进一步保证生成特征的分布一致性,引入正则化回归一致性损失,利用当前主模型的回归头 $R(\cdot)$ 逼近真实标签 y 。则有

$$\mathcal{L}_{lbi} = \frac{1}{n} \sum_{i=1}^n (R(\tilde{h}_{t,i}) - y_i)^2 \quad (2)$$

此外,在任务切换时,利用冻结的生成器 G_{t-1} 对当前生成器 G_t 施加约束,使其在旧任务伪标签 \tilde{y}_{t-1} 条件下输出特征保持一致,从而引入生成器蒸馏损失。即

$$\mathcal{L}_{g-dist} = \frac{1}{m} \times \sum_{j=1}^m \|G_t(z_j, \tilde{y}_{t-1,j}) - G_{t-1}(z_j, \tilde{y}_{t-1,j})\|_2^2 \quad (3)$$

因此,生成器的整体优化目标为

$$\mathcal{L}_G^{total} = \mathcal{L}_G + \lambda_{lbi} \mathcal{L}_{lbi} + \lambda_{dist} \mathcal{L}_{g-dist} \quad (4)$$

其中: λ_{lbi} 、 λ_{dist} 分别为一致性与蒸馏损失的权重系数。

在主模型更新阶段,将生成的旧任务特征与当前批次的真实特征拼接后输入回归头,联合驱动回归预测,计算对应的回放损失(\mathcal{L}_{replay})和当前任务损失(\mathcal{L}_{task})。

尽管统计特征回放能够在特征层面维持旧任务知识,但若仅依赖该机制仍可能导致预测层偏移。为此,SRD进一步引入主模型蒸馏机制:将上一任务训练完成后冻结的特征提取器 $f_{t-1}(\cdot)$ 作为教师网络,当前更新的特征提取器 $f_t(\cdot)$ 作为学生网络;

对于相同输入样本 x_i , 教师网络与学生网络的预测分别为 $f_{i-1}(x_i)$ 与 $f_i(x_i)$ 。蒸馏损失函数定义为

$$\mathcal{L}_{\text{distill}} = \frac{1}{n} \sum_{i=1}^n (f_{i-1}(x_i) - f_i(x_i))^2. \quad (5)$$

该约束保证学生模型在学习新任务分布的同时, 维持对旧任务的预测一致性。最终, 主模型的优化目标由当前任务监督损失、回放损失与蒸馏损失联合构成。即

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{task}} + \mathcal{L}_{\text{replay}} + \alpha_t \cdot \mathcal{L}_{\text{distill}}. \quad (6)$$

其中: α_t 为动态蒸馏系数, 在训练过程中逐步提升, 以避免对新任务学习造成过强干扰。

通过这种联合优化机制, SRD 能在不依赖完整历史数据的条件下, 实现旧知识保留与新知识吸收的平衡, 有效提升了煤近红外光谱建模的长期稳定性与跨批次泛化能力。

1.3 实验设计与评价指标

为验证所提出的 DT-Net+SRD 持续学习框架的有效性, 本文基于数据集 D0、D1 和 D2 设计了多阶段实验流程, 以考察模型在数据分布漂移条件下的适应性与抗遗忘能力。实验流程分为 3 个阶段: 第一阶段以数据集 D0 训练主模型 DT-Net, 获得初始特征提取与回归预测能力; 第二阶段引入数据集 D1, 对模型进行持续更新, 同时启用统计特征生成回放模块与主模型蒸馏约束, 以保证旧任务特征分布的稳定重构; 第三阶段在数据集 D2 上继续训练, 流程同样包括 3 部分: ① 新任务数据上的回归优化; ② 统计特征生成回放提供历史样本代理; ③ 知识蒸馏机制约束主模型预测一致性。最终, 主模型的损失函数由当前任务监督损失、历史特征回放损失和知识蒸馏损失加权构成, 实现了在新任务适应的同时保持旧任务性能。

本文使用均方根误差 (root mean square error, RMSE)、平均绝对误差 (mean absolute error, MAE) 作为评价指标。

RMSE 计算预测值与实际观测值之间差异的平均平方根, 其计算公式为

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2}. \quad (7)$$

MAE 计算预测值与实际观测值之间绝对差异的平均值, 其计算公式为

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i|. \quad (8)$$

其中: n 为样本数量; y_i 和 y'_i 分别表示第 i ($i=1, 2, \dots, n$) 个样本的真实值和预测值。

2 结果与分析

2.1 实验结果分析

为验证 DT-Net+SRD 持续学习框架在煤近红外光谱挥发分建模中的有效性, 本节对不同阶段数据集上的建模结果进行分析, 重点评估其在分布漂移条件下的两项核心能力: ① 保持旧任务预测精度, 以缓解“灾难性遗忘”; ② 引入新任务数据时, 保持对新分布的适应能力, 并避免模型性能显著退化。实验采用阶段性回测策略, 即在完成当前阶段的任务训练后, 对所有已学习任务进行联合评估, 以此量化遗忘程度并检验其跨阶段泛化能力。

表 1 给出了不同阶段模型在各数据集上的预测结果。第一阶段在 D0 上取得 RMSE=1.115 6, MAE=0.814 8 的结构, 展现出良好的初始建模能力。第二阶段引入 D1 后, 模型在新任务上的预测精度为 RMSE=1.036 4, MAE=0.870 0; 同时 D0 对应的误差仅小幅上升至 1.347 9/0.941 0, 说明在轻度分布漂移条件下, SRD 能够有效抑制遗忘并保持较强的适应性。进入第三阶段, 由于 D2 与前两批数据的分布差异显著, 3 个数据集的预测误差整体上升, 其中整体合并测试结果为 RMSE=1.909 9, MAE=1.454 7。尽管如此, 模型在各阶段均未出现性能急剧退化的情况, 表明统计回放与蒸馏机制能够在不同漂移强度下维持相对稳定的表现。在弱漂移场景下, 框架展现出较强的稳定性, 具有良好的知识保持与迁移能力; 在强漂移场景下, 虽然预测精度有所下降, 但整体依然保持可塑性, 能够在新任务上获得一定适应性, 并维持跨阶段的全局稳定性。

表 1 不同阶段数据集预测结果

指标	阶段/数据集	RMSE	MAE
	阶段 1/D0	1.115 6	0.814 8
	阶段 2/D0	1.347 9	0.941 0
	阶段 2/D1	1.036 4	0.870 0
挥发分/%	阶段 3/D0	1.967 3	1.462 4
	阶段 3/D1	1.787 8	1.504 4
	阶段 3/D2	1.861 9	1.393 8
	整体数据集	1.909 9	1.454 7

为进一步验证算法的稳定性及优化特性, 图 5 展示了 DT-Net+SRD 在 3 个阶段任务上的训练与验证损失曲线。可以观察到, 本文方法在各阶段均能较少迭代实现稳定收敛, 训练与验证曲线整体平

滑,说明优化过程稳定可靠。在第三阶段任务(图 5 (c))中,训练曲线出现波动,这是模型在新任务分布下重新平衡旧知识约束与新任务学习目标的结果。随着统计特征回放与知识蒸馏机制逐步达到平衡

后,模型很快收敛并趋于稳定。这一过程表明,本文方法能够在任务切换时实现参数的渐进式平衡,在有效避免“灾难性遗忘”的同时,保持整体收敛性。

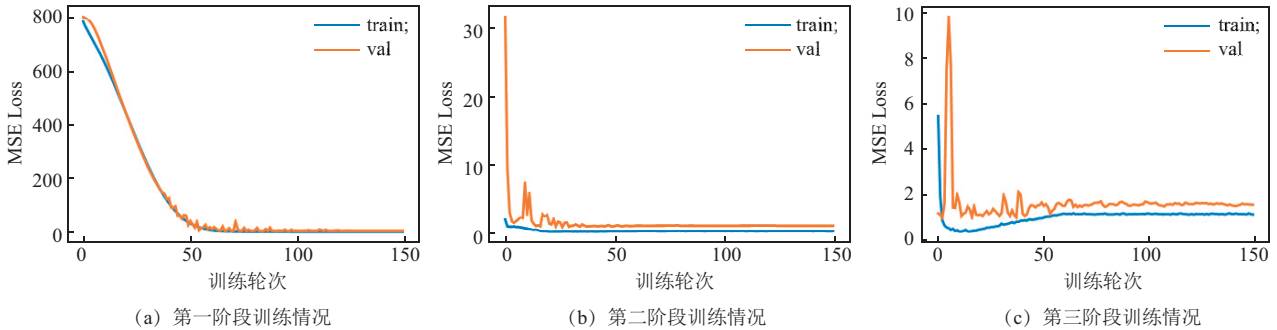


图 5 不同阶段任务的训练与验证收敛曲线

2.2 消融实验分析

为进一步验证 SRD 各组成模块的有效性,本文在 DT-Net 基础上设计了 5 种对比方案:①完整框架 DT-Net+SRD;②仅含基础网络的 DT-Net;③加入统计特征回放但去除主模型蒸馏的 DT-Net+Replay;④加入主模型蒸馏但不包含统计特征回放的 DT-Net+Distill;⑤各数据集单独训练与测试的 Oracle,作为理论上的最优参考。其中,DT-Net+SRD 作为最终模型,用于同时考察统计回放与蒸馏机制的协同作用;Oracle 模型则不涉及持续学习与跨阶段知识保持,因而代表了各阶段任务在独立建模条件下可达到的性能上限。

在持续学习场景中,核心目标是兼顾旧知识保持与新任务适应能力。因此,本节重点分析第三阶段及整体数据集的实验结果。一方面,第三阶段引入的数据集 D2 与前两批次数据存在显著分布差异,更能体现“灾难性遗忘”与分布漂移的影响;另一方面,整体数据集的测试结果能够综合表征模型在跨阶段任务中的全局泛化性能。

图 6 展示了各方法在第三阶段对 D0、D1(旧任务)、D2(新任务)及整体测试集(All)的 RMSE 表现。可以看出,Oracle 在所有数据集上均保持最低 RMSE,代表理论最优性能。相比之下,基础网络 DT-Net 在 D0 与 D1 上出现明显遗忘,在整体数据集上误差较高,说明顺序训练难以适配持续学习场景,稳定性较弱。单一模块的改进效果有限:DT-Net+Replay 在旧任务上较 DT-Net 略有改善,但在 D2 上 RMSE 极低,表明仅依赖回放容易偏向当前任务,反而加剧对旧任务的遗忘;DT-Net+Distill 能在 D0、D1 上维持较低误差,但在 D2 上 RMSE 急剧上升,表明缺少回放支撑时,

单纯蒸馏虽然增强了旧知识的稳定性,却严重削弱了模型对新任务的可塑性。相较之下,DT-Net+SRD 在 D0、D1、D2 上均保持相近的误差水平(无极端波动),并在整体数据集上取得最低 RMSE,展现出旧知识保持与新任务适应之间的良好权衡,以及更强的全局稳定性。消融实验结果表明,统计特征回放与主模型蒸馏在持续学习过程中发挥互补作用:统计特征回放通过近似重构旧任务分布,助力缓解遗忘效应;主模型蒸馏则在预测空间提供一致性约束,提升跨任务学习的稳定性。二者的协同作用使模型在不同分布条件下均表现出较高的稳定性与综合性能,验证了 SRD 设计的有效性。

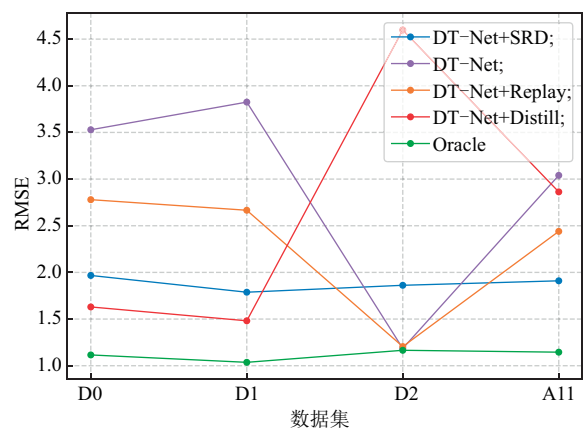


图 6 消融实验对比图

2.3 对比实验分析

消融实验表明,持续学习的核心在于平衡跨阶段知识保持与新任务适应能力。基于此,本节在 DT-Net 基础上进一步选取 4 种具有代表性的持续学习方法作为对比基线:EWC^[26]与 MAS^[27]属于正则化方法,通过对关键参数施加约束以缓解遗忘;

DER^[22]结合样本回放与知识蒸馏策略,是经验回放类方法的典型代表;EFC^[23]则是近年来提出的无样本(exemplar-free)方法,通过高效特征整合实现知识保持。上述方法涵盖正则化、回放学习与无样本特征建模3类核心思路,为全面验证SRD的性能优势提供了参考。

实验结果如表2所示,不同持续学习方法在各批次数据上的表现存在显著差异。MAS和DER在D0、D1上的RMSE较小,但在D2上误差明显升高;EWC和EFC则呈现相反趋势,在D0、D1上误差较大,而在D2上误差相对较小。这一趋势反映出各方法在应对分布漂移时的差异化适配机制。EWC与MAS虽同属参数重要性正则化方法,但由于参数重要性的度量方式不同,二者在稳定性与可塑性上呈现不同取向:EWC基于旧任务参数分布的对数似然曲率衡量参数重要性,具有“局部、稀疏”的特点,对低重要性参数保留较高可塑

性。当模型迁移至D2时,这部分可塑参数可快速贴合新域分布,使模型在D2上误差较低,但同时导致旧任务性能退化更明显;MAS则通过输出函数对参数的灵敏度在更广输入分布上累计重要性,约束更“保守”,因此在D0/D1上表现出更高的稳定性,但在D2上适应性不足、误差上升。相比之下,DER属于样本回放类方法,通过保留部分历史样本并与新任务数据联合训练减缓遗忘。该策略虽提升了基础的新任务适应能力,但当跨批次分布差异过大时,旧样本训练产生的梯度方向与新样本冲突,反而会造成模型优化震荡,最终导致D2上误差显著上升。EFC则是特征整合类方法,通过对不同任务的特征子空间进行线性重组实现特征层对齐。该方法能够在新任务上快速重建特征映射,因此在D2上误差最低,但由于缺乏对先前任务的显式约束,模型在D0、D1上的性能明显下降。

表2 持续学习对比实验结果

指标	模型	RMSE				MAE			
		D0	D1	D2	All	D0	D1	D2	All
挥发分/%	EWC	2.811 9	2.906 5	1.145 9	2.460 0	2.444 8	2.720 4	0.865 4	2.047 2
	MAS	1.199 8	1.132 2	3.772 5	2.307 6	0.904 8	0.962 0	3.047 1	1.682 2
	DER	1.723 9	1.003 3	4.524 8	2.779 8	1.228 3	0.844 0	3.804 6	1.897 8
	EFC	3.115 8	3.405 9	1.098 2	2.774 6	2.675 2	3.195 6	0.820 0	2.271 2
	本文	1.967 3	1.787 8	1.861 9	1.909 9	1.462 4	1.504 4	1.393 8	1.454 7

本文提出的SRD策略在各批次数据及整体测试中均表现平稳,原因在于其兼具统计特征回放与蒸馏一致性约束的双重机制。统计特征回放通过旧任务标签的统计量生成伪标签,为模型提供旧分布的结构化先验信息,从而在特征层维持跨任务的分布连续性;知识蒸馏模块通过特征提取器施加一致性约束,使参数更新过程更趋平滑。两者协同作用,使SRD在持续学习的稳定性与新任务可塑性之间实现高效平衡。

为进一步验证表2中各方法的性能差异源自其方法机制,而非数据划分的偶然性,本文通过改变任务的训练顺序,设计了D1-D2-D0与D2-D0-D1两组置换实验。实验结果如表3所示,从整体趋势来看,各方法在3种训练顺序下的性能变化规律基本一致,说明性能差异主要来源于方法机制与数据分布特性,而非任务划分的偶然性。EWC倾向于在最新任务上获得较低误差,但对早期任务的遗忘程度较高;MAS则呈现相反趋势,在早期任务上保持较高精度,而在新任务上的误差明显增大,说明

其参数约束更为保守,对新分布的适应性不足。DER的性能波动最为显著:当D2位于首位时(D2-D0-D1顺序),模型整体表现较好。这是因为DER采用样本回放机制,可直接利用历史样本进行联合训练,从而在早期阶段有效保持旧分布特征,使D2上的性能稳定;当D2处于中间位置时(D1-D2-D0顺序),模型整体训练效果更趋平衡。旧任务样本回放有助于稳定模型参数,而未阶段任务D0与D1分布相近,减少了分布突变的影响,因此3个任务的性能差异较小;当D2位于末位时(D0-D1-D2顺序),模型在前两个任务中累积了大量相似分布样本记忆,回放样本在梯度优化中占据主导,导致模型过度依赖旧分布特征,难以快速适配新任务,最终对强漂移任务D2的适应性显著下降,误差明显增大。EFC虽通过特征线性整合缓解了部分漂移的影响,但旧任务性能下降速度仍较快,未出现明显的性能趋势改变。

本文提出的SRD模型在3种训练顺序下的整体RMSE分别为1.91,1.79和1.78,不仅波动幅度

表 3 置换实验结果

指标	模型	D1-D2-D0				D2-D0-D1			
		RMSE				RMSE			
		D0	D1	D2	All	D0	D1	D2	All
挥发分 /%	EWC	1.294 1	1.377 7	3.476 4	2.224 8	1.373 0	1.179 7	3.667 1	2.253 0
	MAS	1.324 7	0.966 7	3.643 3	2.314 0	1.134 9	1.356 0	2.963 8	1.970 0
	DER	2.061 1	1.827 9	2.142 7	2.015 7	2.115 4	1.839 9	1.395 3	1.915 7
	EFC	1.323 4	1.275 9	4.255 6	2.597 2	1.408 5	0.920 9	3.792 8	2.331 4
	本文	1.297 6	1.454 1	2.428 6	1.789 6	1.343 6	0.998 5	2.449 6	1.778 0

最小,且在各任务上均维持中低误差水平,未出现极端性能退化的情况。这一结果表明,SRD的统计回放与蒸馏双层约束机制,能够在不同任务顺序与分布条件下,实现稳定的知识迁移与参数更新,从而有效增强模型对任务顺序扰动的鲁棒性。

2.4 噪声实验分析

为验证模型在不同采集工况下的鲁棒性,本文进一步在原始光谱数据上开展噪声扰动实验。考虑到近红外光谱仪的主要噪声来源包括探测器电噪与散射干扰,本研究选取加性高斯噪声作为典型模拟方式,以评估模型在信号干扰条件下的稳定性。

表 4 给出了加性噪声下各方法的预测结果。可以看出,噪声扰动会导致各模型整体误差上升,但

性能变化趋势与无噪声场景保持一致。EWC与EFC在噪声条件下性能波动较大,说明其参数约束与特征整合机制对输入扰动较为敏感。相比之下,MAS、DER与本文提出的SRD策略在噪声条件下均表现出较强的稳定性,整体误差涨幅较小,未出现明显性能退化趋势。从机制层面分析,MAS和DER分别通过基于参数重要性的柔性约束与样本回放机制缓解噪声影响;而SRD在特征空间与预测空间同时施加结构化约束,通过统计特征回放维持旧任务分布的稳定性,并以蒸馏一致性约束平滑输出变化。虽然3种方法的总体抗噪能力相当,但SRD的双层约束机制无需保存原始样本,且在多批次任务间表现出更平衡的稳定性与适应性,凸显了其在工程应用中的可实施性优势。

表 4 加性噪声实验结果

指标	模型	加噪 RMSE				初始 RMSE			
		D0	D1	D2	All	D0	D1	D2	All
挥发分 /%	EWC	3.208 3	3.078 0	1.698 7	2.784 1	2.811 9	2.906 5	1.145 9	2.460 0
	MAS	1.395 4	1.508 2	3.594 1	2.422 2	1.199 8	1.132 2	3.772 5	2.307 6
	DER	1.442 3	1.694 1	4.707 3	2.841 5	1.723 9	1.003 3	4.524 8	2.779 8
	EFC	3.841 8	3.631 6	1.556 4	3.234 2	3.115 8	3.405 9	1.098 2	2.774 6
	本文	2.131 5	1.860 1	2.063 5	2.026 4	1.967 3	1.787 8	1.861 9	1.909 9

3 结论

本文在针对煤炭挥发分近红外光谱建模中,因数据分布漂移导致“灾难性遗忘”的问题,提出了持续学习框架DT-Net+SRD。该方法在特征提取阶段融合密集连接与Transformer结构,增强对光谱复杂峰形与长程依赖关系的表征能力;在模型更新阶段引入统计特征回放与知识蒸馏机制,在不依赖完整历史数据的条件下,实现了旧知识保持与新任务适应。多阶段实验表明,该框架能够有效缓解遗忘现象并保持挥发分预测精度;消融实验、基线对比实验及噪声扰动实验进一步验证了统计回放与蒸馏机

制的协同作用——不仅显著提升了跨阶段学习的稳健性,还在噪声扰动下保持较高的预测稳定性与抗噪鲁棒性,使模型整体性能接近Oracle基准,凸显其在复杂工业场景中实现稳定预测的应用潜力。

参考文献:

- [1] LIN Boqiang, SHI Fengyuan. Coal price, economic growth and electricity consumption in China under the background of energy transition [J]. Energy Policy, 2024, 195: 114400. DOI: 10.1016/j.enpol.2024.114400.
- [2] 顾佰和,于东晖,王琛,等. 进一步深化碳达峰、碳中和

- 战略转型路径的若干思考[J]. 中国科学院院刊, 2024, 39(4): 726-736.
- [3] 邹才能, 陈艳鹏, 熊波, 等. 碳中和目标下中国新能源使命[J]. 中国科学院院刊, 2023, 38(1): 48-58.
- [4] ZHANG Yuanbo, ZHANG Yutao, LI Yaqing, et al. Determination of ignition temperature and kinetics and thermodynamics analysis of high-volatile coal based on differential derivative thermogravimetry [J]. *Energy*, 2022, 240: 122493. DOI: 10.1016/j.energy. 2021. 122493.
- [5] 全国煤炭标准化技术委员会. 煤的工业分析方法: GB/T 212—2008[S]. 北京: 中国标准出版社, 2008.
- [6] SONG Weiran, HOU Zongyu, GU Weilun, et al. Industrial at-line analysis of coal properties using laser-induced breakdown spectroscopy combined with machine learning[J]. *Fuel*, 2021, 306: 121667. DOI: 10.1016/j.fuel. 2021. 121667.
- [7] LIU Ke, HE Chao, ZHU Chenwei, et al. A review of laser-induced breakdown spectroscopy for coal analysis [J]. *TrAC Trends in Analytical Chemistry*, 2021, 143: 116357. DOI: 10.1016/j.trac. 2021. 116357.
- [8] TIAN Zhihui, LI Jiaxuan, WANG Shuqing, et al. Development and industrial application of LIBS-XRF coal quality analyzer by combining PCA and PLS regression methods [J]. *Journal of Analytical Atomic Spectrometry*, 2023, 38(7): 1421-1430.
- [9] GE Lichao, LIU Xiaoyan, FENG Hongcui, et al. The interaction between microwave and coal: A discussion on the state-of-the-art [J]. *Fuel*, 2022, 314: 123140. DOI: 10.1016/j.fuel. 2022. 123140.
- [10] TIAN Jun, LI Ming, TAN Zhiyi, et al. Intelligent non-destructive measurement of coal moisture via microwave spectroscopy and chemometrics [J]. *Chemometrics and Intelligent Laboratory Systems*, 2024, 252: 105175. DOI: 10.1016/j.chemolab. 2024. 105175.
- [11] TSUCHIKAWA S, MA Te, INAGAKI T. Application of near-infrared spectroscopy to agriculture and forestry [J]. *Analytical Sciences*, 2022, 38(4): 635-642.
- [12] QIAO Fengkang, LIU Tingting, LI Zhipeng, et al. Exploration of an apple grading model based on near-infrared spectroscopy [J]. *Academic Journal of engineering and technology science*, 2024, 7(3): 84-88.
- [13] JUNAEDI E C, LESTARI K, MUCHTARIDI M. Infrared spectroscopy technique for quantification of compounds in plant-based medicine and supplement [J]. *Journal of Advanced Pharmaceutical Technology and Research*, 2021, 12(1): 1-7.
- [14] TAO Wei, FAROKHZAD O C. Theranostic nanomedicine in the NIR- II window: classification, fabrication, and biomedical applications [J]. *Chemical reviews*, 2022, 122(6): 5405-5407.
- [15] XU Shifan, XU Zhibin, ZHENG Jiannan, et al. Where does the crude oil originate: The role of near-infrared spectroscopy in accurate source detection [J]. *Memetic Computing*, 2024, 16(3): 429-443.
- [16] SANTOS F D, VIANNA S G T, CUNHA P H P, et al. Characterization of crude oils with a portable NIR spectrometer [J]. *Microchemical Journal*, 2022, 181: 107696. DOI: 10.1016/j.microc. 2022. 107696.
- [17] 张淑利, 翟璐璐, 楚琰, 等. 紫外可见近红外光谱结合偏最小二乘法测定煤中挥发分 [J]. *冶金分析*, 2024, 44(6): 18-24.
- [18] JIANG Daiyu, HU Gang, QI Guanqiu, et al. A fully convolutional neural network-based regression approach for effective chemical composition analysis using near-infrared spectroscopy in cloud [J]. *Journal of Artificial Intelligence and Technology*, 2021, 1(1): 74-82.
- [19] ZOU Liang, QIAO Jiahui, YU Xinhui, et al. Intelligent proximate analysis of coal based on near-infrared spectroscopy and multioutput deep learning [J]. *IEEE Transactions on Artificial Intelligence*, 2023, 5(3): 1398-1410.
- [20] WANG Liyuan, ZHANG Xingxing, SU Hang, 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.
- [21] 张东阳, 陆子轩, 刘军民, 等. 深度模型的持续学习综述: 理论、方法和应用 [J]. *电子与信息学报*, 2024, 46(10): 3849-3878.
- [22] BUZZEGA P, BOSCHINI M, PORRELLO A, et al. Dark experience for general continual learning: A strong, simple baseline [J]. *Advances in neural information processing systems*, 2020, 33: 15920-15930.
- [23] MAGISTRI S, TRINCI T, SOUTIF-CORMERAIS A, et al. Elastic feature consolidation for cold start exemplar-free incremental learning [J]. *arXiv preprint arXiv:2402.03917*, 2024.
- [24] 全国煤炭标准化技术委员会. 煤样的制备方法: GB/T 474—2008[S]. 北京: 中国标准出版社, 2008.
- [25] 姚宗亮, 黄荣, 董爱华, 等. 基于多模态融合和自适应剪枝 Transformer 的脑肿瘤图像分割算法 [J]. *宁夏大学学报(自然科学版)*, 2024, 45(1): 16-24.
- [26] KIRKPATRICK J, PASCANU R, RABINOWITZ N, et al. Overcoming catastrophic forgetting in neural networks [J]. *Proceedings of the national academy of sciences*, 2017, 114(13): 3521-3526.
- [27] ALJUNDI R, BABILONI F, ELHOSEINY M, et al. Memory aware synapses: Learning what (not) to forget [C]//*Proceedings of the European conference on computer vision (ECCV)*. 2018: 139-154.

Continual Learning for Volatile Matter Detection in Coal via Near-Infrared Spectroscopy

WU Zhifeng^{1,2}, CHEN Hailin¹, YE Jinyan³, ZHAO Jinqiu¹, ZOU Liang^{2*}

(1. China Certification & Inspection Group Hebei Co., Ltd., Shijiazhuang 050071, China;

2. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China;

3. Zhanjiang Customs Technology Center, Zhanjiang 524022, China)

Abstract: Volatile matter is a key indicator for evaluating coal quality and combustion characteristics, as its content directly affects combustion efficiency, reaction activity, and process adaptability. Near-infrared spectroscopy is a rapid and non-destructive analytical technique that has been widely applied to the quantitative prediction of coal volatile matter. However, in cross-batch applications, systematic spectral drift often occurs. When new data is used to rebuild the model, the predictive capability for previous samples diminishes, exacerbating the issue of catastrophic forgetting during model updates. To address this challenge, this study introduces a continual-learning modeling method for coal volatile-matter prediction. During the feature extraction stage, a network combining dense connections and self-attention mechanisms is developed to effectively capture both local spectral details and global dependencies. In the model updating stage, a continual learning strategy integrating statistical feature replay and knowledge distillation is implemented to achieve knowledge retention and task adaptability without accessing original historical samples. Experiments conducted with three batches of coal samples demonstrate that the proposed method effectively mitigates catastrophic forgetting and maintains stable prediction performance under distributional shifts. Its overall accuracy approaches that of independent single-task models, thereby providing a feasible technical solution for long-term and stable volatile matter detection in complex industrial environments.

Key words: coal quality; near-infrared spectroscopy; continual learning; feature replay; knowledge distillation

(责任编辑 张 娣)

(上接第 417 页)

Two-Dimensional Organic-Inorganic Hybrid Perovskite Optoelectronic Devices: Mechanisms on Stability Enhancement and Efficiency Improvement and Progress in Applications

SHI Zhenyu^{1,2}, ZHU Chongqin^{3,4*}, YANG Kun^{1,2}, MA Wei^{1,2}

(1. School of Materials and New Energy, Ningxia University, Yinchuan 750021, China;

2. Ningxia Key Laboratory of Photovoltaic Materials, Ningxia University, Yinchuan 750021, China;

3. College of Chemistry, Beijing Normal University, Beijing 100875, China;

4. Key Laboratory of Theoretical and Computational Photochemistry, Ministry of Education, Beijing Normal University, Beijing 100875, China)

Abstract: The introduction of organic spacer cations endows two-dimensional organic-inorganic hybrid perovskites with superior moisture resistance and thermal stability compared to their three-dimensional counterparts, along with a simpler preparation process. These advantages have made them a research hotspot in the field of optoelectronic devices. To explore the research context in the field of two-dimensional organic-inorganic hybrid perovskite optoelectronic devices and to address the critical issue of the synergistic enhancement of stability and efficiency, this review systematically summarizes the core research findings and technological advancements in this field in recent years. It elucidates the structural properties of two-dimensional perovskites and their applications and studies in solar cells, light-emitting diodes, and photodetectors. The results indicate that the synergistic application of organic ligand regulation and heterojunction design is the core pathway to achieve simultaneous enhancement of stability and efficiency in two-dimensional perovskite optoelectronic devices, providing a reference for the performance optimization, technological breakthroughs, and commercialization of such devices in the future. Furthermore, the current bottlenecks in two-dimensional perovskite research are identified, and future research directions in this field are proposed.

Key words: two-dimensional perovskites; solar cells; light-emitting diode; photodetector

(责任编辑 高继红)